2012

Geometry of Partial Differential Equations

Paul Bracken
The University of Texas Rio Grande Valley

Geometry of Partial Differential Equations

Paul Bracken

Department of Mathematics
University of Texas
Edinburg, TX 78540 USA
bracken@utpa.edu

Sometimes a place is very hard
to leave,
But its just not one's destination.
Goethe-Faust I

Contents

1 Introduction to Differential Equations

In the nineteenth century, there arose in the wake of Newton's formulation of the three laws of motion and the development of calculus a great deal of work on the investigation of mathematical problems which had their origins in problems of a physical nature. This provided a wide spectrum of problems which would serve to initiate the study of what is now called the subject of partial differential equations [**1,2**]. Many problems with their origins in geometry appeared to be expressible and amenable to study by means of differential equations as well. It was soon realized that there are links between differential equations and geometric concepts which are being investigated to this day. There are many areas of common interest between the area of differential equations and that of differential geometry, such as the theory of surfaces [**3**] eigenvalue problems and variational problems. One of the purposes here is to elucidate this connection. Differential equations, especially partial differential equations, can encapsulate many fundamental laws of nature. They also appear in the analysis of many diverse problems in science and engineering [**4**]. Linear partial differential equations were studied extensively at first. The attempt to describe a vibrating string was one of the first problems of continuum mechanics which produced a partial differential equation. It was required to develop a general theory, as well as various techniques for finding solutions for these kinds of equations. This process has had a profound impact on many other areas of mathematics as well. For example, the theory of Fourier series had its origin in the study of linear partial differential equations. Partial differential equations have been found essential to the development of the theory of surfaces for example. They appear in the course of the analysis of a wide variety of mathematical, nonphysical based problems as well, such as in the area of analysis. The development of mechanics and the calculus of variations [**5,6**] has also had a great impact on the subject as well since the formalism of these areas generally leads directly to differential equations of various types. The current interest in minmal surfaces and generalized Weierstrass representations exhibits a connection between the study of surfaces on the one hand and the solution of systems of partial differential equations on the other.

Although this process has been ongoing for a long time, a more recent development is the extension of the field to include the study of nonlinear partial differential equations [**6-12**]. This has received impetus from many areas, especially from physics but also from further developments in other areas of mathematics. The solutions of nonlinear equations often have many properties of physical interest which has generated great interest in these types of equations. Of course their influence on the area of geometry has already been noted [**13-15**], but group theory also has come to play an enormous role in their study. The rise of the modern computer has permitted the area of numerical solution of these equations to evolve into its own field of endeavor. This is particularly important in the case of nonlinear equations, where a lot of mathematics must be developed even to begin to obtain solutions [**16-18**].

1.1 Overview of Partial Differential Equations.

A partial differential equation **[4,19]** in terms of a function $u(x, y, \cdots)$ is a relationship between u and its partial derivatives u_x, u_y, u_{xx}, $u_{yy}, \cdots$ which can be written as

$$F(x, y, u, u_x, u_y, u_{xx}, \cdots) = 0. \tag{1.1}$$

Here F represents a given function of M variables, x, y are independent variables and $u(x, y, \cdots)$ is called a dependent variable. The order of a partial differential equation is defined to be the highest order derivative appearing in (1.1). The most general first-order partial differential equation can be written as

$$F(x, y, u, u_x, u_y) = 0. \tag{1.2}$$

Similarly, the most general second-order partial differential equation in two independent variables x, y has the form

$$F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}) = 0. \tag{1.3}$$

A partial differential equation is called linear if it is linear in the unknown function and all its derivatives and the coefficients depend only on the independent variables. It is called quasi-linear if it is linear in the highest-order derivative of the unknown function.

It is convenient to write a partial differential equation in an operator form

$$\mathbf{L}u(\mathbf{x}) = f(\mathbf{x}). \tag{1.4}$$

The operator $\mathbf{L}$ in (1.4) is called a linear operator if it satisfies the property

$$\mathbf{L}(au + bv) = a\mathbf{L}u + b\mathbf{L}v, \tag{1.5}$$

for any two functions u, v and any two constants a and b. Equation (1.3) is called linear if $\mathbf{L}$ is a linear operator. If $f(\mathbf{x}) \equiv 0$, (1.4) is called a homogeneous equation, otherwise, nonhomogeneous. An equation which is not linear is called a nonlinear equation.

A classical solution of (1.1) is an ordinary function $u = u(x, y, \cdots)$ defined in some domain D which is continuously differentiable such that all its partial derivatives involved exist and satisfy (1.1) identically.

However, this notion of classical solution can be extended by relaxing the requirement that u be continuously differentiable over D. The solution $u = u(x, y, \cdots)$ is called a weak, or generalized, solution of (1.1) if u or its partial derivatives are discontinuous in some or all points in D. In the case of only two independent variables x, y, the solution $u(x, y)$ of (1.1) is visualized geometrically as a surface, called an integral surface in the (x, y, u) space.

The general solution of a linear homogeneous ordinary differential equation of order n is a linear combination of n linearly independent solutions with n arbitrary constants. In the case of linear homogeneous partial differential equations of the form (1.4) with $f \equiv 0$, the general solution depends on arbitrary functions rather than arbitrary constants. If we represent this infinite set of solutions by $\{u_1(\mathbf{x}), \cdots, u_n(\mathbf{x}), \cdots\}$, then the infinite linear combinations

$$u(\mathbf{x}) = \sum_{n=1}^{\infty} c_n u_n(\mathbf{x}) \tag{1.6}$$

where c_n are arbitrary constants, in general, may not be again a solution because the infinite series may not be convergent. So, for partial differential equations, the superposition principle may not be true in general. As with linear homogeneous ordinary differential equations, the principle of superposition applies to linear homogeneous partial differential equations and $u(\mathbf{x})$ represents a solution, provided that the infinite series (1.6) is convergent and the operator $\mathbf{L}$ can be applied to the series term by term. Often the general solution has to satisfy other supplementary conditions, usually called initial or boundary conditions. Usually, there are infinitely many solutions and only by specifying the initial or boundary conditions can a specific solution of interest be determined.

1.2 Nonlinear Equations-Basic Concepts.

The most general first-order nonlinear partial differential equation in two independent variables x and y has the form (1.2). The most general second-order nonlinear partial differential equation has the form (1.3). The most general first-order and second-order nonlinear equations in more independent variables can be written down in a similar way.

As discussed already, it is possible to write equations like (1.2) and (1.3) in operator form

$$\mathbf{L}_{\mathbf{x}} u(\mathbf{x}) = f(\mathbf{x}), \tag{1.7}$$

where $\mathbf{L}_{\mathbf{x}}$ is a partial differential operator and $f(\mathbf{x})$ is a given function of two or more variables, and we write $\mathbf{x} = (x, y, \cdots)$ just to highlight the independent variables which would appear in the equation. As mentioned, if $\mathbf{L}_{\mathbf{x}}$ is not a linear operator, (1.7) is called a nonlinear partial differential equation, and an inhomogeneous nonlinear equation if $f(\mathbf{x}) \neq 0$, a homogeneous nonlinear equation if $f(\mathbf{x}) = 0$.

In general, the linear superposition principle can be applied to linear partial differential equations if certain convergence requirements are satisfied. This principle is usually used to find a new solution as a linear combination of a given set of solutions. For nonlinear partial differential equations, the linear superposition principle cannot be applied to generate a new solution. There is no general method of finding analytical solutions of nonlinear partial differential equations since solution methods for linear equations usually do not work. Consequently, numerical techniques are often required, and the development of the modern computer has given a great push to the study of nonlinear equations from all points of view. As in the case of linear equations, questions of uniqueness, existence and stability of solutions of nonlinear equations are of fundamental importance. At this point, it is worth listing some important equations which have appeared in various applications thus far **[20]**.

1. The simplest first-order nonlinear wave equation is given by

$$u_t + c(u)u_x = 0, \qquad x \in \mathbb{R}, \quad t > 0, \tag{1.8}$$

where $c(u)$ is a given function of u. The equation describes the propagation of a nonlinear wave.

2. The nonlinear Klein-Gordon equation is

$$u_{tt} - c^2 \nabla^2 u + V'(u) = 0, \tag{1.9}$$

where c is a constant, and $V'(u)$ is a nonlinear function of u.

3. Burger's equation is

$$u_t + uu_x = \nu u_{xx}, \qquad x \in \mathbb{R}, \qquad t > 0, \tag{1.10}$$

where ν is the kinematic viscosity. This is the simplest nonlinear model equation for diffusive waves in fluid dynamics, introduced to describe one-dimensional turbulence.

4. The Korteweg-de Vries equation **[21]**,

$$u_t + \alpha u u_x + \beta u_{xxx} = 0, \qquad x \in \mathbb{R}, \qquad t > 0, \tag{1.11}$$

where α and β are constants, is a simple and useful model for describing the long time evolution of dispersive wave phenomena in which the steepening effect of the nonlinear term is counterbalanced by dispersion.

5. Many physical systems are often characterized by their extremum property of some associated physical quantity that appears as an integral in a given domain, known as a functional. Such a characterization is a variational principle leading to the Euler-Lagrange equation which optimizes the related functional. This subject has been a great source of various types of differential equations through its applications to mechanics and other subjects over time **[4,5]**.

The classical Euler-Lagrange variational problem is to determine the extreme value of the functional

$$I(u) = \int_a^b F(x, u, u')\, dx \tag{1.12}$$

with the boundary conditions $u(a) = \alpha$ and $u(b) = \beta$, where α, β are given numbers and $u(x)$ belongs to the class $C^2([a, b])$ of functions which have continuous derivatives up to second order in $a \leq x \leq b$ and the integrand F has continuous second derivatives with respect to all of its arguments.

It is assumed that $I(u)$ has an extremum at some $u \in C^2([a, b])$. Then we consider the set of all variations $u + \epsilon v$ for I and arbitrary v belonging to $C^2([a,b])$ such that $v(a) = 0 = v(b)$. Consider the variation δI of the functional $I(u)$

$$\delta I = \int_a^b [F(x, u + \epsilon v, u' + \epsilon v') - F(x, u, v)]\, dx.$$

Taylor expanding with respect to ϵ, there results

$$\delta I = \int_a^b \epsilon\, (v\frac{\partial F}{\partial u} + v'\frac{\partial F}{\partial u'})\, dx + O(\epsilon^2). \tag{1.13}$$

Thus, a necessary condition for the functional $I(u)$ to have an extremum for arbitrary ϵ is

$$0 = \delta I = \int_a^b (v\frac{\partial F}{\partial u} + v'\frac{\partial F}{\partial u'})\, dx.$$

Integrating this by parts,

$$0 = \int_a^b v[\frac{\partial F}{\partial u} - \frac{d}{dx}(\frac{\partial F}{\partial u'})]\, dx + [v\frac{\partial F}{\partial u'}]_a^b. \tag{1.14}$$

Since v is arbitrary with $v(a) = 0 = v(b)$, the last term vanishes and so F satisfies

$$\frac{\partial F}{\partial u} - \frac{d}{dx}(\frac{\partial F}{\partial u'}) = 0. \tag{1.15}$$

Using the fact that

$$d(\frac{\partial F}{\partial u'}) = \frac{\partial}{\partial x}(\frac{\partial F}{\partial u'})\, dx + \frac{\partial}{\partial u}(\frac{\partial F}{\partial u'})\, du + \frac{\partial}{\partial u'}(\frac{\partial F}{\partial u'})\, du',$$

equation (1.15) takes the form,

$$F_u - F_{xu'} - u'F_{uu'} - u''F_{u'u'} = 0. \tag{1.16}$$

This is called the Euler-Lagrange equation for the variational problem involving one independent variable. This is a second-order nonlinear ordinary differential equation for u, provided that $F_{u'u'} \neq 0$ and, hence, there are two arbitrary constants involved in the solution.

According to Hamilton's principle in mechanics, the first variation of the time integral of the Lagrangian $L = L(q_i, \dot{q}_i, t)$ of any dynamical system must be stationary

$$\delta I = \delta \int_{t_1}^{t_2} L(q_i, \dot{q}_i, t)\, dt = 0,$$

where $L = T - V$ is the difference between the kinetic energy, T, and the potential energy, V. Consequently, in this case, the Euler-Lagrange equation (1.16) reduces to

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) = 0, \quad i = 1, 2, \cdots, n. \tag{1.17}$$

In classical mechanics, these equations are universally known as the Lagrange equations of motion.

The Hamiltonian function or Hamiltonian H is defined in terms of the generalized coordinates q_i, generalized momenta $p_i = \partial L/\partial \dot{q}_i$, and L by

$$H = \sum_{i=1}^{n} p_i \dot{q}_i - L = \sum_{i=1}^{n} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L(\mathbf{q}, \dot{\mathbf{q}}). \tag{1.18}$$

It readily follows that

$$\frac{dH}{dt} = \frac{d}{dt}[\sum_{i=1}^{n} p_i \dot{q}_i - L] = \sum_{i=1}^{n} \dot{q}_i (\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i}) = 0.$$

Thus, H is a constant, and hence, the Hamiltonian is a constant of the motion.

In general, the Lagrangian $L = L(q_i, \dot{q}_i, t)$ is a function of q_i, $\dot{q}_i$ and t, where $\dot{q}_i$ enters through the kinetic energy as a quadratic term. The differential of H is

$$\begin{aligned} dH &= \sum_{i=1}^{n} p_i d\dot{q}_i + \sum_{i=1}^{n} \dot{q}_i dp_i - \sum_{i=1}^{n} \frac{\partial L}{\partial q_i} dq_i - \sum_{i=1}^{n} \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt. \\ &= \sum_{i=1}^{n} \dot{q}_i \, dp_i - \sum_{i=1}^{n} \frac{\partial L}{\partial q_i} dq_i - \frac{\partial L}{\partial t} dt, \end{aligned}$$

where $p_i = \partial L/\partial \dot{q}_i$ has been substituted. On the other hand, the differential of the Hamiltonian $H = H(p_i, q_i, t)$ is

$$dH = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} dp_i + \sum_{i=1}^{n} \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt.$$

Equating the coefficients of the two identical expressions, there results

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad -\frac{\partial L}{\partial q_i} = \frac{\partial H}{\partial q_i}, \qquad -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}. \tag{1.19}$$

Invoking Lagrange equation (1.17), the first two equations in (1.19) give

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i}. \tag{1.20}$$

These are the Hamilton canonical equations of motion.

1.3 An Historical Overview Leading to the Present and the Objectives of this Review.

One of the features of nonlinear equations which has given impetus to their study is that these equations possess solutions which can be interpreted as describing nonlinear waves, which are also referred to as solitons. The origins of soliton theory go back to the early part of the nineteenth century. In 1834, John Scott Russell made the remarkable observation of a solitary bump-shaped wave moving down a long canal near Edinburgh in Scotland. This wave seemed to display some of the key features that are now associated with what are called solitons. It wasn't until 1965 that this type of phenomena was rediscovered, in particular, by Kruskal and Zabusky. This was in the context of the Fermi-Pasta-Ulam problem, which had been studied earlier in the century, but whose numerical study had become more tractable on account of the development of the high speed computer. In fact, it was this group that coined the term soliton. In 1895, two Dutch mathematicians, Korteweg and de Vries had derived a nonlinear wave equation now carrying their name and presented in the previous section. This equation has been used to model long wave propagation in a rectangular channel and has a traveling wave solution which resembles the solitary canal wave observed by Russel. It may be said this began a theoretical impetus to the study of these types of equations.

A pair of equations equivalent to the KdV equation appeared even earlier in a work by Boussinesq. It was not until the mid-twentieth century that the equation reappeared in work done by such people as Zabusky and Kruskal **[22]** and Gardner and coworkers **[23,24]**, as well in an analysis of the transmission of hydromagnetic waves. Clearly, the mathematical spectrum of this subject is very wide reaching as far as applications are concerned. There continues to be ongoing and continuing interest in nonlinear equations such as these, and the ones mentioned in Section 1.2. There are many other equations which arise in a diversity of physical systems, such as in the study of solids, liquids and gases. Self-localizing nonlinear excitations are fundamental and are intrinsic features of such systems as quasi-one-dimensional conducting polymers. Perring and Skyrme have applied the sine-Gordon equation to an elementary particle model, in particular, by solving the equation numerically to obtain a numerical solution. The results generated from this equation were found not to disperse and moreover, two solitary waves have been observed to retain their original shapes as well as velocities despite undergoing collisions. As will be seen here, this continues to be a very active area for many reasons outside of physical applications. One of the reasons for this is that there exist numerous interconnections with other areas of mathematics, in particular, the areas of group theory and differential geometry. It has been found, for example, that equations which determine manifolds in three and higher dimensional spaces are closely related to the kinds of partial differential equations that will be discussed here **[25]**. This is especially the case with surfaces immersed in three space **[3]**. These equations not only provide a means for producing such surfaces, but also for describing their evolution in terms of a parameter which can be thought of as time.

The intention here is to introduce and provide a survey to a number of important topics which concern differential equations and their overlap with differential geometry, as well as being of great current interest as well. Although a diverse range of subjects are discussed, it is hoped a panorama of the subject will be developed. The sections are self-contained and should be understandable to a reader with a basic knowledge of differential equations and differential geometry. Some sections are of a pedagogic nature, however some contain new results. Thus, the presentation is carried out with an emphasis on applying concepts and techniques from these areas to the study of a variety of subjects related to both linear and nonlinear differential equations. Some ideas which are used quite a bit are also reviewed in various places, such as vector fields and differential forms [**26,27**]. As well, the notation is standard but may vary somewhat in presentation from section to section, depending on the discussion.

Let us give a brief outline of the subjects which are presented. In Chapter 2, a general introduction to equations of first order is presented. The concept of a vector field as a differential operator and of a one form are introduced and their significance with respect to differential equations is illustrated. The third chapter considers flows of vector fields and introduces a more expanded idea of differential forms in general. Chapter 4 gives an introduction to the concept of geometric distributions and the role vector fields play in that area. The Frobenius Theorem is also introduced. This is a very important result which has a variety of consequences. Chapter 5 discusses Pfaffian systems and gives a number of Theorems relevant to the study of differential equations. The sixth chapter discusses how the symmetry group of an ordinary differential equation can be calculated. A lengthy application illustrating the method to the study of the Duffing-van der Pol oscillator is given in detail. Exterior differential systems and Wahlquist-Estabrook prolongations are considered and leads to the idea of a type of integrability. This is a useful technique for establishing a type of integrability often called Frobenius integrability. Moreover, the theory is fully capable of generating Lax pairs, and Bäcklund transformations can often be determined as well. The overlap of the areas of surface theory and integrable nonlinear equations is touched on next with a novel examination of the sine-Gordon equation. This subject has developed into a very active area of research. A number of applications of Pfaffian systems appears in this context as well. The generation of integrable systems and hierarchies of such systems is also of much interest, and so a long section on the development of a hierarchy is included. Finally, in the last chapter, the subject of heat operators on Riemannian manifolds is briefly discussed. The Laplacian on its own can also provide information concerning the underlying manifold on which it resides, and some final results are given in this area to end.

2 Equations of First Order

Suppose that $\mathbf{x} = (x^1, x^2, \cdots, x^n)$ are $n \geq 2$ independent variables with u a single dependent variable. Moreover $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ denotes the partial derivatives of u which are defined as follows, $p_i = \partial u/\partial x^i$. As already remarked in the previous chapter, equations in which the number of independent variables is greater than one are termed partial differential equations. An equation is of first order if the partial derivatives of highest order that occur are of order one **[4]**. A single partial differential equation of first order in one dependent variable can be written in either of the two forms,

$$F(\mathbf{x}, u, \mathbf{p}) = F(x^1, \cdots, x^n, u, p_1, \cdots, p_n) = 0. \tag{2.1}$$

Definition 2.1 A function $u = \phi(\mathbf{x})$ which is defined and continuously differentiable in a neighborhood of the point $\mathbf{x}_0 = (x_0^1, \cdots, x_0^n)$, is said to be a solution, or integral, of the partial differential equation (2.1) if the substitution $u = \phi(\mathbf{x})$, $p_i = \partial\phi(\mathbf{x})/\partial x^i$ converts (2.1) into an identity in a neighborhood of the point $\mathbf{x}_0$.

In the case of two independent variables, we will use the following notation. The independent variables are denoted x, y, the dependent variable by u, and its first derivatives $p = \partial u/\partial x$ and $q = \partial u/\partial y$. Equation (2.1) is then written

$$F(x, y, u, p, q) = 0. \tag{2.2}$$

A solution $u = \phi(x, y)$ of the differential equation (2.2) defines a surface in the three-dimensional space with Cartesian coordinates x, y and u and therefore it is called an integral surface.

2.1 Linear, Quasi-linear and Nonlinear Equations.

The standard form of a linear partial differential equation of first order is given by

$$a^1(\mathbf{x})p_1 + \cdots + a^n(\mathbf{x})p_n + c(\mathbf{x})u = f(\mathbf{x}). \tag{2.3}$$

that is,

$$a^1(\mathbf{x})\frac{\partial u}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial u}{\partial x^n} + c(\mathbf{x})u = f(\mathbf{x}), \tag{2.4}$$

where $a^i(\mathbf{x})$, $c(\mathbf{x})$ and $f(\mathbf{x})$ are given functions of the independent variables. In particular, (2.3) with $c(\mathbf{x}) = 0$ and $f(\mathbf{x}) = 0$ is given by

$$a^1(\mathbf{x})p_1 + \cdots + a^n(\mathbf{x})p_n = 0, \tag{2.5}$$

and is referred to as a homogeneous linear equation owing to the fact that its left-hand side is a linear form in $\mathbf{p}$. Also, there is no term not involving a derivative and that the derivatives p_i occur in the first power only.

The general quasilinear equation of the first order is

$$a^1(\mathbf{x}, u)p_1 + \cdots + a^n(\mathbf{x}, u)p_n = g(\mathbf{x}, u), \tag{2.6}$$

where a^i and g are given functions of both independent variables. Since (2.6) is linear in $\mathbf{p}$, it is sometimes called a non-homogeneous linear equation. However, (2.6) is nonlinear because the unknown function u is introduced into its coefficients. Consequently, (2.6) is also termed quasi-linear instead of linear. Equations (2.1) which differ from (2.3) and (2.6) are termed nonlinear partial differential equations of first order. When several equations of the form (2.1) are given instead of a single one, they furnish a system of partial differential equations of the first order.

2.2 Integration of Linear Equations.

Consider a system of ordinary differential equations of the first order with $(n-1)$ dependent variables

$$\frac{dy^i}{dx} = f^i(x, y^1, y^2, \cdots, y^{n-1}), \qquad i = 1, \cdots, n-1. \tag{2.7}$$

Its general solution has the form

$$y^i(x) = \phi^i(x, C_1, \cdots, C_{n-1}), \qquad i = 1, \cdots, n-1,$$

whence, upon solving with respect to the constants of integration C_i,

$$\psi_i(x, y^1, y^2, \cdots, y^{n-1}) = C_i, \qquad i = 1, \cdots, n-1. \tag{2.8}$$

The relations (2.8) provide the general integral of the system (2.7). The left-hand side of each relation in (2.8) reduces to a constant when y^1, $y^2, \cdots, y^{n-1}$ are replaced by the coordinates $y^1(x)$, $y^2(x), \cdots, y^{n-1}(x)$ of any solution of the system (2.7). Every single relation in (2.8) is known as a first integral of (2.7) for this reason.

(1) Consider as an example the system

$$\frac{dx}{dt} = x, \qquad \frac{dy}{dt} = y.$$

Integration of the pair yields the general solution

$$x = C_1 e^t, \qquad y = C_2 e^t.$$

Solving these with respect to the constants of integration yields two first integrals,

$$xe^{-t} = C_1, \qquad ye^{-t} = C_2.$$

The set of first integrals (2.8) is not the only representation of the general solution. Indeed, any relation $\Psi(\psi_1, \cdots, \psi_{n-1}) = C$ is a first integral, and hence one can replace the functions ψ by any $n-1$ functionally independent functions $\Psi_i(\psi_1, \cdots, \psi_{n-1})$, $i = 1, \cdots, n-1$.

Definition 2.2. Given a system (2.7), its first integral is a solution of the form

$$\psi(x, y^1, y^2, \cdots, y^{n-1}) = C$$

satisfied for any solution $y^i = y^i(x)$, $i = 1, \cdots, n-1$, where the function ψ is not identically constant. The function ψ keeps a constant value along each solution with the constant C depending on the solution.

The system (2.7) can be rewritten in the form,

$$\frac{dx}{1} = \frac{dy^1}{f^1} = \frac{dy^2}{f^2} = \cdots = \frac{dy^{n-1}}{f^{n-1}}.$$

Since the denominators can be multiplied by any function distinct from zero, one can rewrite these equations, using $\mathbf{x} = (x^1, x^2, \cdots, x^n)$ for the variables $x, y^1, \cdots, y^{n-1}$, in the symmetric form

$$\frac{dx^1}{a^1(\mathbf{x})} = \frac{dx^2}{a^2(\mathbf{x})} = \cdots = \frac{dx^n}{a^n(\mathbf{x})}. \tag{2.9}$$

The term symmetric is due to the fact that the form (2.9) of $n-1$ first-order ordinary differential equations does not specify the independent variable, which may now be any of the n variables x^1, $x^2, \cdots, x^n$. A first integral of the system (2.9) is given by Definition 2.2 and is written

$$\psi(\mathbf{x}) = C. \tag{2.10}$$

Definition 2.3 A set of $n-1$ first integrals

$$\psi_k(\mathbf{x}) = C_k, \qquad k = 1, \cdots, n-1, \tag{2.11}$$

is said to be independent if the functions $\psi_k(\mathbf{x})$ are functionally independent, that is, if there is no relation of the form $F(\psi_1, \cdots, \psi_{n-1}) = 0$.

Any set of $n-1$ independent first integrals represents the general solution of the system (2.9). Since the general solution of a system of $n-1$ first order equations depends exactly on $n-1$ arbitrary constants, one arrives at the following.

Theorem 2.1 A system of $n-1$ first-order ordinary differential equations (2.9) has $n-1$ independent first integrals (2.11). Any other first integral (2.10) of the system (2.9) is expressible in terms of (2.11)

$$\psi = F(\psi_1, \cdots, \psi_{n-1}). \tag{2.12}$$

(2) Consider the system

$$\frac{dx}{yz} = \frac{dy}{xz} = \frac{dz}{xy}.$$

This is rewritten as

$$\frac{dx}{yz} = \frac{dy}{xz}, \qquad \frac{dy}{xy} = \frac{dz}{xy},$$

or multiplying through by z and x, respectively,

$$\frac{dx}{y} = \frac{dy}{x}, \qquad \frac{dy}{z} = \frac{dz}{y}$$

Rewriting these in the form $xdx - ydy = 0$ and $ydy - zdz = 0$ and then integrating, the following two integrals are obtained

$$\psi_1 = x^2 - y^2 = C_1, \qquad \psi_2 = y^2 - z^2 = C_2.$$

Alternatively, the system can be written in the form,

$$\frac{dx}{y} = \frac{dy}{x}, \qquad \frac{dx}{z} = \frac{dz}{x}.$$

This form produces the first integrals

$$\psi_1 = x^2 - y^2 = C_1, \qquad \psi_3 = x^2 - z^2 = C_3,$$

and hence one obtains three different first integrals, $\psi_1 = C_1$, $\psi_2 = C_2$ and $\psi_3 = C_3$. However, these are not independent since $\psi_3 = \psi_1 + \psi_2$, in accordance with the above theorem.

2.3 Homogeneous Linear Partial Differential Equations.

Lemma 2.1. A function $\psi(\mathbf{x}) = \psi(x^1, \cdots, x^n)$ provides a first integral (2.10) of the system (2.9) if and only if it solves the partial differential equation

$$a^1(\mathbf{x})\frac{\partial \psi}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial \psi}{\partial x^n} = 0. \tag{2.13}$$

Proof: Let a function $\psi(\mathbf{x})$ provide a first integral. Since ψ equals a constant for every solution $\mathbf{x} = (x^1, \cdots, x^n)$ of the system (2.9), the directional differential $d\psi$, taken along any integral curve of (2.9), vanishes

$$d\psi = \frac{\partial \psi}{\partial x^1}\, dx^1 + \cdots + \frac{\partial \psi}{\partial x^n}\, dx^n = 0. \tag{2.14}$$

whenever the differential $d\mathbf{x} = (dx^1, \cdots, dx^n)$ is proportional to the vector $\mathbf{a} = (a^1, \cdots, a^n)$ so $d\mathbf{x} = \lambda \mathbf{a}$, where λ is the common value of the ratios dx^i/a^i in (2.9). Consequently, substituting $dx^i = \lambda a^i$ in (2.14), one arrives at (2.13). This latter is satisfied at points $\mathbf{x}$ belonging to any integral curve of the system (2.9). According to the existence theorem, integral curves pass through any point. It follows that (2.13) is satisfied identically in the neighborhood of any generic point $\mathbf{x}$.

Conversely, let $\psi(\mathbf{x})$ be a solution of the partial differential equation (2.13). Since the left-hand side of (2.13) is the directional derivative, $\mathbf{a} \cdot \nabla \psi$, of ψ in the direction $\mathbf{a}$, it follows that the directional differential $d\psi$ along any integral curve of the system (2.9) vanishes. Hence, $\psi(\mathbf{x})$ has a constant value along any integral curve, and $\psi(\mathbf{x}) = C$ is a first integral of (2.9).

Lemma 2.2. Consider the partial differential operator of the first-order

$$X = a^1(\mathbf{x})\frac{\partial}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial}{\partial x^n}. \tag{2.15}$$

Let y^i be new independent variables defined by an invertible transformation

$$y^i = \varphi^i(\mathbf{x}), \qquad i = 1, \cdots, n. \tag{2.16}$$

Then the operator can be written in the new variables in the form

$$\bar{X} = X(\varphi^1)\frac{\partial}{\partial y^1} + \cdots + X(\varphi^n)\frac{\partial}{\partial y^n}, \tag{2.17}$$

where,

$$X(\varphi^i) = a^1(\mathbf{x})\frac{\partial \varphi^i}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial \varphi^i}{\partial x^n}.$$

Proof: The chain rule for partial derivatives gives

$$\frac{\partial}{\partial x^i} = \sum_{k=1}^{n} \frac{\partial \varphi^k}{\partial x^i}\frac{\partial}{\partial y^k}.$$

It is easy to check that the substitution of these expressions into the operator (2.15) transforms it into (2.17).

Theorem 2.2 The general solution to the homogeneous linear partial differential equation

$$X(u) = a^1(\mathbf{x})\frac{\partial u}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial u}{\partial x^n} = 0, \tag{2.18}$$

is given by the formula

$$u = F(\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})), \tag{2.19}$$

where F is an arbitrary function of $n-1$ variables and

$$\psi_1(\mathbf{x}) = C_1, \cdots, \psi_{n-1}(\mathbf{x}) = C_{n-1},$$

is a set of $n-1$ independent first integrals of (2.9) associated with (2.18), namely the characteristic system of equation (2.18),

$$\frac{dx^1}{a^1(\mathbf{x})} = \cdots = \frac{dx^n}{a^n(\mathbf{x})}. \tag{2.20}$$

Proof: The function u defined by (2.19) solves (2.18). For $X(\psi_1) = 0, \cdots, X(\psi_{n-1}) = 0$ by Lemma 2.1, and the equation $X(u) = 0$ follows from the relation

$$X(F(\psi_1, \cdots, \psi_{n-1})) = \frac{\partial F}{\partial \psi_1} X(\psi_1) + \cdots + \frac{\partial F}{\partial \psi_{n-1}} X(\psi_{n-1}).$$

Let us verify that any solution of (2.18) has the form (2.19). We introduce new independent variables as follows

$$y^1 = \psi_1(\mathbf{x}), \cdots, y^{n-1} = \psi_{n-1}(\mathbf{x}), \quad y^n = \phi(\mathbf{x}),$$

where $\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})$ are the left-hand sides of $n-1$ independent first integrals of the characteristic system (2.20), and $\phi(\mathbf{x})$ is any function that is functionally independent of the $\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})$. According to Lemma 2.1, $X(\psi_1) = \cdots = X(\psi_{n-1}) = 0$, whereas $X(\phi) \neq 0$. Now Lemma 2.2 can be used to reduce (2.18) to the form

$$X(u) = X(\phi) \frac{\partial u}{\partial y^n} = 0,$$

and so $\partial u / \partial y^n = 0$. Therefore, the general solution is an arbitrary function of $y^1, \cdots, y^{n-1}$, in accordance with (2.19).

2.4 Non-Homogeneous Equations.

The integration of non-homogeneous linear equations (2.4) is to be discussed now which are of the form,

$$a^1(\mathbf{x}) \frac{\partial u}{\partial x^1} + \cdots + a^n(\mathbf{x}) \frac{\partial u}{\partial x^n} = f(\mathbf{x}). \tag{2.21}$$

In terms of the notation (2.15), equation (2.21) can be written

$$X(u) = f(\mathbf{x}). \tag{2.22}$$

Theorem 2.3. Given a particular solution $u = \varphi(\mathbf{x})$ of the non-homogeneous equation $X(u) = f(\mathbf{x})$, its general solution is obtained by adding to $\varphi(\mathbf{x})$ the general solution of the corresponding homogeneous equation $X(u) = 0$. The general solution of equation (2.21) is given by

$$u = \varphi(\mathbf{x}) + F(\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})), \tag{2.23}$$

where $\varphi(\mathbf{x})$ is any particular solution of (2.21), and $\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})$ are the left-hand sides of any set of $n-1$ independent first integrals of the system of ordinary differential equations (2.20), and F is an arbitrary function.

Proof: Let $X(\varphi(\mathbf{x})) = f(\mathbf{x})$. By setting $u = v + \varphi(\mathbf{x})$, one obtains

$$X(u) = X(v) + X(\varphi(\mathbf{x})) = X(v) + f(\mathbf{x}).$$

It follows that $X(u) = f(\mathbf{x})$ if and only if v satisfies homogeneous equation (2.18), $X(v) = 0$. Hence, after v has been replaced in $u = \varphi(\mathbf{x}) + v$ by (2.19), $v = F(\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x}))$, we arrive at (2.23).

Theorem 2.3 reduces the problem of integration of a non-homogeneous linear partial differential equation (2.21) to that of the associated system of ordinary differential equations (2.20), provided that a particular solution $\varphi(\mathbf{x})$ of (2.21) is known.

(1) Solve an equation of the form (2.21) in the case in which one of the a^i and f depend on a single variable which can be taken as x^1,

$$a^1(x^1)\frac{\partial u}{\partial x^1} + a^2(x^1, \cdots, x^n)\frac{\partial u}{\partial x^2} + \cdots + a^n(x^1, \cdots, x^n)\frac{\partial u}{\partial x^n} = f(x^1). \quad (2.24)$$

A particular solution is readily obtained by letting $u = \varphi(x^1)$. In fact, (2.24) yields the ordinary differential equation

$$a^1(x^1)\frac{d\varphi}{dx^1} = f(x^1).$$

The solution is then obtained by quadrature,

$$\varphi(x^1) = \int \frac{f(x^1)}{a^1(x^1)}\, dx^1.$$

The general solution is provided by using (2.23).

(2) The equation in the independent variables x and y

$$x^2\frac{\partial u}{\partial x} + xy\frac{\partial u}{\partial y} = 1,$$

has the form (2.24) such that $a^1 = x^2$ and $f = 1$. Consequently, a particular solution can be sought which has the form $u = \varphi(x)$. Then the equation in question reduces to the ordinary differential equation $x^2 d\varphi/dx = 1$, whence one easily obtains the particular solution $\varphi = -1/x^2$. The associated system (2.20)

$$\frac{dx}{x^2} = \frac{dy}{xy},$$

has the first integral $\ln\dfrac{y}{x} = \ln C$ or $\psi_1 = \dfrac{y}{x} = C$. The general solution can be written

$$u = \varphi(\mathbf{x}) + F(\psi_1(\mathbf{x})) = -\frac{1}{x^2} + F(\frac{y}{x}).$$

(3) Consider the equation

$$y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = y.$$

Upon dividing by y, this equation takes the form (2.24) with $a^1 = 1$ and $f = 1$. Consequently, assuming $u = \varphi(x)$, one obtains from $d\varphi/dx = 1$ a particular solution $\varphi = x$. Since the general solution of the corresponding homogeneous equation can be found by integrating the system

$$\frac{dx}{a^1(\mathbf{x})} = \frac{dy}{a^2(\mathbf{x})}$$

with $a^1 = y$ and $a^2 = -x$. The solution is given by $\psi_1 = x^2 + y^2 = C_1$, the first integral. Thus, $v = F(x^2 + y^2)$, and so the general solution of the non-homogeneous equation has the form

$$u = x + F(x^2 + y^2).$$

2.5 Quasi-linear Equations and Laplace's Method.

Laplace integrated the equation with two independent variables given by

$$\alpha(x,y)\frac{\partial u}{\partial x} + \beta(x,y)\frac{\partial u}{\partial y} = g(x,y,u), \tag{2.25}$$

by reducing it to an ordinary differential equation by means of an appropriate change of independent variables. Both coefficients α and β are supposed to be different from zero, since otherwise (2.25) is then an ordinary differential equation. Equation (2.25) is rewritten in the new variables $x' = x$ and $y' = \psi(x,y)$ in the form,

$$\alpha\frac{\partial u}{\partial x'} + (\alpha\frac{\partial \psi}{\partial x} + \beta\frac{\partial \psi}{\partial y})\frac{\partial u}{\partial y'} = g(x,y,u).$$

Letting $\psi(x,y)$ be a nonconstant solution of the homogeneous linear equation associated with (2.25),

$$\alpha(x,y)\frac{\partial \psi}{\partial x} + \beta(x,y)\frac{\partial \psi}{\partial y} = 0, \tag{2.26}$$

then using $x' = x$, one arrives at the ordinary differential equation

$$\alpha(x,y)\frac{\partial u}{\partial x} = g(x,y,u), \tag{2.27}$$

where y should be expressed in terms of x and y' by solving $y' = \psi(x,y)$ with respect to y. Thus, (2.25) is reduced to the ordinary differential equation of the first order (2.27). One can use an alternative change of variables, $x' = \psi(x,y)$ and $y' = y$ to obtain, instead of (2.27), the following equation,

$$\beta(x,y)\frac{\partial u}{\partial y} = g(x,y,u), \tag{2.28}$$

where x should be expressed in terms of x' and y from $x' = \psi(x,y)$.

Example 1. Consider the equation

$$y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = 1. \tag{2.29}$$

To solve the homogeneous equation

$$y\frac{\partial\psi}{\partial x} - x\frac{\partial\psi}{\partial y} = 0 \tag{2.30}$$

the characteristic equation $dx/y = -dy/x$ is used, that is, $xdx+ydy = 0$. Integration yields the first integral $x^2 + y^2 = C$. Hence $\psi = x^2 + y^2$ solves (2.30). Consequently, the change of variables $x' = x^2 + y^2$ and $y' = y$ transforms (2.29) into (2.28),

$$x\frac{\partial u}{\partial y} = -1,$$

or upon substitution of $x = \sqrt{x' - y^2}$,

$$\frac{\partial u}{\partial y} = -\frac{1}{\sqrt{x' - y^2}}.$$

After integration with respect to y, the general solution to this equation comes right out,

$$u = -\arcsin(\frac{y}{\sqrt{x'}} + F(x').$$

Using the elementary formula $\arcsin t = \arctan(t/\sqrt{1-t^2})$ and returning to the original variables x and y, the general solution to (2.29) is obtained

$$u = -\arctan(\frac{y}{x}) + F(x^2 + y^2). \tag{2.31}$$

Laplace's method can be extended to equations in terms of many variables $\mathbf{x} = (x^1, \cdots, x^n)$ of the form

$$a^1(\mathbf{x})\frac{\partial u}{\partial x^1} + \cdots + a^n(\mathbf{x})\frac{\partial u}{\partial x^n} = g(\mathbf{x}, u). \tag{2.32}$$

Invoking the operator X defined in (2.15), (2.32) is written

$$X(u) = g(\mathbf{x}, u).$$

Lemma 2.3. Given an operator X as in (2.15), variables $x^{'i}$ can be found by solving the homogeneous equation $X(u) = 0$, such that X reduces to the one-dimensional form

$$X = a(\mathbf{x}')\frac{\partial}{\partial x'^n}, \tag{2.33}$$

with an arbitrary coefficient $a(\mathbf{x}^{'})$.

Proof: It follows that the required variables can be determined by the equations

$$x^{'1} = \psi_1(\mathbf{x}), \cdots, x^{'n-1} = \psi_{n-1}(\mathbf{x}), \quad x^{'n} = \phi(\mathbf{x}), \tag{2.34}$$

where $\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})$ are any $n-1$ functionally independent solutions of the homogeneous equation, $X(\psi_1) = \cdots = X(\psi_{n-1}) = 0$, and $\phi(\mathbf{x})$ is functionally independent of $\psi_1(\mathbf{x}), \cdots, \psi_{n-1}(\mathbf{x})$. In fact, in terms of these variables, the operator X is written

$$X = X(\phi)\frac{\partial}{\partial x'^n}.$$

It has the form (2.33) with $a(\mathbf{x}') = X(\phi(\mathbf{x}))$, where on the right side, $\mathbf{x}$ has to be expressed in terms of the variables $\mathbf{x}'$ by inverting (2.34).

Definition 2.4. Variables (2.34) in which X has the one-dimensional form (2.33), are termed semi-canonical variables for the operator X.

Theorem 2.4. Let $a^n(\mathbf{x}) \neq 0$. Then there exist semi-canonical variables (2.34), such that (2.32) is written as the ordinary differential equation of the first order

$$a^n(\mathbf{x})\frac{\partial u}{\partial x'^n} = g(\mathbf{x}, u). \tag{2.35}$$

Proof: It suffices to choose the following semi-canonical variables (2.34),

$$x^{'1} = \psi_1(\mathbf{x}), \cdots, x^{'n-1} = \psi_{n-1}(\mathbf{x}), \quad x^{'n} = x^n. \tag{2.36}$$

After we have expressed the variables $\mathbf{x}$ in (2.35) in terms of the new variables by inverting (2.36), equation (2.35) becomes an ordinary differential equation with an independent variable $x^{'n}$ where the remaining variables $x^{'1}, \cdots, x^{'n-1}$ are regarded as parameters. Consequently, the constant of integration C will depend on these parameters, $C = F(x^{'1}, \cdots, x^{'n-1})$.

Example 2. Consider the non-homogeneous linear equation,

$$x^1\frac{\partial u}{\partial x^1} + \cdots + x^n\frac{\partial u}{\partial x^n} = 1.$$

Assuming $x^n \neq 0$ and integrating the homogeneous equation, one readily obtains the solution $\psi_1 = x^1/x^n$, $\psi_2 = x^2/x^n, \cdots, \psi_{n-1} = x^{n-1}/x^n$ and hence the semi-canonical variables (2.36)

$$x^{'1} = \frac{x^1}{x^n}, \cdots, x^{'n-1} = \frac{x^{n-1}}{x^n}, \quad x^{'n} = x^n.$$

The ordinary differential equation (2.35) is written

$$x^{'n} = \frac{\partial u}{\partial x'^n} = 1,$$

and has the general solution $u = \ln|x^{'n}| + F(x^{'1}, \cdots, x^{'n-1})$. Therefore, returning to the original variables

$$u = \ln|x^n| + F(\frac{x^1}{x^n}, \cdots, \frac{x^{n-1}}{x^n}).$$

2.6 Reduction to a Homogeneous Linear Equation.

The general quasi-linear equation (2.6) with n independent variables

$$a^1(\mathbf{x},u)\frac{\partial u}{\partial x^1}+\cdots+a^n(\mathbf{x},u)\frac{\partial u}{\partial x^n}=g(\mathbf{x},u), \tag{2.37}$$

in particular, an arbitrary non-homogeneous linear equation (2.4) can be reduced to a homogeneous linear equation with $n+1$ variables as follows.

Define u as a function of $\mathbf{x}=(x^1,\cdots,x^n)$ implicitly by

$$V(x^1,\cdots,x^n,u)=0. \tag{2.38}$$

Now treat V as an unknown function of $n+1$ variables $x^1,\cdots,x^n$ and u. Define

$$D_i=\frac{\partial}{\partial x^i}+p_i\frac{\partial}{\partial u}, \tag{2.39}$$

to be the operator of total differentiation with respect to x^i. It follows from (2.38), upon total differentiation that the following equation results

$$D_iV=\frac{\partial V}{\partial x^i}+p_i\frac{\partial V}{\partial u}=0,$$

whence, solving for p_i,

$$p_i=-\frac{\dfrac{\partial V}{\partial x^i}}{\dfrac{\partial V}{\partial u}},\quad i=1,\cdots,n. \tag{2.40}$$

Replacing p_i in (2.6) by expressions (2.40), one obtains the homogeneous equation

$$X(V)=a^1(\mathbf{x},u)\frac{\partial V}{\partial x^1}+\cdots+a^n(\mathbf{x},u)\frac{\partial V}{\partial x^n}+g(\mathbf{x},u)\frac{\partial V}{\partial u}=0 \tag{2.41}$$

for an unknown function V of $n+1$ variables $x^1,\cdots,x^n$ and u. Applying the theorem for solving homogeneous linear equations, one arrives at the following effective method for solving quasi-linear equations.

Theorem 2.5. Let

$$\psi_1(\mathbf{x},u)=c_1,\cdots,\psi_n(\mathbf{x},u)=c_n$$

be a set of n independent first integrals of the system of equations called the characteristic system for the quasi-linear equation (2.37),

$$\frac{dx^1}{a^1(\mathbf{x},u)}=\frac{dx^2}{a^2(\mathbf{x},u)}=\cdots=\frac{dx^n}{a^n(\mathbf{x},u)}=\frac{du}{g(\mathbf{x},u)}. \tag{2.42}$$

Then the general solution to (2.41) is given by

$$V(\mathbf{x},u) = F(\psi_1(\mathbf{x},u),\cdots,\psi_n(\mathbf{x},u)), \tag{2.43}$$

where F is an arbitrary function of n variables. Consequently, the solution of the quasi-linear equation (2.37) is defined implicitly by (2.38), $V(\mathbf{x},u) = 0$. Provided that $\partial V/\partial u \neq 0$, the solution can be written explicitly $u = \phi(\mathbf{x})$.

Example 3. Apply the method above to (2.29),

$$y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = 1.$$

Here $g(x,y,u) = 1$ and hence the characteristic system (2.42) is written

$$\frac{dx}{y} = -\frac{dy}{x} = \frac{du}{1}.$$

Two independent first integrals of this system must be found. The first equation $xdx + ydy = 0$ gives $x^2 + y^2 = a^2$, where a is constant. By means of this relation, the second equation is rewritten

$$du + \frac{dy}{\sqrt{a^2 - y^2}} = 0,$$

whence, upon integration, $u + \arcsin(y/a) = C$, or $u + \arctan(y/x) = C$. Hence, the two independent first integrals have the form

$$\psi_1 = x^2 + y^2 = C_1, \qquad \psi_2 = u + \arctan(\frac{y}{x}) = C_2.$$

Therefore, the general solution of the corresponding equation (2.41),

$$y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y} + \frac{\partial V}{\partial u} = 0,$$

is given by (2.43),

$$V = F(x^2 + y^2, u + \arctan(\frac{y}{x})).$$

Hence (2.38) is written,

$$F(x^2 + y^2, u + \arctan(\frac{y}{x})) = 0.$$

Under the assumption $\partial V/\partial u \neq 0$, solution (2.31) is obtained by solving the latter equation with respect to u.

Example 4. Consider the quasi-linear equation

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0.$$

The homogeneous linear equation (2.41) for the function $V(t,x,u)$ is

$$X(V) = \frac{\partial V}{\partial t} + u\frac{\partial V}{\partial x} = 0.$$

The characteristic system (2.42) can be written formally as

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

where the last term simply means that this system has the first integral $u = C_1$. By virtue of this first integral, the characteristic system reduces to $dx - C_1 dt = 0$, whence $x - C_1 t = C_2$. Thus there are two first integrals, $u = C_1$ and $x - tu = C_2$. Thus $V = F(u, x - tu)$, and the solution of the equation in question is given implicitly by (2.38),

$$F(u, x - tu) = 0, \qquad u = f(x - tu).$$

2.7 Integral Surfaces of Loci of Characteristic Curves.

Characteristics play a central role in the whole theory of differential equations. The general notion of characteristic curves is amenable to a geometric description in the case of quasi-linear equation (2.6) with two independent variables

$$a^1(x,y,u)p + a^2(x,y,u)q = g(x,y,u), \tag{2.44}$$

where $p = \partial u/\partial x$ and $q = \partial u/\partial y$. Recall some elementary facts from the geometry of surfaces. Consider a surface given in the form $u = \phi(x,y)$ and a point $P(x,y,u)$ on the surface. Let $p = \partial\phi/\partial x$ and $q = \partial\phi/\partial y$. The equation of the tangent plane to the surface at P is given by

$$U - u = p(X - x) + q(Y - y), \tag{2.45}$$

where (X,Y,U) are points on the tangent plane. The equation of a straight line in the direction of a given vector (v^1, v^2, v^3) is

$$\frac{X - x}{v^1} = \frac{Y - y}{v^2} = \frac{U - u}{v^3}, \tag{2.46}$$

where (X,Y,U) are points on the straight line. If the line is perpendicular to the tangent plane to the surface at P, it is called a *normal line* at P. The equation of the normal line to the surface is

$$\frac{X - x}{p} = \frac{Y - y}{q} = \frac{U - u}{-1}. \tag{2.47}$$

Consequently, the line represented by (2.46) lies in the tangent plane (2.45) if the vectors (v^1, v^2, v^3) and $(p, q, -1)$ are orthogonal, so that $v^1 p + v^2 q - v^3 = 0$. Let $u = \phi(x,y)$ be any integral surface of the partial differential equation (2.44). In the above geometric language, (2.44) means that the tangent plane

(2.45) to the integral surface contains the straight line passing through the point of contact $P(x, y, u)$ in the direction of the vector (a^1, a^2, g),

$$\frac{X - x}{a^1(x, y, z)} = \frac{Y - y}{a^2(x, y, z)} = \frac{U - u}{g(x, y, z)}. \tag{2.48}$$

Definition 2.5. Curves which are tangent at each of their points $P(x, y, u)$ to the straight line (2.48) are called *characteristic curves*, or simply characteristics, of the quasi-linear equation (2.44).

Since the tangent vector to a curve is directed along the vector (dx, dy, du), the characteristic curves are determined by the system

$$\frac{dx}{a^1(x, y, u)} = \frac{dy}{a^2(x, y, u)} = \frac{du}{g(x, y, u)}, \tag{2.49}$$

obtained from (2.48) by merely letting (X, Y, U) tend to (x, y, u).

Theorem 2.6. Every integral surface of the quasi-linear equation (2.44) is generated by a one-parameter family of characteristic curves.

Proof: Consider two independent first integrals of system (2.49)

$$\psi_1(x, y, u) = a, \qquad \psi_2(x, y, u) = b. \tag{2.50}$$

Equations (2.50) define a two-parameter family of characteristic curves. Let us single out a one-parameter family by subjecting the parameters a and b to any relation $F(a, b) = 0$. Eliminating a and b from the latter relation and from (2.50), one obtains a surface given by

$$F(\psi_1(x, y, u), \psi_2(x, y, u)) = 0.$$

This surface, generated by the one-parameter family of characteristics, has the form (2.38), (2.43) and hence it is an integral surface for (2.44). Since F is an arbitrary function, any integral surface can be obtained by this construction.

The characteristics of the general quasi-linear equation (2.37) are determined by the characteristic system (2.42). Thus a function $u = \phi(x^1, \cdots, x^n)$ is a solution of equation (2.37) if and only if it is formed by a family of characteristics depending on $n - 1$ parameters.

In the particular case of (2.44), where $g = 0$ and the coefficients a^1 and a^2 do not depend on u, in the case of the homogenous linear equation

$$a^1(x, y)\frac{\partial u}{\partial x} + a^2(x, y)\frac{\partial u}{\partial y} = 0, \tag{2.51}$$

the characteristics are given by (2.50) of the form $\Psi(x, y) = a$, $u = b$, where the first relation provides a first integral of the equation

$$\frac{dx}{a^1(x, y)} = \frac{dy}{a^2(x, y)}. \tag{2.52}$$

It follows that the characteristic curves are merely cuts of cylinders protracted along the u-axis with directrices $\psi(x, y) = a$ by planes u equal to a constant, parallel to the xy-plane.

For example, for the following equation

$$y\frac{\partial u}{\partial x} - x\frac{\partial u}{\partial y} = 0,$$

the characteristics are concentric circles $x^2 + y^2 = C$.

2.8 Nonlinear Equations.

Consider the general equation (2.2) of first order

$$F(x, y, u, p, q) = 0, \tag{2.53}$$

with two independent variables. Its solution involves, in general, an arbitrary function. The fundamental result due to Lagrange states however that it suffices to know a solution depending on only two variables. Then all other integrals of (2.53) can be derived from such a solution through differentiation and elimination of parameters.

Definition 2.6. The complete integral of equation (2.53) is a solution depending on two arbitrary constants

$$u = \phi(x, y, a, b). \tag{2.54}$$

This means that relation (2.53) becomes an identity in x, y, a, b whenever u and p, q are replaced by $u = \phi(x, y, a, b)$ and

$$p = \phi_x(x, y, a, b), \qquad q = \phi_y(x, y, a, b), \tag{2.55}$$

respectively, where $\phi_x = \partial\phi/\partial x$ and $\phi_y = \partial\phi/\partial y$. Furthermore, it is assumed that elimination of the parameters a and b from the relations (2.54) and (2.55) leads precisely to (2.53).

Theorem 2.7. Given a complete integral (2.54), let the parameters a and b undergo an arbitrary relation $b = \sigma(a)$. Let

$$u = f_\sigma(x, y) \tag{2.56}$$

be the envelope of the one-parameter family of integral surfaces,

$$u = \phi(x, y, a, \sigma(a)). \tag{2.57}$$

Then (2.56) is an integral surface for equation (2.53). The subscript σ in (2.56) indicates that the solution depends upon choice of function σ.

Proof: The envelope (2.56) of the family of surfaces (2.57) is obtained by eliminating the parameter a from equation (2.57) and the equation

$$[\frac{\partial\phi(x,y,a,b)}{\partial a} + \frac{\partial\phi(x,y,a,b)}{\partial b}\sigma'(a)]_{b=\sigma(a)} = 0. \tag{2.58}$$

By definition, the envelope has the same p and q along the curve of contact with the enveloped surface. Consequently, the envelope (2.56) of integral surfaces (2.57) is also an integral surface of (2.53).

Definition 2.7. The general integral is the set of all particular solutions (2.57) obtained for all possible relations $b = \sigma(a)$ between the two parameters. Hence, the general integral involves indirectly an arbitrary function $\sigma(a)$.

Definition 2.8. The singular integral is the envelope of the family of integral surfaces (2.54) depending on two parameters. It is obtained by eliminating a and b from equation (2.54) and the equations

$$\frac{\partial\phi(x,y,a,b)}{\partial a} = 0, \qquad \frac{\partial\phi(x,y,a,b)}{\partial b} = 0,$$

provided that this elimination is possible.

For equations (2.1) with any number of variables, the name complete integral is given to a solution involving as many parameters as there are independent variables.

Definition 2.9. The complete integral of the equation

$$F(x^1, \cdots, x^n, u, p_1, \cdots, p_n) = 0, \tag{2.59}$$

is a solution

$$u = \phi(x^1, \cdots, x^n, a_1, \cdots, a_n) \tag{2.60}$$

containing n arbitrary parameters a_i and such that elimination of the parameters from equation (2.60) and the equations

$$p_i = \frac{\partial\phi(x^1, \cdots, x^n, a_1, \cdots, a_n)}{\partial x^i}, \quad i = 1, \cdots, n \tag{2.61}$$

leads precisely to the differential equation (2.59).

2.9 Completely Integrable Systems. The Lagrange-Charpit Method.

According to what has just been said, one can find all solutions of equation (2.53) by calculating its complete integral. To solve the latter problem, one can use Lagrange and Charpit's method. The method is based on the following notion of completely integrable systems.

Let u be an unknown function of n variables $\mathbf{x} = (x^1, \cdots, x^n)$, and let $f_1(\mathbf{x}, u), \cdots, f_n(\mathbf{x}, u)$ be given functions of $x^1, \cdots, x^n$ and u. Consider the system of partial differential equations of the first order

$$\frac{\partial u}{\partial x^1} = f_1(\mathbf{x}, u), \cdots, \frac{\partial u}{\partial x^n} = f_n(\mathbf{x}, u), \tag{2.62}$$

and define an associated one-form ω by

$$\omega = f_1(\mathbf{x}, u)\, dx^1 + \cdots + f_n(\mathbf{x}, u)\, dx^n. \tag{2.63}$$

Definition 2.10. The system (2.62) is said to be completely integrable if, for any solution $u = u(\mathbf{x})$ of (2.62), the form ω is exact, specifically,

$$f_1(\mathbf{x}, u)\, dx^1 + \cdots + f_n(\mathbf{x}, u)\, dx^n = du. \tag{2.64}$$

Form ω is exact if and only if it is closed. Hence, the condition of complete integrability is $d\omega = 0 \mod (2.62)$. Invoking the exterior differentiation formula, one arrives at the following theorem.

Theorem 2.8 The system (2.63) is completely integrable if and only if the following $n(n-1)/2$ equations are satisfied identically in $x^1, \cdots, x^n, u$,

$$\frac{\partial f_i}{\partial x^k} + \frac{\partial f_i}{\partial u} f_k = \frac{\partial f_k}{\partial x^i} + \frac{\partial f_k}{\partial u} f_i, \qquad i, k = 1, \cdots, n. \tag{2.65}$$

The canonical idea of Lagrange and Charpit's method is to find an auxiliary differential equation,

$$\Phi(x, y, u, p, q) = a, \tag{2.66}$$

a equal to a constant, such that equations (2.53) and (2.66) can be solved in the form $p = f_1(x, y, u, a)$, $q = f_2(x, y, u, a)$, to provide a completely integrable system (2.62),

$$\frac{\partial u}{\partial x} = f_1(x, y, u, a), \qquad \frac{\partial u}{\partial y} = f_2(x, y, u, a). \tag{2.67}$$

The general solution of this system is obtained by integration of ordinary differential equations, and contains an arbitrary constant of integration, b. Thus upon solving the system (2.67), one obtains a complete integral $u = \Phi(x, y, a, b)$ of equation (2.53).

Construction of the auxiliary equation (2.66) requires the following calculations. The values of the partial derivatives $\partial p/\partial y$, $\partial p/\partial u$, $\partial q/\partial z$ and $\partial q/\partial u$, are obtained from (2.53) and (2.66) by differentiation with respect to x and y and elimination. The test for complete integrability is obtained by substituting these values in the integrability condition (2.65)

$$\frac{\partial p}{\partial y} + q\frac{\partial p}{\partial u} = \frac{\partial q}{\partial x} + p\frac{\partial q}{\partial u},$$

and can be written explicitly in terms of the functions F and Φ as the following linear partial differential equation in five independent variables x, y, u, p, q

$$P\frac{\partial \Phi}{\partial x} + Q\frac{\partial \Phi}{\partial y} + (pP + qQ)\frac{\partial \Phi}{\partial u} - (X + pU)\frac{\partial \Phi}{\partial p} - (Y + qU)\frac{\partial \Phi}{\partial q} = 0. \tag{2.68}$$

Here, the functions X, Y, U, P and Q are defined by

$$X = \frac{\partial F}{\partial x}, \quad Y = \frac{\partial F}{\partial y}, \quad U = \frac{\partial F}{\partial u}, \quad P = \frac{\partial F}{\partial p}, \quad Q = \frac{\partial F}{\partial q}.$$

To integrate (2.68), one needs first integrals of the characteristic system,

$$\frac{dx}{P} = \frac{dy}{Q} = \frac{du}{pP + qQ} = -\frac{dp}{X + pU} = -\frac{dq}{Y + qU}. \tag{2.69}$$

However, the method requires a knowledge of one first integral (2.66) only.

Find a complete integral and investigate the singular and general integrals of the nonlinear equation,

$$\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} - u = 0. \tag{2.70}$$

by applying the Lagrange-Charpit method. Here $F = pq + xp + yq - u$, so using the definitions,

$$X = p, \quad Y = q, \quad U = -1, \quad P = x + q, \quad Q = y + p.$$

From these, it follows that $X + pU = Y + qU = 0$, and hence (2.68) is written,

$$(x+q)\frac{\partial \Phi}{\partial x} + (y+p)\frac{\partial \Phi}{\partial y} + (xp + yq + 2pq)\frac{\partial \Phi}{\partial u} = 0.$$

One of its simple solutions is, for example, $\Phi = p$. Consequently, we take the first integral (2.66) in the form $p = a$. The equations

$$p = a, \qquad u + xp + yq + pq$$

yield $p = a$, $q = (u - ax)/(y + a)$. Hence, the completely integrable system (2.67) has the form,

$$\frac{\partial u}{\partial x} = a, \qquad \frac{\partial u}{\partial y} = \frac{u - ax}{y + a}.$$

Integrating the first equation and substituting the resulting formula $u = ax + v(y)$ into the second one, it is found that $v'(y) = v/(y + a)$. Integration yields $v = b(y + a)$. Thus, the following complete integral for (2.70) is found

$$u = ax + by + ab. \tag{2.71}$$

This complete integral comprises a two-parameter family of planes. These planes envelop the singular integral obtained by eliminating a and b from the relations,

$$u = ax + by + ab, \quad x + b = 0, \quad y + a = 0.$$

Hence, the singular integral is the hyperboloid,

$$u = -xy.$$

The general integral is the envelope of the one parameter family of planes obtained by setting $b = \sigma(a)$ in the complete integral (2.71). This envelope is represented parametrically by equations (2.57)-(2.58) containing an arbitrary function $\sigma(a)$

$$u = ax + (y + a)\sigma(a), \quad x + \sigma(a) + (y + a)\sigma'(a) = 0. \tag{2.72}$$

It is clearly tangent to the hyperboloid $u = -xy$ representing the singular integral.

Particular solutions can be obtained by specifying $\sigma(a)$ in (2.72). For example, let us take $\sigma(a) = 1/a$. Then equations (2.72) are written

$$u = 1 + ax + \frac{y}{a}, \qquad x - \frac{y}{a^2} = 0.$$

Expressing the parameter a from the second equation and substituting its value $a = \sqrt{y/x}$ into the first equation, a particular solution of (2.70) is obtained,

$$u = 12\sqrt{xy}.$$

2.10 Solution of Cauchy's Problem via Complete Integrals.

The Cauchy problem for equation (2.53), namely $F(x, y, u, p, q) = 0$, is that of determining an integral surface passing through a given curve γ. Cauchy's problem has, in general, a unique solution just as in the case of ordinary differential equations.

Let an initial curve γ be given parametrically by

$$x = x_0(s), \qquad y = y_0(s), \qquad u = u_0(s),$$

and let a complete integral (2.54), $u = \phi(x, y, a, b)$, of system (2.53) be known. Introduce the function

$$W(s, a, b) = u_0(s) - \phi(x_0(s), y_0(s), a, b) \tag{2.73}$$

obtained by considering the complete integral on the initial curve γ. Elimination of the parameter s from the equations

$$W(s, a, b) = 0, \qquad \frac{\partial W(s, a, b)}{\partial s} = 0 \tag{2.74}$$

provides a relation $b = \sigma(a)$, a one-parameter family of integral surfaces (2.57). The envelope of this family passes through the curve γ by construction, and it is an integral surface by construction. Hence, it provides the solution of the Cauchy problem in question.

For example, consider again (2.70), $pq + xp + yq - u = 0$, and the aim is to find its integral surface passing through the parabola $x = 0$, $u = y^2$. The complete integral above $u = ax + by + ab$ can be used. The initial curve can be written parametrically as $x = 0$, $y = s$, $u = s^2$, and (2.73) yields $W(s, a, b) = s^2 - bs - ab$. Consequently, (2.74) yield

$$s^2 - bs - ab = 0, \qquad 2s - b = 0,$$

and eliminating s one obtains $b = -4a$. The solution of the Cauchy problem is obtained now by eliminating a from the equations $u = ax - 4ay - 4a^2$, $x - 4y - 8a = 0$ and has the form $u = (x - 4y)^2/16$.

3 Flows, Vector Fields and Differential Forms.

The subject of the previous chapter can be extended much further by introducing the idea of a vector field and a differential form on a smooth manifold [**1,19,28**]. First, let $U \subset \mathbb{R}^n$ be open. A vector field on U is a smooth map

$$X : U \to \mathbb{R}^n. \tag{3.1}$$

Consider the corresponding equation,

$$y' = X(y), \qquad y(0) = x, \tag{3.2}$$

with $x \in U$. A curve which solves (3.2) is called an integral curve of the vector field X. It is also called an orbit. For t fixed, write

$$y = y(t, x) = \mathcal{F}_X^t(x). \tag{3.3}$$

The mapping $\mathcal{F}_X^t$, which is locally defined, maps a subdomain of U to U. It is called the flow generated by the vector field X.

The vector field X defines a differential operator on scalar functions

$$\mathcal{L}_X f(x) = \lim_{h \to 0} \frac{[f(\mathcal{F}_X^h x) - f(x)]}{h} = \frac{d}{dt} f(\mathcal{F}_X^t x)|_{t=0}. \tag{3.4}$$

The following notation

$$\mathcal{L}_X f(x) = Xf \tag{3.5}$$

is also commonly used. Thus, X is applied to f as a first order differential operator.

If we apply the chain rule to (3.4) and use (3.2),

$$\mathcal{L}_X f(x) = X(x) \cdot \nabla f(x) = \sum_j a_j(x) \frac{\partial f}{\partial x_j} \tag{3.6}$$

if $X = \sum_j a_j(x) e_j$, where e_j is the standard basis of $\mathbb{R}^n$. Note that X is a derivation, that is, a map on $C^\infty(U)$, linear over $\mathbb{R}$, which satisfies

$$X(fg) = X(f)g + fX(g). \tag{3.7}$$

Conversely, any derivation on $C^\infty(U)$ defines a vector field, that is, it has the form

$$X = \sum_j a_j(x) \frac{\partial}{\partial x_j}. \tag{3.8}$$

Proposition 3.1. If X is a derivation on $C^\infty(U)$, then X has the form (3.8).

If $F : Y \to W$ is a diffeomorphism between two open domains in $\mathbb{R}^n$, or between two smooth manifolds, and Y is a vector field on W, a vector field $F_* Y$ on V is defined so that

$$\mathcal{F}^t_{F_* Y} = F^{-1} \circ \mathcal{F}^t_Y \circ F. \tag{3.9}$$

If $U \subset \mathbb{R}^n$ is open and X is a vector field on U defining a flow $\mathcal{F}^t$, then for a vector field Y, $\mathcal{F}^t_*$ is defined on most of U, for small $|t|$, the Lie derivative can be defined as

$$\mathcal{L}_X Y = \lim_{h \to 0} \frac{(\mathcal{F}^h_* Y - Y)}{h} = \frac{d}{dt} \mathcal{F}^t_* Y|_{t=0}, \tag{3.10}$$

as a vector field on U. Another natural construction is the operator theoretic bracket

$$[X, Y] = XY - YX, \tag{3.11}$$

where the vector fields X and Y are regarded as first-order differential operators on $C^\infty(U)$. Now (3.11) defines a derivation on $C^\infty(U)$, hence a vector field on U.

Theorem 3.1. If X and Y are smooth vector fields, then

$$\mathcal{L}_X Y = [X, Y]. \tag{3.12}$$

Corollary 3.1. If X and Y are smooth vector fields on U, then

$$\frac{d}{dt} \mathcal{F}^t_{X\,*} Y = \mathcal{F}^t_{X\,*} [X, Y], \tag{3.13}$$

for all t.

Let $G : U \to V$ be a diffeomorphism. A characterization of $G_* Y$ is given in terms of the flow it generates. One has

$$\mathcal{F}^t_Y \circ G = G \circ \mathcal{F}^t_{G_* Y}. \tag{3.14}$$

The proof of this is a direct consequence of the chain rule. As a special case, there is the following.

Proposition 3.2. If $G_* Y = Y$, then $\mathcal{F}^t_Y \circ G = G \circ \mathcal{F}^t_Y$.

From this statement, the following condition for a pair of flows to commute can be derived. Let X and Y be vector fields on U.

Proposition 3.3. Let X and Y commute as differential operators

$$[X, Y] = 0, \tag{3.15}$$

then locally, $\mathcal{F}^s_X$ and $\mathcal{F}^t_Y$ commute; in other words, for any $p_0 \in U$, there exists a $\delta > 0$ such that for $|s|, |t| < \delta$,

$$\mathcal{F}^s_X \mathcal{F}^t_Y p_0 = \mathcal{F}^t_Y \mathcal{F}^s_X p_0. \tag{3.16}$$

Proof: By Proposition 3.2, it suffices to show that $\mathcal{F}^s_{X*} Y = Y$. This clearly holds at $s = 0$. By (3.13) it follows that

$$\frac{d}{ds} \mathcal{F}^s_{X*} Y = \mathcal{F}^s_{X*} [X, Y],$$

which vanishes if $[X, Y] = 0$ when (3.15) holds.

There is a notion which is complementary to that of a vector field, namely, a differential form. It is especially useful to make constructions that depend as little as possible on a particular choice of coordinate system. The use of differential forms is one mathematical tool for this purpose.

Consider the idea of a one-form; formally a one-form on a set $\Omega \subset \mathbb{R}^n$ is written

$$\alpha = \sum_j a_j(x)\, dx^j. \tag{3.17}$$

It is not hard to modify these definitions to apply to a general manifold. Suppose $F : \mathcal{O} \to \Omega$ is a smooth map, $\mathcal{O} \subset \mathbb{R}^m$, open. The pull-back $F^*\alpha$ is a one-form on $\mathcal{O}$ defined by

$$F^*\alpha = \sum_{j,k} a_j(F(y))\, \frac{\partial F_j}{\partial y_k} dy_k. \tag{3.18}$$

The usual change of variable for integrals gives

$$\int_\gamma \alpha = \int_\sigma F^*\, \alpha,$$

if γ is the curve $F \circ \sigma$.

If $F : \mathcal{O} \to \Omega$ is a diffeomorphism and

$$X = \sum_j b^j(x) \frac{\partial}{\partial x_j}, \tag{3.19}$$

is a vector field on Ω, a pairing between one-forms and vector fields on Ω is defined by

$$\langle X, \alpha \rangle = \sum_j b^j(x) a_j(x), \tag{3.20}$$

and a simple calculation yields,

$$\langle F_* X, F^* \alpha \rangle = \langle X, \alpha \rangle \circ F.$$

Thus, a one-form on Ω is characterized at each point $p \in \Omega$ as a linear transformation of vectors at p to $\mathbb{R}$.

In general, a k-form α on Ω can be regarded as a k-multilinear map on vector fields

$$\alpha(X_1, \cdots, X_k) \in C^\infty(\Omega),$$

with the further condition of antisymmetry,

$$\alpha(X_1, \cdots, X_j, \cdots, X_s, \cdots, X_k) = -\alpha(X_1, \cdots, X_s, \cdots, X_j, \cdots, X_k).$$

If $1 \le j_1 < \cdots < j_k \le n$, $j = (j_1, \cdots, j_k)$, it is customary to set

$$\alpha = \sum_j a_j(x)\, dx_{j_1} \wedge \cdots \wedge dx_{j_k}, \tag{3.21}$$

where

$$a_j(x) = \alpha(\partial_{j_1}, \cdots, \partial_{j_k}), \qquad \partial_j = \frac{\partial}{\partial x_j}. \tag{3.22}$$

In order to express the statement that α is a k-form on Ω, it is customary to write $\alpha \in \Lambda^k(\Omega)$.

If $F : \mathcal{O} \to \Omega$ is a smooth map, the pull-back $F^*\alpha$ of a k-form α specified by (3.21) is defined to be

$$F^*\alpha = \sum_j a_j(F(y))\, (F^* dx_{j_1}) \wedge \cdots \wedge (F^* dx_{j_k}), \tag{3.23}$$

where

$$F^*\, dx_j = \sum_s \frac{\partial F_j}{\partial y_s}\, dy_s.$$

A smooth map $F : \mathcal{O} \to \Omega$ between two open subsets of $\mathbb{R}^n$ preserves orientation if $DF(y)$ is everywhere positive. The object called an orientation on Ω can be identified as an equivalence class of nowhere-vanishing n-forms on Ω, where two such forms are equivalent if one is a multiple of another by a positive function in $C^\infty(\Omega)$.

More generally, if S is an n-dimensional manifold with an orientation, the image of an open set $\mathcal{O} \subset \mathbb{R}^n$ by $\varphi : \mathcal{O} \to S$, carrying the natural orientation of $\mathcal{O}$, we can set,

$$\int_S \alpha = \int_{\mathcal{O}} \varphi^*\, \alpha,$$

for an n-form α on S. If it takes several coordinate patches to cover S, as with many manifolds, define $\int_S \alpha$ by writing α as a sum of forms, each supported on a single patch.

It can be shown that this definition of $\int_S \alpha$ is independent of the choice of coordinate system on S, as long as the orientation of S is respected. An important operator on forms is the exterior derivative,

$$d : \Lambda^k(\Omega) \to \Lambda^{k+1}(\Omega),$$

for $\alpha \in \Lambda^k(\Omega)$ defined by (3.21) such that, with j defined above as $j = (j_1, \cdots, j_k)$,

$$d\alpha = \sum_{j,s} \frac{\partial a_j}{\partial x_s}\, dx_s \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_k}. \tag{3.24}$$

The exterior derivative has the following important property under pull-backs

$$F^*(d\alpha) = dF^*\, \alpha, \tag{3.25}$$

if $\alpha \in \Lambda^k(\Omega)$ and $F : \mathcal{O} \to \Omega$ is a smooth map. From (3.24), it is clear that

$$d(d\alpha) = 0,$$

must hold for any differential form α. If $d\alpha = 0$, then α is said to be closed; if $\alpha = d\beta$ for some $\beta \in \Lambda^{k-1}(\Omega)$, it is said that α is exact.

Let F_t be any smooth family of diffeomorphisms from M to $F_t(M) \subset M$. Define vector fields X_t on $F_t(M)$ by

$$\frac{d}{dt} F_t(x) = X_t(F_t(x)). \tag{3.26}$$

Then it easily follows that, for $\alpha \in \Lambda^k M$,

$$\frac{d}{dt} F_t^* \alpha = F_t^* \mathcal{L}_{X_t} \alpha = F_t^* [d(\alpha \lrcorner X_t) + d\alpha \lrcorner X_t]. \tag{3.27}$$

In particular, if α is closed, then with F_t diffeomorphisms and $0 \le t \le 1$,

$$F_1^* \alpha - F_0^* \alpha = d\beta, \qquad \beta = \int_0^1 F_t^*(\alpha \lrcorner X_t)\, dt. \tag{3.28}$$

Theorem 3.2. (Poincaré Lemma) If B is the unit ball in $\mathbb{R}^n$, centered at 0, $\alpha \in \Lambda^k(B)$, $k > 0$, and $d\alpha = 0$, then $\alpha = d\beta$ for some $\beta \in \Lambda^{k-1}(B)$.

Proof: Consider the family of maps $F_t : B \to B$ given by $F_t(x) = tx$. For $0 < t \le 1$, these are diffeomeorphisms and (3.27) applies. Note that

$$F_1^* \alpha = \alpha, \qquad F_0^* \alpha = 0.$$

A simple limiting argument shows that (3.28) remains valid, so $\alpha = d\beta$ with

$$\beta = \int_0^1 F_t^*(\alpha \lrcorner V) t^{-1}\, dt,$$

where $V = r\frac{\partial}{\partial r} = \sum_j x_j \frac{\partial}{\partial x_j}$. Since $F_0^* = 0$, the apparent singularity in the integral is removable.

Now (3.28) can be generalized to the case in which $F_t : M \to N$ is a smooth family of maps, not necessarily diffeomorphisms. Then (3.26) does not work to define X_t as a vector field, however

$$\frac{d}{dt} F_t(x) = Z(t,x), \qquad Z(t,x) \in T_{F_t(x)} N.$$

Based on (3.28), it can be seen that

$$F^*(\alpha \lrcorner X_t)(Y_1, \cdots, Y_{k-1}) = \alpha(F_t(x))(X_t, DF_t(x)Y_1, \cdots, DF_t(x)Y_{k-1}),$$

and X_t can be replaced by $Z(t,x)$. Hence, in this more general case, if α is closed, there results

$$F_1^* \alpha - F_0^* \alpha = d\beta, \qquad \beta = \int_0^1 \gamma_t\, dt, \tag{3.29}$$

where, at $x \in M$,

$$B(Y_1, \cdots, Y_{k-1}) = \alpha(F_t(x))(Z(t,x), DF_t(x)Y_1, \cdots, DF_t(x)Y_{k-1}).$$

Differential forms are not only a fundamental tool in analysis, but also have many important applications to topology. Here are some basic propositions which illustrate this statement.

Proposition 3.4. If $\bar{M}$ is a compact, oriented manifold with nonempty boundary ∂M, there is no continuous retraction $\varphi : \bar{M} \to \partial M$.

Proof: A retraction φ satisfies $\varphi \circ j(x) = x$, where $j : \partial M \to \bar{M}$ is the natural inclusion. By a simple approximation, if there were a continuous retraction, there would be a smooth one, so suppose φ is smooth. Select $\omega \in \Lambda^{n-1}(\partial M)$ to be the volume form on ∂M, endowed with some Riemannian metric, so $\int_{\partial M} \omega > 0$. Apply Stokes' theorem to $\alpha = \varphi^* \omega$. If φ is a retraction, it is the case that $j^* \varphi^* \omega = \omega$, consequently

$$\int_{\partial M} \omega = \int_M d\varphi^* \omega.$$

However, $d\varphi^* \omega = \varphi^* d\omega = 0$, so the integral is zero. This is a contradiction, so there can be no retraction.

The Brower fixed-point theorem is a consequence of this.

Theorem 3.3. If $F : B \to B$ is a continuous map on the closed unit ball in $\mathbb{R}^n$, then F has a fixed point.

Proof: The claim can be stated as $F(x) = x$ for some $x \in B$. If not, define $\varphi(x)$ to be the endpoint of the ray from $F(x)$ to x, continued until it hits $\partial B = S^{n-1}$. It is clear that φ would be a retraction, which contradicts Proposition 3.4.

An even-dimensional sphere cannot have a smooth nonvanishing vector field.

Proposition 3.5. There is no smooth nonvanishing vector field on S^n if n is even.

Proof: If X were such a vector field, it could be arranged that it have unit length, so $X : S^n \to S^n$, with $X(v) \perp v$ for $v \in S^n \subset \mathbb{R}^{n+1}$. Thus, there is a unique unit-speed geodesic γ_v from v to $X(v)$ of length $\pi/2$. Define a smooth family of maps $F_t : S^n \to S^n$ by $F_t(v) = \gamma_v(t)$. Thus $F_0(v) = v$, $F_{\pi/2}(v) = X(v)$, and $F_\pi = A$ the antipodal map, $A(v) = -v$. From (3.29), it is deduced that $A^*\omega - \omega = d\beta$ is exact, where ω is the volume form on S^n. Hence, by Stokes theorem,

$$\int_{S^n} A^* \omega = \int_{S^n} \omega. \tag{3.30}$$

On the other hand, it is straightforward that $A^* \omega = (-1)^{n+1} \omega$, so the integral in (3.30) is consistent only when n is odd.

The existence of n-forms on a compact, oriented n-dimensional manifold M which are not exact, but are closed, is an important part of the proof of these propositions. The following is an important counterpoint to the Poincaré lemma.

Proposition 3.6. If M is a compact, connected, oriented manifold of dimension n and $\alpha \in \Lambda^n M$, then $\alpha = d\beta$ for some $\beta \in \Lambda^{n-1}(M)$ if and only if

$$\int_M \alpha = 0.$$

Some of these topological results can be extended by using the idea of the degree of a map between compact, oriented surfaces. Let X and Y be compact, oriented, n-dimensional surfaces. To define the degree of a smooth map $F : X \to Y$, assume that Y is connected. Pick $\omega \in \Lambda^n Y$ such that

$$\int_Y \omega = 1, \tag{3.31}$$

and define Deg(F) to be

$$\text{Deg}\,(F) = \int_X F^*\omega. \tag{3.32}$$

The following Lemma states that Deg$\,(F)$ is well defined by (3.32).

Lemma 3.1. The quantity in (3.32) is independent of the choice of ω, as long as (3.31) holds.

Proof: Pick $\omega_1 \in \Lambda^n Y$ which satisfies $\int_\gamma \omega_1 = 1$, so $\int_\gamma \omega - \omega_1 = 0$. By Proposition 3.6, it must be that

$$\omega - \omega_1 = d\alpha,$$

for some $\alpha \in \Lambda^{n-1}$. Thus

$$\int_X F^*\omega - \int_X F^*\omega_1 = \int_X dF^*\alpha = 0,$$

which proves the Lemma.

Proposition 3.7. If F_0 and F_1 are homotopic, then Deg$\,(F_0) =$ Deg(F_1).

Proof: If F_0 and F_1 are homotopic, then $F_0^*\omega - F_1^*\omega$ is exact, so we can write $d\beta$, and it follows that $\int_X d\beta = 0$.

There is another formula for the degree of a map. A point $y_0 \in Y$ is called a regular value of F provided that, for each $x \in X$ satisfying $F(x) = y_0$, $DF(x) : T_xX \to T_{y_0}X$ is an isomorphism. Suppose X and Y are endowed with volume elements ω_X and ω_Y, respectively. If $DF(x)$ is invertible, define $JF(x) \in \mathbb{R}\setminus 0$ by $F^*(\omega_Y) = JF(x)\omega_X$. Clearly, the sign of $JF(x)$ is independent of the choices of ω_X and ω_Y, as long as they determine the given orientations of Y.

Proposition 3.8. If y_0 is a regular value of F, then

$$\text{Deg}\,(F) = \sum \{sgn\, JF(x_j) : F(x_j) = y_0\}. \tag{3.33}$$

Proof: Pick $\omega \in \Lambda^n Y$ which satisfies $\int_\gamma \omega = 1$ with support in a small neighborhood of y_0. Then $F^*\omega$ will be a sum $\sum \omega_j$, such that ω_j is supported in a small neighborhood of x_j and $\int \omega_j = \pm 1$ as $sgn\, JF(x_j) = \pm 1$.

Proposition 3.9. Let $\bar{M}$ be a compact, oriented manifold with boundary. Assume that dim $M = n+1$. Given a smooth map $F : \bar{M} \to Y$, let $f = F|_{\partial M} : \partial M \to Y$. Then,

$$\mathrm{Deg}(f) = 0.$$

Proof: Applying Stokes' Theorem to $\alpha = F^*\omega$, there results

$$\int_{\partial M} F^*\omega = \int_M dF^*\omega.$$

However, $dF^*\omega = F^* d\omega$ and $d\omega = 0$ if dim $Y = n$.

An easy corollary of this is another proof of Brower's no-retraction theorem.

Corollary 3.2. If $\bar{M}$ is a compact, oriented manifold with nonempty boundary ∂M, then there is no smooth retraction $\varphi : \bar{M} \to \partial M$.

Proof: Without loss of generality, it can be assumed that $\bar{M}$ is connected. If there were a retraction, then $\partial M = \varphi(M)$ must also be connected, so the previous proposition applies. However, we would then have, for the map $id = \varphi|_{\partial M}$, the contradiction that its degree is both 0 and 1.

4 Geometric Distributions

4.1 Distributions of Vector Fields.

Starting with ideas introduced in the previous chapter, let us proceed to look at the subject of partial differential equations from the perspective of vector fields [**28,29**]. Consider equations which are linear homogeneous first order partial differential equations. Suppose there is one unknown f, so these have the form,

$$X^i \partial_i f = 0. \tag{4.1}$$

Systems of such equations can also be considered. In this case, an f is to be found which simultaneously satisfies k equations

$$X_1^i \partial_i f = 0, \qquad \cdots, \qquad X_k^i \partial_i f = 0. \tag{4.2}$$

Now we generalize to a search for a function on a manifold. In some sense, these are not partial differential equations at all, since when possible, solutions are obtained by means of ordinary differential equations and flows of vector fields, as already described.

Suppose we write $X_\alpha = X_\alpha^i \partial_i$, $\alpha = 1, \cdots, k$, so we seek functions annihilated by the k vector fields $X_1, \cdots, X_k$,

$$X_\alpha f = 0.$$

If the number of linearly independent $X_\alpha(m)$ varies as a function of m, the problem is referred to as degenerate. Nondegeneracy is assumed, so $X_1(m), \cdots, X_h(m)$ are the maximum number of linearly independent $X_\alpha(m)$ at m. By continuity, $X_1(x), \cdots, X_h(x)$ are linearly independent for all points x in some neighborhood U of m. It follows that $X_\alpha(x) = \sum_{s=1}^h F_\alpha^s(x) X_s(x)$ for each $x \in U$, $\alpha = h+1, \cdots, k$. Thus if $X_s f = 0$, then $X_\alpha f = 0$. Consequently we can always reduce locally to the linearly independent number of equations h.

Now $X_1, \cdots, X_h$ is a local basis of the system in a neighborhood U, and h is called the dimension of the system. If we move to another point p outside U, then $X_1, \cdots, X_h$ may become linearly dependent, but in some neighborhood V of p, some other h of X_α will be a local basis. In the intersection, we may have two or more local bases. In fact, if $Y_\alpha = \sum_{s=1}^h G_\alpha^s X_s$, $\alpha = 1, \cdots, h$, where (G_α^β) is an array of C^∞ functions on U with nonzero determinant at each point, then $Y_\alpha f = 0$ have the same solutions on U as $X_\alpha f = 0$, so Y_α should be considered a local basis. Thus, one way to solve such a system is to choose a local basis Y_α as simple as possible.

For example, if $h = 1$, then $X_1(m) \neq 0$, there are coordinates x^i at m such that $X_1 = \partial_1$, so the equations in terms of the x^i coordinates become $\partial_1 f = 0$. Hence a solution is given by any function not dependent on x^1, that is, a function of $x^2, \cdots, x^d$. For $h = 2$, if $X_\alpha f = 0$ for $\alpha = 1, \cdots, h$, then

$[X_\alpha, X_\beta]f = 0$ for $\alpha, \beta = 1, \cdots, h$. As a consequence, if $h = 2$ and $[X_1, X_2]$ is linearly independent of the local basis X_1, X_2, then the system $X_1 f = 0$, $X_2 f = 0$ for which $h = 2$, does not have more solutions than the system $X_1 f = 0$, $X_2 f = 0$, $[X_1, X_2]f = 0$, for which $h = 3$. Thus, the number of variables on which f depends is determined not only by h, but also by the relation of X_α to each other.

Now concentrate on the subspaces spanned by the $X_\alpha(m)$. The set of all tangents at m is the tangent space at m which can also be written $M_m = T_m(M)$. Thus, we have assigned to every m an h-dimensional subspace, $D(m)$ of M_m. If $X_\alpha f = 0$ for every α, then for every $t \in D(m)$, t is a linear combination of $X_\alpha(m)$ so that $tf = c^\alpha X_\alpha(m) f = (c^\alpha X_\alpha f)m = 0$. Conversely, if $tf = 0$ for every $t \in D(m)$, and every m, then $(X_\alpha f)m = X_\alpha(m) f = 0$, for all α, m, since $X_\alpha(m) \in D(m)$. Hence, the problem of finding a function annihilated by all vectors in $D(m)$ is equivalent to the solution of the system of partial differential equations.

A function D which assigns to each $m \in M$ an h-dimensional subspace $D(m)$ of M_m is called an h-dimensional distribution on M, and is C^∞ if for every $m \in M$, there is a neighborhood U of m and C^∞ vector fields $X_1, \cdots, X_h$ such that for every $p \in U$, $X_1(p), \cdots, X_h(p)$ is a basis of $D(p)$, called a local basis for D at m.

A C^∞ distribution D is *involutive* if for all $X, Y \in D$, we have $[X, Y] \in D$.

Proposition 4.1. A C^∞ distribution D is involutive iff for every local basis $X_1, \cdots, X_h$, the brackets $[X_\alpha, X_\beta]$ are linear combinations of the X_γ, that is, there are C^∞ functions $F^\gamma_{\alpha\beta}$ such that $[X_\alpha, X_\beta] = F^\gamma_{\alpha\beta} X_\gamma$.

Proof: If D is involutive, then $[X_\alpha, X_\beta] \in D$ and hence $[X_\alpha, X_\beta]$ can be expressed as a linear combination of the local basis $X_1, \cdots, X_h$. The coefficients of these linear combinations are clearly C^∞.

If $[X_\alpha, X_\beta] = F^\gamma_{\alpha\beta} X_\gamma$, then for $X, Y \in D$ we may write $X = G^\alpha X_\alpha$, $Y = H^\alpha X_\alpha$, where the G^α and H^α are C^∞ functions. Then

$$[X, Y] = [G^\alpha X_\alpha, H^\beta X_\beta] = G^\alpha (X_\alpha H^\beta) X_\beta - H^\beta (X_\beta G^\alpha) X_\alpha + G^\alpha H^\beta F^\gamma_{\alpha\beta} X_\gamma,$$

which is certainly an element of D.

Consider the following two examples.

(**i**) Let $X = z\partial_y - y\partial_z$, $Y = x\partial_z - z\partial_x$ and $Z = y\partial_x - x\partial_y$, all restricted to $M = \mathbb{R}^3 - \{0\}$. At any $m \in M$, X, Y and Z span a two-dimensional subspace $D(m)$ of M_m. Then D may be described directly by the fact that $D(m)$ is the subspace on M_m normal to the line in E^3 through 0 and m, where E^3 is $\mathbb{R}^3$ with the usual Euclidean metric. Since the brackets are given by $[X, Y] = Z$, $[Y, Z] = X$ and $[Z, X] = Y$, this distribution is involutive.

(**ii**) The distribution on $\mathbb{R}^d$ with local basis $\partial_1, \cdots, \partial_h$ is involutive since $[\partial_\alpha, \partial_\beta] = 0 \in D$.

In fact, another way to state Frobenius' theorem is that locally every involutive distribution has precisely this form: for an involutive distribution, there exist coordinates at each point such that $\partial_1, \cdots, \partial_h$ is a local basis of D.

An integrable submanifold of D is a submanifold N of M such that for every $x \in N$, the tangent space of N at x is contained in $D(x)$, $N_x \subset D(x)$. If $X \in D$ and $X(m) \neq 0$, then the range of an integral curve γ of X is a one-dimensional integral submanifold if γ is defined on an open interval.

An h-dimensional distribution is completely integrable if there is an h-dimensional integral submanifold through each $m \in M$. The one-dimensional C^∞ distributions are completely integrable, since the local basis field will always have integral curves. The distribution of the first example (i) above is completely integrable, since there is a central sphere through each point. Not every two-dimensional C^∞ distribution is integrable since, as an example, the vector fields ∂_x and $\partial_y + x\partial_z$ on $\mathbb{R}^3$ span a two-dimensional distribution. However, $[\partial_x, \partial_y + x\partial_z] = \partial_z$ does not belong to this distribution. The following proposition then says this distribution and also others are not completely integrable.

Proposition 4.2. A completely integrable C^∞ distribution is involutive.

A solution function, or first integral of D is a C^∞ function f such that for every p in the domain of f and every $t \in D(p)$, $tf = 0$, so $D(p)$ annihilates f or df annihilates $D(p)$. If f is a solution function such that $df_p \neq 0$, that is, p is not a critical point of f, then the level hypersurface $f = c$, where $c = f(p)$, is a $(d-1)$-dimensional submanifold M_1 in a neighborhood of p on which $df \neq 0$. The tangent spaces of M_1 are the subspaces of the tangent spaces of M on which $df = 0$, and since $df(D(p)) = 0$ for every $p \in M_1$, $D(p) \subset (M_1)_p$. Thus D also defines an h-dimensional distribution D_1 on M_1.

Proposition 4.3. Let D be a C^∞ h-dimensional distribution. Suppose $f_1, \cdots, f_{d-h}$ are solution functions such that the df_i are linearly independent at some $m \in M$. Then, there are coordinates x^i at m such that $x^{h+i} = f_i$, $i = 1, \cdots, d-h$. For any such coordinates, $\partial_1, \cdots, \partial_h$ is a local basis for D, and the coordinate slices $f_i = c^i$, $i = 1, \cdots, d-h$, are h-dimensional integral submanifolds of D. Finally, if D is restricted to such a coordinate neighborhood, it is involutive.

In the first example above, $f = r$ can be verified to be a first integral, where $r^2 = x^2 + y^2 + z^2$ and $r > 0$. Since $d - h = 1$, any coordinate system of the form x^1, x^2, r gives ∂_1, ∂_2 as a local basis for D. This is true for spherical polar coordinates. The level surfaces $r = c$ are the central spheres of $\mathbb{R}^3$, which are integral submanifolds. Every function of r is a first integral, and conversely, every first integral is a function of r.

In the following two examples, systems of the form (4.2) will be established based on the given vector fields.

(**1**) The subspace of $T(E_3)$ spanned by the vector fields $V_1 = x\partial_y - y\partial_x$ and $V_2 = y\partial_z - z\partial_y$ forms a Lie subalgebra of $T(E_3)$. The equation $V_1 f = 0$ is $xf_y - yf_x = 0$, which can be solved by characteristics, namely,

$$\frac{dx}{-y} = \frac{dy}{x},$$

with solution $x^2 + y^2 = C$, while $V_2 f = 0$ implies that $f = \phi(x, y^2 + z^2)$. To obtain a common integral of the pair, we solve

$$V_2(g_1)\frac{\partial \psi}{\partial g_1} + V_2(g_2)\frac{\partial \psi}{\partial g_2} = 0,$$

where $g_1 = x^2+y^2$ and $g_2 = z$. This equation takes the form $-2yz\psi_{g_1} + y\psi_{g_2} = 0$ so $dg_1 = -2zdz$ hence $g_1 + z^2 = C$. Therefore, the common integrals of $V_1 f = V_2 f = 0$ are given by

$$f = \rho(x^2 + y^2 + z^2),$$

where the function ρ is at least C^1.

(2) Here is a second example of the form (4.2) which is defined by the vector fields

$$X = 9y\partial_x - 4x\partial_y, \qquad Y = x\partial_x + y\partial_y + 2(z+1)\partial_z.$$

Solve first the equation $Xf = 0$ to obtain that

$$2x^2 + \frac{9}{2}y^2 = C.$$

so this yields

$$f = \psi(2x^2 + \frac{9}{2}y^2, z) = \psi(g_1, g_2).$$

It is next required to solve

$$Y(g_1)\frac{\partial \psi}{\partial g_1} + Y(g_2)\frac{\partial \psi}{\partial g_2} = 0.$$

The coefficients in this are given by $Y(g_1) = 4x^2 + 9y^2$ and $Y(g_2) = 2(z+1)$, therefore,

$$(4x^2 + 9y^2)\frac{\partial \psi}{\partial g_1} + 2(z+1)\frac{\partial \psi}{\partial g_2} = 0.$$

To solve this equation, write the characteristic equation,

$$\frac{dg_1}{4x^2 + 9y^2} = \frac{dz}{2(z+1)}.$$

Integrating and solving, there results

$$\frac{g_1}{z+1} = C, \qquad g_1 = 2x^2 + \frac{9}{2}y^2.$$

Therefore, the common integrals of $Xf = Yf = 0$ are given by

$$f = \rho(\frac{2x^2 + \frac{9}{2}y^2}{z+1}),$$

where function ρ is at least C^1.

4.2 The Frobenius Theorem.

For smooth functions $f_1, \cdots, f_{m-k} : \mathbb{R}^m \to \mathbb{R}$ with linearly independent differentials, the equations

$$M^k(c_1, \cdots, c_{m-k}) = \{x \in \mathbb{R}^m : f_1(x) = c_1, \cdots, f_{m-k}(x) = c_{m-k}\} \tag{4.3}$$

define a smooth k-dimensional manifold. This nonlinear system of equations can be linearized by going to the tangent bundle. It is seen that this manifold can be described by the system

$$TM^k(c_1, \cdots, c_{m-k}) = \{\nu \in T\mathbb{R}^m : df_1(\nu) = 0, \cdots, df_{m-k}(\nu) = 0\}. \tag{4.4}$$

These equations determine at each point in $\mathbb{R}^m$ a k-dimensional subspace of the tangent space to $\mathbb{R}^m$. The resulting family of subspaces can also be described by systems of one-forms, $\omega_1, \cdots, \omega_{m-k}$. For example, if (h_{ij}) is a matrix of functions with nowhere vanishing determinant, the one-forms $\omega_i = \Sigma_{j=1}^{m-k} h_{ij}\, df_j$ satisfy

$$TM^k(c_1, \cdots, c_{m-k}) = \{\nu \in T\mathbb{R}^m : \omega_1(\nu) = 0, \cdots, \omega_{m-k}(\nu) = 0\}.$$

The level surfaces cannot however be recovered from the knowledge of the forms ω_i alone. In general, the problem arises under such conditions as linearly independent one-forms $\omega_1, \cdots, \omega_{m-k}$ describe a family of k-dimensional manifolds by means of the system of Pfaffian equations

$$\omega_1 = 0, \cdots, \omega_{m-k} = 0. \tag{4.5}$$

Frobenius' theorem provides a complete answer to this question.

Definition 4.1. A k-dimensional *geometric distribution* on M^m is a family $\mathcal{E}^k = \{E^k(x)\}$ consisting of k-dimensional subspaces $E^k(x) \subset T_xM^m$ in the tangent spaces to M^m depending smoothly on the points in the following sense: For each point $x_0 \in M^m$, there exists a neighborhood $x_0 \in U \subset M^m$ and vector fields $X_1, \cdots, X_k$ defined on U such that $E^k(x)$ coincides with the linear hull of the vectors $X_1(x), \cdots, X_k(x)$ at every point $x \in U$.

Here are two examples which are important.

(1) Every nowhere vanishing vector field X on M^m induces a one-dimensional distribution $E^1(x)$ formed by all multiples of the vector $X(x)$. Conversely, every one-dimensional distribution $\mathcal{E}^1$ is locally determined by a nowhere vanishing vector field.

(2) The linearly independent 1-forms $\omega_1, \cdots, \omega_{m-k}$ on M^m determine a k-dimensional distribution by

$$E^k(x) = \{\nu \in T_xM^m : \omega_1(\nu) = \cdots = \omega_{m-k}(\nu) = 0\}.$$

In analogy to the integral curve for a vector field, we now introduce the notion of an integral manifold for a distribution in the next definition.

Definition 4.2. Let $\mathcal{E}^k$ be a k-dimensional distribution on M^m. A k-dimensional submanifold $N^k \subset M^m$ is called an integral manifold of $\mathcal{E}^k$ if the tangent spaces of N^k coincide with the spaces of the distributions

$$T_x N^k = E^k(x), \qquad x \in N^k.$$

Definition 4.3. For the k-distribution $\mathcal{E}^k$ defined by linearly independent one-forms $\omega_1, \cdots, \omega_{m-k}$ as in example **(2)** above, a submanifold $i : N^k \to M^m$ is an integral manifold of $\mathcal{E}^k$ if and only if the restriction of the forms to N^k vanishes,

$$i^*(\omega_1) = \cdots = i^*(\omega_{m-k}) = 0.$$

The local existence theorem for solutions of ordinary differential equations can be formulated in the following way.

Proposition 4.4. Every one-dimensional geometric distribution is integrable.

(3) Consider the nowhere vanishing one-form $\omega = xdy + dz$ on $\mathbb{R}^3$ together with the two-dimensional distribution determined by ω, namely $E^2 = \{\nu \in T\mathbb{R}^3 : \omega(\nu) = 0\}$. It will be proved that this distribution is not integrable.

Suppose $\mathcal{E}^2$ is integrable. Then it must be that there exists an open set $W \subset \mathbb{R}^2$ as well as a smooth map $f : W \to \mathbb{R}^3$ such that $f^*(\omega) = 0$ and also (rank) $(D(f)) \equiv 2$. For example, choose for f a chart of the integral manifold. In the coordinates of $\mathbb{R}^3$, the map $f = (f_1, f_2, f_3)$ consists of three functions, and the condition $f^*(\omega) = 0$ is expressed on $W \subset \mathbb{R}^2$ as

$$f_1 \, df_2 + df_3 = 0.$$

Differentiating this expression, there results

$$df_1 \wedge df_2 = 0.$$

Moreover, forming the exterior product of the same expression with df_2, we have

$$df_2 \wedge df_3 = 0.$$

Finally, wedge df_1 on $f_1 \, df_2 + df_3 = 0$, whence $f_1 \, df_1 \wedge df_2 + df_1 \wedge df_3 = df_1 \wedge df_3 = 0$. It has been shown that all twofold products vanish $df_i \wedge df_j = 0$. This result contradicts the assumption that the differential $D(f)$ of $f : W \to \mathbb{R}^3$ has the maximal rank two. Therefore, the two-dimensional distribution in $\mathbb{R}^3$ defined by $\omega = x \, dy + dz$ cannot have an integral manifold.

For higher-dimensional distributions on a manifold, the following problem thus arises, which has been addressed in one way already. Under which conditions do these turn out to be integrable? The answer to this question is supplied by the Frobenius theorem and the notion of an involutive distribution is needed.

Definition 4.4. A distribution $\mathcal{E}^k$ on the manifold M^m is called involutive if, for every pair of vector fields V, W on M^m whose values $V(x)$, $W(x) \in$

$E^k(x)$ at each point belong to the distribution, the commutator $[V,W](x) \in E^k(x)$ again has values in $\mathcal{E}^k$.

Theorem 4.1. (Frobenius' Theorem-Second Version) Let $\mathcal{E}^k$ be a k-dimensional distribution on the manifold M^m defined by $(m-k)$ linearly independent one-forms $\omega_1, \cdots, \omega_{m-k}$

$$\mathcal{E}^k = \{\nu \in TM^m : \omega_1(\nu) = \cdots = \omega_{m-k}(\nu) = 0\}.$$

Then the following conditions are equivalent :

(1) $\mathcal{E}^k$ is integrable,

(2) $\mathcal{E}^k$ is involutive,

(3) for every point $x_0 \in M^m$, there exist a neighborhood such that $x_0 \in U \subset M^m$ and one-forms θ_{ij} defined on U such that

$$d\omega_i = \sum_{j=1}^{m-k} \theta_{ij} \wedge \omega_j,$$

for $1 \leq i \leq m-k$.

(4) for all indices $1 \leq i \leq m-k$, the following exterior products vanish

$$d\omega_i \wedge (\omega_1 \wedge \cdots \wedge \omega_{m-k}) = 0.$$

Condition (4) occurring in Frobenius's theorem is called the *integrability condition* for the geometric distribution or the corresponding Pfaffian system.

If condition (3) holds for distribution $\mathcal{E}^k$, then

$$d\omega_i(V,W) = \sum_{j=1}^{m-k} \theta_{ij} \wedge \omega_j(V,W) = 0.$$

This implies that $\omega_i([V,W]) = 0$, hence all one-forms $\omega_1, \cdots, \omega_{m-k}$ vanish on the commutator $[V,W]$. Therefore, this vector field takes values in $\mathcal{E}^k$, which means the distribution $\mathcal{E}^k$ is involutive.

A one-dimensional distribution is determined by $m-1$ one-forms $\omega_1, \cdots, \omega_{m-1}$, and then $d\omega_i \wedge (\omega_1 \wedge \cdots \wedge \omega_{m-1})$ is an $(m+1)$-form on the m-dimensional manifold. This has to be zero for trivial reasons, and the integrability condition of the Frobenius theorem is automatically satisifed.

The implication (3) implies (1), which is at the core of the Frobenius Theorem, is a local statement. It is a direct consequence of Theorem 4.2, which will not be proved here.

Theorem 4.2. Let $\omega_1, \cdots, \omega_{m-k}$ be linearly independent one-forms on an open subset $M^m \subset \mathbb{R}^m$ such that

$$d\omega_i = \sum_{j=1}^{m-k} \theta_{ij} \wedge \omega_j,$$

for certain one-forms θ_{ij}. Then there exists at each point $x \in M^m$ a neighborhood $U \subset M^m$ of x and functions h_{ij} and f_j defined on U satisfying

$$\omega_i = \sum_{i=1}^{m-k} h_{ij} df_j.$$

Proof (3) $\Longleftrightarrow$ (4): Suppose that there exist local one-forms θ_{ij} such that $d\omega_i = \sum_{j=1}^{m-k} \theta_{ij} \wedge \omega_j$. Then

$$d\omega_i \wedge (\omega_1 \wedge \cdots \wedge \omega_{m-k}) = \sum_{j=1}^{m-k} \theta_{ij} \wedge \omega_j \wedge (\omega_1 \wedge \cdots \wedge \omega_{m-k}) = 0,$$

since the exterior square of any one-form vanishes.

Conversely, assume that condition (4) is satisfied. In a neighborhood $U \subset M^m$ of the point x_0, extend the family of linearly independent one-forms $\omega_1, \cdots, \omega_{m-k}$ by adding one-forms $\eta_1, \cdots, \eta_k$ so that the combined family $\{\omega_1, \cdots, \omega_{m-k}, \eta_1, \cdots, \eta_k\}$ forms a basis for $\Lambda_x^1(M^m)$ at each point x of U. The two-form $d\omega_i$ $(1 \leq i \leq m-k)$ can thus be represented as

$$d\omega_i = \sum_{\alpha,\beta=1}^{m-k} C_{\alpha\beta}\, \omega_\alpha \wedge \omega_\beta + \sum_{\alpha=1}^{m-k} \sum_{j=1}^{k} P_{\alpha j}\, \omega_\alpha \wedge \eta_j + \sum_{j,l=1}^{k} Q_{jl}\, \eta_j \wedge \eta_l,$$

where $C_{\alpha\beta}$, $P_{\alpha j}$, Q_{jl}. The condition $d\omega_i \wedge (\omega_1 \wedge \cdots \wedge \omega_{m-k}) = 0$ implies that

$$\sum_{j,l=1}^{k} Q_{jl}\, \eta_j \wedge \eta_l \wedge (\omega_1 \wedge \cdots \omega_{m-k}) = 0.$$

Therefore, all coefficients $Q_{jl} = 0$. If we define the one-forms

$$\theta_{i\alpha} = -\sum_{\beta=1}^{m-k} C_{\alpha\beta} \omega_\beta - \sum_{j=1}^{k} P_{\alpha j}\, \eta_j,$$

the exterior derivative $d\omega_i$ takes the desired form

$$d\omega_i = \sum_{\alpha=1}^{m-k} \theta_{i\alpha} \wedge \omega_\alpha.$$

Proof (3) $\Longrightarrow$ (2): For any two vector fields V and W with values in the distribution $\mathcal{E}^k$, it follows that $\omega_i(V) = \omega_i(W) = 0$. Moreover, from

$$d\omega_i(V, W) = V(\omega_i(W)) - W(\omega_i(V)) - \omega_i([V, W])$$

it follows that

$$d\omega_i(V, W) = -\omega_i([V, W]).$$

Proof (3) $\Longrightarrow$ (1): Let $\mathcal{E}^k$ be a distribution with the property stated in (3). By Theorem 4.2, the forms ω_i can be represented in a neighborhood U of an arbitrary point $x_0 \in M^m$ as $\omega_i = \sum_{j=1}^{m-k} h_{ij} df_j$ for certain functions. By assumption, the one-forms $\omega_1, \cdots, \omega_{m-k}$ are linearly independent. Thus the differentials $df_1, \cdots, df_{m-k}$ are linearly independent as well and $N^k = \{x \in U : f_1(x) = f_1(x_0), \cdots, f_{m-k}(x) = f_{m-k}(x_0)\}$ is a submanifold containing $x_0 \in M^m$, At an arbitrary point $x \in N^k$, we determine the tangent space $T_x N^k = \{v \in TM^m : df_j(\nu) = 0, j = 1, \cdots m-k\} \subset \{\nu \in TM^m : \omega_j(\nu) = 0, j = 1, \cdots, m-k\} = E^k(x)$. For dimensional reasons, the vector spaces coincide, so N^k is an integral manifold of distribution $\mathcal{E}^k$ through $x_0 \in M^m$, and thus the integrability of the distribution $\mathcal{E}^k$ is proved.

4.3 Some Applications of the Frobenius Theorem.

The simplest case is that of an $(m-1)$-dimensional distribution $\mathcal{E}^{m-1}$ on an m-dimensional manifold M^m. If $\mathcal{E}^{m-1}$ is defined by one nowhere vanishing one-form ω, the integrability of the distribution reduces to the condition that the three-form $d\omega \wedge \omega$ vanishes,

$$d\omega \wedge \omega = 0.$$

The method to explicitly integrate this $(m-1)$-dimensional Pfaffian system is based on looking for a so-called integrating factor and an application of Poincaré 's lemma.

Definition 4.5. An integrating factor for the one-form ω is a nowhere vanishing function $f : M^m \to \mathbb{R}$ such that the one-form $f \cdot \omega$ is closed

$$d(f \cdot \omega) = 0.$$

Theorem 4.3. Let ω be a nowhere vanishing one-form on the manifold M^m.

(i) If there exists an integrating factor for ω, then $d\omega \wedge \omega = 0$. In this case, the distribution $\mathcal{E}^{m-1}$ is integrable.

(ii) If $d\omega \wedge \omega = 0$, then there exists an integrating factor for the one-form ω in a neighborhood of each point in M^m.

(iii) Locally, the integral manifolds of the distribution $\mathcal{E}^{m-1}$ are the level surfaces of the function g determined from the integrating factor f by the equations

$$d(f \cdot \omega) = 0, \qquad f \cdot \omega = dg.$$

Proof: The equation $d(f\omega) = 0$ implies that $df \wedge \omega + f \cdot d\omega = 0$. Multiplying this equation once again by the one-form ω leads to $f d\omega \wedge \omega = 0$. Since $f \neq 0$, there results $d\omega \wedge \omega = 0$ as a necessary condition for the existence of an integrating factor. If on the other hand $d\omega \wedge \omega = 0$, then the existence of an integrating factor follows immediately from Theorem 4.2.

In dimension $m = 2$, the three form $d\omega \wedge \omega = 0$ vanishes for purely algebraic reasons. In this case, the following corollary results.

Corollary 4.1. Every nowhere vanishing one-form on a two-dimensional manifold locally has an integrating factor.

Example 4.1. Consider in $\mathbb{R}^2$ the differential equation

$$P(t,x) + Q(t,x)\dot{x} = 0.$$

Near a point $(t_0, x_0) \in \mathbb{R}^2$ at which P and Q do not vanish simultaneously, the following one-form can be written

$$\omega = P\,dt + Q\,dx.$$

Its integrating factor will be $f(t,x)$. The equivalent differential equation

$$(fP)(t,x) + (fQ)(t,x)\dot{x} = 0$$

is called the total differential equation, and the solution curves are implicitly determined by the equation

$$g(t,x) = C,$$

where C is a constant such that $dg = f\omega$. Frobenius's theorem now claims that it is always possible to solve the original differential equation. It does not however provide an algorithm for finding the integrating factor. In simple cases, this may be computed directly. If functions $F(t)$ and $G(t)$ can be found depending only on the variables t and x and satisfying

$$\frac{\partial P(t,x)}{\partial x} - \frac{\partial Q(t,x)}{\partial t} = Q(t,x)F(t) - P(t,x)G(x),$$

then

$$f(t,x) = e^{\int F(t)\,dt} e^{\int G(t)\,dx},$$

is an integrating factor.

Example 4.2. Consider the differential equation

$$(2t^2 + 3tx - 4t)\dot{x} + (3x - 2tx - 3x^2) = 0,$$

so we can set $P = 3x - 2tx - 3x^2$ and $Q = 2t^2 + 3tx - 4t$, then

$$\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial t} = -6t - 9x + 7.$$

and

$$(2t^2 + 3tx - 4t)F(t) - (3x - 2tx - 3x2)G(x)$$
$$= (2t^2 + 3tx - 4t)\frac{2}{t} + (3x - 2tx - 3x^2)\frac{5}{x} = 7 - 9x - 6t.$$

Therefore

$$\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial t} = Q(t,x)F(t) - P(t,x)G(x) = 7 - 9x - 6t,$$

provided that $F(t) = 2/t$ and $G(x) = -5/x$, which implies there is an integrating factor of the form $f = t^2 x^{-5}$. The solutions of this differential equation computed by means of this integrating factor are curves described by the equation

$$t^3 x^{-4} - \frac{1}{2} t^4 x^{-4} - t^3 x^{-3} = C.$$

Assume now that on the manifold M^m, a Riemannian metric g is given, as well as an $(m-1)$-dimensional distribution $\mathcal{E}^{m-1}$ described by the one-form ω. Denote the vector field associated with the one-form by means of the Riemannian metric by W. This field is uniquely determined by either of the two equivalent equations

$$*\omega = W \lrcorner dv(M^m), \qquad \omega(V) = g(V, W), \qquad \omega(W) = g(W, W) = ||W||^2.$$

The assignment $V(f)(x)$ determines a linear functional on the tangent space, hence there exists a vector $\mathrm{grad}(f)(x) \in T_x M^m$ such that

$$V(f)(x) = g(\mathrm{grad}(f)(x), V(x))$$

holds for all vector fields. By using the vector field W, the distribution $\mathcal{E}^{m-1}$ can be described in the following way

$$\mathcal{E}^{m-1} = \{\nu \in TM^m : g(\nu, W) = 0\}.$$

Thus, W is orthogonal to each integral manifold N^{m-1} of the distribution. Normalizing the length of the vector field W to one, the volume form of each integral manifold is given by means of

$$dv(N^{m-1}) = \frac{1}{||W||} W \lrcorner\, dv(M^m). \tag{4.6}$$

The volume form is written here as $dv(M^m)$ rather than dv to include the space referred to, in this instance M^m. If $\{x_i\}$ is a coordinate system on M^m then

$$dv(M^n) = \sqrt{|g_{ij}|}\, dx_1 \wedge \cdots \wedge dx_m.$$

The behavior of the integral manifold $N^{m-1} \subset M^m$ of the distribution can be studied under the flow $\Phi_t : M^m \to M^m$ of the vector field W. To initiate this, compute first the Lie derivative of the one-form ω with respect to the associated vector field W. Since $\omega(W) = ||W||^2$ and $\mathcal{L}_W(\omega) = d(W \lrcorner \omega) + W \lrcorner (d\omega)$, we obtain

$$(\mathcal{L}_W(\omega))V = V(\omega(W)) + d\omega(W, V) = V(\omega(W)) + W(\omega(V)) - V(\omega(W)) - \omega([W, V])$$
$$= W(\omega(V)) - \omega([W, V]). \tag{4.7}$$

Here use is made of the well-known result $d\omega(V, W) = V(\omega(W)) - W(\omega(V)) - \omega([V, W])$. In particular, for a vector field V tangent to the distribution, this simplifies to

$$(\mathcal{L}_W(\omega))V = -\omega([W,V]) - g(W,[W,V]).$$

There is then the following theorem.

Theorem 4.4. The flow Φ_t of the vector field W maps an integral manifold of the distribution $\mathcal{E}^{m-1}$ to another integral manifold if and only if

$$V(||W||^2) + d\omega(W,V) = -g(W,[W,V]) = 0, \tag{4.8}$$

for every vector field V on M^m with values in $\mathcal{E}^{m-1}$.

Corollary 4.2. Let the distribution $\mathcal{E}^{m-1}$ be defined by the closed one-form ω. Then the flow of the dual vector field W transforms integral manifolds into integral manifolds if the length $||W||$ is constant on every connected integral manifold.

Corollary 4.3. If the distribution $\mathcal{E}^{m-1}$ is defined by a one-form ω of constant length, and if, moreover, the flow of the dual vector field W transforms integral manifolds into integral manifolds, then $d\omega = 0$. In this case $\mathcal{E}^{m-1}$ locally consists of level surfaces of a function whose gradient has constant length.

Proof: Theorem 4.4 implies that $W \lrcorner d\omega = 0$. At a point $x \in M^m$, choose an orthonormal basis $e_1, \cdots, e_m$ in the tangent space so that W is proportional to e_1, that is $W = ae_1$. Denoting by $\sigma_1, \cdots, \sigma_m$ the dual basis, the form $\omega = a\sigma_1$ is proportional to σ_1. Write the two-form $d\omega$ in this basis,

$$d\omega = \sum_{i<j} B_{ij}\, \sigma_i \wedge \sigma_j.$$

Since $d\omega \wedge \omega = 0$, the two-form $d\omega$ only contains the $m-1$ summands $d\omega = b_{12}\sigma_1 \wedge \sigma_2 + \cdots + b_{1m}\sigma_1 \wedge \sigma_m$. However, the condition $0 = W \lrcorner d\omega = a(b_{12}\sigma_2 + \cdots + b_{1m}\sigma_m)$ implies that ω is a closed form.

Next, the infinitesimal volume change of a compact integral manifold N^{m-1} can be computed under the flow Φ_t of the vector field W. The Lie derivative of the volume form dN^{m-1} is given by

$$\mathcal{L}_W(dv(N^{m-1})) = \mathcal{L}_W(\frac{1}{||W||} W \lrcorner dv(M^m)) = W \lrcorner d(\frac{1}{||W||} W \lrcorner dv(M^m))$$

$$= (\operatorname{div} W - \frac{1}{2} W \ln(||W||^2))\, dv(N^{m-1}).$$

Based on this, the following Theorem can be formulated.

Theorem 4.5. The derivative of the volume change of a compact integral manifold N^{m-1} of the distribution $\mathcal{E}^{m-1}$ under the flow of the vector field W is given by the formula

$$\frac{d}{dt}(\operatorname{vol}(\Phi_t(N^{m-1})))|_{t=0} = \int_{N^{m-1}} (\operatorname{div}(W) - \frac{1}{2} W \ln ||W||^2) dv(N^{m-1}). \tag{4.9}$$

If the distribution $\mathcal{E}^{m-1}$ consists of the level surfaces of a function $f : M^m \to \mathbb{R}$, and the gradient of this function is chosen as the vector field, $W = \operatorname{grad}(f)$, then $\operatorname{div}(W) = \Delta(f)$, which is the Laplacian of f.

Corollary 4.4. The volume change of a compact level surface N^{m-1} of the function $f : M^m \to \mathbb{R}$ under the flow of the gradient vector field is given by the formula

$$\frac{d}{dt}(\mathrm{vol}(\Phi_t(N^{m-1})))|_{t=0} = \int_{N^{m-1}} (\nabla f - \frac{1}{2}\mathrm{grad}(f)\ \ln ||\mathrm{grad}(f)||^2)\, dv(N^{m-1}). \tag{4.10}$$

In all these formulas, the Laplacian, the divergence and the gradient are taken with respect to the manifold.

As an example, consider a function $f : M^m \to \mathbb{R}$ and assume there is another function $\mu : M^m \to \mathbb{R}$ such that

$$d(||\mathrm{grad}(f)||^2) = 2\mu\, df.$$

By Theorems 4.4 and 4.5, the flow of the gradient vector field grad(f) maps level surfaces of f to level surfaces, and the volume change is described by

$$\frac{d}{dt}(\mathrm{vol}(\Phi_t(N^{m-1})))|_{t=0} = \int_{N^{m-1}} (\nabla(f) - \mu)\, dv(N^{m-1}).$$

The spheres $S^{m-1}(R) \subset \mathbb{R}^m$ are level surfaces of the function $f(x) = ||x||^2$, and it follows that $||\mathrm{grad}(f)||^2 = 4||x||^2$ as well as $\nabla(f) = 2m$. Hence, $\mu = 2$ and $\nabla f - \mu = 2(m-1)$ is constant. For the flow, $\Phi_t(x) = e^{2t}x$, we obtain the following differential equation describing the evolution of the volume

$$\frac{d}{dt}\,(\mathrm{vol}(\Phi_t(S^{m-1}))) = 2(m-1)\mathrm{vol}(\Phi_t(S^{m-1})).$$

The second application of Frobenius Theorem plays an important role in surface theory.

Theorem 4.6. Let $\Omega = \{\omega_{ij}\}$ be a $(k \times k)$ matrix of one-forms defined on a neighborhood of $0 \in \mathbb{R}^m$, and let A_0 be an invertible $(k \times k)$ matrix. In a connected neighborhood $O \in V$, there exists a $(k \times k)$ matrix $A = (f_{ij})$ of functions which satisfy

$$\Omega = dA \cdot A^{-1}, \qquad A(0) = A_0, \tag{4.11}$$

if and only if

$$d\Omega = \Omega \wedge \Omega. \tag{4.12}$$

In this case, the matrix A is uniquely determined. If, in addition, Ω is an anti-symmetric matrix $\Omega + \Omega^t = 0$, and A_0 is an orthogonal matrix so that $A_0 \cdot A_0^t = I$, then the solution $A(x)$ is also orthogonal at each point of the set V.

Proof: The condition $d\Omega = \Omega \wedge \Omega$ is necessary for the solvability of the equation $\Omega = dA \cdot A^{-1}$. In fact, $dA = \Omega \cdot A$, so

$$0 = ddA = d(\Omega A) = d\Omega\, A - \Omega \wedge dA = (d\Omega - \Omega \wedge \Omega)A,$$

The matrix A is invertible, therefore it follows that $d\Omega - \Omega\wedge\Omega = 0$. Uniqueness of the solution follows, since for two solutions $A(x)$ and $B(x)$, there results,

$$dB^{-1} = -B^{-1}\cdot(dB)\cdot B^{-1}.$$

Hence the differential $d(B^{-1}\cdot A)$ vanishes,

$$d(B^{-1}\cdot A) = d(B^{-1})\cdot A + B^{-1}dA = -B^{-1}\cdot(dB)\cdot B^{-1}A + B^{-1}\cdot\Omega\cdot A$$

$$= -B^{-1}\Omega B\cdot B^{-1}A + B^{-1}\Omega A = 0.$$

Therefore, $B^{-1}\cdot A$ is constant, and at $x = 0$, it is equal to the unit matrix. This implies that $A(x) = B(x)$ for all $x \in V$. Now the existence of a solution will be proved under the condition $d\Omega = \Omega\wedge\Omega$. To this end, consider the following $(k\times k)$ matrix of one-forms on the space $\mathbb{R}^m\times\mathbb{R}^{k^2}$ with coordinates $(x^1,\cdots,x^m,z^{ij})$ so

$$\Lambda = dZ - \Omega\cdot Z = (dz^{ij} - \sum_{r=1}^{k}\omega_{ir}\,z^{rj}).$$

Using the relation $d\Omega = \Omega\wedge\Omega$, we obtain

$$d\Lambda = ddZ - d\Omega\wedge Z + \Omega\wedge dZ = -\Omega\wedge\Omega\wedge Z + \Omega\wedge dZ$$

$$= -\Omega\wedge\Omega\wedge Z + \Omega\wedge(\Lambda + \Omega\wedge Z) = \Omega\wedge\Lambda.$$

The system of forms $dz^{ij} - \sum_{r=1}^{k}\omega_{ir}z^{rj}$ is linearly independent in $\mathbb{R}^m\times\mathbb{R}^{k^2}$, so by Frobenius' theorem, there exists an m-dimensional integral manifold $M^m\subset\mathbb{R}^m\times\mathbb{R}^{k^2}$ through $(0,A_0)\in\mathbb{R}^m\times\mathbb{R}^{k^2}$. The tangent space $T_{(0,A_0)}M^m$ to this integral manifold has only the null vector in common with $\mathbb{R}^{k^2}$, $T_{(0,A_0)}M^m\cap\mathbb{R}^{k^2} = \{0\}$. This follows directly from the shape of the form Λ, since the tangent space to M^m is determined by $\Lambda = 0$. Then the integral manifold M^m is the graph of a map $A: W\to\mathbb{R}^{k^2}$ defined on an open set $O\in W\subset\mathbb{R}^m$ satisfying the initial condition $A(0) = A_0$. Using

$$A^*(\Lambda) = A^*(dZ - \Omega Z) = dA - \Omega A,$$

it is seen that the $(k\times k)$-matrix A is the solution of the differential equations sought after. The remaining statements follow from the result

$$d(A^t\cdot A) = (dA)^t\cdot A + A^t\cdot dA = A^t\Omega^t A + A^t\Omega A.$$

In fact, it may be concluded that if Ω is an antisymmetric matrix, and $dA = \Omega\cdot A$ is a solution of the differential equation, then it can be concluded that $d(A^t\cdot A) = 0$.

5 Pfaffian Systems.

This is a view of Pfaffian systems which is somewhat more abstract than the previous section, but closely related to it. All manifolds, maps, differential forms, vector fields and other geometric structures will be assumed to be C^∞ **[30-33]**.

Definition 5.1. An exterior differential system on an n-dimensional manifold N^n is a differential ideal I of the ring $\Omega(M^n)$ of exterior differential forms on M^n, that is, an ideal of $\Omega(M^n)$ which is closed under exterior differentiation. A *Pfaffian system* is an exterior differential system generated by one-forms as a differential ideal.

An m-dimensional integral manifold of I is an m-dimensional immersed submanifold $h : W^m \to M^n$ such that

$$h^*\omega = 0,$$

for all $\omega \in I$.

Here are two examples which address integral manifolds.

$(v.i)$ On $\mathbb{R}^3$ with coordinates (x, y, p), consider the Pfaffian system

$$I = \{dy - pdx, dp \wedge dx\}.$$

The one-dimensional integral manifolds of I are the curves $(x(t), y(t), p(t))$ such that $y' - px' = 0$. The curves for which $x' \neq 0$ may be reparametrized in the form $(x, f(x), f'(x))$, where f is an arbitrary function. The one-dimensional integral manifolds of I thus depend on one arbitrary function of one variable.

$(v.ii)$ On $\mathbb{R}^5$ with coordinates (x, y, u, p, q), consider a function $F : \mathbb{R}^5 \to \mathbb{R}$ such that $F_p \neq 0$, $F_q \neq 0$. Consider the exterior differential system

$$I = \{F, du - pdx - qdy, dF, dx \wedge dp + dy \wedge dq\}.$$

The integral manifolds of I, which are two-dimensional surfaces $(x(s,t), y(s,t), u(s,t), p(s,t), q(s,t))$ such that

$$F(x(s,t), y(s,t), u(s,t), p(s,t), q(s,t)) = 0,$$

$$u_s - px_s - qy_s = 0, \qquad u_t - px_t - qy_t = 0.$$

If $|\partial(x,y)/\partial(s,t)| \neq 0$, then the integral surfaces may be reparametrized in the form $(x, y, u(x,y),\ p(x,y), q(x,y))$ where $p = u_x$, $q = u_y$ and $u(x,y)$ is a solution of the first-order partial differential equation $F(x, y, u, u_x, u_y) = 0$. In local coordinates then, the condition that an immersion defines an integral manifold of an exterior differential system is expressed as a system of differential equations on the component maps of the immersion in the corresponding local coordinate charts.

Some basic existence theorems will be stated for local integral manifolds of Pfaffian systems. The general principle behind these existence theorems is to establish local normal forms in which the integral manifolds are manifest. The proofs of these theorems are based on the fundamental theorems for the existence, uniqueness and smooth dependence on initial conditions of solutions of ordinary differential equations.

Consider a Pfaffian system I generated as a differential ideal by $s \leq n$ linearly independent one-forms

$$\omega^a = \sum_{i=1}^{n} A_i^a(x^1, \cdots, x^n)\, dx^i, \qquad 1 \leq a \leq s. \tag{5.1}$$

Written out in local coordinates, the integral manifolds of a Pfaffian system I correspond to the solutions of a system of partial differential equations. Indeed, if the immersion f is given locally in a domain U of $\mathbb{R}^p$ with coordinates $u^\alpha, 1 \leq \alpha \leq p$, by an n-tuple of functions $(f^1(u^1, \cdots, u^p), \cdots, f^n(u^1, \cdots, u^p))$ satisfying $\operatorname{rank}(\partial f^i/\partial u^\alpha) = p$ in U, then $f^* \omega^a = 0$ has the form,

$$\sum_{i=1}^{n} \sum_{\alpha=1}^{p} A_i^\alpha(f^1(u^1, \cdots, u^p), \cdots, f^n(u^1, \cdots, u^p)) \frac{\partial f}{\partial u^\alpha} = 0, \qquad 1 \leq a \leq s. \tag{5.2}$$

The simplest existence theorem for integral manifolds of a Pfaffian system is the Frobenius theorem, which was introduced in the previous section. The last part of Theorem 4, which is more relevant for Pfaffian systems, is repeated here.

Theorem 5.1. Let I be a Pfaffian system generated by linearly independent one-forms ω^a, $1 \leq a \leq s$, satisfying

$$d\omega^a \wedge \omega^1 \wedge \cdots \omega^s = 0, \qquad 1 \leq a \leq s. \tag{5.3}$$

There exist local coordinates $(u^1, \cdots, u^n)$ such that I is generated by the differentials $du^1, \cdots, du^s$.

$(v.iii)$ On $\mathbb{R}^4$ with local coordinates (x, y, z, u), minus the locus $x + z = 0$, $y = 0$, $u = 0$, consider the Pfaffian system $I = \{\omega^1, \omega^2, d\omega^1, d\omega^2\}$, where

$$\omega^1 = u^2(x+z)(dx+dz) + u^2(dy + udu), \qquad \omega^2 = y^4(dy + udu).$$

The integrability conditions (5.3) are satisfied and we have

$$I = \{d(x+z), d(2y+u^2)\},$$

so that the two-dimensional integral manifolds are the surfaces obtained by taking the intersection of the 3-planes $x+z = c_1$ with the parabolic 3-cylinders $2y = u^2 = c_2$, where c_1 and c_2 are arbitrary constants.

If I is an s-dimensional Pfaffian system satisfying the Frobenius conditions (5.3), there exist, at least locally, $(n-s)$-dimensional integral manifolds of I given by $u^1 = c^1, \cdots, u^s = c^s$, where $c^1, \cdots, c^s$ are arbitrary real constants. These integral manifolds are of maximal dimension. A Pfaffian system generated by linearly independent one-forms ω^j is said to be completely integrable if it satisfies (5.3), and the u^j are called first integrals of the completely integrable system I.

There is a classical construction due to E. Cartan, for obtaining a minimal set of coordinates in which to express the generators of a Pfaffian system. Let Char (I) denote the system of Cauchy characteristic vector fields of I, defined by Char $(I) = \{X|X \in I^{\perp}, X \lrcorner dI \subset I\}$. The Cartan system of I, denoted $C(I)$, is the dual Pfaffian system defined by $C(I) =$Char $(I)^{\perp}$. The *class* of a Pfaffian system is by definition the dimension of its Cartan system. The Cartan system $C(I)$ of any Pfaffian system I is always completely integrable. The first integrals of $C(I)$ provide the required minimal set of coordinates.

Theorem 5.2. (Cartan) Let I be a Pfaffian system of class r and let $\{w^1, \cdots, w^r\}$ denote a set of first integrals of the Cartan system $C(I)$. There exists a small neighborhood U with local coordinates $(w^1, \cdots, w^r; y^{r+1}, \cdots, y^n)$ such that I is generated in U by one-forms in $w^1, \cdots, w^r$ and their differentials.

The Frobenius theorem leads to a normal form in which the integral manifolds of a completely integrable Pfaffian system are manifest. There are similar normal form results which apply to Pfaffian systems which are not completely integrable, but whose structure equations are of a special type. Consider the simplest result, known as the solution to the Pfaff problem. Let I be a Pfaffian system generated as a differential ideal by a single one-form ω. The rank of $I = \{\omega, d\omega\}$ is the integer r defined by

$$(d\omega)^r \wedge \omega \neq 0, \qquad (d\omega)^{r+1} \wedge \omega = 0. \tag{5.4}$$

Theorem 5.3. If rank $\{\omega, d\omega\} = r$, then class $\{\omega\} = 2r+1$ and there exist local coordinates $(z, p_1, \cdots, p_r, x^1, \cdots, x^r, u^{2r+2}, \cdots, u^n)$ such that

$$\{\omega, d\omega\} = \{dz - \sum_{i=1}^{r} p_i dx^i, \sum_{i=1}^{r} dp_i \wedge dx^i\}. \tag{5.5}$$

Unlike the case in which I is completely integrable, the integral manifolds of $I = \{\omega, d\omega\}$ depend now on one arbitrary function of r variables. They can be put in the form $z = f(x^1, \cdots, x^n), p_i = f_{x^i}, 1 \leq i \leq r$. ♣

$(v.iii)$ On $\mathbb{R}^4$ with coordinates (x, y, z, u), minus the locus $y(x+y^2) = 0$, consider the Pfaff system $I = \{\omega, d\omega\}$, where

$$\omega = (x+y^2)y^2\, dz - y(yz + u^2(x+y^2)^2)\, dx + (u^2 x(x+y^2)^2 - 2y^3 z)\, dy.$$

Then $d\omega \wedge d\omega \wedge \omega = 0$ and $I = \{dZ - PdX, dP \wedge dX\}$, where

$$X = \frac{x}{y}, \quad Z = \frac{z}{x+y^2}, \qquad P = u^2.$$

The one-dimensional integral manifolds of I thus depend on one arbitrary function of one variable and are given by

$$\frac{z}{x+y^2} = f(\frac{x}{y}), \qquad u^2 = f'(\frac{x}{y}).$$

A slightly stronger result which can be thought of as the solution to a relative version of the Pfaff problem can be proved. It goes like this.

Theorem 5.4. Let I be an $(s+1)$-dimensional Pfaffian system of the form

$$I = \{\theta^1, \cdots, \theta^s, d\theta^1, \cdots, d\theta^s, \omega, d\omega\},$$

where $J = \{\theta^1, \cdots, \theta^s, d\theta^1, \cdots, d\theta^s\}$ is completely integrable and the rank r of $I = \{\omega, d\omega\}$ relative to J, defined by

$$(d\omega)^r \wedge \omega \not\equiv 0 \mod J, \qquad (d\omega)^{r+1} \wedge \omega \equiv 0 \mod J, \tag{5.6}$$

is constant.

There exist local coordinates

$$(z, p_1, \cdots, p_r, x^1, \cdots, x^r, w^1, \cdots, w^s, u^{2r+s+1}, \cdots, u^n)$$

in which

$$I = \{dz - \sum_{i=1}^r p_i\, dx^i, \sum_{i=1}^r dp_i \wedge dx^i, dw^1, \cdots, dw^s\}. \tag{5.7}$$

♣

The local integral manifolds of I now depend on one arbitrary function of r variables and s arbitrary constants. They are parametrized by $z = f(x^1, \cdots, x^n)$, $p_i = f_{x^i}$, $w^\alpha = c^\alpha$, $1 \le i \le r$, $1 \le \alpha \le s$.

$(v.iv)$ The solution of the relative Pfaff problem has interesting applications to the classification and normal forms problem for partial differential equations. Consider the system

$$F(x,y,z,w,\frac{\partial z}{\partial x},\frac{\partial z}{\partial y},\frac{\partial w}{\partial x},\frac{\partial w}{\partial y}) = 0, \qquad G(x,y,z,w,\frac{\partial z}{\partial x},\frac{\partial z}{\partial y},\frac{\partial w}{\partial x},\frac{\partial w}{\partial y}) = 0, \tag{5.8}$$

of two first-order partial differential equations for maps $(z,w) : \mathbb{R}^2 \to \mathbb{R}^2$.

The system (5.8) is said to be parabolic if and only if the matrix

$$\begin{pmatrix} \frac{\partial(F,G)}{\partial(m,n)} & \frac{1}{2}(\frac{\partial(F,G)}{\partial(q,m)} - \frac{\partial(F,G)}{\partial(p,n)}) \\ \frac{1}{2}(\frac{\partial(F,G)}{\partial(q,m)} - \frac{\partial(F,G)}{\partial(p,n)}) & \frac{\partial(F,G)}{\partial(p,q)} \end{pmatrix} \tag{5.9}$$

has rank one. The first step is to formulate system (5.8) geometrically as a Pfaffian system. Consider the bundle $J^1(\mathbb{R}^2,\mathbb{R}^2)$ of 1-jets of maps from $\mathbb{R}^2$ to $\mathbb{R}^2$ and use local coordinates (x,y,z,w,p,q,m,n) in which the Pfaffian system $\Omega^1_{cont}(\mathbb{R}^2,\mathbb{R}^2)$ of contact 1-forms is the differential ideal generated by

$$\theta^1 = dz - p\, dx - q\, dy, \qquad \theta^2 = dw - m\, dx - n\, dy. \tag{5.10}$$

Thus if $f : \mathbb{R}^2 \to J^0(\mathbb{R}^2, \mathbb{R}^2) : (x, y) \to (x, y, z(x, y), w(x, y))$ is a section, then

$$p(j^1 f) = \frac{\partial z}{\partial x}, \quad q(j^1 f) = \frac{\partial z}{\partial y}, \quad m(j^1 f) = \frac{\partial w}{\partial x}, \quad n(j^1 f) = \frac{\partial w}{\partial y}.$$

It is assumed that the equations

$$F(x, y, z, w, p, q, m, n) = 0, \qquad G(x, y, z, w, p, q, m, n) = 0, \tag{5.11}$$

corresponding to the partial differential equations (5.8) give rise to a six-dimensional submanifold $i : \Sigma_6 \to J^1(\mathbb{R}^2, \mathbb{R}^2)$. It can be shown that the local sections $f : \mathbb{R}^2 \to \Sigma_6$ which are the integral manifolds of the Pfaffian system $\mathcal{I}$ obtained by pulling back the contact system $\Omega^1_{cont}(\mathbb{R}^2, \mathbb{R}^2)$ to Σ_6 are in one-to-one correspondence with the solutions of (5.8). Moreover, it can be shown that if (5.8) is parabolic and $\mathcal{I}$ has a one-dimensional first derived system, then there exist generators π_1 and π_2 such that $\mathcal{I} = \{\pi_1, \pi_2, d\pi_1, d\pi_2\}$ and the following structure equations are valid,

$$d\pi_1 = 0, \qquad d\pi_2 = \omega_3 \wedge \omega_5 + \omega_4 \wedge \omega_6, \qquad \mathrm{mod}\ \{\pi_1, \pi_2\}. \tag{5.12}$$

Theorem 5.5. Every partial differential equation system (5.8) of parabolic type whose associated Pfaffian system I has a one-dimensional first derived system can be locally transformed to the normal form

$$\frac{\partial z}{\partial x} = 0, \qquad \frac{\partial z}{\partial y} = 0,$$

by a contact transformation.

Proof: The hypotheses of the theorem allow us to apply the Cartan-Von Weber theorem to conclude that $I^{(1)} = \{\pi_1\}$ must be completely integrable. We can therefore apply the solution to the relative Pfaff problem to argue that there exist local coordinates (x, y, z, w, m, n) such that

$$\{\pi_1, \pi_2\} = \{dw - m\,dx - n\,dy, dz\}. \qquad \clubsuit$$

Another normal form result is the Goursat normal form, of which the relative version is presented.

Theorem 5.6. Let $\mathcal{I}$ be an $(r + s)$-dimensional Pfaffian system of codimension two, given by

$$\mathcal{I} = \{\omega^1, \cdots, \omega^r, \theta^1, \cdots, \theta^s, d\omega^1, \cdots, d\omega^r, d\theta^1, \cdots, d\theta^s\} \tag{5.13}$$

where $\mathcal{I} = \{\theta^1, \cdots, \theta^s, d\theta^1, \cdots, d\theta^s\}$ is completely integrable. Suppose that there exist one-forms α and π such that $\alpha \neq 0$, $\pi \neq 0$, $\mathrm{mod}\ \mathcal{I}$ and such that the following structure equations are valid

$$\begin{aligned}
d\omega^1 &\equiv \omega^2 \wedge \pi && \mathrm{mod}\ \{\omega^1, \theta^1, \cdots, \theta^s\} \\
&\vdots \\
d\omega^i &\equiv \omega^{i+1} \wedge \pi && \mathrm{mod}\ \{\omega^1, \cdots, \omega^i, \theta^1, \cdots, \theta^s\}, \quad 1 \le i \le r-1 \\
&\vdots \\
d\omega^r &\equiv \alpha \wedge \pi && \mathrm{mod}\ \{\omega^1, \cdots, \omega^r, \theta^1, \cdots, \theta^s\} = I
\end{aligned}$$

There exist local coordinates $(x, y, y', \cdots, y^{(r)}, w^1, \cdots, w^s, u^{r+s+3}, \cdots, u^n)$ such that

$$\mathcal{I} = \{dy - y'\,dx, \cdots, dy^{(r-1)} - y^{(r)}dx, dy' \wedge dx, \cdots, dy^{(r)} \wedge dx, dw^1, \cdots, dw^s\}. \tag{5.14}$$

The local integral manifolds of $\mathcal{I}$ depend on one arbitrary function of one variable and r arbitrary constants. They can be generically parametrized as

$$y = f(x), \quad y' = f'(x), \quad \cdots, y^{(s)} = f^{(s)}(x), u^1 = c^1, \cdots, u^r = c^r.$$

Finally, an introduction to Cauchy characteristics is given. To set the stage some background results are needed. Let V here be a real vector space of dimension n and V^* its dual space. An element $x \in V$ is called a vector and an element $\omega \in V^*$ a covector such that between V and V^* there exists a pairing such that $\langle x, \omega \rangle$, $x \in V$, $\omega \in V^*$, is a real number. Over V there is the exterior algebra, which is a graded algebra

$$\Lambda(V) = \Lambda^0(V) \oplus \Lambda^1(V) \oplus \cdots \oplus \Lambda^n(V),$$

with $\Lambda^0(V) = \mathbb{R}$, $\Lambda^1(V) = V$. In the same way, there is over V^* an exterior algebra

$$\Lambda(V^*) = \Lambda^0(V^*) \oplus \Lambda^1(V^*) \oplus \cdots \oplus \Lambda^n(V^*),$$

with $\Lambda^0(V^*) = \mathbb{R}$, $\Lambda^1(V^*) = V^*$. If e_i is a basis of V and ω^k its dual basis, then $\langle e_i, \omega^k \rangle = \delta_i^k$, $1 \leq i, k \leq n$.

Given an ideal $I \subset \Lambda(V^*)$, it is desired to determine the smallest subspace $W^* \subset V^*$ such that I is generated, as an ideal, by the set S of elements of $\Lambda(W^*)$. Define

$$A(I) = \{x \in V | x \lrcorner I \subset I\},$$

where the last condition means that $x \lrcorner \alpha \in I$, for all $\alpha \in I$. Now $A(I)$ is clearly a subspace of V, and is referred to as the Cauchy characteristic space of I. Its annihilator

$$\mathcal{C}(I) = A(I)^{\perp} \subset V^*,$$

will be called the retracting subspace of I.

Proposition 5.1. (*Retraction Theorem.*) Let I be an ideal of $\Lambda(V^*)$. Its retracting subspace $\mathcal{C}(I)$ is the smallest subspace of V^* such that $\Lambda(\mathcal{C}(I))$ contains a set S of elements generating I as an ideal. The set S also generates an ideal J in $\Lambda(\mathcal{C}I))$ to be called a retracting ideal of I. There exists a mapping

$$\Delta : \Lambda(V^*) \to \Lambda(\mathcal{C}(I))$$

of graded algebras such that $\Delta(I) = J$.

The Frobenius Theorem shows that a completely integrable system takes a very simple form upon a proper choice of local coordinates. Given any exterior differential system, one can ask the question whether there is a coordinate system such that the system is generated by forms in a smaller number of these coordinates. This question is answered by the Cauchy characteristics, and its algebraic basis is the Retraction theorem.

Let I be a differential ideal. A vector field ξ such that $\xi \lrcorner I \subset I$ is called a Cauchy characteristic vector field of I. At a point $x \in M$ we define

$$A(I)_x = \{\xi_x \in T_x M | \xi_x \lrcorner I_x \subset I_x\}, \tag{5.15}$$

and $C(I)_x = A(I)_x^\perp \subset T_x^* M$. Call $C(I)_x$ the retracting space at x and $\dim C(I)_x$ the class of I at x. We have now a family of ideals I_x depending on the parameter $x \in M$. When restricting to a point x we have a purely algebraic situation.

Proposition 5.2. If ξ, η are Cauchy characteristic vector fields of a differential ideal $\mathcal{I}$, their Lie bracket $[\xi, \eta]$ is as well.

Proof. Let L_ξ be the Lie derivative defined by ξ. Then L_ξ takes the form

$$L_\xi = d(\xi \lrcorner) + (\xi \lrcorner) d.$$

Since $\mathcal{I}$ is closed, $d\mathcal{I} \subset \mathcal{I}$. If ξ is a characteristic vector field, then $\xi \lrcorner \mathcal{I} \subset \mathcal{I}$. It follows that $L_\xi \mathcal{I} \subset \mathcal{I}$. This follows from the identity

$$[L_\xi, \eta \lrcorner] = L_\xi \eta \lrcorner - \eta \lrcorner L_\xi = [\xi, \eta] \lrcorner, \tag{5.16}$$

which is valid for any two vector fields ξ, η.

Theorem 5.7. Let $\mathcal{I}$ be a finitely generated differential ideal whose retracting space $C(\mathcal{I})$ has constant dimension $s = n - r$. Then there is a neighborhood in which there are coordinates $(x^1, \cdots, x^r; y^1, \cdots, y^s)$ such that $\mathcal{I}$ has a set of generators that are forms in $y^1, \cdots, y^s$ and their differentials.

Proof: By the Frobenius condition the differential system defined by $C(\mathcal{I})$ is completely integrable. Choose coordinates $(x^1, \cdots, x^r; y^1, \cdots, y^s)$ so that the foliation so defined is given by

$$y^\sigma = c^\sigma, \qquad 1 \le \sigma \le s.$$

The c^σ are constants. By Proposition 5.1, $\mathcal{I}$ has a set of generators which are forms in dy^σ, $1 \le \sigma \le s$. Their coefficients may involve x^ρ, $1 \le \rho \le r$. The theorem follows when it is shown a new set of generators for $\mathcal{I}$ can be chosen which are forms in the y^σ coordinates in which x^ρ do not enter. To exclude the trivial case, we suppose the ideal $\mathcal{I}$ is a proper ideal, so that it contains no non-zero functions.

Let $\mathcal{I}_q$ be the set of q-forms in $\mathcal{I}$, $q = 1, 2, \cdots$. Let $\varphi^1, \cdots, \varphi^p$ be linearly independent one-forms in $\mathcal{I}_1$ such that any form in $\mathcal{I}_1$ is their linear combination. Since $\mathcal{I}$ is closed, $d\varphi^i \in \mathcal{I}$, $1 \le i \le p$. For a fixed p it holds that $\partial / \partial x^\rho \in A(\mathcal{I})$, which implies

$$\frac{\partial}{\partial x^\rho} \lrcorner d\varphi^i = L_{\partial/\partial x^\rho} \varphi^i \in \mathcal{I}_1,$$

since the left-hand side is of degree one. It follows that

$$\frac{\partial \varphi^i}{\partial x^\rho} = L_{\partial/\partial x^\rho}\varphi^i = \sum_j a^i_j \varphi_j, \quad 1 \leq i, j \leq p \tag{5.17}$$

where the left-hand side stands for the form obtained from φ^i by taking the partial derivatives of the coefficients with respect to x^ρ.

For this fixed ρ, regard x^ρ as the variable and the remaining quantities $x^1, \cdots, x^{\rho-1}$, $x^{\rho+1}, \cdots, x^r, y^1, \cdots, y^s$ as parameters. Write the system of ordinary differential equations

$$\frac{dz^i}{dx^\rho} = \sum_j a^i_j z^j, \qquad 1 \leq i, j \leq p. \tag{5.18}$$

Let $z^i_{(k)}$, $1 \leq k \leq p$, be a fundamental system of solutions, such that

$$\det(z^i_{(k)}) \neq 0.$$

Now replace φ^i by the $\tilde{\varphi}^k$ defined by

$$\varphi^i = \sum z^i_{(k)} \tilde{\varphi}^k. \tag{5.19}$$

Differentiating (5.19) with respect to x^ρ and using (5.17) and (5.18), there results

$$\frac{\partial \tilde{\varphi}^k}{\partial x^\rho} = 0,$$

so $\tilde{\varphi}^k$ does not involve x^ρ. Applying the same process to the other x 's, a set of generators of $\mathcal{I}_1$ is reached which are forms in y^σ.

Suppose this process is continued for $\mathcal{I}_1, \cdots, \mathcal{I}_{q-1}$, so that they consist of forms in y^σ. Let $\mathcal{J}_{q-1}$ be the ideal generated by $\mathcal{I}_1, \cdots, \mathcal{I}_{q-1}$. Let $\psi^\alpha \in \mathcal{I}_q$, $1 \leq \alpha \leq r$, be linearly independent $\mod \mathcal{J}_{q-1}$, such that any q-form of $\mathcal{I}_q$ is congruent $\mod \mathcal{J}_{q-1}$ to a linear combination of them. Thus, such forms include

$$\frac{\partial}{\partial x^\rho} \lrcorner d\psi^\alpha = L_{\partial/\partial x^\rho}\psi^\alpha.$$

Hence

$$\frac{\partial \psi^\alpha}{\partial x^\rho} \equiv \sum b^\alpha_\beta \psi^\beta, \qquad \mod \mathcal{J}_{q-1}, \quad 1 \leq \alpha, \beta \leq r.$$

Using the argument above, replace the ψ^α by $\tilde{\psi}^\beta$ such that $\partial \tilde{\psi}^\alpha / \partial x^p \in \mathcal{I}_{q-1}$, which means that

$$\frac{\partial \tilde{\psi}^\alpha}{\partial x^\rho} = \sum \eta^\alpha_h \wedge \omega^\alpha_h,$$

where $\eta^\alpha_h \in \mathcal{I}_1 \cup \cdots \cup \mathcal{I}_{q-1}$ and are therefore forms in y^σ. Let θ^α_h be defined by

$$\frac{\partial \theta^\alpha_h}{\partial x^\rho} = \omega^\alpha_h.$$

Then the forms

$$\hat{\psi}^\alpha = \tilde{\psi}^\alpha - \sum_h \eta_h^\alpha \wedge \theta_h^\alpha$$

do not involve x^ρ, and can be used to replace ψ^α. Applying this process to all x^ρ, $1 \leq \rho \leq r$, a set of generators for $\mathcal{I}_q$ is found which are forms in y^σ only. ♣

Definition: The leaves defined by the distribution $A(\mathcal{I})$ are called Cauchy characteristics.

Notice that generally r is zero, so that a differential system generally does not have Cauchy characteristics. The theorem just shown allows us to locally reduce a differential ideal to a system in which there are no extraneous variables in the sense that all coordinates are needed to express $\mathcal{I}$ in any coordinate system.

Corollary 5.1. Let $f : M \to M'$ be a fibration with vertical distribution $V \subset T(M)$ with connected fibers over $x \in M'$ given by $(ker\, f_*)_x$. Then a form α on M is the pull-back $f^*\alpha'$ of a form α' on M' if and only if

$$v \lrcorner \alpha = 0, \qquad v \lrcorner d\alpha = 0 \quad \forall \quad v \in V.$$

Proof: By the substitution theorem, there are local coordinates such that

$$f x^1, \cdots, x^p, x^{p+1}, \cdots, x^N) = (x^1, \cdots, x^p).$$

As such

$$V = (\frac{\partial}{\partial x^{p+1}}, \cdots, \frac{\partial}{\partial x^N}).$$

Now setting $\mathcal{I} = (\alpha)$, it is seen that $V \subset A(\mathcal{I})$. Therefore, by Theorem 2.2, there exists a generator for $\mathcal{I}$ independent of $(x^{p+1}, \cdots, x^N)$, and hence of the form $F^*\alpha''$ with $\alpha'' \in M'$. Thus there is a function μ such that

$$\mu\alpha = f^* \, \alpha''.$$

Since

$$0 = v \lrcorner (d\mu \wedge \alpha'' + \mu \, d\alpha'') = v(\mu)\alpha'', \quad \forall \quad v \in V,$$

it is seen that μ is independent of $(x^{p+1}, \cdots, x^N)$ and hence $\mu = \lambda \circ f$ for some function λ defined on M'. Setting $\alpha' = \frac{1}{\lambda}\alpha''$ the result that $\alpha = f^*(\alpha')$ follows.

Now these results are applied to the first-order partial differential equation

$$F(x^i, z, \frac{\partial z}{\partial x^i}) = 0, \qquad 1 \leq i \leq n. \tag{5.20}$$

This equation can be formulated as an exterior differential system. To (5.20) are added the exterior derivatives,

$$
\begin{aligned}
&F(x^i, z, p_i) = 0, \\
&dz - \textstyle\sum_i p_i \, dx^i = 0, \\
&\textstyle\sum_i (F_{x^i} + F_z p_i) \, dx^i + \sum_i F_{p_i} \, dp_i = 0, \\
&\textstyle\sum_i dx^i \wedge dp_i = 0.
\end{aligned} \tag{5.21}
$$

These equations are in the $(2n+1)$-dimensional space (x^i, z, p_i). The corresponding differential ideal is generated by the left-hand members of (5.21).

To determine the space $A(\mathcal{I})$, consider the vector

$$
\xi = \sum_i u^i \frac{\partial}{\partial x^i} + u \frac{\partial}{\partial z} + \sum_i v_i \frac{\partial}{\partial p_i}, \tag{5.22}
$$

and then express the condition that the interior product $\xi \lrcorner$ keeps the ideal $\mathcal{I}$ stable. Doing so gives the result

$$
\begin{aligned}
&u - \textstyle\sum_i p_i u^i = 0, \\
&\textstyle\sum_i (F_{x^i} + F_z p_i) u^i + F_{p_i} v_i = 0, \\
&\textstyle\sum_i (u^i \, dp_i - v_i \, dx^i) = 0.
\end{aligned} \tag{5.23}
$$

Comparing the last equation of (5.23) with the third equation of (5.21), it follows that

$$
u^i = \lambda F_{p_i}, \qquad v_i = -\lambda (F_{x^i} + F_z p_i). \tag{5.24}
$$

The first equation of (5.23) then gives

$$
u = \lambda \sum_i p_i F_{p_i}. \tag{5.25}
$$

The parameter λ is arbitrary, so (5.24) and (5.25) show that $\dim A(\mathcal{I}) = 1$. Thus, the characteristic vectors at each point form a one-dimensional space. This fundamental fact is the key to the theory of partial differential equations of first order. The characteristic curves in the space (x^i, z, p_i), or characteristic strips in the classical terminology, are the integral curves of the differential system,

$$
\frac{dx^i}{F_{p_i}} = -\frac{dp_i}{F_{x^i} + p_i F_z} = \frac{dz}{\sum_i p_i F_{p_i}}. \tag{5.26}
$$

These are the equations of Charpit and Lagrange. To construct an integral manifold of dimension n, it suffices to take an $(n-1)$-dimensional integral transverse to the Cauchy characteristic vector field, or noncharacteristic data in classical terminology, and draw the characteristic strips through its points. To put it another way, an n-dimensional integral manifold is generated by characteristic strips.

Note that points in (x^i, p_i)-space may be thought of as hyperplanes $\sum_i p_i dx^i = 0$ in the tangent spaces $T_x(\mathbb{R}^n)$. A curve in (x^i, z, p_i)-space projects to a curve in (x^i, p_i) space, which is geometrically a one-parameter family of tangent hyperplanes. This is the meaning of the terminology 'strips'.

Consider the example of the initial value problem for the partial differential equation

$$u\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 1, \tag{5.27}$$

with initial data given along $y = 0$ by $u(x,0) = \sqrt{x}$.

Let us introduce natural coordinates in $J^2(2,1)$ by selecting (x, y, u, p, q). This initial data $D : \mathbb{R} \to \mathbb{R}^2 \times \mathbb{R}$ where $D(x) = (x, 0, \sqrt{x})$ is extended to a map $\delta : \mathbb{R} \to J^2(2,1)$ such that the image satisfies the equation and the strip condition

$$0 = \delta^*(du - p\,dx - q\,dy) = \frac{1}{2\sqrt{x}}\,dx - p\,dx.$$

Here $p = 1/2\sqrt{x}$ and $q = 1 - up = 1/2$ and δ is unique. In general, there are several choices of δ due to non-linearity of the equation. The extended data becomes

$$\delta(x) = (x, 0, x, \frac{1}{2\sqrt{x}}, \frac{1}{2}).$$

If we parametrize the equation by $i : \sigma \to J^1(2,1)$ where $i(x, y, u, p) = (x, y, u, p, 1 - up)$, then the data can be pulled back to a map $\varDelta : \mathbb{R} \to \varSigma$, where $\varDelta(s) = (s, 0, \sqrt{s}, \frac{1}{2\sqrt{s}})$.

The Cauchy characteristic vector field is found by calculating $u^1 = \lambda F_{p_1} = \lambda u$, $u^2 = \lambda$, $v_1 = -\lambda(F_{x^1} + F_u p_1) = -\lambda p_1^2$, $v_2 = -\lambda(F_{x^2} + F_u p_2) = -\lambda p_1 p_2$, therefore,

$$X = u\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial u} - p^2\frac{\partial}{\partial p} - pq\frac{\partial}{\partial q}.$$

On a strip where q is constant, the Cauchy characteristic vector field is

$$X = u\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial u} - p^2\frac{\partial}{\partial p},$$

and the corresponding flow is given by

$$\frac{dx}{dt} = u, \quad \frac{dy}{dt} = 1, \quad \frac{du}{dt} = 1, \quad \frac{dp}{dt} = -p^2.$$

Integrating these equations, the solution for the given data representing the union of charcteristic curves along the data is given by

$$x = \frac{1}{2}t^2 + \sqrt{s}t + s, \qquad y = t, \qquad u = t + \sqrt{s}.$$

Eliminating s and t gives an implicit equation for $z(x, y)$ given by

$$u^2 + uy = x - \frac{1}{2}y^2.$$

Clearly, upon differentiating

$$u\frac{\partial u}{\partial x} = \frac{u}{2u - y}, \qquad \frac{\partial u}{\partial y} = \frac{u - y}{2u - y},$$

and these derivatives satisfy (5.27).

6 Introduction to Symmetry Calculations for Ordinary Differential Equations

6.1 General Theory

Ordinary differential equations of the form

$$y^{(n)} = \omega(x, y, y', \cdots, y^{(n-1)}), \tag{6.1}$$

will be considered where $y^{(k)}$ is the k-th derivative of y with respect to x. The symmetry condition will be stated and then the general results will be applied to study a specific nonlinear ordinary differential equation **[34,35]**. A symmetry of (6.1) is a diffeomorphism that maps the set of solutions of the equation to itself. Any diffeomorphism

$$\Gamma : (x, y) \to (\hat{x}, \hat{y}),$$

maps smooth planar curves to smooth planar curves. The action of Γ on the plane induces an action on the derivatives as well. This is the mapping

$$\Gamma : (x, y, y', \cdots, y^{(n)}) \to (\hat{x}, \hat{y}, \hat{y}', \cdots, \hat{y}^{(n)}), \qquad y^{(k)} = \frac{d^k \hat{y}}{d\hat{x}^k}, \quad k = 1, \cdots, n.$$

This mapping is called the n-th prolongation of Γ. The new variables depend on a parameter ϵ. The functions $\hat{y}^{(k)}$ are calculated recursively by means of the chain rule

$$\hat{y}^{(k)} = \frac{d\hat{y}^{(k-1)}}{d\hat{x}} = \frac{D_x \hat{y}^{(k-1)}}{D_x \hat{x}}, \quad \hat{y}^{(0)} = \hat{y}_0. \tag{6.2}$$

Here D_x is the total derivative with respect to x,

$$D_x = \partial_x + y' \partial_y + y'' \partial_{y'} + \cdots .$$

The symmetry condition for (6.1) is given by

$$\hat{y}^{(n)} = \omega(\hat{x}, \hat{y}, \hat{y}', \cdots, \hat{y}^{(n-1)}), \tag{6.3}$$

when (6.1) holds, and the functions $\hat{y}^{(k)}$ are given by (6.2). The symmetry condition (6.3) is nonlinear. Lie symmetries are obtained by linearizing (6.3) about $\epsilon = 0$. For ϵ sufficiently close to zero, the prolonged Lie symmetries are of the form,

$$\begin{aligned} \hat{x} &= x + \epsilon \xi + O(\epsilon^2), \\ \hat{y} &= y + \epsilon \eta + O(\epsilon^2), \\ y^{(k)} &= y^{(k)} + \epsilon \eta^{(k)} + O(\epsilon^2), \quad k \geq 1. \end{aligned} \tag{6.4}$$

Substitute (6.4) into the symmetry condition (6.3) and the order ϵ terms yield the linearized symmetry condition

$$\eta^{(n)} = \xi\omega_x + \eta\omega_y + \eta^{(1)}\omega_{y'} + \cdots + \eta^{(n-1)}\omega_{y^{(n-1)}}, \tag{6.5}$$

when (6.1) holds. The $\hat{y}^{(k)}$ can be derived recursively from (6.2),

$$\hat{y}^{(1} = \frac{D_x\hat{y}}{D_x\hat{x}} = \frac{y' + \epsilon D_x\eta + O(\epsilon^2)}{1 + \epsilon D_x\xi + O(\epsilon^2)} = y' + \epsilon(D_x\eta - y' D_x\xi) + O(\epsilon^2),$$
$$y^{(k)} = \frac{y^{(k)} + \epsilon D_x\eta^{(k-1)} + O(\epsilon^2)}{1 + \epsilon D_x\xi + O(\epsilon^2)} \tag{6.6}$$

The results in (6.6) give

$$\eta^{(1)} = D_x\eta - y'\, D_x\xi, \tag{6.7}$$

and

$$\eta^{(k)}(x, y, y', \cdots, y^{(k)}) = D_x\eta^{(k-1)} - y^{(k)} D_x\xi. \tag{6.8}$$

The functions ξ, η and $\eta^{(k)}$ can be written in terms of the characteristic $Q = \eta - y'\xi$ as

$$\xi = -Qy', \quad \eta = Q - y'Q_{y'}, \quad \eta^{(k)} = D_x^k Q - y^{(k+1)}Q_{y'}, \quad k \geq 1.$$

In order to find the symmetry group G admitted by a differential equation with generator

$$X = \xi(x, y)\partial_x + \eta(x, y)\,\partial_y, \tag{6.9}$$

the prolonged generator is introduced

$$X^{(n)} = \xi\partial_x + \eta\,\partial_y + \eta^{(1)}\partial_{y'} + \cdots + \eta^{(n)}\partial_{y^{(n)}}. \tag{6.10}$$

The prolonged infinitesimal generator can be used to write the linearized symmetry condition in the following form

$$X^{(n)}(y^{(n)} - \omega(x, y, y', \cdots, y^{(n-1)})) = 0,$$

when (6.1) holds. Consider diffeomorphisms of the form

$$(\hat{x}, \hat{y}) = (\hat{x}(x, y), \hat{y}(x, y))$$

corresponding to a symmetry. This type of symmetry is called a point transformation. Any point transformation that is also a symmetry is called a point symmetry. To find the Lie point symmetries of an equation of the form (6.1), the $\eta^{(k)}$ must be calculated. The functions ξ and η depend on x and y and therefore, (6.7) and (6.8) give

$$\eta^{(1)} = \eta_x + (\eta_y - \xi_x)y' - \xi_y y'^2, \tag{6.11}$$

$$\eta^{(2)} = \eta_{xx} + (2\eta_{xy} - \xi_{xx})y' + (\eta_{yy} - 2\xi_{xy})y'^2 - \xi_{yy}y'^3 + (\eta_y - 2\xi_x - 3\xi_y y')y'', \tag{6.12}$$

$$\eta^{(3)} = \eta_{xxx} + (3\eta_{xxy} - \xi_{xxx})y' + 3(\eta_{xyy} - \xi_{xxy})y'^2 + (\eta_{yyy} - 3\xi_{xyy})y'^3$$
$$-\xi_{yyy}y'^4 + 3(\eta_{xy} - \xi_{xx} + (\eta_{yy} - 3\xi_{xy})y' - 2\xi_{yy}y'^2)y'' - 3\xi_y y''^2$$
$$+(\eta_y - 3\xi_x - 4\xi_y y')y''.$$

Consider the case of second-order ordinary differential equations described by the equation

$$y'' = F(x, y, y'). \tag{6.13}$$

The linearized symmetry condition is obtained by substituting (6.11) and (6.12) into (6.5) and then replacing y'' by $F(x, y, y')$. This results in

$$\eta_{xx} + (2\eta_{xy} - \xi_{xx})y' + (\eta_{yy} - 2\xi_{xy})y'^2 - \xi_{yy}y'^3 + (\eta_y - 2\xi_x - 3\xi_y y')F$$
$$= \xi F_x + \eta F_y + (\eta_x + (\eta_y - \xi_x)y' - \xi_y y'^2)F_{y'}. \tag{6.14}$$

Both ξ and η are independent of y' and so (6.14) can be decomposed into a system of partial differential equations. These are the determining equations for Lie point symmetries.

6.2 Application to a Nonlinear Equation

As an application of the theory, a general nonlinear oscillator system will be presented [**36-39**]. The general form of the equation is

$$y'' + (\delta + \beta y^m)y' - \mu y + \alpha y^{m+1} = 0. \tag{6.15}$$

Differentiation is with respect to the independent variable x and all the coefficients δ, β, μ and α are real. Equation (6.15) is frequently referred to as the Duffing-van der Pol oscillator. When we set $\alpha = 0$ and $m = 2$, it becomes the van der Pol oscillator

$$y'' + (\delta + \beta y^2)y' - \mu y = 0.$$

Consider (6.13) written in the form (6.13),

$$y'' = -(\delta + \beta y^m)y' + \mu y - \alpha y^{m+1} = F(x, y, y'). \tag{6.16}$$

The integrability of this equation can be studied by using the Lie theory of differential equations. In order to obtain the symmetry group G which is admitted by an equation with infinitesimal operator (6.9), it is required to obtain an infinitesimal operator X_1 such that it annihilates the equation

$$X_1(y'' + (\delta + \beta y^m)y' - \mu y + \alpha y^{m+1}) = 0. \tag{6.17}$$

The operator X_1 is given by

$$X_1 = \xi(x, y)\frac{\partial}{\partial x} + \eta(x, y)\frac{\partial}{\partial y} + A(x, y, y')\frac{\partial}{\partial y'} + B(x, y, y', y'')\frac{\partial}{\partial y''} \tag{6.18}$$

and $A(x,y,y')$ and $B(x,y,y',y'')$ are obtained from $\eta^{(1)}$ and $\eta^{(2)}$ as follows,

$$A(x,y,y') = \eta_x + y'(\eta_y - \xi_x) - y^{'2}\xi_y,$$

$$B(x,y,y',y'') = \eta_{xx} + y'(2\eta_{xy} - \xi_{xx}) + y^{'2}(\eta_{yy} - 2\xi_{xy}) - y^{'3}\xi_{yy} \qquad (6.19)$$

$$+y''(\eta_y - 2\xi_x - 3y'\xi_y).$$

All $\xi(x,y)$ and $\eta(x,y)$ that verify (6.10) generate infinitesimal operators X as in (6.17) which make up the symmetries of the differential equation. Moreover, it is known that one symmetry can be used to reduce by one the order of an equation. Thus (6.16) will be integrated only if $\xi(x,y)$ and $\eta(x,y)$ are such that they generate two linearly independent infinitesimal operators.

To obtain the determining equations, we write (6.17) in the form

$$\eta_{xx} + (2\eta_{xy} - \xi_{xx})y' + (\eta_{yy} - 2\xi_{xy})y^{'2} - \xi_{yy}y^{'3}$$

$$= (2\xi_x - \eta_y + 3\xi_y y')F + \xi F_x + \eta F_y + [\eta_x + (\eta_y - \xi_x)y' - \xi_y y^{'2}]F_{y'}. \quad (6.20)$$

For this equation,

$$F_y = -\beta m y^{m-1} y' + \mu - \alpha(m+1)y^m, \qquad F_{y'} = -\delta - \beta y^m.$$

Both ξ and η are independent of y' and therefore (6.20) can be pulled apart according to the powers of y' into a system of partial differential equations which are the determining equations for the Lie point symmetries. Writing first the power of y' on the left and then the corresponding coefficient, we have

$$0 \;\; \eta_{xx} = \mu\eta - \delta\eta_x + \mu(2\xi_x - \eta_y)y - (\beta\eta_x + \alpha(m+1)\eta)y^m + \alpha(\eta_y - 2\xi_2)y^{m+1},$$

$$1 \qquad 2\eta_{xy} - \xi_{xx} = -\delta\xi_x + 3\mu\xi_y y - \beta m\eta y^{m-1} - \beta\xi_x y^m - 3\alpha\xi_y y^{m+1}$$

$$2 \qquad \eta_{yy} - 2\xi_{xy} = -2\delta\xi_y - 2\beta\xi_y y^m,$$

$$3 \qquad \xi_{yy} = 0. \qquad (6.21)$$

The equation obtained from the coefficient of $y^{'3}$ implies that

$$\xi = a(x)y + b(x). \qquad (6.22)$$

Differentiating this with respect to x and y and substituting into the $y^{'2}$ equation gives

$$\eta_{yy} - 2a'(x) = -2\delta a(x) - 2\beta a(x) y^m.$$

Integrating this twice assuming that $m \neq -1, -2$, it is found that

$$\eta = a'(x)y^2 - \delta a(x)y^2 - \frac{2}{(m+1)(m+2)}\beta a(x) y^{m+2} + c(x)y + d(x). \quad (6.23)$$

Suppose that $m = -1$, then integrating once, we have

$$\eta_y = 2a'(x)y - 2\delta a(x)y - 2\beta a(x)\ln y + c(x),$$

and so η is given by

$$\eta = a'(x)y^2 - \delta a(x)y^2 - 2\beta a(x)y\ln y + 2\beta a(x)y + c(x)y + d(x). \tag{6.24}$$

On the other hand, if $m = -2$, the integral can be done to give

$$\eta = a'(x)y^2 - \delta a(x)y^2 + 2\beta a(x)\ln y + c(x)y + d(x). \tag{6.25}$$

The functions $a(x)$, $b(x)$, $c(x)$ and $d(x)$ are arbitrary constants of integration at this point. Putting (6.22) and (6.23) into the y^0 equation, it reduces to a polynomial of degree $2m+2$ in the y variable. It is zero if and only if the following system of equation holds. These equations are obtained by equating coefficients of powers of y to zero. Below, the power is given first and then the equation:

$$\begin{aligned}
&2m+2 && \beta a'(x) - \alpha a(x) = 0,\\
&m+2 && \frac{2\beta\mu}{m+2}a + \frac{2\beta\delta}{(m+1)(m+2)}a' + \frac{2\beta a''}{(m+1)(m+2)}\\
& && -\alpha(m+1)a' + \alpha(m-1)a\delta - \beta a'' + \beta\delta a' = 0,\\
&m+1 && m\alpha c(x) + \beta c'(x) + 2\alpha b'(x) = 0,\\
&m && \beta d'(x) + \alpha(m+1)d(x) = 0,\\
&2 && a'''(x) - \mu a'(x) - \delta^2 a'(x) - \delta\mu a(x) = 0,\\
&1 && c''(x) + \delta c'(x) - 2\mu b'(x) = 0,\\
&0 && d''(x) - \mu d(x) + \delta d'(x) = 0.
\end{aligned} \tag{6.26}$$

The y' equation in turn yields the system of equations

$$\begin{aligned}
&2m+1 && 2m\beta^2 a(x) = 0,\\
&m+1 && 4\frac{\beta}{m+1}a'(x) - m\beta a'(x) + m\beta\delta a(x) - \beta a'(x) - 3\alpha a(x) = 0,\\
&m && \beta(mc(x) + b'(x)) = 0,\\
&m-1 && m\beta d(x) = 0,\\
&1 && \delta a'(x) + \mu a(x) - a''(x) = 0,\\
&0 && 2c'(x) + \delta b'(x) - b''(x) = 0.
\end{aligned} \tag{6.27}$$

It is assumed that $\alpha \neq 0$ and $\beta \neq 0$ in these equations. Then from (6.27), the y^{2m+1} equation yields $a(x) = 0$ and from the y^{m-1} equation it is clear that $d(x) = 0$. Substituting these results into both (6.26) and (6.27), these equations reduce to the set of four,

$$\begin{aligned} &m\alpha c(x) + \beta c'(x) + 2\alpha b'(x) = 0, \\ &c''(x) + \delta c'(x) - 2\mu b'(x) = 0, \\ &mc(x) + b'(x) = 0, \\ &2c'(x) + \delta b'(x) - b''(x) = 0. \end{aligned} \tag{6.28}$$

From the third equation in (6.28), it follows that $mc(x) = -b'(x)$. Substituting this into the first, an equation entirely in terms of $b(x)$ is obtained,

$$b''(x) - \frac{m\alpha}{\beta} b'(x) = 0. \tag{6.29}$$

This can be solved for $b(x)$, which in turn yields $c(x)$,

$$b(x) = -\frac{c_0 \beta}{\alpha} e^{\alpha m x/\beta} + b_0, \qquad c(x) = c_0 e^{\alpha m x/\beta}. \tag{6.30}$$

In (6.30), both b_0 and c_0 are arbitrary constants. Consider the following two cases.

(1) The first case is that with $c_0 = 0$. In this case, b_0 can be an arbitrary constant so we take $b \equiv 1$ and $c \equiv 0$, which implies that

$$\xi = 1, \qquad \eta = 0, \qquad \chi_1 = \partial_x. \tag{6.31}$$

In this case, only one infinitesimal operator is obtained, $\chi_1 = \partial_x$.

(2) Two symmetries result if it is assumed that $c_0 \neq 0$. The general solutions for $b(x)$ and $c(x)$ can be put in the remaining two equations in (6.28) to yield the system

$$(\frac{\alpha m}{\beta})^2 + \delta \frac{\alpha m}{\beta} + 2\mu m = 0, \qquad 2\frac{\alpha m}{\beta} - m\delta + \frac{\alpha m^2}{\beta} = 0. \tag{6.32}$$

Simplifying these, we have

$$m = \frac{\beta\delta}{\alpha} - 2, \qquad \frac{\alpha^2 m}{\beta^2} = -\frac{\delta\alpha}{\beta} - 2\mu.$$

Eliminating m the following result is obtained

$$\delta = \frac{\alpha}{\beta} - \frac{\mu\beta}{\alpha}. \tag{6.33}$$

This is a constraint equation which must hold between the four parameters which appear in (6.16). Since b_0 and c_0 are arbitrary constants, assume first that $b_0 = 0$ and $c_0 = 1$ so the solution becomes

$$b(x) = -\frac{\beta}{\alpha} e^{\alpha m x/\beta}, \qquad c(x) = e^{\alpha m x/\beta}. \tag{6.34}$$

Consequently, the functions which appear in the generator are

$$\xi = -\frac{\beta}{\alpha} e^{\alpha m x/\beta}, \qquad \eta = c(x) y = e^{\alpha m x/\beta}\, y.$$

In this case, two infinitesimal generators result

$$\chi_1 = \partial_x, \qquad \chi_2 = -\frac{\beta}{\alpha} e^{\alpha m x/\beta} \partial_x + e^{\alpha m x/\beta} y \partial_y. \tag{6.35}$$

Every infinitesimal generator is of the form

$$\chi = C_1 \chi_1 + C_2 \chi_2.$$

Equation (6.15) is completely integrable only when the constraint $\delta = \frac{\alpha}{\beta} - \mu \frac{\beta}{\alpha}$ holds. Otherwise, the oscillator is only partially integrable.

Now let us look at the problem of reduction to canonical variables. If an ordinary differential equation admits an infinitesimal generator, there exist a pair of variables

$$t = f(x, y), \qquad u = g(x, y),$$

called canonical variables such that f and g ($g \neq 0$) are arbitrary particular solutions of the first-order linear partial differential equations

$$\xi(x, y) \frac{\partial f}{\partial x} + \eta(x, y) \frac{\partial f}{\partial y} = \sigma, \qquad \xi(x, y) \frac{\partial g}{\partial x} + \eta(x, y) \frac{\partial g}{\partial y} = 0, \tag{6.36}$$

where σ is a nonzero constant and can be arbitrarily chosen. If the general solution of the characteristic equation

$$\frac{dx}{\xi(x, y)} = \frac{dy}{\eta(x, y)}$$

has the form $U(x, y) = C$, where C is an arbitrary constant, then the general solutions of (6.36) are expressed as

$$f(x, y) = \sigma \int \frac{dx}{\xi(x, U)} + \Phi_1(U), \quad g(x, y) = \Phi_2(U), \quad U = U(x, y), \tag{6.37}$$

where $\Phi_1(U)$ and $\Phi_2(U)$ are arbitrary functions, and U is regarded as a parameter at some point. With $k = \alpha m/\beta$ and ξ, η from χ_2 in (6.35),

$$-\frac{\alpha}{\beta} e^{-kx}\, dx = e^{-kx} \frac{dy}{y}.$$

Integrating, we have

$$U(x,y) = \ln y + \frac{\alpha}{\beta} x = C,$$

which means that f is given by integrating

$$f(x,y) = -m \int \frac{\alpha}{\beta e^{kx}}\, dx + \Phi_1(U) = e^{-\alpha m x/\beta}. \tag{6.38}$$

Here $\Phi_1(U)$ has been set to zero. Now $U(x,y)$ takes the form

$$U(x,y) = \ln(y e^{\alpha x/\beta}).$$

Consequently, taking Φ_2 to be the exponential function,

$$g(x,y) = \Phi_2(U) = y e^{\alpha x/\beta}. \tag{6.39}$$

Since $t = f(x,y)$ and $u = g(x,y)$, (6.38) and (6.39) are equivalent to

$$x = -\frac{\beta}{\alpha m} \ln\, t, \qquad y = u t^{1/m}. \tag{6.40}$$

Under this nonlinear transformation, the derivatives of y are found to be

$$\frac{\partial y}{\partial x} = -\frac{\alpha}{\beta}(m u_t t^{\frac{1}{m}+1} + u t^{\frac{1}{m}}),$$

$$\frac{\partial^2 y}{\partial x^2} = \frac{\partial t}{\partial x}\frac{\partial}{\partial t}(-\frac{\alpha}{\beta}(m u_t t^{\frac{1}{m}+1} + u t^{\frac{1}{m}})) = \frac{\alpha^2}{\beta^2}(m^2 u_{tt} t^{\frac{1}{m}+2} + m u_t (m+2) t^{\frac{1}{m}+1} + u t^{\frac{1}{m}}).$$

Returning to the oscillator equation (6.15) and substituting y and these derivatives, it becomes

$$\frac{\alpha^2}{\beta^2}(m^2 u_{tt} t^{\frac{1}{m}+2} + m u_t (m+2) t^{\frac{1}{m}+1} + u t^{\frac{1}{m}}) - \frac{\alpha}{\beta}(\delta + \beta u^m t)(m u_t t^{\frac{1}{m}+1} + u t^{\frac{1}{m}})$$
$$-\mu u t^{\frac{1}{m}} + \alpha u^{m+1} t^{1+\frac{1}{m}} = 0.$$

Collecting powers of t in this, we obtain

$$(\frac{\alpha^2 m^2}{\beta^2} u_{tt} - \alpha m u^m u_t) t^2 + (\frac{\alpha^2}{\beta^2} m(m+2) u_t - \frac{\alpha}{\beta}\delta m u_t - \alpha u^{m+1} + \alpha u^{m+1}) t$$
$$+(\frac{\alpha^2}{\beta^2} u - \frac{\alpha}{\beta}\delta u - \mu u) = 0.$$

Equating the coefficients of t to zero, the following three equations are produced,

$$\frac{\alpha m}{\beta^2} u_{tt} - u^m u_t = 0, \quad \frac{\alpha}{\beta}(m+2) u_t - \delta u_t = 0, \quad \frac{\alpha^2}{\beta^2} - \frac{\alpha}{\beta}\delta = \mu.$$

The third equation is just the parametric equation (6.33). In the second it results in

$$-\mu\frac{\beta^2}{\alpha^2} = m+1. \tag{6.41}$$

This is another constraint equation, however, it is a linear combination of two previous constraints. Finally, the first implies that

$$\frac{\alpha m}{\beta^2}u_{tt} = \frac{1}{m+1}(u^{m+1})_t.$$

Integrating both sides with respect to t yields,

$$u_t = \frac{\beta^2}{\alpha m(m+1)}u^{m+1} + I. \tag{6.42}$$

Now the reverse transformation is done,

$$\frac{\partial u}{\partial t} = -(\frac{\beta}{\alpha m}y_x + \frac{y}{m})e^{\frac{\alpha}{\beta}(m+1)x}.$$

On account of (6.42) and $u^{m+1} = y^{m+1}e^{\frac{\alpha}{\beta}(m+1)x}$, it is found that

$$(y_x + \frac{\alpha}{\beta}y + \frac{\alpha\beta}{m+1}y^{m+1})e^{\frac{\alpha}{\beta}(m+1)x} + I = 0, \tag{6.43}$$

where I is an arbitrary constant.

Suppose now that $\beta = 0$ however $\alpha \neq 0$ so that equation (6.15) reduces to

$$y'' + \delta y' - \mu y + \alpha y^{m+1} = 0. \tag{6.44}$$

The third equation of (6.28) is multiplied by β, which is zero here and so only three equations remain,

$$mc(x) + 2b'(x) = 0,$$

$$c''(x) + \delta c'(x) - 2\mu b'(x) = 0,$$

$$2c'(x) - b''(x) + \delta b'(x) = 0.$$

The first two have the solutions

$$b(x) = -c_0\frac{m+4}{2\delta}e^{\frac{\delta m}{m+4}x} + b_0, \qquad c(x) = c_0 e^{\frac{\delta m}{m+4}x}. \tag{6.45}$$

Again, there are two cases to consider.

(1) Suppose that $c_0 = 0$ and $b_0 = 1$, which is equivalent to $\xi = 1$, $\eta = 0$. Hence the only infinitesimal operator in this case is $\chi_1 = \partial_x$.

(2) Suppose that $c_0 \neq 0$, in which case the third equation implies a constraint relation on μ, namely,

$$\mu=-\frac{2m+4}{(m+4)^2}\delta^2. \tag{6.46}$$

It suffices to pick $b_0=0$ and $c_0=1$ so that

$$b(x)=-\frac{(m+4)}{2\delta}e^{\frac{\delta m}{m+4}x}, \qquad c(x)=e^{\frac{\delta m}{m+4}x},$$

which in turn implies that

$$\xi=-\frac{m+4}{2\delta}e^{\frac{\delta m}{m+4}x}, \qquad \eta=e^{\frac{\delta m}{m+4}x}y.$$

This solution leads to the following generator

$$\chi_2=-\frac{m+4}{2\delta}e^{\frac{\delta m}{m+4}x}\partial_x+e^{\frac{\delta m}{m+4}x}y\partial_y.$$

Thus, when (6.46) holds, the oscillator is completely integrable and solutions in terms of known functions can be written down.

As before, the functions f and g can be calculated

$$U(x,y)=\ln(ye^{\frac{2\delta}{m+4}x}), \qquad f(x,y)=e^{-\frac{\delta m}{m+4}x}, \qquad g(x,y)=ye^{\frac{2\delta}{m+4}x}.$$

Since $t=f(x,y)$ and $u=g(x,y)$, these are equivalent to the parametric form,

$$x=-\frac{m+4}{\delta m}\ln t, \qquad y=ut^{\frac{2}{m}}.$$

Using this nonlinear transformation, we calculate that

$$\frac{\partial y}{\partial x}=-\frac{\delta m}{m+4}u_t t^{\frac{m+2}{m}}-\frac{2\delta}{m+4}ut^{\frac{2}{m}},$$

$$\frac{\partial^2 y}{\partial x^2}=(\frac{\delta m}{m+4})^2 u_{tt}t^{2\frac{m+1}{m}}+\frac{\delta^2 m}{m+4}t^{\frac{m+2}{m}}u_t+4(\frac{\delta}{m+4})^2 t^{\frac{2}{m}}u.$$

Substituting these into the equation, it becomes

$$((\frac{2\delta}{m+4})^2-2\frac{\delta^2}{m+4}-\mu)ut^{\frac{2}{m}}+(\frac{m(m+2)}{(m+4)^2}\delta^2+\frac{2m}{(m+4)^2}\delta^2-\frac{m}{m+4}\delta^2)u_t t^{\frac{m+2}{m}}$$
$$+((\frac{m\delta}{m+4})^2u_{tt}+\alpha u^{m+1})t^{\frac{2m+2}{m}}=0.$$

The coefficient of the first term is just the condition (6.46), the second vanishes and the third reduces to

$$(\frac{m\delta}{m+4})^2u_{tt}=-\alpha u^{m+1}. \tag{6.47}$$

This is integrated to obtain

$$u_t^2 = -2(\frac{m+4}{m\delta})^2 \frac{\alpha}{m+2} u^{m+2} + I.$$

Now

$$\frac{\partial u}{\partial t} = -(\frac{m+4}{\delta m} y' - \frac{2}{m} y) e^{\frac{m+2}{m+4}\delta x}.$$

and so substituting this into u_t^2, a first integral of the equation is obtained when (6.46) holds,

$$[(\frac{m+4}{\delta m})^2 y'^2 + (\frac{2}{m})^2 y^2 + 4\frac{m+4}{m^2\delta} yy' + 2\frac{\alpha(m+4)^2}{m^2\delta^2(m+2)} y^{m+2}] e^{2\frac{\delta(m+2)}{m+4}x} = I. \tag{6.48}$$

Two special cases of this first integral will be mentioned now.

The choice $\beta = 0$ and $m = 1$ leads to the damped Helmholtz oscillator equation

$$y'' + \delta y' - \mu y + \alpha y^2 = 0.$$

Putting $m = 1$ in (6.48), a first integral of the Helmholtz oscillator is obtained,

$$(\frac{25}{\delta^2} y'^2 + 4y^2 + \frac{20}{\delta} yy' + \frac{50\alpha}{3\delta^2} y^3) e^{\frac{6}{5}\delta x} = I,$$

provided that $\mu = -\frac{6}{25}\delta^2$.

Picking $\beta = 0$ and $m = 2$ leads to the damped Duffing equation

$$y'' + \delta y' - \mu y + \alpha y^3 = 0.$$

Substituting $m = 2$ into (6.48), a first integral of the Duffing equation results,

$$(\frac{9}{\delta^2} y'^2 + y^2 + \frac{6}{\delta} yy' + \frac{9\alpha}{2\delta^2} y^4) e^{\frac{4}{3}\delta x} = I,$$

provided that $\mu = -\frac{2}{9}\delta^2$.

7 Exterior Differential Systems and Wahlquist-Estabrook Prolongations.

7.1 Introduction.

Once nonlinear terms are included in linear dispersive equations, solitary waves can result which can be stable enough to persist indefinitely. It is well known that many important nonlinear evolution equations which have numerous applications in mathematical physics appear as sufficient conditions for the integrability of systems of linear partial differential equations of first order, and such systems are referred to as integrable [**40,41**].

This provides an excellent opportunity to develop a very useful, interesting application for many of the theoretical topics which have been introduced so far [**42**].

To introduce the basic ideas, there exists an immersion map, or submanifold, of the differentiable manifold spanned by all the variables of a specified differential system. Independent variables constitute a subset which span the submanifold, so a solution of a partial differential system can be thought of as a map of those independent variables onto the rest of the variables, or dependent variables. Thus, by requiring a given set of differential forms which constitute a differential system to be identically zero when so restricted, certain immersion maps can be distinguished, in particular the nonlinear equation itself. The fact that the equation results is reflected in the choice of the differential system in the first place. They are solutions as well of a particular set of coupled first-order partial differential equations. Such submanifolds have come to called integral submanifolds of the set of differential forms.

A very important extension of this idea is the Cartan-Wahlquist-Estabrook prolongation technique. Wahlquist and Estabrook [**43-46**] first constructed prolongations for the KdV and other systems. This procedure produces a nonclosed Lie algebra of vector fields in general which are defined on fibres above the base manifold that supports the exterior differential system defining the equation. Moreover, the vanishing of the curvature form of a Cartan-Ehresmann connection is the necessary and sufficient condition for the existence of the prolongation. The prolongations which this method generates have many useful applications, such as generating Lax pairs and Bäcklund transformations. Several applications of this theory to a variety of nonlinear equations have already been done [**47-51**]. An application to a coupled system of equations will be presented at the end of this section.

7.2 Theoretical Introduction.

Let M be a manifold which has dimension m. The case in which $M = \mathbb{R}^m$ with coordinates $(u_1, u_2, \cdots, u_m)$ is of particular interest here. Let a closed exterior differential system be defined,

$$\xi^1 = 0, \cdots, \xi^l = 0. \tag{7.1}$$

Let $\mathcal{I}$ be the ideal generated by the forms $\{\xi^i\}_1^l$,

$$\mathcal{I} = \{\omega = \sum_{i=1}^{l} \eta^i \wedge \xi^i, \xi^i \in \Lambda(M)\}. \tag{7.2}$$

The exterior system is chosen such that the solutions of an evolution equation correspond with two-dimensional integral manifolds of the differential system It must also be closed, so $d\mathcal{I} \subset \mathcal{I}$, so that the differential system is integrable.

Definition 7.1. A Cartan-Ehresmann connection in the fibre bundle $(\tilde{M}, \tilde{\pi}, M)$ is a system of one-forms $\tilde{\omega}^i$, $i = 1, \cdots, k$ in $T^*(\tilde{M})$ with the property that the mapping $\tilde{\pi}_*$ from the vector space $H_{\tilde{m}} = \{\tilde{X} \in T_{\tilde{m}} | \tilde{\omega}^i(\tilde{X}) = 0, i = 1, \cdots, k\}$ to the tangent space T_m is a bijection.

A version of the Wahlquist-Estabrook method will be presented. Suppose M is a manifold as above with a projection map $\pi : M \to \mathbb{R}^2$ which is defined by $\pi(u_1, \cdots, u_m) = (u_1, u_2)$ and $\mathcal{I}(\xi^i)$ the differential ideal of forms on M generated by the $\{\xi^i\}_1^l$. It will be the case that $\mathcal{I}$ are chosen such that $d\mathcal{I} \subset \mathcal{I}$. As a consequence, the Frobenius Theorem indicates that the Pfaff system $\{\xi^i\}$ is completely integrable. In applications which involve the nonlinear equations discussed here, u_1, u_2 will be identified with the independent x, t variables of the equation and u_3 the dependent variable. If more than one equation is considered, further dependent variables can be introduced, as in the case considered at the end of this chapter. The system $\{\xi^i\}$ is constructed in such a way that solutions $u = u(x, t)$ of an evolution equation correspond to the two-dimensional transversal integral manifolds.

Suppose $N \subset \mathbb{R}^2$ is coordinatized by the variables (x, t) and $\pi : M \to N$ such that $s : N \to M$ is a cross section of π. The integral manifolds can then be written as sections S in M specified by

$$s(x, t) = (x, t, u(x, t), \cdots, u_m(x, t)). \tag{7.3}$$

A bundle can now be constructed based on M so that $\tilde{M} = M \times \mathbb{R}^n$ and $B = (\tilde{M}, \tilde{\pi}, M)$. The $\mathbb{R}^n$ factor is coordinatized by means of the coordinates $\mathbf{y} = (y_1, \cdots, y_n)$, and whose number may be left undetermined at this point. The $\mathbf{y}$ will be referred to as the prolongation variables; everything so far can be lifted up to $\tilde{M}$. Thus, consider the exterior differential system in $\tilde{M}$ specified by

$$\tilde{\xi}^i = \tilde{\pi}^* \xi^i = 0, \quad i = 1, \cdots, l, \quad \tilde{\omega}^j = 0, \quad j = 1, \cdots, n. \tag{7.4}$$

The forms $\{\tilde{\omega}^j\}$ have been included in the set in order to specify a type of Cartan-Ehresmann connection on the bundle B. System (7.4) is called a Cartan prolongation if it is closed and whenever S is a transversal solution of $\mathcal{I}$ there should also exist a transverse local solution $\tilde{S}$ of (7.4) with $\tilde{\pi}(\tilde{S}) = S$. There is a Theorem to the effect that (7.4) is a Cartan prolongation of $\mathcal{I}$ if and only if (7.4) is closed.

There are a number of ways to state the definition of a connection, for example, Definition 7.1. Also, a Cartan-Ehresmann connection on B can be regarded as a field H of horizontal contact elements on $\tilde{M}$ which is supplementary to the field V of the π-vertical contact elements. Also H is assumed complete, so every complete vector field X on M has a complete horizontal lift $\tilde{X}$ on M. The ideal $\tilde{I}$ of differential forms on $\tilde{M}$, which is generated by $\tilde{\pi}^* \mathcal{I} \cup H^*$ determines on $\tilde{M}$ the exterior differential system, which we continue to write as $\{\xi^i = 0\}$ in the following. Thus H^* is the set of one-forms on $\tilde{M}$ which vanish on the field H.

It remains to specify the form of the connection explicitly. The method here lets us keep the dimension of the space of $\mathbf{y}$ variables undetermined until a representation for the algebra is fixed at the end. In terms of the coordinates of the bundle, the connection forms are designated to have the general form

$$\tilde{\omega}^k = dy^k - F^k(u_1, \cdots, u_m, \mathbf{y})\, dt - G^k(u_1, \cdots, u_m, \mathbf{y})\, dx \equiv dy^k - \eta^k, \quad (7.5)$$

where $k = 1, \cdots, n$. The idea then is to include the forms $\tilde{\omega}^k$ so as to enlarge the initial differential ideal of forms.

The integrability condition requires that the prolonged ideal $\{\xi^i, \tilde{\omega}^k\}$ remains closed. This implies that the exterior derivatives of the $\tilde{\omega}^k$ can be expressed in the form,

$$d\tilde{\omega}^k = \sum_{j=1}^{l} f^{kj} \xi^j + \sum_{j=1}^{n} \eta^{kj} \wedge \tilde{\omega}_j. \quad (7.6)$$

The f^{kj} in (7.5) represent dependent functions of the bundle coordinates and the η^{ki} represent a matrix of one-forms.

For a connection such as (7.5), the prolongation condition can be expressed equivalently using the summation convention over repeated indices as follows,

$$-d\eta^i = \frac{\partial \eta^i}{\partial y^j} \wedge (dy^j - \eta^j), \qquad \mod \tilde{\pi}^*(\mathcal{I}). \quad (7.7)$$

This result can be rewritten by using the identity

$$d\eta^i = d_M \eta^i - (\frac{\partial \eta^i}{\partial y^j}) \wedge dy^j.$$

Here d_M refers to differentiation with respect to the variables of the base manifold. The prolongation condition then becomes

$$d_M \eta^i - (\frac{\partial \eta^i}{\partial y^j}) \wedge \eta^j = 0, \qquad \mod \tilde{\pi}^*(\mathcal{I}). \quad (7.8)$$

Introduce the vertical valued one-form as well as the following definitions

$$\eta = \eta^i \frac{\partial}{\partial y^i}, \qquad d\eta = (d_M \eta^i) \frac{\partial}{\partial y^i}, \quad [\eta, \tau] = (\eta^j \wedge \frac{\partial \tau^i}{\partial y^j} + \tau^j \wedge \frac{\partial \eta^i}{\partial y^j}) \frac{\partial}{\partial y^i}.$$

The prolongation condition then takes the form

$$d\eta + \frac{1}{2}[\eta, \eta] = 0, \qquad \mod \tilde{\pi}^*(\mathcal{I}). \tag{7.9}$$

A particular version of connection form (7.5) which will allow the dimension of the **y** prolongation variables to remain unspecified until a representation is specified for the algebra and is suited to writing Lax pairs is

$$\tilde{\Omega}^k = dy^k - \eta^k = dy^k - \sum_{i=1}^{n} F^{ki}(\mathbf{u})y^i \, dt - \sum_{i=1}^{n} G^{ki}(\mathbf{u})y^i \, dx. \tag{7.10}$$

The commutator can be simplified using (7.9). It is given explicitly as

$$[\eta, \eta] = (G^{ji}F^{\nu j}y^i \, dx \wedge dt + F^{ji}G^{\nu j}y^i \, dt \wedge dx$$

$$+F^{ji}G^{\nu j}y^i \, dt \wedge dx + G^{ji}F^{\nu j}y^i \, dx \wedge dt)\frac{\partial}{\partial y^\nu} = 2[F, G]^{\nu i}y^i \frac{\partial}{\partial y^\nu} \, dx \wedge dt. \tag{7.11}$$

The prolongation condition takes the form

$$(\frac{\partial F^{\nu i}}{\partial u_j} \, du_j \wedge dt + \frac{\partial G^{\nu i}}{\partial u_j} \, du_j \wedge dx)y^i \frac{\partial}{\partial y^\nu} + [F, G]^{\nu i} yi \frac{\partial}{\partial y^\nu} \, dx \wedge dt = 0, \qquad \mod \tilde{\pi}^*(\mathcal{I}). \tag{7.12}$$

If the ideal of forms is specified by the system of two forms $\{\xi^i\}$ closed over $\mathcal{I}$, then (7.12) takes the equivalent form

$$(\frac{\partial F^{\nu i}}{\partial u_j} \, du_j \wedge dt + \frac{\partial G^{\nu i}}{\partial u_j} \, du_j \wedge dx) + [F, G]^{\nu i} \, dx \wedge dt \equiv \lambda_j^{\nu i} \xi^j. \tag{7.13}$$

The objective in any given case is to produce the forms $\{\xi^j\}$ which generate the differential ideal $\mathcal{I}$ relevant to the equation and then solve (7.13) for the components of the connection $F^{\nu i}$ and $G^{\nu i}$. In effect, the following theorem has been established.

Theorem 7.1. Each prolongation of Pfaffian system $\{\xi^i = 0\}$ which corresponds to a nonlinear equation on the integral manifold by a Cartan-Ehresmann connection determines a geometrical realization of a Wahlquist-Estabrook partial Lie algebra $\mathcal{L}$ by solving (7.13). Conversely, every geometrical realization of $\mathcal{L}$ corresponds to such a prolongation by constructing (7.5). Moreover, on a two-dimensional solution submanifold of the differential ideal, the one-forms are annihilated and there exists the differential Lax pair given by

$$\mathbf{y}_x = -F\mathbf{y}, \qquad \mathbf{y}_t = -G\mathbf{y}. \tag{7.14}$$

The results of Theorem 7.1 are of use in making Lax pairs once F and G have been determined. Bäcklund transformationscan also be found from these results.

7.3 An Application to a Coupled KdV System.

The intention in this section is to apply the formalism just discussed to a nontrivial system of coupled nonlinear equations. The prolongation structure of a coupled KdV system will be studied. It will be shown that the prolongation structure of the system can be determined, so the system is integrable. A matrix spectral problem can be constructed as well.

Define the manifold $M = \mathbb{R}^8(x,t,u,v,p,q,z,r)$ over which the exterior differential system

$$\begin{aligned}
\xi^1 &= du \wedge dt - p\, dx \wedge dt, \\
\xi^2 &= dp \wedge dt - q\, dx \wedge dt, \\
\xi^3 &= dv \wedge dt - z\, dx \wedge dt, \\
\xi^4 &= dz \wedge dt - r\, dx \wedge dt, \\
\xi^5 &= du \wedge dx + (-vz + \tfrac{5}{4}pv + \tfrac{5}{4}uz - \tfrac{7}{4}up)\, dt \wedge dx - dq \wedge dt, \\
\xi^6 &= dv \wedge dx + (-\tfrac{5}{2}up + 2pv + 2uz - \tfrac{7}{4}vz)\, dt \wedge dx - dr \wedge dt.
\end{aligned} \tag{7.15}$$

By straightforward differentiation, it is determined that

$$\begin{aligned}
d\xi^1 &= dx \wedge \xi^2, \\
d\xi^2 &= -dx \wedge \xi^5, \\
d\xi^3 &= dx \wedge \xi^4, \\
d\xi^4 &= -dx \wedge \xi^6, \\
d\xi^5 &= dx \wedge [-(\tfrac{5}{4}z + \tfrac{7}{4}p)\xi^1 + (\tfrac{5}{4}v + \tfrac{7}{4}u)\xi^2 + (\tfrac{5}{4}p - z)\xi^3 - (v + \tfrac{5}{4}u)\xi^4] \\
d\xi^6 &= dx \wedge [(2z - \tfrac{5}{2}p)\xi^1 + (2v - \tfrac{5}{2}u)\xi^2 + (2p - \tfrac{7}{4}z)\xi^3 + (2u - \tfrac{7}{4}v)\xi^4].
\end{aligned} \tag{7.16}$$

Therefore, the ideal defined by $\mathcal{I} = \{\omega = \sum_{i=1}^{6} \eta^i \wedge \xi^i : \eta^i \in \Lambda(M)\}$ is therefore closed, so $d\mathcal{I} \subset \mathcal{I}$. Let s be a section of the projection $\pi(x,t,u,v,p,q,z,r) = (x,t)$, then the transverse integral manifolds are given by

$$s(x,t) = (x,t,u(x,t),v(x,t),\cdots).$$

On system (7.15), this gives

$$0 = \xi^1|_S = s^*\xi^1 = (u_x - p)\, dx \wedge dt,$$

$$0 = \xi^2|_S = s^*\xi^2 = (p_x - q)\, dx \wedge dt,$$

$$0 = \xi^3|_S = s^*\xi^3 = (v_x - z)\,dx \wedge dt,$$

$$0 = \xi^4|_S = s^*\xi^4 = (z_x - r)\,dx \wedge dt,$$

$$0 = \xi_5|_S = s^*\xi^5 = (u_t - vz + \tfrac{5}{4}pv + \tfrac{5}{4}uz - \tfrac{7}{4}up + q_x)dt \wedge dx, \tag{7.17}$$

$$0 = \xi_6|_S = s^*\xi^6 = (v_t - \tfrac{5}{2}up + 2pv + 2uz - \tfrac{7}{4}vz + r_x)\,dt \wedge dx.$$

On the transverse integral manifold, it follows that

$$p = u_x, \quad q = p_x = u_{xx}, \quad z = v_x, \quad r = z_x = v_{xx}. \tag{7.18}$$

These transverse integral manifolds give solutions to the coupled KdV system given by,

$$u_t + u_{xxx} - \tfrac{7}{4}uu_x + \tfrac{5}{4}u_x v + \tfrac{5}{4}uv_x - vv_x = 0,$$
$$v_t + v_{xxx} - \tfrac{5}{2}uu_x - \tfrac{7}{4}vv_x + 2u_x v + 2uv_x = 0. \tag{7.19}$$

Substituting (7.15) into (7.13), there results the expression,

$$F_u\,du \wedge dt + F_v\,dv \wedge + F_p\,dp \wedge dt + F_q dq \wedge dt + F_z dz \wedge dt + F_r\,dr \wedge dt$$

$$G_u\,du \wedge dx + G_v\,dv \wedge dx + G_p dp \wedge dx + G_q\,dq \wedge dx + G_z\,dz \wedge dx + G_r\,dr \wedge dx$$

$$+[F,G]\,dx\wedge dt = \lambda_1(du\wedge dt - p\,dx\wedge dt) + \lambda_2(dp\wedge dt - q dx\wedge dt) + \lambda_3(dv\wedge dt - z\,dx\wedge dt)$$

$$+\lambda_4(dz\wedge dt - r\,dx\wedge dt) + \lambda_5(du\wedge dx + (-vz + \frac{5}{4}pv + \frac{5}{4}uz - \frac{7}{4}up)dt\wedge dx - dq\wedge dt)$$

$$+\lambda_6(dv \wedge dx + (-\frac{5}{2}up - \frac{7}{4}vz + 2pv + 2uz)\,dt \wedge dx - dr \wedge dt).$$

Equating the coefficients of the two-forms on both sides gives the system,

$$F_u = \lambda_1, \quad F_v = \lambda_3, \quad F_p = \lambda_2, \quad F_q = -\lambda_5, \quad F_z = \lambda_4, \quad F_r = -\lambda_6,$$

$$G_u = \lambda_5, \qquad G_v = \lambda_6, \qquad G_p = 0, \qquad G_q = 0, \qquad G_z = 0, \qquad G_r = 0,$$

$$[F,G] = -p\lambda_1 - q\lambda_2 - z\lambda_3 - r\lambda_4 - \lambda_5(-vz + \frac{5}{4}pv + \frac{5}{4}uz - \frac{7}{4}up)$$

$$-\lambda_6(-\frac{5}{2}up - \frac{7}{4}vz + 2pv + 2uz).$$

Eliminating the set of λ_i from this set of equations, it reduces to

$$G_p = G_q = G_z = G_r = 0, \qquad F_q + G_u = 0, \qquad F_r + G_v = 0,$$

$$pF_u + qF_p + zF_v + rF_z - F_q(-vz + \frac{5}{4}pv + \frac{5}{4}uz - \frac{7}{4}up) - F_r(2pv + 2uz - \frac{5}{2}up - \frac{7}{4}rz)$$
$$+[F,G] = 0. \tag{7.20}$$

The first four derivatives of G ensure that G is independent of p, q, z, r, but can depend on u and v. Thus we can take G of the form

$$G = uX_1 + vX_2 + X_3. \tag{7.21}$$

The X_i are noncommuting generators which do not depend on the bundle coordinates and may be given an explicit representation at the end by embedding in an algebra. By means of the next two equations, the derivatives $F_q = -X_1$ and $F_r = -X_2$ are determined, and so F has the general structure

$$F = -qX_1 - rX_2 + H(u, v, p, z). \tag{7.22}$$

Substituting (7.21) and (7.22) into the final equation of system (7.20), it remains to solve the expression

$$pH_u + qH_p + zH_v + rH_z + X_1(\tfrac{5}{4}pv + \tfrac{5}{4}uz - vz - \tfrac{7}{4}up)$$
$$+X_2(2pv + 2uz - \tfrac{5}{2}up - \tfrac{7}{4}vz) - qv[X_2, X_1] - q[X_3, X_1] - ur[X_1, X_2]$$
$$-r[X_3, X_2] + [uX_1 + rX_2 + X_3, H] = 0. \tag{7.23}$$

The first step is to obtain the coefficients of q and r and set them to zero. This gives the pair of equations $H_p + v[X_1, X_2] - [X_3, X_1] = 0$ and $H_z - u[X_1, X_2] - [X_3, X_2] = 0$. Integrating these two results, H is found to be of the form,

$$H = -vp[X_1, X_2] + p[X_3, X_1] + uz[X_1, X_2] + z[X_3, X_2] + \tilde{H}(u, v).$$

Define the new generators $X_4 = [X_1, X_2]$, $X_5 = [X_3, X_1]$ and $X_6 = [X_3, X_2]$ so that H takes the form,

$$H = -pvX_4 + pX_5 + uzX_4 + zX_6 + \tilde{H}(u, v). \tag{7.24}$$

At this point H in (7.24) can be replaced back in (7.23). This gives two equations to determine $\tilde{H}$ by isolating the coefficients of p and z. Doing so gives two coupled equations,

$$\tilde{H}_u + [-\frac{7}{4}X_1 - \frac{5}{2}X_2 + [X_1, X_5])u + [X_3, X_5] + (\frac{5}{4}X_1 + 2X_2 - [X_3, X_4] + [X_2, X_5])v = 0,$$

$$\tilde{H}_v + (\frac{5}{4}X_1 + 2X_2 + [X_3, X_4] + [X_1, X_6])u + [X_3, X_6](-X_1 - \frac{7}{4}X_2 + [X_2, X_6])v = 0.$$

To get a unique solution, these two equations must be rendered compatible. To accomplish this, new generators X_8, X_9 and X_{10} are introduced which satisfy,

$$2X_8 - \tfrac{7}{4}X_1 - \tfrac{5}{2}X_2 + [X_1, X_5] = 0, \qquad 2X_9 - X_1 - \tfrac{7}{4}X_2 + [X_2, X_6] = 0,$$

$$X_{10} + \tfrac{5}{4}X_1 + 2X_2 - [X_3, X_4] + [X_2, X_5] = 0, \qquad X_{10} + \tfrac{5}{4}X_1 + 2X_2 + [X_3, X_4] + [X_1, X_6] = 0.$$

Moreover, defining $X_{11} = [X_5, X_3]$ and $X_{12} = [X_6, X_3]$, the two equations above integrate to the form

$$\tilde{H} = X_7 + u^2 X_8 + v^2 X_9 + uvX_{10} - uX_{11} - vX_{12}. \tag{7.25}$$

In (7.25), X_7 is an arbitrary generator obtained from the integration. Substituting (7.25) back into (7.23), the coefficients of the remaining powers of u and v specify all the remaining brackets. To make the presentation condensed, all of the brackets for the algebra $\mathcal{A}$, or integrability conditions, obtained thus far will be summarized below,

$$\begin{gathered} X_4 = [X_1, X_2], \quad X_5 = [X_3, X_1], \quad X_6 = [X_3, X_2], \\ [X_1, X_8] = [X_2, X_9] = [X_1, X_4] = [X_2, X_4] = [X_3, X_7] = 0, \\ [X_1, X_7] + [X_3, X_{11}] = 0, \quad [X_2, X_7] + [X_3, X_{12}] = 0, \quad [X_3, X_8] + [X_1, X_{11}] = 0, \\ [X_1, X_9] + [X_2, X_{10}] = 0, \\ [X_3, X_9] + [X_2, X_{12}] = 0, \quad [X_1, X_{10}] + [X_2, X_8] = 0, \\ [X_3, X_{10}] + [X_1, X_{12}] + [X_2, X_{11}] = 0, \\ X_{11} = [X_5, X_3], \qquad X_{12} = [X_6, X_3], \\ \frac{5}{4}X_1 + 2X_2 + [X_2, X_5] - [X_3, X_4] + X_{10} = 0, \quad 2X_8 - \frac{7}{4}X_1 - \frac{5}{2}X_2 + [X_2, X_5] = 0, \\ 2X_9 - X_1 - \frac{7}{4}X_2 + [X_2, X_6] = 0, \quad \frac{5}{4}X_1 + 2X_2 + [X_1, X_6] + [X_3, X_4] + X_{10} = 0. \end{gathered} \tag{7.26}$$

These generators X_i determine an incomplete Lie algebra called the prolongation algebra. An explicit matrix Lax representation can be found in terms of $n \times n$ matrices. Moreover, the final form for F is given by

$$\begin{aligned} F = -qX_1 - rX_2 - pvX_4 + pX_5 + uzX_4 + zX_6 + X_7 + u^2 X_8 + v^2 X_9 \\ + uvX_{10} - uX_{11} - vX_{12}. \end{aligned} \tag{7.27}$$

With F and G given by (7.21) and (7.27), the structure of a Lax pair is given in Theorem 7.1.

A spectral parameter can be introduced by making use of the scale symmetry of the coupled pair. The symmetry is given by

$$x \to \lambda^{-1} x, \qquad t \to \lambda^{-3} t, \qquad u \to \lambda^2 u, \quad v \to \lambda^2 v,$$

which leads to an automorphism of the prolongation algebra. To establish a matrix representation of the algebra $\mathcal{A}$, the approach of Dodd and Fordy **[52]** can be used to embed the elements of $\mathcal{A}$ in a simple Lie algebra. The generators $\{X_1, \cdots, X_{10}\}$ which span $\mathcal{A}$ will be taken as

$$X_1 = \begin{pmatrix} 0\,0\,0\,0 \\ 1\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,1\,0 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 1 & \frac{1}{12} & 0 & 0 \\ \lambda & 0 & 0 & 0 \\ 0 & \frac{5}{12} & 0 & -\frac{1}{8} \\ 0 & 0 & \lambda & 0 \end{pmatrix},$$

$$X_4 = \begin{pmatrix} 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \\ 0\,0\,0\,0 \end{pmatrix} \quad X_5 = \begin{pmatrix} \frac{1}{12} & 0 & 0 & 0 \\ 0 & -\frac{1}{12} & 0 & 0 \\ \frac{5}{12} & 0 & -\frac{1}{8} & 0 \\ 0 & -\frac{5}{12} & 0 & \frac{1}{8} \end{pmatrix}, \quad X_6 = \begin{pmatrix} -\frac{1}{24} & 0 & 0 & 0 \\ 0 & \frac{1}{24} & 0 & 0 \\ -\frac{1}{3} & 0 & \frac{1}{8}0 & 0 \\ 0 & \frac{1}{3} & 0 & -\frac{1}{8} \end{pmatrix},$$

$$X_7 = \begin{pmatrix} 0 & -\frac{\lambda}{36} & 0 & 0 \\ -\frac{\lambda^2}{3} & 0 & 0 & 0 \\ 0 & \frac{5\lambda}{72} & 0 & -\frac{\lambda}{16} \\ -\frac{5\lambda^2}{3} & 0 & \frac{\lambda^2}{2} & 0 \end{pmatrix}, \quad X_8 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{5}{6} & 0 & -\frac{1}{4} & 0 \end{pmatrix}, \quad X_9 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ \frac{1}{24} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{7}{12} & 0 & -\frac{1}{4} & 0 \end{pmatrix}, \tag{7.28}$$

$$X_{10} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{4}{3} & 0 & \frac{1}{2} & 0 \end{pmatrix}, \quad X_{11} = \begin{pmatrix} 0 & \frac{1}{72} & 0 & 0 \\ -\frac{\lambda}{6} & 0 & 0 & 0 \\ 0 & -\frac{5}{144} & 0 & \frac{1}{32} \\ -\frac{5\lambda}{6} & 0 & \frac{\lambda}{4} & 0 \end{pmatrix}, \quad X_{12} = \begin{pmatrix} 0 & -\frac{1}{144} & 0 & 0 \\ \frac{\lambda}{12} & 0 & 0 & 0 \\ 0 & \frac{7}{144} & 0 & -\frac{1}{32} \\ \frac{2\lambda}{3} & 0 & -\frac{\lambda}{4} & 0 \end{pmatrix}.$$

Clearly, coupled KdV system (7.19) will have a nontrivial prolongation structure under the representation. Consequently, the dimension of the space of $\mathbf{y}$ variables is four so that

$$\mathbf{y} = \begin{pmatrix} y^1 \\ y^2 \\ y^3 \\ y^4 \end{pmatrix}$$

A matrix representation for F and G can be obtained by substituting (7.28) into equations (7.21) and (7.27). For G, it is found that

$$G = \begin{pmatrix} 0 & \frac{1}{12} & 0 & 0 \\ u - \frac{v}{2} + \lambda & 0 & 0 & 0 \\ 0 & \frac{5}{12} & 0 & -\frac{1}{8} \\ v & 0 & u - v + \lambda & 0 \end{pmatrix}. \tag{7.29}$$

The matrix form of F in (7.27) under algebra (7.28) is given by

$$F = \begin{pmatrix} \frac{p}{12} - \frac{z}{24} & -\frac{u}{72} + \frac{v}{144} - \frac{\lambda}{36} & 0 & 0 \\ F_{12} & -\frac{p}{12} + \frac{z}{24} & 0 & 0 \\ \frac{5}{12}p - \frac{z}{3} & \frac{5}{144}u - \frac{7}{144}v + \frac{5\lambda}{72} & -\frac{p}{8} + \frac{z}{8} & -\frac{\lambda}{16} - \frac{u}{32} + \frac{v}{32} \\ F_{14} & -\frac{5}{12}p + \frac{z}{3} & F_{34} & \frac{p}{8} - \frac{z}{8} \end{pmatrix}. \tag{7.30}$$

In (7.30), p, q, z and r can be replaced by the results in (7.17). and

$$F_{12} = -u_{xx} + \frac{1}{2}v_{xx} + \frac{1}{24}(2u - v)^2 - \frac{\lambda}{12}(v - 2u + 4\lambda),$$

$$F_{14} = \frac{5}{6}u^2 - v_{xx} + \frac{7}{12}v^2 - \frac{4}{3}uv + \frac{\lambda}{6}(5u - 4v - 10\lambda),$$

$$F_{34} = v_{xx} - u_{xx} + \frac{\lambda}{4}(-u + v + 2\lambda) - \frac{1}{4}(u - v)^2.$$

The Lax pair is composed of the equations specified in (7.14).

8 Surfaces of Arbitrary Constant Negative Gaussian Curvature and the Related Sine-Gordon Equation

8.1 Introduction.

A great deal of the development of the ideas pertaining to nonlinear equations, surfaces and solitons had their origins in investigations concerning the sine-Gordon equation **[2,4]**. The study of surfaces with constant Gaussian curvature dates back to E. Bour **[3]**, who in 1862 generated the particular form

$$\omega_{uv} = \frac{1}{\rho^2} \sin(\omega), \tag{8.1}$$

where $K = -1/\rho^2$. Although much has been written with regard to this system, it seems to invariably return to the case in which $K = -1$. Surfaces with constant Gaussian curvature are of great interest **[51]**. The intention here is to produce an equation similar in form to (8.1) of sine-Gordon type such that K for the corresponding surface is negative but arbitrary. Such an equation will be referred to as a deformed sine-Gordon equation and the discussion can be thought to pertain to any two-dimensional manifold which can be embedded in $\mathbb{R}^3$. In fact, any compact, smooth two-manifold can be embedded smoothly in $\mathbb{R}^3$. This enables the use of the natural metric $\langle , \rangle$ on $\mathbb{R}^3$ so that lengths can be calculated as well as angles between normals in order that the formalism of a line congruence can be invoked. A two parameter family of lines in $\mathbb{R}^3$ or $\mathbb{R}^{2+1}$ forms a line congruence, and all normal lines of a surface form a line congruence called a normal line congruence. A line congruence can be expressed by writing $\mathbf{Y} = \mathbf{X}(u,v) + \lambda \mathbf{q}(u,v)$, $\langle \mathbf{q}\,, \mathbf{q} \rangle = 1$. For fixed parameters u, v, this represents a straight line passing through $\mathbf{X}(u,v)$ in the direction $\mathbf{q}(u,v)$. This then is a two parameter family of straight lines, or a line congruence. This idea appears in a formulation of Bäcklund's theorem which will be invoked to aid in establishing the claims which are formulated, as well as the fundamental equations for a two-manifold or surface.

Suppose that S and S^* are two focal surfaces of a line congruence, and PP^* is the line in the congruence and the common tangent line of the two surfaces, so $P \in S$ and $P^* \in S^*$. Suppose that e_3, e_3^* are the normal vectors at points P and P^* to S and S^*, respectively. Finally, let τ be the angle between e_3 and e_3^*, so $\langle e_3, e_3^* \rangle = \cos\tau$, and let l be the distance between P and P^*. Proposition 8.1 will be invoked as needed.

Proposition 8.1. (Bäcklund's Theorem) Suppose that S and S^* are two focal surfaces of a pseudo-spherical congruence in $\mathbb{R}^3$, the distance between the corresponding points is constant and denoted l. The angle between the corresponding normals is a constant τ. Then these two focal surfaces S and S^* have the same negative constant Gaussian curvature

$$K = -\frac{\sin^2 \tau}{l^2}. \tag{8.2}$$

Thus, from any solution of the sine-Gordon or deformed sine-Gordon equation, a corresponding surface of negative constant curvature can be obtained. It is the latter case that is elucidated here.

On the other hand, from the Bäcklund theorem, it is known that two focal surfaces of a pseudospherical congruence are surfaces with the same negative constant curvature. These two focal surfaces will correspond to two solutions of the deformed sine-Gordon equation to appear. It will be seen that a relation can be established between the two solutions from the Bäcklund theorem, or equivalently, from the correspondence between two focal surfaces of a pseudo-spherical line congruence. This will be enough to give a Bäcklund transformation for this new deformed sine-Gordon equation.

8.2 Development of the Equation and Bäcklund Transformation.

Suppose S and S^* are two focal surfaces with arbitrary constant negative curvature K such that $\{P, e_1, e_2, e_3\}$ is a frame [**13**] corresponding to coordinates of surface S with

$$\omega_1 = \cos\frac{\alpha}{2}\,du \qquad \omega_2 = \sin\frac{\alpha}{2}\,dv,$$

$$\omega_{13} = \sin\frac{\alpha}{2}\,du \qquad \omega_{23} = -\cos\frac{\alpha}{2}\,du, \tag{8.3}$$

$$\omega_{12} = \tfrac{1}{2}(\alpha_v\,du + \alpha_u\,du) = -\omega_{21}.$$

These forms completely specify the set dr, de_1, de_2, de_3 in the fundamental equations given that $\omega_{ij} + \omega_{ji} = 0$.

Suppose

$$x^* = x + le = x + l(\cos\vartheta e_1 + \sin\vartheta e_2), \tag{8.4}$$

form a pseudo-spherical line congruence and ϑ is to be specified. In (8.4), x and x^* correspond to the surfaces S and S^*, l is the distance between the corresponding points P and P^* on the surfaces S and S^*, e is in the direction of PP^* and ϑ is the angle between e and e_1. Suppose S corresponds to a solution α of the deformed sine-Gordon equation to be obtained and α' a second solution. The fundamental equations for S are given by

$$dx = \omega_1 e_1 + \omega_2 e_2, \qquad \omega_3 = 0,$$

$$de_1 = \omega_{12}e_2 + \omega_{13}e_3, \quad de_2 = \omega_{21}e_1 + \omega_{23}e_3, \quad de_3 = \omega_{31}e_1 + \omega_{32}e_2. \tag{8.5}$$

The fundamental equations for S^* are the same as (8.5), but with star on each quantity. By exterior differentiation of (8.4), it is found that

$$dx^* = dx + l(\cos\vartheta\,de_1 + \sin\vartheta\,de_2) + l(-\sin\vartheta\,e_1 + \cos\vartheta\,e_2)\,d\vartheta. \tag{8.6}$$

Using (8.3) in (8.5) and then substituting this into (8.6), there results,

$$dx^* = [\cos\frac{\alpha}{2}\,du - l\sin\vartheta\,d\vartheta - \frac{1}{2}l\sin\vartheta(\alpha_v\,du + \alpha_u\,dv)]e_1$$
$$+[\sin\frac{\alpha}{2}\,dv + \frac{1}{2}l\cos\vartheta(\alpha_v\,du + \alpha_u\,dv) + l\cos\vartheta\,d\vartheta]e_2$$
$$+[l\sin\frac{\alpha}{2}\cos\vartheta\,du - l\cos\frac{\alpha}{2}\sin\vartheta\,dv]e_3. \tag{8.7}$$

Due to the fact that e_3^* has to be perpendicular to e_1 with respect to $\langle\,,\rangle$ and have a constant angle τ with respect to e_3, the unit normal of S^* at P^* takes the form

$$e_3^* = \sin\tau\sin\vartheta\,e_1 - \sin\tau\cos\vartheta\,e_2 + \cos\vartheta\,e_3. \tag{8.8}$$

Since e_3^* is the normal vector of S^*, with respect to the usual metric on $\mathbb{R}^3$

$$\langle dx^*, e_3^*\rangle = 0. \tag{8.9}$$

Calculating the left-hand side of (8.9) and simplifying, the following result is obtained

$$l\sin\tau\,d\vartheta + \frac{1}{2}l\sin\tau(\alpha_v\,du + \alpha_u\,dv)$$
$$-\sin\tau(\cos\frac{\alpha}{2}\sin\theta\,du - \sin\frac{\alpha}{2}\cos\vartheta\,dv) - l\cos\tau(\sin\frac{\alpha}{2}\cos\vartheta\,du - \cos\frac{\alpha}{2}\sin\vartheta\,dv) = 0. \tag{8.10}$$

Now ϑ is specified by considering the case in which

$$\vartheta = \frac{\alpha'}{2},$$

and since the orthogonality condition holds and du, dv are independent differentials, the coefficients in (8.7) can be equated to zero giving

$$\tfrac{1}{2}l\sin\tau\,(\alpha'_u + \alpha_v) = \sin\tau\cos(\frac{\alpha}{2})\sin(\frac{\alpha'}{2}) + l\cos\tau\sin(\frac{\alpha}{2})\cos(\frac{\alpha'}{2}),$$
$$\tfrac{1}{2}l\sin\tau(\alpha'_v + \alpha_u) = -\sin\tau\sin(\frac{\alpha}{2})\cos(\frac{\alpha'}{2}) - l\cos\tau\cos(\frac{\alpha}{2})\sin(\frac{\alpha'}{2}). \tag{8.11}$$

No restrictions have been placed on the value of K up to this point. To give system (8.9) in another form, let us introduce a set of new variables σ, η defined to be

$$\sigma = \frac{1}{2}(u+v), \qquad \eta = \frac{1}{2}(u-v). \tag{8.12}$$

In terms of the variables (8.12), upon adding and subtracting the pair of equations in (8.11) and, using standard trigonometric identities, they simplify to

$$\tfrac{1}{2}l\sin\tau(\alpha+\alpha')_\sigma = \sin\tau\sin(\frac{\alpha'-\alpha}{2}) + l\cos\tau\sin(\frac{\alpha-\alpha'}{2}),$$
$$\tfrac{1}{2}l\sin\tau(\alpha'-\alpha)_\eta = \sin\tau\sin(\frac{\alpha'+\alpha}{2}) + l\cos\tau\sin(\frac{\alpha+\alpha'}{2}). \tag{8.13}$$

Introducing constants C_1 and C_2 to denote the pair of constants

$$C_1 = \frac{\sin\tau - l\cos\tau}{l\sin\tau}, \qquad C_2 = \frac{\sin\tau + l\cos\tau}{l\sin\tau}, \tag{8.14}$$

it is clear that (8.13) can be expressed in the equivalent form

$$(\alpha' + \alpha)_\sigma = 2C_1 \sin(\frac{\alpha' - \alpha}{2}), \qquad (\alpha' - \alpha)_\eta = 2C_2 \sin(\frac{\alpha' + \alpha}{2}). \tag{8.15}$$

System (8.15) will be compatible provided that the quantities α and α' satisfy a specific nonlinear equation. To obtain this equation, the compatibility condition for system (8.15) must be worked out. Differentiating, we obtain

$$\begin{aligned} (\alpha + \alpha')_{\sigma\eta} &= 2C_1C_2 \cos(\frac{\alpha' - \alpha}{2}) \sin(\frac{\alpha' + \alpha}{2}), \\ (\alpha' - \alpha)_{\eta\sigma} &= 2C_1C_2 \cos(\frac{\alpha' + \alpha}{2}) \sin(\frac{\alpha' - \alpha}{2}). \end{aligned} \tag{8.16}$$

Adding and subtracting the two equations in (8.16) and invoking trigonometric identities, it is found that both α and α' satisfy an identical deformed sine-Gordon equation, namely,

$$\psi_{\eta\sigma} = C_1C_2 \sin(\psi), \qquad \psi = \alpha, \alpha'. \tag{8.17}$$

Equation (8.17) can be expressed in terms of the variables u and v as follows

$$\psi_{uu} - \psi_{vv} = \frac{1}{2} C_1C_2 \sin(\psi). \tag{8.18}$$

It should be remarked that, based on (8.14), the combination C_1C_2 is not in general related in a straightforward way to K. The case in which $\sin\tau/l = 1$ can be considered separately. This corresponds to the case in which $K = -1$ so that

$$C_1 = \frac{1 - \cos\tau}{\sin\tau}, \qquad C_2 = \frac{1 + \cos\tau}{\sin\tau}.$$

In this case, it is easy to determine that

$$C_1C_2 = \frac{1 - \cos^2\tau}{\sin^2\tau} = 1. \tag{8.19}$$

Therefore, corresponding to the case $K = -1$, upon setting $\beta = C_1$ and using (8.19) to get C_2, it is useful to note that system (8.15) assumes the usual form,

$$(\alpha' + \alpha)_\sigma = 2\beta \sin(\frac{\alpha' - \alpha}{2}), \qquad (\alpha' - \alpha)_\eta = \frac{2}{\beta} \sin(\frac{\alpha' + \alpha}{2}). \tag{8.20}$$

Let us make a summary of what has been done up to now. It has been seen that Bäcklund's theorem has the following implications. Suppose S is a surface in

$\mathbb{R}^3$ with negative, constant Gaussian curvature (8.2) such that $l > 0$ and $\tau \neq n\pi$ are constants. Let $e_0 \in T_{P_0}M$ be a unit vector which is not in the principle direction. Then there exists a unique surface S^* and a pseudo-spherical line congruence $\{PP^*\}$ where $P \in S$ and $P^* \in S^*$ satisfy $PP_0^* = le_0$, and τ is the angle between the normal direction of S at P and S^* at P^*. The content of the new results is Proposition 2.

Proposition 8.2: A surface of arbitrary constant negative curvature (8.2) is determined by any nontrivial solution to (8.17)-(8.18) combined with the fundamental surface equations (8.5).

8.3 Calculation of Solutions and Formulation of the Theorem of Permutability.

The system described by (8.15) is in fact a Bäcklund transformation for the equation (8.18). Given a particular solution to (2.16), it will be shown that (8.15) can be used to obtain a new solution to (8.18). Since $\alpha = 0$ is a solution to (8.17)-(8.18) for any C_1, C_2, substituting into (8.15) with $\alpha' = \alpha_1$, we have

$$\partial_\sigma \alpha_1 = 2C_1 \sin(\frac{\alpha_1}{2}), \qquad \partial_\eta \alpha_1 = 2C_2 \sin(\frac{\alpha_1}{2}). \tag{8.21}$$

Introduce two new variables s, t which are defined such that $s = C_1\sigma$ and $t = C_2\eta$ so that system (8.21) takes the form

$$\frac{\partial \alpha_1}{\partial s} = 2\sin\frac{\alpha_1}{2}, \qquad \frac{\partial \alpha_1}{\partial t} = 2\sin\frac{\alpha_1}{2}. \tag{8.22}$$

Since the derivatives in (8.22) are the same, it follows that α_1 must have the form $\alpha_1 = \alpha_1(s+t)$. To determine the form of the new solution α_1 corresponding to $\alpha = 0$ explicitly, let us write the first equation in (8.22) as

$$\frac{\partial \alpha_1}{\partial s} = 2\sin(\frac{\alpha_1}{2}) = 4\sin\frac{\alpha_1}{4}\cos\frac{\alpha_1}{4} = 4\tan\frac{\alpha_1}{4}\cos\frac{\alpha_1}{4}. \tag{8.23}$$

This equation is equivalent to the form,

$$\frac{\partial}{\partial s}\tan\frac{\alpha_1}{4} = \tan\frac{\alpha_1}{4}. \tag{8.24}$$

In this form, the equation may be easily integrated with the help of (8.22) to give the solution

$$\tan\frac{\alpha_1}{4} = C\,\exp(s+t), \tag{8.25}$$

where C is an arbitrary real constant. It is now straightforward to transform back to the (u, v) variables from the (s, t) variables to yield

$$\tan\frac{\alpha_1}{4} = C\exp[C_1\sigma + C_2\eta] = C\exp[\frac{1}{2}C_1(u+v) + \frac{1}{2}C_2(u-v)]. \tag{8.26}$$

Therefore a new solution to deformed sine-Gordon equation (8.17) has been found starting with the $\alpha = 0$ solution applying (8.15) and integrating. Let us summarize it in the form,

$$\alpha_1 = 4\tan^{-1}\{C\exp[\frac{1}{2}(C_1 + C_2)u + \frac{1}{2}(C_1 - C_2)v]\}.$$

Other solutions to (8.17) can be constructed along similar lines.

According to the theorem of permutability, the application of two successive Bäcklund transformations commutes. To express it more quantitatively, if two successive Bäcklund transformations with distinct parameters λ_1 and λ_2 map a given solution α_0 through intermediate solutions to either α_{12} or α_{21}, the order in which this is done is irrelevant and in fact $\alpha_{12} = \alpha_{21}$. If the intermediate solutions are denoted α_1 and α_2, then making use of the η equation in (8.15) and identifying the Bäcklund parameter as the constant which appears on the right, the scheme described can be expressed in the form

$$(\alpha_1 - \alpha_0)_\eta = 2\lambda_1 \sin(\frac{\alpha_1 + \alpha_0}{2}),\ (\alpha_{12} - \alpha_1)_\eta = 2\lambda_2 \sin(\frac{\alpha_{12} + \alpha_1}{2}),$$

$$(\alpha_2 - \alpha_0)_\eta = 2\lambda_2 \sin(\frac{\alpha_2 + \alpha_0}{2}),\ (\alpha_{12} - \alpha_2)_\eta = 2\lambda_1 \sin(\frac{\alpha_{12} + \alpha_2}{2}).$$

In fact, all the derivative terms can be eliminated from these equations. Subtracting the first two and the last two pairwise, and then subtracting these two resulting equations produces the result,

$$\lambda_2(\sin(\frac{\alpha_{12} + \alpha_1}{2}) - \sin(\frac{\alpha_2 + \alpha_0}{2})) + \lambda_1(\sin(\frac{\alpha_1 + \alpha_0}{2}) - \sin(\frac{\alpha_{12} + \alpha_2}{2})) = 0. \tag{8.27}$$

By making use of standard trigonometric identities, it is possible to render this in the following concise form,

$$(\lambda_2 - \lambda_1)\tan(\frac{\alpha_{12} - \alpha_0}{4}) = (\lambda_1 + \lambda_2)\tan(\frac{\alpha_2 - \alpha_0}{4}). \tag{8.28}$$

The usual result for the sine-Gordon equation is obtained. The theorem of permutability allows the construction algebraically of a second order solution, and the procedure can be carried out order by order.

To conclude, it has been seen here that the sine-Gordon equation has been generalized to accommodate cases of arbitrary Gaussian curvature, and a Bäcklund transformation has been calculated as well as applying it to generate a solution. Further solutions can be produced from it using the theorem of permutability.

8.4 Bäcklund Transformation for the Sine-Gordon Equation.

Historically the sine-Gordon equation has played an important role in the development of the Bäcklund transformation. In fact, by using the prolongation ideas of the last chapter, a Bäcklund tranformation has been obtained by combining the prolongation results with the idea of a Mauer-Cartan algebra. Due to its significance, it is worth discussing some related but perhaps somewhat unconventional methods for generating Bäcklund transformations for this and other nonlinear equations [**53**]. There is an important relationship between the non-linear partial differential equations which have soliton solutions, such as the sine-Gordon and Korteweg-de Vries equations and the group $SL(2,\mathbb{R})$ of 2×2 real matrices with determinant one. This relationship is deepened by the fact that it helps to explain some of the features of these soliton equations, namely they have Bäcklund transformations and may be solved by the Inverse Scattering method. In fact, by using the prolongation ideas of the last chapter, a Bäcklund transformation has been obtained for these equations by combining the prolongation results with the idea of a Mauer-Cartan algebra. The indication of such a relationship is the existence of three one-forms $\{\theta^1, \theta^2, \theta^3\}$ on the space of independent variables x and t of the equation, whose coefficients depend on the dependent variable and its partial derivatives, and which satisfy

$$d\theta^1 = -\theta^2 \wedge \theta^3, \qquad d\theta^2 = -2\,\theta^1 \wedge \theta^2, \qquad d\theta^3 = 2\theta^1 \wedge \theta^3. \tag{8.29}$$

These are formally the same as the Maurer-Cartan equations satisfied by the left-invariant one-forms $\{\omega^1, \omega^2, \omega^3\}$ of $SL(2,\mathbb{R})$ but with the θ^i defined on a two-dimensional space. The Frobenius Theorem implies there must be a map from some open subset of $\mathbb{R}^2$ to $SL(2,\mathbb{R})$ under which the one-forms $\{\omega^i\}$ pull back to $\{\theta^i\}$.

Thus let

$$X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \tag{8.30}$$

where $\det X = 1$, so that X is a general element of $SL(2,\mathbb{R})$. Then $X^{-1}dX$ considered as a matrix of one-forms, takes its values in the Lie algebra of $SL(2,\mathbb{R})$. Moreover, if

$$X^{-1}dX = \begin{pmatrix} \omega^1 & \omega^2 \\ \omega^3 & -\omega^1 \end{pmatrix}$$

then the $\{\omega^i\}$ are the left-invariant forms of $SL(2,\mathbb{R})$. Thus there is a local $SL(2,\mathbb{R})$-valued function G on $\mathbb{R}^2$ such that,

$$G^{-1}dG = \begin{pmatrix} \theta^1 & \theta^2 \\ \theta^3 & -\theta^1 \end{pmatrix}.$$

This $sl(2,\mathbb{R})$-valued one-form on $\mathbb{R}^2$ will be represented as Θ, so the expression above is $dG = G\Theta$. Thus, each row of matrix G satisfies,

$$dr = r\theta^1 + s\theta^3, \qquad ds = r\theta^2 - s\theta^1.$$

Consequently, the Maurer-Cartan equations for the forms $\{\theta^i\}$ may be written $d\Theta + \Theta \wedge \Theta = 0$. The form Θ may be regarded as defining a connection on a principal $SL(2,\mathbb{R})$ bundle over $\mathbb{R}^2$. The nonlinear equation expresses the fact that the curvature of the connection $d\Theta + \Theta \wedge \Theta$ must vanish, hence $\Theta = G^{-1}dG$ is pure gauge.

Bäcklund transformations fit into this picture. Any element of $SL(2,\mathbb{R})$ can be expressed uniquely as the product of an upper-triangular matrix-valued function T and a rotation matrix-valued function R as $G = T\,R^{-1}$. It follows from this expression that

$$TR^{-1}\Theta = G\Theta = dG = dT\,R^{-1} - TR^{-1}\,dR\,R^{-1},$$

hence,

$$T^{-1}\,dT = R^{-1}\Theta R + R^{-1}\,dR.$$

Now $T^{-1}\,dT$ is itself upper triangular, however, the lower left corner element on the right is not as the following calculation shows. The matrices R and Θ may be written in the form,

$$R = \begin{pmatrix} \cos\frac{v}{2} & \sin\frac{v}{2} \\ -\sin\frac{v}{2} & \cos\frac{v}{2} \end{pmatrix}, \qquad \Theta = \begin{pmatrix} \theta^1 & \theta^2 \\ \theta^3 & \theta^1 \end{pmatrix}.$$

Therefore,

$$R^{-1} = \begin{pmatrix} \cos\frac{v}{2} & -\sin\frac{v}{2} \\ \sin\frac{v}{2} & \cos\frac{v}{2} \end{pmatrix}, \qquad dR = \frac{1}{2}\begin{pmatrix} -\sin\frac{v}{2} & \cos\frac{v}{2} \\ -\cos\frac{v}{2} & -\sin\frac{v}{2} \end{pmatrix}.$$

With these matrices, the two terms on the right of $T^{-1}dT$ can be calculated. First, the term $R^{-1}\,dR$ is given by

$$R^{-1}\,dR = \frac{1}{2}\begin{pmatrix} 0 & dv \\ -dv & 0 \end{pmatrix}$$

and then $R^{-1}\Theta R$ takes the form,

$$R^{-1}\Theta R$$
$$= \frac{1}{2}\begin{pmatrix} 2\cos v\,\theta^1 - \sin v(\theta^2+\theta^3) & 2\sin v\,\theta^1 + (\theta^2-\theta^3) \\ 2\sin v\theta^1 - (\theta^2-\theta^3) + \cos v\,(\theta^2+\theta^3) & -2\cos v\theta^1 + \sin v\,(\theta^2+\theta^3) \end{pmatrix}$$

The lower left-hand corner of $R^{-1}\,dR + R^{-1}\Theta R$ is determined to be

$$-dv + 2\sin v\,\theta^1 - (\theta^2-\theta^3) + \cos v\,(\theta^2+\theta^3).$$

Since this must match the corresponding element in $T^{-1}dT$, which is upper-triangular, this result can be equated to zero to produce the following equation,

$$dv + \theta^2 - \theta^3 = 2\sin v\,\theta^1 + \cos v(\theta^2+\theta^3). \tag{8.31}$$

This is equivalent to two first-order partial differential equations for v in terms of u and its partial derivatives, which are the Bäcklund transformation equations associated with the soliton equation satisfied by u. The Bäcklund equations may therefore be interpreted as the equations satisfied by the angle in the rotation part of G. Alternatively, the transformation $\Theta \to R^{-1}\,dR + R^{-1}\Theta R$ may be viewed as a gauge transformation of Θ. If this viewpoint is taken, the Bäcklund transformation is the gauge transformation which makes Θ upper triangular.

These equations may be written in another form as follows. Comparison of G and TR^{-1} leads to the equation $\tau = \tan(v/2) = r/s$. Therefore,

$$d\tau = \frac{1}{s}\,dr - \frac{r}{s^2}\,ds = 2\tau\theta^1 - \tau^2\theta^2 + \theta^3.$$

This form is equivalent to the previous version. Here τ can be thought of as a pseudo-potential

There is a certain amount of freedom in the choice of $\{\theta^i\}$, and two different Θ which lead to the same equation are typically related by means of a gauge transformation and possibly a change of coordinates. As an example, let us make a particular choice for the $\{\theta^i\}$ and then substitute into (8.31). The forms are given as

$$\theta^1 = \frac{1}{2}\cos\frac{u}{2}(\eta\,dx + \frac{1}{\eta}\,dt),$$

$$\theta^2 = (\frac{1}{4}u_x - \frac{\eta}{2}\sin\frac{u}{2})\,dx - (\frac{1}{4}u_t - \frac{1}{2\eta}\sin\frac{u}{2})\,dt,$$

$$\theta^3 = -(\frac{1}{4}u_x + \frac{\eta}{2}\sin\frac{u}{2})\,dx + (\frac{1}{4}u_t + \frac{1}{2\eta}\sin\frac{u}{2})\,dt.$$

Substituting these forms into (8.31), the following result appears

$$dv + \frac{u_x}{2}\,dx - \frac{u_t}{2}\,dt = \sin v\cos\frac{u}{2}\,(\eta\,dx + \frac{1}{\eta}\,dt) + \cos v(-\eta\sin\frac{u}{2}\,dx + \frac{1}{\eta}\,\sin\frac{u}{2}\,dt).$$

It remains to equate the coefficients of dx and dt on both sides and then use a well known idntity to obtain the Bäcklund transformation

$$v_x = -\frac{1}{2}u_x + \eta\sin(v - \frac{u}{2}), \qquad v_t = \frac{1}{2}u_t + \frac{1}{\eta}\sin(v + \frac{u}{2}). \tag{8.32}$$

8.5 Another Application of Differential Systems to Bäcklund Correspondences.

A quite specific type of differential ideal often appears when Cartan's method of moving frames is applied to classical problems of differential geometry. These ideals can be generated by sets of two-forms and have a canonical structure inasmuch as they are expressed in an anholonomic basis of one-forms in which all of their terms have constant coefficients **[54]**. Such ideals are usually derived by specializing the closure relations that are fulfilled by the left or right invariant one-forms in the space of Lie groups. They can be analyzed for invariances, conservation laws, Bäcklund correspondences and integral manifolds classified by computing Cartan's local algebraic characters. Depending on how dependent and independent sets of variables are introduced, elegant sets of coupled nonlinear partial differential equations, such as the sine-Gordon equation can emerge.

Consider a six parameter Lie group that is built upon the three-parameter rotation group $O(3)$, In terms of six basis one-forms ω_i, and ϕ_i with $i = 1, 2, 3$, the following set of two-forms generates a closed differential ideal $\mathcal{I}$

$$\vartheta_1 = d\omega_1 + \omega_2 \wedge \omega_3, \; \vartheta_4 = d\phi_1 + \omega_2 \wedge \phi_3 - \omega_3 \wedge \phi_2,$$
$$\vartheta_2 = d\omega_2 + \omega_3 \wedge \omega_1, \; \vartheta_5 = d\phi_2 + \omega_3 \wedge \phi_1 - \omega_1 \wedge \phi_3, \tag{8.33}$$
$$\vartheta_3 = d\omega_3 + \omega_1 \wedge \omega_2, \; \vartheta_6 = d\phi_3 + \omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1.$$

This is a differential ideal and in fact it can be shown that it is a closed differential ideal over the entire set of forms $\{\vartheta_i\}_1^6$.

Theorem 8.1. The exterior derivatives of the system $\{\vartheta_i\}$ in (8.33) are calculated to be,

$$\begin{aligned}
d\vartheta_1 &= \vartheta_2 \wedge \omega_3 + \vartheta_3 \wedge \omega_2,\\
d\vartheta_2 &= \vartheta_3 \wedge \omega_1 + \vartheta_1 \wedge \omega_3,\\
d\vartheta_3 &= \vartheta_1 \wedge \omega_2 + \vartheta_2 \wedge \omega_1,\\
d\vartheta_4 &= \vartheta_2 \wedge \phi_3 + \vartheta_6 \wedge \omega_2 - \vartheta_3 \wedge \phi_2 - \vartheta_5 \wedge \omega_3,\\
d\vartheta_5 &= -\vartheta_1 \wedge \phi_3 + \vartheta_3 \wedge \phi_1 + \vartheta_4 \wedge \omega_3 - \vartheta_6 \wedge \omega_1,\\
d\vartheta_6 &= \vartheta_1 \wedge \phi_2 - \vartheta_2 \wedge \phi_1 - \vartheta_4 \wedge \omega_2 + \vartheta_5 \wedge \omega_1.
\end{aligned} \tag{8.34}$$

Proof: To produce these results, differentiate each of the ϑ_i in (8.33) and eliminate the known exterior derivatives. The case which involves ϑ_1 is illustrated explicitly,

$$d\vartheta_1 = d\omega_2 \wedge \omega_3 - \omega_2 \wedge d\omega_3 = (\vartheta_2 - \omega_3 \wedge \omega_1) \wedge \omega_3 - \omega_2 \wedge (\vartheta_3 - \omega_1 \wedge \omega_2) = \vartheta_2 \wedge \omega_3 + \vartheta_3 \wedge \omega_2.$$

The others follow a similar procedure.

If, in a space of six dimensions, the forms ϑ_i in (8.33) were to vanish identically, the basis one-forms ω_i, ϕ_i could be called left-invariant, and the space identified at least locally with the group space. A set of canonical structure constants for the group can be read off (8.33).

It may be that there exists a matrix Ω composed of the ω_i, ϕ_j such that the integrability conditions are expressed as the vanishing of a matrix of two-forms,

$$\Theta = 0, \qquad \Theta = d\Omega - \Omega \wedge \Omega,$$

which gives, by construction, the nonlinear equation to be solved.

This differential ideal can be enlarged in the following way. Introducing the two auxiliary variables f and g, which are referred to as prolongation variables or pseudopotentials, consider the differential ideal augmented by means of the following two additional differential forms,

$$\alpha_1 = df - fg\omega_1 + (1+f^2)\omega_2 - g\omega_3, \qquad \alpha_2 = dg - (1+g^2)\omega_1 + fg\omega_2 + f\omega_3. \tag{8.35}$$

Solving (8.35) for df and dg, we obtain

$$df = \alpha_1 + fg\omega_1 - (1+f^2)\omega_2 + g\omega_3, \qquad dg = \alpha_2 + (1+g^2)\omega_1 - fg\omega_2 - f\omega_3.$$

Theorem 8.2. The exterior derivatives of $\{\alpha_1, \alpha_2\}$ in (8.35) take the form

$$d\alpha_1 = -g\alpha_1 \wedge \omega_1 - f\alpha_2 \wedge \omega_1 + 2f\alpha_1 \wedge \omega_2 - \alpha_2 \wedge \omega_3,$$

$$d\alpha_2 = g\alpha_1 \wedge \omega_2 + \alpha_1 \wedge \omega_3 - 2g\alpha_2 \wedge \omega_1 + f\alpha_2 \wedge \omega_2.$$

Proof: This is a long calculation which utilizes the results for the forms df and dg. For example,

$$d\alpha_2 = -2g\,dg \wedge \omega_1 - (1+g^2)d\omega_1 + g\,df \wedge \omega_2 + fgd\omega_2 + df \wedge \omega_3 + f\,d\omega_3$$

$$= -2g(\alpha_2 + (1+g^2)\omega_1 - fg\omega_2 - f\omega_3) \wedge \omega_1 + (1+g^2)\omega_2 \wedge \omega_3 + g(\alpha_1 + fg\omega_1$$
$$-(1+f^2)\omega_2 + g\omega_3) \wedge \omega_2 + f(\alpha_2 + (1+g^2)\omega_1 - fg\omega_2 - f\omega_3) \wedge \omega_2 - fg\omega_3 \wedge \omega_1$$
$$+(\alpha_1 + fg\omega_1 - (1+f^2)\omega_2 + g\omega_3) \wedge \omega_3 - f\omega_1 \wedge \omega_2.$$
$$= -2g\,\alpha_2 \wedge \omega_1 + g\alpha_1 \wedge \omega_2 + f\alpha_2 \wedge \omega_2 + \alpha_1 \wedge \omega_3.$$

By putting identically equal to zero one or more basic one-forms in the set (8.33) of a group, one immediately obtains members of an interesting class of nonlinear systems generated by two-forms having constant coefficients and closed under exterior differentiation. These are often referred to as canonical systems. For the remainder of this section, such a system is considered. At this point, we set $\phi_3 = 0$ in (8.33) and call the resulting ideal $\tilde{\mathcal{I}}$ such that

$$\begin{aligned}
\tilde{\vartheta}_1 &= d\omega_1 + \omega_2 \wedge \omega_3,\\
\tilde{\vartheta}_2 &= d\omega_2 + \omega_3 \wedge \omega_1,\\
\tilde{\vartheta}_3 &= d\omega_3 + \omega_1 \wedge \omega_2,\\
\tilde{\vartheta}_4 &= d\phi_1 - \omega_3 \wedge \phi_2,\\
\tilde{\vartheta}_5 &= d\phi_2 + \omega_3 \wedge \phi_1,\\
\tilde{\vartheta}_6 &= \omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1.
\end{aligned} \tag{8.36}$$

Using Theorem 8.2, two auxilliary variables f, g can be introduced to enlarge the ideal to the form

$$d\phi_1 - f(\omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1) - \omega_2 \wedge \phi_2,$$
$$d\phi_2 - g(\omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1) + \omega_3 \wedge \phi_1,$$
$$d\omega_1 + \omega_2 \wedge \omega_3,$$
$$d\omega_2 + \omega_3 \wedge \omega_1,$$
$$d\omega_3 + \omega_1 \wedge \omega_2,$$
$$df - fg\omega_1 + (1 + f^2)\omega_2 - g\omega_3,$$
$$dg - (1 + g^2)\omega_1 + fg\omega_2 + f\omega_3,$$
$$\omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1.$$

Finally, the immersion ideal can be written for surfaces of constant negative curvature as

$$\begin{aligned} \tilde{\vartheta}_1 &= d\omega_1 + \omega_2 \wedge \omega_3, \\ \tilde{\vartheta}_2 &= d\omega_2 + \omega_3 \wedge \omega_1, \\ \tilde{\vartheta}_3 &= d\omega_3 + \omega_1 \wedge \omega_2, \\ \tilde{\vartheta}_4 &= d\phi_1 - \omega_3 \wedge \phi_2, \\ \tilde{\vartheta}_5 &= d\phi_2 + \omega_3 \wedge \phi_1, \\ &\omega_1 \wedge \phi_2 - \omega_2 \wedge \phi_1, \\ &\omega_1 \wedge \omega_2 + \phi_1 \wedge \phi_2. \end{aligned} \tag{8.37}$$

This remains closed and so forms a canonical system. The final step is to observe that, because this ideal has a Cauchy characteristic vector, a prolongation one-form can be added to the ideal without introducing any prolongation variable. The one-form introduced by Chern and Terng introduced into the study of the Sine-Gordon equation is given by

$$\Psi = \phi_2 + \sin\tau\, \omega_3 + \cos\tau\, \omega_2,$$

where τ is an arbitrary parameter. Call this augmented ideal $\tilde{\mathcal{I}}_B$. Thus there will be a Bäcklund correspondence between solutions of these subideals.

Knowing that the search for two-dimensional integral submanifolds of $\tilde{\mathcal{I}}_B$ is a well-posed problem, by picking suitable coordinates the resulting partial differential equations can be examined, in particular the sine-Gordon equation is studied. In the preceding, the view has been adopted that $\{\tilde{\vartheta}_i\}$ can be set identically equal to zero. This defines a basis set in a space, and the remaining algebraic forms can be treated as fields there. Now going in a slightly different direction, choose an algebraically degenerate coordinate representation of the basis forms in terms of two independent variables u_1 and u_2 with corresponding basis one-forms du_1 and du_2 such that the second set of forms in $\tilde{\mathcal{I}}_B$ is identically zero, and the partial differential equations arise from the first set, $\{\tilde{\vartheta}_i\}$.

Introduce now the following set of one-forms

$$\phi_1 = \cos\psi \cos\alpha \, du_1 + \sin\psi \sin\alpha \, du_2,$$

$$\begin{aligned} \phi_2 &= -\cos\psi \sin\alpha \, du_1 + \sin\psi \cos\alpha \, du_2, \\ \omega_1 &= -\sin\psi \sin\alpha \, du_1 - \cos\psi \cos\alpha \, du_2, \\ \omega_2 &= -\sin\psi \cos\alpha \, du_1 + \cos\psi \sin\alpha \, du_2, \end{aligned} \tag{8.38}$$

$$\sin\tau \omega_3 = (\cos\psi \sin\alpha + \cos\tau \sin\psi \cos\alpha) \, du_1 - (\sin\psi \cos\alpha + \cos\tau \cos\psi \sin\alpha) \, du_2.$$

Theorem 8.3. With respect to the basis set of forms (8.38), all of the $\{\tilde{\vartheta}_i\}_1^5$ can be expressed in the form of linear combinations of the two-forms χ_1 and χ_2 defined by

$$\begin{aligned} \chi_1 &= (d\psi - d\alpha) \wedge (du_1 + du_2) + \frac{1+\cos\tau}{\sin\tau} \sin(\psi+\alpha) \, du_1 \wedge du_2, \\ \chi_2 &= (d\psi + d\alpha) \wedge (du_1 - du_2) - \frac{1-\cos\tau}{\sin\tau} \sin(\psi-\alpha) \, du_1 \wedge du_2. \end{aligned} \tag{8.39}$$

Proof: The calculations for two of them will be outlined. By straightforward exterior differentiation,

$$d\omega_1 = -\sin(\psi+\alpha)(d\psi+d\alpha)\wedge(du_1-du_2) + \sin(\psi-\alpha)(d\psi-d\alpha)\wedge(du_1+du_2).$$

Using trigonometric identities, it follows as well that,

$$\sin\tau \, \omega_2 \wedge \omega_3 = \sin(\psi+\alpha)\sin(\psi-\alpha) \, du_1 \wedge du_2.$$

Therefore, $\tilde{\vartheta}_1$ is given by,

$$\tilde{\vartheta}_1 = d\omega_1 + \omega_2 \wedge \omega_3 = -\frac{1}{2}\sin(\psi+\alpha)(d\psi+d\alpha)\wedge(du_1-du_2)$$

$$+\frac{1}{2}\sin(\psi-\alpha)(d\psi-d\alpha)\wedge(du_1-du_2) + \frac{1}{\sin\tau}\sin(\psi+\alpha)\sin(\psi-\alpha)\, du_1 \wedge du_2$$

$$= -\frac{1}{2}\sin(\psi+\alpha)\chi_1 - \frac{1}{2}\sin(\psi-\alpha)\chi_2.$$

Similarly, differentiation of ω_2 gives,

$$2d\omega_2 = -\cos(\psi+\alpha)(d\psi+d\alpha)\wedge(du_1-du_2) - \cos(\psi-\alpha)(d\psi-d\alpha)\wedge(du_1+du_2),$$

and

$$\sin\tau\,\omega_3\wedge\omega_1=-(\sin\alpha\cos\alpha+\cos\tau\cos\psi\sin\psi)\,du_1\wedge du_2.$$

Therefore,

$$\tilde{\vartheta}_2=d\omega_2+\omega_3\wedge\omega_1=-\frac{1}{2}\cos(\psi+\alpha)(d\psi+d\alpha)\wedge(du_1\wedge du_2)$$

$$-\frac{1}{2}\cos(\psi-\alpha)(d\psi-d\alpha)\wedge(du_1+du_2)-\frac{1}{\sin\tau}(\sin\alpha\cos\alpha+\cos\tau\cos\psi\sin\psi)\,du_1\wedge du_2$$

$$=-\frac{1}{2}\cos(\psi-\alpha)\chi_1-\frac{1}{2}\cos(\psi+\alpha)\chi_2.$$

Similar linear combinations can be found for the other cases.

Now let us take the two-forms χ_i to an integral manifold and obtain the equation implied by the representation (8.38). As an intermediate step, a Bäcklund transformation for the equation will be generated. First of all,

$$d\psi-d\alpha=(\psi_{u_1}-\alpha_{u_1}-\psi_{u_2}+\alpha_{u_2})\,du_1\wedge du_2,$$

$$d\psi+d\alpha=-(\psi_{u_1}+\alpha_{u_1}+\psi_{u_2}+\alpha_{u_2})\,du_1\wedge du_2.$$

Equating the coefficient of $du_1\wedge du_2$ to zero in each χ_i in (8.39) and then using these results, the following pair of first order equations results,

$$\begin{aligned}\psi_{u_1}-\alpha_{u_1}-\psi_{u_2}+\alpha_{u_2}+\frac{1+\cos\tau}{\sin\tau}\sin(\psi+\alpha)&=0,\\ \psi_{u_1}+\alpha_{u_1}+\psi_{u_2}+\alpha_{u_2}+\frac{1-\cos\tau}{\sin\tau}\sin(\psi-\alpha)&=0.\end{aligned}\tag{8.40}$$

This is one version of the Bäcklund. To make everything more concise, introduce the following pair of variables

$$\beta=\psi-\alpha,\qquad\gamma=\psi+\alpha.$$

In terms of these new variables, (8.40) is transformed into the pair of equations,

$$\begin{aligned}\beta_{u_1}-\beta_{u_2}+\frac{1+\cos\tau}{\sin\tau}\sin\gamma&=0,\\ \gamma_{u_1}+\gamma_{u_2}+\frac{1-\cos\tau}{\sin\tau}\sin\beta&=0.\end{aligned}\tag{8.41}$$

Each of the equations in (8.41) will be differentiated in turn with respect to u_1 and u_2. The first equation in β yields

$$\beta_{u_1u_1}-\beta_{u_2u_1}+\frac{1+\cos\tau}{\sin\tau}\cos\gamma\,\gamma_{u_1}=0,$$

$$\beta_{u_1u_2}-\beta_{u_2u_2}+\frac{1+\cos\tau}{\sin\tau}\cos\gamma\,\gamma_{u_2}=0.$$

Adding this pair of equations, the $\beta_{u_1 u_2}$ terms cancel, and we are left with,

$$\beta_{u_1 u_1} - \beta_{u_2 u_2} + \frac{1+\cos\tau}{\sin\tau} \cos\gamma (\gamma_{u_1} + \gamma_{u_2}) = 0. \tag{8.42}$$

Replacing $\gamma_{u_1} + \gamma_{u_2}$ from the second equation of (8.41) into (8.42), we get,

$$\beta_{u_1 u_1} - \beta_{u_2 u_2} - \sin\beta \cos\gamma = 0. \tag{8.43}$$

Following exactly the same procedure for the second equation of (8.41) produces the result for γ,

$$\gamma_{u_1 u_1} - \gamma_{u_2 u_2} - \cos\beta \sin\gamma = 0. \tag{8.44}$$

Adding the equations in (8.43) and (8.44) together, a version of the sine-Gordon equation in the new variable $\beta + \theta = 2\psi$ is obtained

$$(\beta + \gamma)_{u_1 u_1} - (\beta + \gamma)_{u_2 u_2} = \sin(\beta + \gamma). \tag{8.45}$$

Moreover, there is a Bäcklund correspondence given by (8.41).

9 Generation of Integrable Systems and Hierarchies.

Completely integrable systems arise from many different areas of physics and mathematics and constitute a very active area of investigation at the moment. Although somewhat more algebraic in scope than the geometric topics of the previous chapters, it is of both theoretical and practical value to find as many new integrable systems as possible and to expand in depth their algebraic and geometric properties. This has been a subject of considerable interest recently **[55]**. The idea here is to illustrate how Lie algebras play a role in generating integrable systems. In fact, it will be shown that these ideas lead to the production of hierarchies of equations. Thus, on the basis of the gradation and decomposition of an element of the algebra $\mathcal{A}$, the solvability in $V \in \mathcal{A}$ is shown for the equation $V_x = [U, V]$ for a class of elements $U \in \mathcal{A}$. The solution V together with the gradation of $\mathcal{A}$ leads then to an element which is modified further by the introduction of a modification quantity and leads to a new hierarchy of integrable systems **[56]**.

9.1 Definition of the Algebra.

Consider a loop algebra $\tilde{\mathcal{A}}$, that is, the affine Lie algebra without center given as

$$\tilde{\mathcal{A}} = \mathcal{A} \otimes \mathbb{C}(\lambda, \lambda^{-1}).$$

A basis for $\tilde{\mathcal{A}}$ will be taken as follows

$$\{h(n), e(n), f(n) | n \in \mathbb{Z}\}. \tag{9.1}$$

where $\mathbb{Z}$ stands for the set of integers and $x(n) = x \otimes \lambda^n$ for $x \in \mathcal{A}$. Finally, h, e and f represent a basis of $\mathcal{A}$ which has a matrix representation of the form

$$h = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{9.2}$$

From these definitions, it follows that

$$[h, e] = e, \qquad [h, f] = -f, \qquad [e, f] = 2h. \tag{9.3}$$

Moreover, since $x(n) = x \otimes \lambda^n$, these brackets imply that

$$[h(m), e(n)] = e(m + n),$$
$$[h(m), f(n)] = -f(m + n), \qquad [e(m), f(n)] = 2h(m + n). \tag{9.4}$$

The basis set $\{h(n), e(n), f(n) | n \in \mathbb{Z}\}$ can be transformed into the following related set

$$\{h(n), e_+(n), e_-(n) | n \in \mathbb{Z}\},$$
$$e_\pm(n) = \tfrac{1}{2}(e(n-1) \pm f(n)), \tag{9.5}$$

$$e(n) = e_+(n+1) + e_-(n+1), \qquad f(n) = e_+(n) - e_-(n).$$

The commutators of the algebra elements in (9.5) are given by

$$[h(m), e_\pm(n)] = \frac{1}{2}[h(m), e(n-1) \pm f(n)] = \frac{1}{2}[h(m), e(n-1)] \pm \frac{1}{2}[h(m), f(n)]$$

$$= \frac{1}{2}(e(m+n-1) \mp f(m+n)) = e_\mp(m+n).$$

$$[e_-(m), e_+(n)] = h(m+n-1).$$

Introduce the following gradation on the basis elements $\{h(n), e_+(n), e_-(n)\}$ with deg $h(n) = 2n$ and deg $e_\pm(n) = 2n-1$. From the bracket relations above, it follows that

$$\deg[h(m), e_\pm(n)] = \deg h(m) + \deg\, e_\pm(n),$$

$$\deg[e_-(m), e_+(n)] = \deg e_-(m) + \deg\, e_+(n).$$

In this way, the loop algebra $\tilde{\mathcal{A}}$ is made into a graded Lie algebra with respect to the gradation given above. For a summation

$$M = \sum_{m \in \mathbb{Z}} \mathbb{C}\, h(m) + \mathbb{C}\, e_+(m) + \mathbb{C}\, e_-(m),$$

in $\tilde{\mathcal{A}}$, we write

$$M_+ = \sum_{m \geq 0} \mathbb{C}\, h(m) + \mathbb{C}\, e_+(m) + \mathbb{C}\, e_-(m),$$

and $M_- = M - M_+$. In these expressions $\mathbb{C}$ represents some scalars taken from the complex field, or in other words, M_+ is obtained from M by keeping all terms of gradations greater than or equal to one. To any Lie algebra $\mathcal{L}$ and an element $\Lambda \in \mathcal{L}$, we have

$$K = \ker \mathrm{ad}\Lambda = \{x | x \in \mathcal{L}, [\Lambda, x] = 0\},$$

$$K^\perp = \Im\, \mathrm{ad}\Lambda = \{x | x \in \mathcal{L}, \exists\, y \in \mathcal{L}, x = [\Lambda, y]\}.$$

In particular, for $\mathcal{L} = \tilde{\mathcal{A}}$ and

$$\Lambda = 2e_+(1),$$

it is clear that, from the bracket relations,

$$K = \sum_{n \in \mathbb{Z}} \mathbb{C}\, e_+(n), \qquad K^\perp = \sum_{n \in \mathbb{Z}} (\mathbb{C}\, e_-(n) + \mathbb{C}\, h(n)).$$

Hence for the above specific choice, the following result is valid.

Lemma 9.1. (i) K is a commutative subalgebra. (ii) $\mathcal{L} = K \oplus K^\perp$.

9.2 A Hierarchy of Evolution Equations.

Begin with the linear problem defined to be,

$$\Psi_x = U\Psi, \qquad \Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix},$$
$$U = \Lambda + Q, \qquad \Lambda = 2e_+(1), \qquad Q = ue_+(0) + ve_-(0), \tag{9.6}$$

where $u = u(x,t)$ and $v = v(x,t)$ are two independent potentials, in other words, the dependent variables in the equations of the intended hierarchy. Note that we have,

$$\deg \Lambda = 1, \qquad \deg Q = -1.$$

Suppose a, b and c are scalar-valued but depend on u, v and the parameter λ, in the following way $a = a(u,v,\lambda)$, $b = b(u,v,\lambda)$, $c = c(u,v,\lambda)$. Let V be given by

$$V = ah(0) + be_+(0) + ce_-(0) \tag{9.7}$$

be a solution of the equation

$$V_x = [U, V]. \tag{9.8}$$

Observing that $e_+(n) = e_+(0)\,\lambda^n$ and then using the definition of U given in (9.6),

$$U = (2\lambda + u)e_+(0) + ve_-(0), \tag{9.9}$$

then equation (9.8) becomes,

$$a_x h(0) + b_x e_+(0) + c_x e_-(0) = [2\lambda + u)e_+(0) + ve_-(0), ah(0) + be_+(0) + ce_-(0)]$$

$$= (2\lambda+u)a[e_+(0), h(0)] + av[e_-(0), h(0)] + bv[e_-(0), e_+(0)] + (2\lambda+u)c[e_+(0), e_-(0)]$$
$$= -a(2\lambda + u)e_-(0) - ave_+(0) + bvh(0)\lambda^{-1} - c(2\lambda + u)h(0)\lambda^{-1}.$$

Equating coefficients of $h(0)$ and $e_\pm(0)$ on both sides of this result, we obtain

$$a_x = bv\lambda^{-1} - 2c - cu\lambda^{-1}, \qquad b_x = -av, \qquad c_x = -2a\lambda - au. \tag{9.10}$$

Let a, b and c have the expansions in λ given by

$$a = \sum_{m\ge 0} a_m \lambda^{-m}, \qquad b = \sum_{m\ge 0} b_m \lambda^{-m}, \qquad c = \sum_{m\ge 0} c_m \lambda^{-m}. \tag{9.11}$$

Substituting these into the first equation of (9.10), it is found that

$$a_{0,x} + \sum_{m\ge 1} a_{m,x}\lambda^{-m} = v\sum_{m\ge 1} b_{m-1}\lambda^{-m} - 2c_0 - 2\sum_{m\ge 1} c_m\lambda^{-m} - u\sum_{m\ge 1} c_{m-1}\lambda^{-m}.$$

Equating powers of λ on both sides of this result generates the system of equations

$$a_{0,x} = -2c_0, \quad a_{m,x} = vb_{m-1} - 2c_m - uc_{m-1}, \quad m \geq 1. \tag{9.12}$$

Substituting (9.11) into the second equation, we get

$$\sum_{m\geq 0} b_{m,x}\lambda^{-m} = -v \sum_{m\geq 0} a_m \lambda^{-m}.$$

This relation implies that

$$b_{m,x} = -va_m. \tag{9.13}$$

Finally, doing the same thing with the last equation of (9.10) gives

$$\sum_{m\geq 0} c_{m,x}\lambda^{-m} = -2a_0\lambda - 2\sum_{m\geq 0} a_{m+1}\lambda^{-m} - u\sum_{m\geq 0} a_m\lambda,$$

and therefore,

$$a_0 = 0, \qquad c_{m,x} = -2a_{m+1} - ua_m. \tag{9.14}$$

Equivalently, going from the third to the first, each of these can be solved for a_{m+1}, b_{m+1} and c_{m+1} respectively, to give the coupled system,

$$a_{m+1} = \frac{1}{2}(-ua_m - c_{m,x}),$$

$$b_{m+1} = -\partial^{-1}(va_{m+1}), \qquad c_{m+1} = \frac{1}{2}(vb_m - uc_m - a_{m+1,x}), \tag{9.15}$$

introducing the notation $\partial \equiv \partial/\partial x$ and ∂^{-1} is the inverse operation, that is integration.

Starting from the initial values $a_0 = 0$, $b_0 = \beta$ and $c_0 = 0$, where β is a constant, a_{m+1} is calculated first after a_m, b_m and c_m have been recurrently calculated. From the last two recursions, b_{m+1} and c_{m+1} are obtained. The integral constant occurring in the calculation of b_{m+1} can be taken to be zero. The first few a_i, b_i and c_i are given next setting $u_i = \partial^i u$ and $v_i = \partial^i v$,

$$a_0 = 0, \qquad b_0 = \beta, \qquad c_0 = 0,$$

$$a_1 = 0, \qquad b_1 = 0, \qquad c_1 = \frac{1}{2}\beta v,$$

$$a_2 = -\frac{\beta}{4}v_1, \qquad b_2 = \frac{\beta}{8}v^2, \qquad c_2 = \frac{\beta}{8}(v_2 - 2uv).$$

Now take this set of equations (9.15) and eliminate all a_m. To do this, begin by solving the second for a_{m+1}

$$a_{m+1} = -\frac{1}{v}\partial b_{m+1}.$$

Substituting this into the first equation of (9.15), a_m can be entirely eliminated to yield,

$$-\partial b_{m+1} = \frac{1}{2} u\partial b_m - \frac{1}{2} v\partial c_m.$$

In integrated form, this reads,

$$-b_{m+1} = \frac{1}{2}\partial^{-1}(u\partial b_m) - \frac{1}{2}\partial^{-1}(v\partial c_m).$$

Similarly, substituting these results into the third equation gives,

$$c_{m+1} = \frac{1}{2}(vb_m - uc_m + \partial(\frac{1}{v}\partial b_{m+1})) = \frac{1}{4}(2v - \partial(\frac{u}{v})\partial)b_m - \frac{1}{4}(2u - \partial^2)c_m.$$

It will turn out to be useful in what follows to put these results in the form of a single matrix equation, namely,

$$\begin{pmatrix} -b_{n+1} \\ c_{n+1} \end{pmatrix} = -\frac{1}{4}\begin{pmatrix} 2\partial^{-1}u\partial & 2\partial^{-1}v\partial \\ 2v - \partial(\frac{u}{v})\partial & 2u - \partial^2 \end{pmatrix}\begin{pmatrix} -b_n \\ c_n \end{pmatrix}. \tag{9.16}$$

Using the expansions (9.11) in V given in (9.7), V takes the form,

$$\begin{aligned} V &= \sum_{m\geq 0} a_m h(0)\lambda^{-m} + \sum_{m\geq 0} b_m e_+(0)\lambda^{-m} + \sum_{m\geq 0} c_m e_-(0)\lambda^{-m} \\ &= \sum_{m\geq 0} (a_m h(-m) + b_m e_+(-m) + c_m e_-(-m)). \end{aligned}$$

From $V_x = [U, V]$ it follows that,

$$-(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+] = (\lambda^n V)_{-,x} - [U, (\lambda^n V_-].$$

Observing that the terms on the left-hand side are of gradation greater than or equal -2, while the terms on the right-hand side are of gradation less than or equal to -1, it is concluded that both sides contain only terms of gradation -1 and -2, so that

$$-(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+] \in \{\mathbb{C}e_+(0) + \mathbb{C}e_-(0) + \mathbb{C}h(-1)\}.$$

In fact, the quantity on the left-hand side can be calculated exactly in terms of the elements of the algebra.

Theorem 9.1.

$$-(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+] = (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0).$$

Proof: Write out U and $(\lambda^n V)_+$ explicitly in terms of $e_+(0)$, $e_-(0)$ and $h(0)$, there results,

$$U = 2e_+(0)\lambda + ue_+(0) + ve_-(0),$$

$$(\lambda^n V)_+ = \sum_{m=0}^{n} (a_m h(0)\lambda^{n-m} + b_m e_+(0)\lambda^{n-m} + c_m e_-(0)\lambda^{n-m}).$$

These results go into the bracket, which can be expressed as

$$[2e_+(0)\lambda + ue_+(0) + ve_-(0), \sum_{m=0}^{n} (a_m h(0)\lambda^{n-m} + b_m e_+(0)\lambda^{n-m} + c_m e_-(0)\lambda^{n-m})]$$

$$= -2\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m+1} - 2\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m+1} + u\sum_{m=0}^{n} a_m(-e_-(0))\lambda^{n-m}$$
$$-u\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m} + v\sum_{m=0}^{n} a_m(-e_+(0))\lambda^{n-m} + v\sum_{m=0}^{n} b_m h(-1)\lambda^{n-m}.$$

Now substitute the result for this bracket into the equation

$$-(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+]$$

$$= -\sum_{m=0}^{n} (a_{m,x} h(0)\lambda^{n-m} + b_{m,x} e_+(0)\lambda^{n-m} + c_{m,x} e_-(0)\lambda^{n-m})$$
$$-2\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m+1} - 2\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m-1} - u\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m}$$
$$-u\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m} - v\sum_{m=0}^{n} a_m e_+(0)\lambda^{n-m} + v\sum_{m=0}^{n} b_m h(-1)\lambda^{n-m}.$$

Substituting the known derivatives $a_{m,x}$, $b_{m,x}$ and $c_{m,x}$, this becomes,

$$-\sum_{m=0}^{n}((-2c_m - uc_{m-1} + vb_{m-1})h(0)\lambda^{n-m} - a_m ve_+(0)\lambda^{n-m}$$
$$+(-2a_{m+1} - a_m u)e_-(0)\lambda^{n-m}) - 2\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m+1} - 2\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m+1}$$
$$-u\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m} - u\sum_{m=0}^{n} c_m h(-1)\lambda^{n-m} - v\sum_{m=0}^{n} a_m e_+(0)\lambda^{n-m}$$
$$+v\sum_{m=0}^{n} b_m h(-1)\lambda^{n-m}.$$
$$= u\sum_{m=0}^{n} c_{m-1} h(0)\lambda^{n-m} - u\sum_{m=0}^{n} c_m h(0)\lambda^{n-m-1} - v\sum_{m=0}^{n} b_{m-1} h(0)\lambda^{n-m}$$
$$+v\sum_{m=0}^{n} b_m h(0)\lambda^{n-m-1} - 2\sum_{m=0}^{n} a_{m+1} e_-(0)\lambda^{n-m} - 2\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m+1}$$

$$= u\sum_{m=0}^{n-1} c_m h(0)\lambda^{n-m-1} - u\sum_{m=0}^{n} c_m h(0)\lambda^{n-m-1} - v\sum_{m=0}^{n-1} b_m h(0)\lambda^{n-m+1}$$

$$+v\sum_{m=0}^{n} b_m h(0)\lambda^{n-m-1} + 2\sum_{m=0}^{n-1} a_m e_-(0)\lambda^{n-m+1} - 2\sum_{m=0}^{n} a_m e_-(0)\lambda^{n-m+1}$$

$$= -uc_n h(0)\lambda^{-1} + vb_n h(0)\lambda^{-1} + 2a_{n+1}e_-(0)$$

$$= (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0).$$

Thus, by direct calculation, it has been shown that

$$-(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+] = (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0).$$

This is exactly the statement given in the Theorem. ♣

To cancel the term $h(-1)$ in this result, let us consider introducing the additional term

$$\chi \equiv (\frac{u}{v}c_n - b_n)\, e_+(0). \tag{9.17}$$

Next define the modified operator

$$V^{(n)} = (\lambda^n V)_+ + \chi. \tag{9.18}$$

Using (9.18), the following can be calculated,

$$-V_x^{(n)} + [U, V^{(n)}] = -(\lambda^n V)_{+,x} - \chi_x + [U, (\lambda^n V)_+ + \chi]$$

$$= -(\lambda^n V)_{+,x} + [U, (\lambda^n V)_+] - \chi_x + [U, \chi]$$

$$= (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0) - (\frac{u}{v}c_n - b_n)_x e_+(0)$$

$$+[2e_+(0)\lambda + ue_+(0) + ve_-(0), (\frac{u}{v}c_n - b_n)e_+(0)]$$

$$= (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0) - (\frac{u}{v}c_n - b_n)_x e_+(0) + (uc_n - vb_n)[e_-(0), e_+(0)]$$

$$= (vb_n - uc_n)h(-1) + 2a_{n+1}e_-(0) - (\frac{u}{v}c_n - b_n)_x e_+(0) + (uc_n - vb_n)h(-1).$$

Therefore,

$$-V_x^{(n)} + [U, V(n)] = 2a_{n+1}e_-(0) - (\frac{u}{v}c_n - b_n)_x e_+(0).$$

Now with U given by (9.9), the t derivative of U is

$$U_t = u_t e_+(0) + v_t e_-(0).$$

Substituting U_t into

$$U_t - V_x^{(n)} + [U, V^{(n)}] = 0, \tag{9.19}$$

then (9.19) becomes,

$$u_t e_+(0) + v_t e_-(0) = (\frac{u}{v} c_n - b_n)_x e_+(0) - 2a_{n+1} e_-(0). \tag{9.20}$$

Equating the coefficients of $e_+(0)$ and $e_-(0)$ on both sides of (9.20) and using (9.15), the following hierarchy of equations comes out,

$$u_t = (\frac{u}{v} c_n - b_n)_x, \qquad v_t = -2a_{n+1} = u a_n + c_{n,x}. \tag{9.21}$$

The relation (9.19) is exactly the integrability condition for the following pair of linear problems

$$\Psi_x = U\Psi, \qquad \Psi_t = V^{(n)}\Psi. \tag{9.22}$$

Let us now take the three basic equations

$$a_{m,x} = -2c_m - u c_{m-1} + v b_{m-1}, \qquad b_{m,x} = -a_m v, \qquad c_{m,x} = -2a_{m+1} - a_m u,$$

and eliminate the quantities a_j in order to develop a recursion relation. First of all,

$$b_{n+1} = -\partial^{-1} v a_{n+1} = \frac{1}{2}\partial^{-1}(uva_n - 2va_{n+1} - uva_n) = \frac{1}{2}\partial^{-1}(-u\partial b_n + v\partial c_n).$$

Next, we get

$$c_{n+1} = \frac{1}{2}(vb_n - uc_n - \partial a_{n+1}) = \frac{1}{2}(vb_n - uc_n + \frac{1}{2}(-\partial(\frac{u}{v}(-a_n v)) + \partial(-2a_{n+1} - a_n u)))$$

$$= \frac{1}{2}(vb_n - uc_n + \frac{1}{2}(-\partial(\frac{u}{v}\partial b_n) + \partial^2 c_n)) = \frac{1}{4}(2vb_n - \partial(\frac{u}{v}\partial b_n) - 2uc_n + \partial^2 c_n).$$

These two results imply the following recurrence relations for the quantities b and c which in matrix form read,

$$\begin{pmatrix} -b_{n+1} \\ c_{n+1} \end{pmatrix} = -\frac{1}{4}\begin{pmatrix} 2\partial^{-1}u\partial & 2\partial^{-1}v\partial \\ 2v - \partial(\frac{u}{v})\partial & 2u - \partial^2 \end{pmatrix}\begin{pmatrix} -b_n \\ c_n \end{pmatrix} = L\begin{pmatrix} -b_n \\ c_n \end{pmatrix}. \tag{9.23}$$

This last equation (9.23) also serves to define the matrix operator L. Beginning with the equality relation in (9.23), apply the operator L to both sides of it to give

$$\begin{pmatrix} -b_{n+2} \\ c_{n+2} \end{pmatrix} = L\begin{pmatrix} -b_{n+1} \\ c_{n+1} \end{pmatrix} = L^2\begin{pmatrix} -b_n \\ c_n \end{pmatrix}.$$

Applying L a total of m times to the basic relation produces,

$$\begin{pmatrix} -b_{n+m} \\ c_{n+m} \end{pmatrix} = L^m\begin{pmatrix} -b_n \\ c_n \end{pmatrix}. \tag{9.24}$$

Therefore, with J defined to be

$$J = \begin{pmatrix} \partial & \partial(\frac{u}{v}) \\ (\frac{u}{v})\partial & \partial \end{pmatrix},$$

it follows that

$$J\begin{pmatrix}-b_{n+m}\\ c_{n+m}\end{pmatrix}=\begin{pmatrix}\partial((\frac{u}{v})c_{n+m}-b_{n+m})\\ ua_{n+m}+\partial c_{n+m}\end{pmatrix}=\begin{pmatrix}\partial((\frac{u}{v})c_{n+m}-b_{n+m})\\ -2a_{n+m+1}\end{pmatrix}=\begin{pmatrix}\tilde{u}\\ \tilde{v}\end{pmatrix}_t.$$

It can be verified that

$$J^*=-J, \qquad JL=L^*J, \tag{9.25}$$

where A^* represents the formal conjugation of the matrix differential operator A, that is, $(A_{ij})^*=(A^*_{ji})$ and $(\sum a_i\partial^i)^*=\sum(-\partial)^i a_i$ for scalars a_i. The existence of the operators J and L, which satisfy condition (9.25), is essential to the establishment of the Hamiltonian structure of the hierarchy of equations later.

Taking $n=1$ and $\beta=2$ in recursion (9.21) we obtain that

$$\begin{pmatrix}u_t\\ v_t\end{pmatrix}=\begin{pmatrix}\partial & \partial(\frac{u}{v})\\ (\frac{u}{v})\partial & \partial\end{pmatrix}\begin{pmatrix}0\\ v\end{pmatrix}=\begin{pmatrix}\partial u\\ \partial v\end{pmatrix}$$

Taking $n=2$ and $\beta=8$, we get,

$$\begin{pmatrix}u_t\\ v_t\end{pmatrix}=-\frac{1}{4}\begin{pmatrix}\partial & \partial(\frac{u}{v})\\ (\frac{u}{v})\partial & \partial\end{pmatrix}\begin{pmatrix}2\partial^{-1}u\partial & 2\partial^{-1}v\partial\\ 2v-\partial(\frac{u}{v})\partial & 2u-\partial^2\end{pmatrix}\begin{pmatrix}0\\ 4v\end{pmatrix}$$

$$=\begin{pmatrix}\partial((\frac{u}{v})\partial^2 v-v^2-2u^2)\\ \partial^3 v-2\partial u\, v-4u\partial v\end{pmatrix}.$$

This implies the following coupled pair of nonlinear evolution equations for u and v,

$$u_t=(\frac{u}{v}v_{xx}-2u^2-v^2)_x, \qquad v_t=v_{xxx}-2u_xv-4uv_x. \tag{9.26}$$

If $u=v$ in (9.26), the equations reduce to,

$$u_t=u_{xxx}-3(u^2)_x, \qquad u_t=u_{xxx}-2u_xu-4uu_x.$$

Both of these are equivalent to $u_t=(u_{xx}-3u^2)_x$. If $u=-v$, then

$$u_t=(u_{xx}-3u^2)_x, \qquad -u_t=-u_{xxx}+2u_xu+4uu_x.$$

Thus, in both cases, when $u=\pm v$ the nonlinear evolution equations reduce to the KdV equation.

Setting $u=vw$ in (9.26), there results the pair

$$(vw)_t=(wv_{xx}-2v^2w^2-v^2)_x, \qquad v_t=v_{xxx}-2(vw)_x-4vwv_x.$$

The first equations is $wv_t + vw_t = (wv_{xx})_x - 2(v^2w^2)_x - (v^2)_x$, so substituting for v_t, this becomes

$$wv_{xxx} - 2vw(vw)_x - 2w^2(v^2)_x + vw_t = vv_xv_{xx} + wv_{xxx} - 4(vw)_xvw - (v^2)_x.$$

This simplifies to

$$vw_t = w_xw_{xx} - v^2(w^2)_x + (w^2-1)(v^2)_x.$$

To summarize these two equations, we have

$$v_t = v_{xxx} - 6wvv_x - 2w_xv^2, \qquad w_t = \frac{1}{v}w_xv_{xx} - 2vww_x + 2(w^2-1)v_x.$$

9.3 A Set of Conserved Densities.

A set of conserved densities will be derived which are common to the entire hierarchy of equations,

$$u_t = (\frac{u}{v}c_n - b_n)_x, \qquad v_t = -2a_{n+1}. \tag{9.27}$$

To carry this out, it is required to write U given by (9.9) in matrix form. Based on the representation of (9.2) and (9.5), we have

$$e_+(1) = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ \lambda & 0 \end{pmatrix}, \qquad e_+(0) = \frac{1}{2}\begin{pmatrix} 0 & \lambda^{-1} \\ 1 & 0 \end{pmatrix}, \qquad e_-(0) = \frac{1}{2}\begin{pmatrix} 0 & \lambda^{-1} \\ -1 & 0 \end{pmatrix}. \tag{9.28}$$

Moreover, define y to be

$$y = \frac{\psi_2}{\psi_1}. \tag{9.29}$$

To obtain an expression for y_x, it is necessary to calculate $\psi_{1,x}$ and $\psi_{2,x}$ from the associated linear problem in (9.6). To calculate U explicitly, the matrix forms for $e_+(1)$, $e_+(0)$ and $e_-(0)$ are substituted into the expression for U giving

$$U = 2e_+(1) + ue_+(0) + ve_-(0)$$

$$= \begin{pmatrix} 0 & 1 + \frac{1}{2\lambda}(u+v) \\ \lambda + \frac{1}{2}(u-v) & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 + \frac{1}{2\lambda}u_+ \\ \lambda + \frac{1}{2}u_- & 0 \end{pmatrix}. \tag{9.30}$$

In (9.30), $u_\pm$ is defined to be $u_\pm = (u \pm v)/2$. Therefore, the required derivatives of ψ_1 and ψ_2 in terms of $u_\pm$ are given by

$$\psi_{1,x} = (1 + \frac{u_+}{2\lambda})\psi_2, \qquad \psi_{2,x} = (\lambda + \frac{u_-}{2})\psi_1.$$

Then differentiating y given by (9.29) with respect to x, we obtain,

$$y_x = \frac{\psi_{2,x}}{\psi_1} - \frac{\psi_2}{\psi_1^2}\psi_{1,x} = \lambda + \frac{u_-}{2} - (1 + \frac{u_+}{2\lambda})y^2. \tag{9.31}$$

To carry out expansions, it is convenient to introduce $\mu^2 = \lambda$, so from (9.31), y must satisfy the equation

$$y_x - (\mu^2 + u_-) + (1 + u_+ \mu^{-2}) y^2 = 0. \tag{9.32}$$

Expanding y in inverse powers of μ,

$$y = \mu + \sum_{i=0}^{\infty} y_i \mu^{-i}, \tag{9.33}$$

and it follows from this that,

$$y_x = \sum_{i=0}^{\infty} y_{i,x} \mu^{-i}, \qquad y^2 = \mu^2 + 2\mu \sum_{i=0}^{\infty} y_i \mu^{-i} + (\sum_{i=0}^{\infty} y_i \mu^{-i})^2,$$

$$(\sum_{i=0}^{\infty} y_i \mu^{-i})^2 = \sum_{n=0}^{\infty} (\sum_{k=0}^{n} y_k y_{n-k}) \mu^{-n} \equiv \sum_{n=0}^{\infty} c_n \mu^{-n}.$$

The Cauchy product formula has been used to obtain a formula for the c_n. Substituting this information into (9.31), it takes the form,

$$\sum_{i=0}^{\infty} y_{i,x} \mu^{-i} - \mu^2 - u_- + \mu^2 + 2\mu \sum_{i=0}^{\infty} y_i \mu^{-i+1} + (\sum_{i=0}^{\infty} y_i \mu^{-i})^2 + u_+$$

$$+2u_+ \sum_{i=0}^{\infty} y_i \mu^{-i-1} + u_+ \mu^{-2} (\sum_{i=0}^{\infty} y_i \mu^{-i})^2 = 0.$$

Since $u_+ + u_- = u$ and $u_+ - u_- = v$, using the Cauchy product formula and reindexing the sums, it is found that

$$\sum_{i=0}^{\infty} y_{i,x} \mu^{-i} + v + 2 \sum_{i=-1}^{\infty} y_{i+1} \mu^{-i} + 2u_+ \sum_{i=1}^{\infty} y_{i-1} \mu^{-i} + \sum_{i=0}^{\infty} (\sum_{k=0}^{i} y_k y_{i-k}) \mu^{-i}$$

$$+u_+ \sum_{i=2}^{\infty} \sum_{k=0}^{i-2} (y_k y_{i-2-k}) \mu^{-i} = 0. \tag{9.34}$$

Expanding out the first terms, this equation takes the form

$$y_{0,x} + y_{1,x} \mu^{-1} + \sum_{i=2}^{\infty} y_{i,x} \mu^{-i} + v$$

$$+2y_0 \mu + 2y_1 + 2y_2 \mu^{-1} + 2 \sum_{i=2}^{\infty} y_{i+1} \mu^{-i} + 2u_+ y_0 \mu^{-1} + 2u_+ \sum_{i=2}^{\infty} y_{i-1} \mu^{-i}$$

$$+y_0^2+2y_0y_1\mu^{-1}+\sum_{i=2}^{\infty}(\sum_{k=0}^{i}y_ky_{i-k})\mu^{-i}+u_+\sum_{i=2}^{\infty}\sum_{k=0}^{i-2}(y_ky_{i-k-2})\mu^{-i}=0.$$

Equating the coefficients of the three lowest powers $(-1,0,1)$ of μ in this to zero, the following system of three equations is obtained,

$$y_{1,x}+2y_2+2y_0y_1=0, \qquad y_{0,x}+v+2y_1+y_0^2=0, \qquad 2y_0=0. \tag{9.35}$$

Substituting $y_0=0$ back into the first two equations of (9.35), it is clear that this system reduces to,

$$y_0=0, \qquad y_1=-\frac{1}{2}v, \qquad y_2=-\frac{1}{2}y_{1,x}=\frac{1}{4}v_x. \tag{9.36}$$

Imposing these results, what is left of the general recursion above is given by

$$\sum_{n=2}^{\infty}(y_{n,x}+2y_{n+1}+2u_+y_{n-1}+(\sum_{k=0}^{n}y_ky_{n-k})+u_+(\sum_{k=0}^{n-2}y_ky_{n-k-2}))\mu^{-n}=0.$$

Requiring that the coefficient of each remaining power of μ^{-n} in this vanish for each $n\geq 2$ implies the following recursion relation,

$$y_{n,x}+2y_{n+1}+2u_+y_{n-1}+\sum_{k=0}^{n}y_ky_{n-k}+u_+\sum_{k=0}^{n-2}y_ky_{n-k-2}=0, \quad n\geq 2. \tag{9.37}$$

As an example, put $n=2$ in (9.37). Then given (9.36), the result becomes

$$y_{2,x}+2y_3+2u_+y_1+(y_0y_2+y_1^2+y_2y_0)+u_+y_0^2=0.$$

Hence, solving for $2y_3$ in this gives,

$$2y_3=-y_{2,x}-2u_+y_1-y_1^2=\frac{1}{8}(-v_{xx}+2uv+v^2).$$

Now the conserved densities $\tilde{H}_i$ can be derived for the hierarchy (9.27). The generating function for $\tilde{H}_i$ is given by

$$\tilde{H}=(1+u_+\lambda^{-1})y=(1+u_+\mu^{-2})y. \tag{9.38}$$

This must match exactly the expansion

$$\tilde{H}=\mu+\sum_{i=1}^{\infty}\tilde{H}_i\mu^{-i}. \tag{9.39}$$

Substituting (9.33) into (9.38) so the coefficients of the two expressions can be compared, we obtain

$$\tilde{H}=\mu+\sum_{i=0}^{\infty}y_i\mu^{-i}+u_+\mu^{-1}+u_+\sum_{i=2}^{\infty}y_i\mu^{-i-2}$$

$$= \mu + (y_1 + u_+)\mu^{-1} + \sum_{i=2}^{\infty} (y_i + u_+ y_{i-2})\mu^{-i}. \tag{9.40}$$

Equating the coefficients of like powers of μ on both sides of (9.39) and (9.40), the following results are obtained,

$$\tilde{H}_1 = y_1 + u_+ = \frac{1}{2}u,$$

$$\tilde{H}_2 = y_2 + \frac{1}{2}(u+v)y_0 = \frac{1}{4}v_x,$$

$$\tilde{H}_3 = y_3 + \frac{1}{2}(u+v)y_1 = -\frac{1}{8}v_{xx} - \frac{1}{8}v^2.$$

9.4 Hamiltonian Structures.

Let the Hamiltonian be given by (9.38) and (9.32). The constrained variational calculus will be applied to deduce the equation which K and S should satisfy. Proceeding in this direction, K and S are given as

$$K = \frac{\delta \tilde{H}}{\delta u}, \qquad S = \frac{\delta \tilde{H}}{\delta v},$$

To calculate K and S, introduce the Lagrange multipliers θ_1 and θ_2, and form the sum

$$W = \tilde{H} + \theta_1(\tilde{H} - (1 + u_+\lambda^{-1})y) + \theta_2(y_x - (\lambda + u_-) + \tilde{H}y). \tag{9.41}$$

Evaluating $\delta W/\delta \tilde{H}$ and $\delta W/\delta y$ and then setting them to zero, two equations result,

$$1 + \theta_1 + y\theta_2 = 0, \qquad -\theta_1(1 + u_+\lambda^{-1}) - \theta_{2x} + \theta_2 \tilde{H} = 0.$$

Differentiating $\tilde{H}$, K and S result,

$$K = \frac{\delta \tilde{H}}{\delta u} = \frac{\delta W}{\delta u} = -\frac{1}{2}(\lambda^{-1}\theta_1 y + \theta_2), \quad S = \frac{\delta \tilde{H}}{\delta v} = \frac{\delta W}{\delta v} = -\frac{1}{2}(\lambda^{-1}\theta_1 y - \theta_2).$$

From the equation $1 + \theta_1 + y\theta_2 = 0$, we have an expression for $\theta_1 = -1 - y\theta_2$. Using θ_1 in this form, it can be eliminated from K and S in favor of θ_2

$$K = \frac{1}{2}\lambda^{-1}y(1 + y\theta_2) - \frac{1}{2}\theta_2, \qquad S = \frac{1}{2}\lambda^{-1}y(1 + y\theta_2) + \frac{1}{2}\theta_2. \tag{9.42}$$

The equation $-\theta_1(1 + u_+\lambda^{-1}) - \theta_{2x} + \theta_2\tilde{H} = 0$ yields an expression for θ_{2x}. Therefore,

$$\theta_{2x} = (1 + y\theta_2)(1 + u_+\lambda^{-1}) + \theta_2(1 + u_+\lambda^{-1})y = (1 + 2y\theta_2)(1 + u_+\lambda^{-1}),$$

$$y_x = \lambda + u_- - (1 + u_+\lambda^{-1})y^2. \tag{9.43}$$

With θ_{2x} and y_x from (9.43), both K and S as well as their derivatives can be represented in terms of y and θ_2. This has just been done for K and S. For the derivatives, this can be done as well. Differentiating K, we have

$$K_x = \frac{1}{2}\lambda^{-1} y_x (1 + y\theta_2) + \frac{1}{2}\lambda^{-1} y(y_x\theta_2 + y\theta_{2x}) - \frac{1}{2}\theta_{2x},$$

Replacing y_x, θ_{2x} and $u_\pm$ by their expressions preceding, we obtain,

$$K_x = -\frac{v}{2\lambda}(1 + 2\theta_2 y).$$

Similarly, differentiating S, there results,

$$S_x = \frac{1}{2}\lambda^{-1} y_x (1 + y\theta_2) + \frac{1}{2}\lambda^{-1} y(y_x\theta_2 + y\theta_{2x}) + \frac{1}{2}\theta_{2x} = \frac{1}{2}(2 + \frac{u}{\lambda})(1 + 2\theta_2 y).$$

Summarizing the results for K_x and S_x,

$$K_x = -\frac{v}{2\lambda}(1 + 2\theta_2 y), \qquad S_x = \frac{1}{2}(2 + \frac{u}{\lambda})(1 + 2\theta_2 y). \tag{9.44}$$

Based on the results for K, S and K_x, S_x, the following quantities are calculated,

$$\begin{aligned} 2(v_x K + u_x S) &= v_x \lambda^{-1} y(1 + y\theta_2) - v_x\theta_2 + u_x \lambda^{-1} y(1 + y\theta_2) + u_x\theta_2 \\ &= (u_x - v_x)\theta_2 + (u_x + v_x)\lambda^{-1} y(1 + y\theta_2). \end{aligned} \tag{9.45}$$

Similarly,

$$\begin{aligned} 2(vK_x + uS_x) &= -v^2\lambda^{-1}(1 + 2\theta_2 y) + u(2 + u\lambda^{-1})(1 + 2\theta_2 y) \\ &= (2u + \frac{u^2}{\lambda} - \frac{v^2}{\lambda})(1 + 2\theta_2 y). \end{aligned} \tag{9.46}$$

Also, we can form

$$\begin{aligned} 4\lambda S_x = (2u + 4\lambda)(1 + 2y\theta_2), \quad \frac{u}{v}K_x + S_x &= -\frac{u}{2\lambda v}(1 + 2\theta_2 y) + \frac{1}{2}(2 + \frac{u}{\lambda})(1 + 2\theta_2 y) \\ &= 1 + 2\theta_2 y. \end{aligned} \tag{9.47}$$

Differentiating the second equation in (9.47), we obtain

$$\partial(\frac{u}{v}K_x + S_x) = 2\theta_{2x} y + 2\theta_2 y_x$$

$$= 2(1 + 2y\theta_2)(1 + u_+\lambda^{-1})y + 2\theta_2(\lambda + u_- - (1 + u_+\lambda^{-1})y^2)$$

Differentiating this once more, the second derivative is found to be

$$\partial^2(\frac{u}{v}K_x + S_x)$$

$$= 4(y_x\theta_2 + y\theta_{2x})(1 + u_+\lambda^{-1})y + (1 + 2y\theta_2)(u_x + v_x)\lambda^{-1}y$$
$$+2(1 + 2y\theta_2)(1 + u_+\lambda^{-1})y_x + 2\theta_{2x}(\lambda + u_- - (1 + u_+\lambda^{-1})y^2)$$
$$+\theta_2(u_x - v_x - \lambda^{-1}(u_x + v_x)y^2 - 2(1 + u_+\lambda^{-1})2yy_x)$$
$$= \frac{1}{\lambda}(2\lambda+u+v)(2\lambda+u-v)(1+2y\theta_2)+\theta_2(u_x-v_x)+\frac{1}{\lambda}(1+y^2\theta_2)(u_x+v_x). \quad (9.48)$$

Based on these results, the following expressions can be calculated,

$$(2\partial v\partial^{-1} - \partial^2(\frac{u}{v}))K_x + (2\partial u\partial^{-1} - \partial^2 + 4\lambda)S_x$$
$$= 2(v_xK + u_xS) + 2(vK_x + uS_x) + 4\lambda S_x - \partial^2(\frac{u}{v}K_x + S_x) = 0.$$

A similar calculation yields

$$(2u + 4\lambda)K_x + 2vS_x = -v\lambda^{-1}(u + 2\lambda)(1 + 2\theta_2 y) + v(2 + \frac{u}{\lambda})(1 + 2\theta_2 y)$$
$$= (-\frac{v}{\lambda}(u + 2\lambda) + v(2 + u\lambda^{-1}))(1 + 2\theta_2 y) = 0.$$

These last two results can be summarized in the form of a 2×2 matrix. The system is

$$\begin{pmatrix} 2u + 4\lambda & 2v \\ 2\partial v\partial^{-1} - \partial^2(\frac{u}{v}) & 2\partial u\partial^{-1} - \partial^2 + 4\lambda \end{pmatrix} \begin{pmatrix} K_x \\ S_x \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (9.49)$$

Theorem 9.2. The matrix system defined by (9.49) is equivalent to the following matrix equation given by

$$\partial[-\begin{pmatrix} 2\partial^{-1}u\partial & 2\partial^{-1}v\partial \\ 2v - \partial(\frac{u}{v})\partial & 2u - \partial^2 \end{pmatrix} - \begin{pmatrix} 4\lambda & 0 \\ 0 & 4\lambda \end{pmatrix}] \begin{pmatrix} K \\ S \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (9.50)$$

Proof: It suffices to verify that (9.50) holds one row at a time. The first row of (9.50) can be written out in the form,

$$-\partial(2\partial^{-1}u\partial K + 2\partial^{-1}v\partial S + 4\lambda K) = 0.$$

Bringing the operator through the bracket, this assumes the form

$$2uK_x + 2vS_x + 4\lambda K_x = 0.$$

This is exactly the first row of system (9.49).

The second row of (9.50) is given by

$$-\partial(2vK - \partial(\frac{u}{v})K_x + 2uS - \partial^2 S + 4\lambda S_x) = 0.$$

Taking the operator through the bracket, this assumes the form,

$$2(\partial v)\partial^{-1}K_x - \partial^2(\frac{u}{v})K_x + 2\partial(u\partial^{-1}S_x) - \partial^2 S + 4\lambda S_x = 0.$$

In fact, this equation is exactly the second row implied by (9.49), and this completes the proof. ♣

If L is the operator defined in (9.23), then the matrix problem (9.50) in Theorem 9.2 may be expressed very succinctly in the form,

$$\partial(L-\lambda)\begin{pmatrix} K \\ S \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{9.51}$$

Equation (9.51) implies that

$$(L-\mu^2)\begin{pmatrix} K \\ S \end{pmatrix} = -\mu\begin{pmatrix} \tilde{K}_1 \\ \tilde{S}_1 \end{pmatrix} - \mu^2\begin{pmatrix} \tilde{K}_2 \\ \tilde{S}_2 \end{pmatrix}, \tag{9.52}$$

where $\tilde{K}_i$ and $\tilde{S}_i$, $i = 1, 2$ are to be determined. Since $\tilde{H}$ has an expansion of the form (9.39), it follows that,

$$K = \frac{\delta\tilde{H}}{\delta u} = \sum_{i=1}^{\infty}\frac{\delta\tilde{H}_i}{\delta u}\mu^{-i} = \sum_{i=1}^{\infty}\tilde{K}_i\mu^{-i}, \quad S = \frac{\delta\tilde{H}}{\delta v} = \sum_{i=1}^{\infty}\frac{\delta\tilde{H}_i}{\delta v}\mu^{-i} = \sum_{i=1}^{\infty}\tilde{S}_i\mu^{-i}. \tag{9.53}$$

Substituting (9.53) into (9.52), the corresponding relations for the $\tilde{K}_i$ and $\tilde{S}_i$ can be obtained,

$$L\begin{pmatrix} \sum_{i=1}^{\infty}\tilde{K}_i\mu^{-i} \\ \sum_{i=1}^{\infty}\tilde{S}_i\mu^{-i} \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{\infty}\tilde{K}_i\mu^{-i+2} \\ \sum_{i=1}^{\infty}\tilde{S}_i\mu^{-i+2} \end{pmatrix} - \mu\begin{pmatrix} \tilde{K}_1 \\ \tilde{S}_1 \end{pmatrix} - \mu^2\begin{pmatrix} \tilde{K}_2 \\ \tilde{S}_2 \end{pmatrix},$$

that is,

$$L\begin{pmatrix} \sum_{i=1}^{\infty}\tilde{K}_i\mu^{-i} \\ \sum_{i=1}^{\infty}\tilde{S}_i\mu^{-i} \end{pmatrix} = \mu\begin{pmatrix} \tilde{K}_1 \\ \tilde{S}_1 \end{pmatrix} + \begin{pmatrix} \tilde{K}_0 \\ \tilde{S}_0 \end{pmatrix} + \begin{pmatrix} \sum_{i=3}^{\infty}\tilde{K}_i\mu^{-i+2} \\ \sum_{i=3}^{\infty}\tilde{S}_i\mu^{-i+2} \end{pmatrix} - \mu\begin{pmatrix} \tilde{K}_1 \\ \tilde{S}_1 \end{pmatrix} - \mu^2\begin{pmatrix} \tilde{K}_2 \\ \tilde{S}_2 \end{pmatrix}.$$

The term in μ cancels on the right hand side, and we are left with simply,

$$L\begin{pmatrix} \sum_{i=1}^{\infty}\tilde{K}_i\mu{-i} \\ \sum_{i=1}^{\infty}\tilde{S}_i\mu^{-i} \end{pmatrix} = -\mu^2\begin{pmatrix} \tilde{K}_2 \\ \tilde{S}_2 \end{pmatrix} + \begin{pmatrix} \tilde{K}_0 \\ \tilde{S}_0 \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^{\infty}\tilde{K}_{i+2}\mu^{-i} \\ \sum_{i=1}^{\infty}\tilde{S}_{i+2}\mu^{-i} \end{pmatrix}.$$

In order for these to possibly be able to match from both sides, it follows that $\tilde{K}_0 = \tilde{S}_0 = 0$ and $\tilde{K}_2 = \tilde{S}_2 = 0$.

Definition 9.1. The notation $f \sim g$ implies that $f - g \sim \partial h$ for some polynomial h of u_i and v_i, so $f \sim 0$ means $f = \partial h$, that is, f is a total derivative.

The results so far are leading to a claim about the properties of the $\tilde{H}_{2k}$, which we state now.

Theorem 9.3. The set of $\tilde{H}_{2k}$ satisfy

$$\tilde{H}_{2k} \sim 0.$$

Proof: Since $\tilde{H}_2 \sim 0$, it follows of course that $\tilde{K}_2 = \tilde{S}_2 = 0$, and thus

$$\tilde{K}_{2k} = \tilde{S}_{2k} = 0, \qquad k \geq 1.$$

This is equivalent to $\tilde{H}_{2k} \sim 0$. This leaves precisely

$$L \begin{pmatrix} \tilde{K}_i \\ \tilde{S}_i \end{pmatrix} = \begin{pmatrix} \tilde{K}_{i+2} \\ \tilde{S}_{i+2} \end{pmatrix}, \quad i = 1, 2, \cdots .$$

Alternatively, this in turn implies that $\tilde{K}_{2k} = \tilde{S}_{2k} = 0$, $k \geq 1$, which is equivalent to $\tilde{H}_{2k} \sim 0$. ♣

One can determine as well

$$\begin{pmatrix} \tilde{K}_1 \\ \tilde{S}_1 \end{pmatrix} = -\frac{1}{2\beta} \begin{pmatrix} -b_0 \\ c_0 \end{pmatrix}. \tag{9.54}$$

Replacing i by $2k-1$, the recursion becomes

$$L \begin{pmatrix} \tilde{K}_{2k-1} \\ \tilde{S}_{2k-1} \end{pmatrix} = \begin{pmatrix} \tilde{K}_{2k+1} \\ \tilde{S}_{2k+1} \end{pmatrix}. \tag{9.55}$$

Setting

$$K_n = \tilde{K}_{2n-1}, \qquad S_n = \tilde{S}_{2n-1},$$

recursion (9.55) can be written in the form,

$$L \begin{pmatrix} K_n \\ S_n \end{pmatrix} = \begin{pmatrix} K_{n+1} \\ S_{n+1} \end{pmatrix}, \qquad n \geq 1,$$

$$K = \mu \sum_{i=1}^{\infty} K_i \lambda^{-i}, \qquad S = \mu \sum_{i=1}^{\infty} S_i \lambda^{-i}.$$

Comparing this with (9.23), it can be seen that both pairs (K_n, S_n) and $(-b_n, c_n)$ satisfy the same homogeneous recurrence relations, and moreover (9.54) holds for their initial values. It may then be concluded that

$$\begin{pmatrix} K_n \\ S_n \end{pmatrix} = -\frac{1}{2\beta} \begin{pmatrix} -b_n \\ c_n \end{pmatrix} = L^n \begin{pmatrix} \frac{1}{2} \\ 0 \end{pmatrix}.$$

From this, it is deduced that b_n, c_n and hence a_n are polynomials in u_i and v_i.

Therefore, the hierarchy of equations (9.27) is a pure differential one in spite of the fact that the operator L used in the derivation of the hierarchy is intego-differential, a fact common to many integrable nonlinear evolution equations.

Thus with matrix J defined just below (9.24) the hierarchy assumes the form

$$\begin{pmatrix} u \\ v \end{pmatrix}_t = J \begin{pmatrix} -b_n \\ c_n \end{pmatrix} = J \begin{pmatrix} \dfrac{\delta H_n}{\delta u} \\ \dfrac{\delta H_n}{\delta v} \end{pmatrix}, \qquad H_n = -2\beta \tilde{H}_{2n-1}. \tag{9.56}$$

Introduce the inner product defined by

$$\langle A, B \rangle = \int A^T B \, dx.$$

By means of this inner product, the following generalized Poisson bracket can be introduced

$$\{G, H\} = \langle \frac{\delta G}{\delta q}, J \frac{\delta H}{\delta q} \rangle = (\frac{\delta G}{\delta q})^T J \frac{\delta H}{\delta q} = (\frac{\delta G}{\delta u}, \frac{\delta G}{\delta v}) J (\frac{\delta H}{\delta u}, \frac{\delta H}{\delta v})^T$$

$$= (\frac{\delta G}{\delta u}, \frac{\delta G}{\delta v}) J \begin{pmatrix} \dfrac{\delta H}{\delta u} \\ \dfrac{\delta H}{\delta v} \end{pmatrix}.$$

Since $J^* = -J$, $JL = L^* J$, the following bracket can be worked out,

$$-\frac{1}{4\beta^2} \{H_n, H_m\} = \langle \begin{pmatrix} K_n \\ S_n \end{pmatrix}, J \begin{pmatrix} K_m \\ S_m \end{pmatrix} \rangle = \langle \begin{pmatrix} K_n \\ S_n \end{pmatrix}, JL \begin{pmatrix} K_{m-1} \\ S_{m-1} \end{pmatrix} \rangle$$

$$= \langle \begin{pmatrix} K_n \\ S_n \end{pmatrix}, L^* J \begin{pmatrix} K_{m-1} \\ S_{m-1} \end{pmatrix} \rangle = \langle L \begin{pmatrix} K_n \\ S_n \end{pmatrix}, J \begin{pmatrix} K_{m-1} \\ S_{m-1} \end{pmatrix} \rangle$$

$$= \langle \begin{pmatrix} K_{n+1} \\ S_{n+1} \end{pmatrix}, J \begin{pmatrix} K_{m-1} \\ S_{m-1} \end{pmatrix} \rangle = -\frac{1}{4\beta^2} \{H_{n+1}, H_{m-1}\}.$$

This immediately implies that

$$\{H_n, H_m\} = \{H_{n+1}, H_{m-1}\}. \tag{9.57}$$

The equality in effect represents equality with respect to the equivalence classes, so $f = g$ is interpreted as $f = g \pmod{\partial h}$ or $f \sim g$, as in Definition 9.1. Applying the result above repeatedly, it follows that

$$\{H_n, H_m\} = \{H_m, H_n\}.$$

However, the bracket clearly satisfies $\{H_n, H_m\} = -\{H_m, H_n\}$ due to the antisymmetry of the operator J. Therefore, it is the case that the H_n satisfy

$$\{H_n, H_m\} = \int (\frac{\delta H_n}{\delta q})^T J (\frac{\delta H_m}{\delta q}) \, dx = 0. \tag{9.58}$$

Thus, it has been proved that the conserved densities H_n are pairwise in involution. Using the time variable t_k to distinguish the equations of the hierarchy, we write,

$$q_{t_k} = J\,\frac{\delta H_k}{\delta q}. \tag{9.59}$$

Consequently,

$$(H_n)_{t_k} = (\frac{\delta H_n}{\delta q})^T\, q_{t_k} = \langle \frac{\delta H_n}{\delta q}, J\frac{\delta H_k}{\delta q} \rangle = \{H_n, H_k\} = 0. \tag{9.60}$$

The result in (9.60) implies that the $\{H_n\}$ are common conserved densities for the entire hierarchy (9.59).

10 The Laplacian and Heat Operators on Riemannian Manifolds.

Let (M,g) be a Riemannian manifold and let $C^{1,2}(M)$ be the space of functions $f:(0,\infty)\times M\to\mathbb{R}$, which are continuous on $[0,\infty)\times M$, C^1-differentiable in the first variable, and C^2-differentiable in the second variable. The Laplacian written out in a coordinate system $\{x_i\}$ is given by

$$\Delta f = -\frac{1}{\sqrt{g}}\frac{\partial}{\partial x_j}(\sqrt{g}g^{ij}\frac{\partial f}{\partial x_i}). \tag{10.1}$$

Definition 10.1. The operator $P=\dfrac{\partial}{\partial t}+\Delta$ defined on the space $C^{1,2}(M)$ is called the heat operator on (M,g).

In order to invert the heat operator, it is required to study the fundamental solution [**57,58**].

Definition 10.2. A fundamental solution K for the heat operator $P=\partial_t+\Delta_y$ is a function $K:M\times M\times(0,\infty)\to\mathbb{R}$ which has the following properties:

(i) $K\in C(M\times M\times(0,\infty))$, C^2 in first variable, C^1 in second.

(ii) $(\dfrac{\partial}{\partial t}+\Delta_y)K(x,y,t)=0$ for all $t>0$.

(iii) $\lim_{t\searrow 0}K(x,y,t)=\delta(x-y),\quad \forall x\in M$,

where δ_x is the Dirac distribution centered at x and the limit (iii) is considered in the distribution sense,

$$\lim_{t\searrow 0}\int_M K(x,y,t)\phi(x)dv(x)=\phi(y),\quad \forall\phi\in C_0(M),\,\forall\, x\in M,$$

and $C_0(M)$ denotes the set of smooth functions with compact support, and if $\{x_i\}$ is a coordinate system for M, the volume form is

$$dv(x)=\sqrt{|g_{ij}(x)|}\,dx_1\wedge\cdots\wedge dx_n.$$

10.1 The Heat Operator on Compact Manifolds.

Let (M,g) be a compact Riemannian manifold. An inner product on M can be defined as

$$(f,g)_0=\int_M fg\,dv(x). \tag{10.2}$$

Set $||f||_{L^2}=(f,f)^{1/2}$, then the space $L^2(M)$ is obtained from $\mathcal{F}(M)=\{f: M\to\mathbb{R}, f\in C^\infty\}$ by invoking completeness with respect to the norm $||\cdot||_{L^2}$.

The real numbers λ for which there is a non-zero smooth function f such that $\Delta f=\lambda f$ are called eigenvalues [**57**]. Let $V_\lambda(M,g)=\{f:M\to\mathbb{R}:\Delta f=\lambda f\}$ be the vector space of the eigenfunctions together with the zero function. The number $m_\lambda=\dim V_\lambda(M,g)$ is called the multiplicity of λ.

The fundamental solution of P will be found here for the case of a compact manifold. Hence, the following spectral theorem holds for the Laplace operator on Riemannian manifolds.

Theorem 10.1. (i) The eigenvalues are nonnegative and form a countable infinite set

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \lambda_3 < \cdots ,$$

such that $\lambda_k \to +\infty$ as $k \to +\infty$. Moreover, the series

$$\sum_{k\geq 1} \frac{1}{\lambda_k^2}$$

converges.

(ii) Each eigenvalue λ_k has finite multiplicity m_k. The eigenspaces $V_{\lambda_k}(M,g)$ and $V_{\lambda_j}(M,g)$, $k \neq j$ are orthogonal with respect to the inner product $(,)_0$.

(iii) By means of the Gram-Schmidt procedure beginning with the set of eigenfunctions, a complete orthonormal system of eigenfunctions may be obtained $\{f_{kj} : k \in \mathbb{N}, j = 1, \cdots , m_k\}$ such that

$$h = \sum_{k=0}^{\infty} \sum_{j=1}^{m_k} a_{kj} f_{kj}, \qquad \forall\, h \in L^2(M),$$

where $a_{kj} = (h, f_{kj})_0$. In particular, the Parseval identity holds,

$$||h||_0^2 = \sum_{k=0}^{\infty} \sum_{j=1}^{m_k} (h, f_{kj})_0^2.$$

The following result provides a formula for the fundamental solution on a compact Riemannian manifold.

Proposition 10.1. Let $\{\varphi_i : i \in \mathbb{N}\}$ be a complete orthonormal system of eigenfunctions for the Laplace operator on a compact Riemannian manifold (M, g) such that

$$\lambda_0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \cdots .$$

Then the fundamental solution is given by

$$K(x, y, t) = \sum_{i=0}^{\infty} e^{-\lambda_i t} \varphi_i(x)\varphi_i(y). \tag{10.3}$$

Proof: Since the system $\{\varphi_i : i \in \mathbb{N}\}$ is an orthonormal basis of the Hilbert space $L^2(M)$, the existence of a fundamental solution is assumed for fixed x and t. Thus,

$$K(x, \cdot, t) = \sum_{i=0}^{\infty} \rho_i(x, t)\varphi_i,$$

where

$$\rho_i(x,t) = \int_M K(x,y,t)\varphi_i(y)\,dv(y)$$

Differentiating ρ_i with respect to t yields,

$$\frac{\partial \rho_i}{\partial t} = \int_M \frac{\partial K}{\partial t}(x,y,t)\varphi_i(y)dv(y) = (\frac{\partial K}{\partial t},\varphi_i)_0 = -(\Delta_y K,\varphi_i) = -(K,\Delta_y\varphi_i)_0$$

$$= -\lambda_i(K,\varphi_i)_0 = -\lambda_i\rho_i.$$

This is a differential equation which can be solved to yield $\rho_i(x,t) = c_i(x)e^{-\lambda_i t}$. The functions $c_i(x)$ satisfy

$$\lim_{t\searrow 0}\rho_i(,x,t) = \lim_{t\searrow 0}\int_M K(x,y,t)\varphi_i(y)\,dv(y) = \int_M \delta(y-x)\varphi_i(y)dv(y) = \varphi_i(x).$$

On the other side,

$$\lim_{t\searrow 0}\rho_i(x,t) = c_i(x),$$

and hence $c_i(x) = \varphi_i(x)$. This proves (10.3).

♣

This proof assumes the existence of a fundamental solution for the heat operator. The series $\sum_{i=0}^{\infty}\rho_i(x,t)\varphi_i(y)$ is pointwise convergent on $(0,\infty)\times M\times M$ and its sum is $K(x,y,t)$.

It may be of interest to solve the initial value problem for the heat operator: Given a continuous function $g\in C^0(M)$, find a function $f\in C^{1,2}(M)$ such that

(I) $\quad (\frac{\partial}{\partial t}+\Delta)f = 0,$

(II) $\quad \lim_{t\searrow 0} f(x,t) = g(x), \quad \forall x\in M.$

Proposition 10.2. The solution to the initial value problem (I)-(II) is given by the expression

$$f(x,t) = \int_M K(x,y,t)g(y)\,dv(y), \tag{10.4}$$

where K is given by (10.3).

Proof:

$$\frac{\partial}{\partial t}f(x,t) = \frac{\partial}{\partial t}\int_M \sum_{i=0}^{\infty} e^{-\lambda_i t}\,\varphi_i(x)\varphi_i(y)g(y)\,dv(y).$$

The Laplacian of f with respect to x is,

$$\Delta_x f(x,t) = \Delta_x \int_M \sum_{i=0}^{\infty} e^{-\lambda_i t}\varphi_i(x)\varphi_i(y)g(y)\,dv(y)$$

$$= \int_M \sum_{i=0}^{\infty} e^{-\lambda_i t}\Delta_x\,\varphi_i(x)\varphi_i(y)g(y)\,dv(y) = \int_M \sum_{i=0}^{\infty}\lambda_i e^{-\lambda_i t}\varphi_i(x)\varphi_i(y)g(y)dv(y).$$

Adding these two expressions gives the heat equation,

$$(\frac{\partial}{\partial t} + \Delta) f = 0.$$

It remains to be shown that

$$\lim_{t \searrow 0} f(x,t) = g(x).$$

Using the third part of Definition 10.2, it follows that

$$\lim_{t \searrow 0} f(x,t) = \lim_{t \searrow 0} \int_M K(x,y,t) g(y)\, dv(y) = \int_M \lim_{t \searrow 0} K(x,y,t) g(y) dv(y)$$

$$= \int_M \delta(y-x) g(y)\, dv(y) = g(x).$$

10.2 Heat Kernel on Radially Symmetric Spaces.

It has been seen that $\mathbb{R}^n$ equipped with the standard metric is a radially symmetric space. This means that the scalar curvature of the geodesic sphere depends only on its radius. It is known that the fundamental solution in this case is given by,

$$K(x,y,t) = (4\pi t)^{-n/2}\, e^{-\frac{|x-y|^2}{4t}}, \qquad t > 0. \tag{10.5}$$

The result (10.5) is a product between the volume function $v(t) = t^{-n/2}$ and an exponential function which has exponent

$$-\frac{|x-y|^2}{4t} = -\frac{1}{2} S.$$

Here, S is the classical action between the points x and y within time t.

Lemma 10.1. For any smooth function φ on a Riemannian manifold (M,g), it is the case that

$$\Delta e^{\varphi} = e^{\varphi} (\Delta\varphi - |\nabla\varphi|^2). \tag{10.6}$$

Proof: It is shown that $\nabla e^{\varphi} = e^{\varphi} \nabla\varphi$. This arises out of the definition of the gradient. For any vector field X,

$$g(\nabla e^{\varphi}, X) = X(e^{\varphi}) = \sum_i X^i \partial_{x_i} e^{\varphi} = e^{\varphi} X(\varphi)$$

$$= e^{\varphi} g(\nabla\varphi, X) = g(e^{\varphi}\nabla\varphi, X).$$

Using the formula,

$$\mathrm{div}(fX) = f \mathrm{div}(X) + g(\nabla f, X),$$

it follows by taking $f = e^{\varphi}$, $X = \nabla\varphi$, that,

$$\Delta e^{\varphi} = -\mathrm{div}(\nabla e^{\varphi}) = -\mathrm{div}(e^{\varphi}\nabla\varphi) = -e^{\varphi}\mathrm{div}\nabla\varphi - g(\nabla e^{\varphi}, \nabla\varphi).$$

Since φ is a function on M, $\Delta f = -\mathrm{div}(\nabla f)$, so $-\Delta\varphi = \mathrm{div}\nabla\varphi$, and thus,

$$\Delta e^{\varphi} = e^{\varphi}\Delta\varphi - g(e^{\varphi}\nabla\varphi, \nabla\varphi) = e^{\varphi}(\Delta\varphi - g(\nabla\varphi, \nabla\varphi)) = e^{\varphi}(\Delta\varphi - |\nabla\varphi|^2).$$

♣

Let $d = d(x_0, x)$ be the Riemannian distance between the points x_0 and $x \in M$. Let

$$f = \frac{1}{2}d^2(x_0, x).$$

Now the distance function satisfies $|\nabla d^2|^2 = 4d^2$, hence the function f satisfies the equation

$$|\nabla f|^2 = 2f.$$

The classical action starting at x_0 is written as

$$S = S(x_0, x, t) = \frac{d^2(x_0, x)}{2t} = \frac{f}{t}. \tag{10.7}$$

Then it follows that,

$$|\nabla S|^2 = |\nabla \frac{f}{t}|^2 = \frac{1}{t^2}|\nabla f|^2 = \frac{2f}{t^2} = \frac{2S}{t} = 2E,$$

where $E = \frac{1}{2t^2}d^2(x_0, x)$ is the energy.

A fundamental solution of the form,

$$K(x_0, x, t) = V(t)e^{kS} \tag{10.8}$$

will be sought where $k \in \mathbb{R}$ is a constant, $V(t)$ is a differentiable function and S is the action in (10.7). Differentiating K and applying the Hamilton-Jacobi equation $S_t = -E$, it follows that

$$\frac{\partial K}{\partial t} = \frac{\partial V}{\partial t}e^{kS} + kVe^{kS}\frac{\partial S}{\partial t} = e^{kS}(\frac{\partial V}{\partial t} - kVE).$$

Lemma 10.1 implies that,

$$\Delta(V(t)e^{kS}) = V(t)e^{kS}(\Delta(kS) - |\nabla kS|^2) = e^{kS}V(t)(k\Delta S - 2k^2E).$$

Therefore,

$$(\frac{\partial}{\partial t} + \Delta)(V(t)e^{kS}) = e^{kS}\frac{\partial V}{\partial t} - ke^{kS}VE + e^{kS}V(t)(k\Delta S - 2k^2E)$$

$$= e^{kS}V(t)(\frac{V'(t)}{V(t)} + k\Delta S - (2k+1)kE).$$

Choose $k = -1/2$ and let $V(t)$ satisfy the equation

$$\frac{V_t}{V} + k\Delta S = 0,$$

which can be rewritten in the form,

$$V_t = \frac{1}{2}\Delta S V(t).$$

As the manifold (M, g) is radially symmetric, ΔS is a function of only the t variable. Thus there is a function $h(t)$ such that,

$$h(t) = \frac{1}{2}\Delta S = \frac{n-1}{2}\alpha(t),$$

where $\alpha(t) = \alpha(c(t))$ is the mean scalar curvature of the geodesic sphere centered at x_0 with radius t. The solution is given by

$$V(t) = V(t_0)e^{\int_{t_0}^{t} h(u)\, du}.$$

Theorem 10.2. Let (M, g) be a radially symmetric space about the point $x_0 \in M$. Then the fundamental solution for the heat operator is given by

$$K(x_0, x, t) = CV(t)e^{-\frac{1}{2}S} = CV(t)e^{-\frac{d^2(x_0,x)}{4t}}, \tag{10.9}$$

where $V(t)$ is the solution of $V_t = \frac{1}{2}\Delta S V(t)$ under the condition that $\lim_{t\searrow 0} t^{n/2}V(t) = 1$ and

$$\frac{1}{C} = 2^n \int_0^\infty e^{-y^2}\omega(x_0, y)\, dy \tag{10.10}$$

where ω is defined as

$$\operatorname{vol}\mathcal{S}(x_0, 2\sqrt{t}y) \approx (2\sqrt{t})^n\omega(x_0, y), \qquad t \searrow 0, \tag{10.11}$$

and $\mathcal{S}(x_0, 2\sqrt{t}y)$ is a geodesic sphere centered at x_0.

This result leads to a nice proof of part (iii) of Definition 10.2 by taking $y = d(x_0, x)/2\sqrt{t}$ and $x \in d^{-1}(2\sqrt{t}y) = \mathcal{S}(x_0, 2\sqrt{t}y)$, the geodesic sphere centered at x_0. As ϕ has compact support, set $D = $ support (ϕ). Then let $\delta = \max_{x\in D} d(x_0, x)$ and $y \in [0, \delta/2\sqrt{t}]$. Using (10.11), the required limit can be calculated,

$$\lim_{t\searrow 0}\int_M K(x_0, x, t)\phi(x)dv(x) = C\lim_{t\searrow 0} V(t)\int_M e^{-\frac{d^2(x_0,x)}{4t}}\phi(x)dv(x)$$

$$= C\lim_{t\searrow 0} V(t)\int_0^{\delta/2\sqrt{t}}\int_{\mathcal{S}(x_0,2\sqrt{t}y)} e^{-y^2}\phi(x)d\sigma_x dy$$

$$= C \lim_{t \searrow 0} V(t) \int_0^{\delta/2\sqrt{t}} e^{-y^2} \phi(x_t) \mathrm{vol}\mathcal{S}(x_0, 2\sqrt{t}y)\, dy$$

$$= C \lim_{t \searrow 0} 2^n t^{n/2} V(t) \phi(x_t) \int_0^{\infty} e^{-y^2} \omega(x_0, y)\, dy = \phi(x_0).$$

Here Fubini's theorem has been applied and the mean value theorem for integrals to obtain $x_t \in \mathcal{S}(x_0, 2\sqrt{t}y)$.

10.3 Heat Kernel for the Casimir Operator.

The Casimir operator treated here is defined as an elliptic operator given by

$$\Delta_{cas} = \frac{1}{2}(X_1^2 + X_2^2 + T^2),$$

where X_1, X_2 and T are vector fields given by

$$X_1 = \partial_1 + x_2 \partial_t, \qquad X_2 = \partial_2 - 2x_2 \partial_t, \qquad T = \partial_t.$$

These are left invariant vector fields with respect to the Heisenberg group law defined as $(x, t) \circ (x', t') = (x + x', t + t' + 2x_1' x_2 - 2x_1 x_2')$.

Theorem 10.3. There is a constant c such that the fundamental solution for the operator $\partial_\tau - \Delta_{cas}$ is

$$K(y, s, x, t, \tau) = K(0, 0, (y, s)^{-1} \circ (x, t), \tau),$$

where $\circ$ is the Heisenberg group law defined above and

$$K(0, 0, x, t, \tau) = \frac{2c}{\sinh(2\tau)}\, e^{-\frac{1}{2}(-it + \frac{\tau}{2}|x|^2 \coth(2\tau))},$$

and $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

10.4 Heat Kernel for Operators with Potential.

The action and volume functions shall now be computed explicitly for some heat operators with potential. This procedure will produce closed form solutions.

(**1**) The first operator to be considered is $\partial_t - \partial_x^2 \pm b^2 x^2$. Start with the operator

$$L = \frac{d^2}{dx^2} - a^2 x^2,$$

in which $a \in \mathbb{R}_+$ is a nonnegative real parameter. Now associate the Hamiltonian function to this as half the principle symbol, that is,

$$H(\xi, x) = \frac{1}{2}(\xi^2 - a^2 x^2).$$

The Hamiltonian system of equations is given by

$$\dot{x} = H_\xi = \xi, \qquad \dot{\xi} = -H_x = a^2 x.$$

It is required to find the geodesic between the points $x_0, x \in \mathbb{R}$, so $x(s)$ will satisfy the boundary problem,

$$\ddot{x} = a^2 x, \qquad x(0) = x_0, \qquad x(t) = x.$$

Based on the Hamiltonian, conservation of energy implies that

$$\frac{1}{2}\dot{x}(s)^2 - \frac{1}{2}a^2 x(s)^2 = E, \tag{10.12}$$

This can be used to obtain an ordinary differential equation for the solution,

$$\frac{dx}{ds} = \sqrt{2E + a^2 x^2}.$$

Integrating this between $s = 0$ and $s = t$, with $x(0) = x_0$ and $x(t) = x$ yields,

$$\int_{x_0}^{x} \frac{du}{\sqrt{2E + a^2 u^2}} = t.$$

The integral is given by

$$\frac{ax}{\sqrt{2E}} = \frac{ax_0}{\sqrt{2E}}\cosh(at) + \sqrt{1 + \frac{a^2 x_0^2}{2E}}\sinh(at).$$

To solve this for E, we write

$$\frac{a(x - x_0\cosh(at))}{\sinh(at)} = \sqrt{2E + a^2 x_0^2}.$$

Squaring both sides, this can be solved for E,

$$2E = \frac{a^2(x - x_0\cosh(at))^2}{\sinh^2(at)} - a^2 x_0^2 = \frac{a^2(x^2 + x_0^2 - 2xx_0\cosh(at))}{\sinh^2(at)}.$$

Proposition 10.3. The energy along a geodesic derived from the Hamiltonian $H(\xi, x)$ between the points x_0 and x is

$$E = \frac{a^2(x^2 + x_0^2 - 2xx_0\cosh(at))}{2\sinh^2(at)}. \tag{10.13}$$

The energy along a geodesic derived from the hamiltonian $H(\xi, x)$ joining the origin $x_0 = 0$ and x is given by

$$E = \frac{a^2 x^2}{2\sinh^2(at)}. \tag{10.14}$$

In the limit $a \to 0$, the Euclidean energy is obtained

$$\lim_{a \to 0} E = \frac{(x - x_0)^2}{2t^2}.$$

The action can now be determined. Let $S = S(x_0, x, t)$ be the action for initial point x_0 and final point x, within time t. The action satisfies the Hamilton-Jacobi equation

$$\partial_t S + H(\nabla S) = 0.$$

Since,

$$H = \frac{1}{2}(\xi^2 - a^2 x^2) = \frac{1}{2}\dot{x}^2 - \frac{1}{2}a^2 x^2 = E,$$

and hence $\partial_t S = -E$. Using E from (10.13), we obtain the differential equation for S,

$$\frac{\partial S}{\partial t} = -\frac{a^2(x^2 + x_0^2 - 2xx_0 \cosh(at))}{2\sinh^2(at)}.$$

Integrating both sides of this with respect to t gives,

$$S(x_0, x, t) = \frac{a}{2}[(x^2 + x_0^2)\coth(at) - \frac{2xx_0}{\sinh(at)}]. \tag{10.15}$$

Note also that in the limit a tends to zero,

$$\lim_{a \to 0} S(x_0, x, t) = \frac{(x - x_0)^2}{2t},$$

the Euclidean action is obtained.

Lemma 10.2. (a) $(\partial_x S)^2 = a^2 x^2 + 2E$, $\qquad (b)$ $\partial_x^2 S = a\coth(at)$.

Proof: (a) Differentiation of $S(x_0, x, t)$ with respect to x yields,

$$\partial_x S = \frac{a}{\sinh(at)}(x\cosh(at) - x_0).$$

Squaring both sides gives,

$$(\partial_x S)^2 = \frac{a^2(x^2\cosh^2(at) + x_0^2 - 2xx_0\cosh(at))}{\sinh^2(at)}$$

$$= a^2x^2 + \frac{a^2(x^2 + x_0^2 - 2xx_0\cosh(at))}{\sinh^2(at)} = a^2x^2 + 2E.$$

(b) Differentiating $\partial_x S$ again gives

$$\partial_x^2 S = \frac{a}{\sinh(at)}\cosh(at) = a\coth(at).$$

A fundamental solution of the form

$$K(x_0, x, t) = V(t)e^{kS(x_0,x,t)} \tag{10.16}$$

is sought, where $V(t)$ satisfies a volume function equation and k is a real constant. On account of Lemma 10.3, we have the derivatives,

$$\partial_t K = V'(t)e^{kS} + V(t)ke^{kS}\partial_t S = e^{kS}(V'(t) - kV(t)E),$$

$$\partial_x e^{kS} = ke^{kS}\partial_x S,$$

$$\partial_x^2 e^{kS} = k^2 e^{kS}(\partial_x S)^2 + ke^{kS}\partial_x^2 S = ke^{kS}[k(a^2x^2 + 2E) + a\coth(at)].$$

The heat kernel shall be found by using a multiplier method. Let

$$P = \partial_t - \partial_x^2 + \alpha a x^2, \tag{10.17}$$

where α is a real multiplier, which will be determined such that $PK(x_0, x, t) = 0$ for any $t > 0$. Consequently,

$$PK(x_0, x, t)$$

$$= e^{kS}(V'(t) - kEV(T)) - ke^{kS}(k(a^2x^2 + 2E) + a\coth(at))V(t) + \alpha a^2 x^2 e^{kS}V(t)$$

$$= e^{kS}V(t)[\frac{V'(t)}{V(t)} - kE - k^2(a^2x^2 + 2E) - ka\coth(at) + \alpha a^2x^2]$$

$$= e^{kS}V(t)[\frac{V'(t)}{V(t)} - kE(2k+1) + (\alpha - k^2)a^2x^2 - ka\coth(at)].$$

In order to eliminate the middle two terms in the bracket on the right, we choose $k = -1/2$ and $\alpha = 1/4$. Let $b = a/2 > 0$ so that the operator P becomes,

$$P = \partial_t - \partial_x^2 + b^2x^2,$$

and

$$PK(x_0, x, t) = K(x_0, x, t)(\frac{V'(t)}{V(t)} + b\coth(2bt)).$$

The function $V(t)$ is chosen to satisfy the equation,

$$\frac{V'(t)}{V(t)} = -b\coth(2bt), \qquad t > 0.$$

Integration of this equation yields,

$$\ln V(t) = -\frac{1}{2}\ln(\sinh(2bt)), \qquad V(t) = \frac{C}{\sqrt{\sinh(2bt)}}.$$

Using the action given by (10.15), the fundamental solution for K takes the form,

$$K(x_0,x,t)=\frac{C}{\sqrt{\sinh(2bt)}}e^{-\frac{1}{4t}\frac{2bt}{\sinh(2bt)}}[(x^2+x_0^2)\cosh(2bt)-2xx_0]. \quad (10.18)$$

The constant C can be determined by investigating the limit $b\to 0$, such that the operator P becomes the usual one-dimensional operator $\partial_t-\partial_x^2$. As $2bt/\sinh(2bt)\to 1$, the above fundamental solution takes the form,

$$K(x_0,x,t)\sim\frac{C}{\sqrt{2bt}}e^{\frac{1}{4t}(x-x_0)^2},\qquad b\to 0.$$

By comparison with the fundamental solution for the usual heat operator, which is

$$\frac{1}{\sqrt{4\pi t}}e^{\frac{1}{4t}(x-x_0)^2},$$

it is found that $C=\sqrt{\frac{b}{2\pi}}$, and the following result can be stated.

Proposition 10.4. Let $b>0$. The fundamental solution for the operator $P=\partial_t-\partial_x^2+b^2x^2$ is

$$K(x_0,x,t)=\frac{1}{\sqrt{4\pi t}}\sqrt{\frac{2bt}{\sinh(2bt)}}e^{-\frac{bt}{2t\sinh(2bt)}[(x^2+x_0^2)\cosh(2bt)-2xx_0]},\qquad t>0. \quad (10.19)$$

The computations are similar for the case in which $b=-i\beta$. Using $\cosh(i\beta t)=\cos(\beta t)$ and $\sinh(2i\beta t)=i\sin(2\beta t)$, a dual theorem results.

Proposition 10.5. Let $\beta\geq 0$. The fundamental solution for the operator $P=\partial_t-\partial_x^2-\beta^2x^2$ is

$$K(x_0,x,t)=\frac{1}{\sqrt{4\pi t}}\sqrt{\frac{2\beta t}{\sin(2\beta t)}}e^{-\frac{\beta t}{2t\sin(2\beta t)}[(x^2+x_0^2)\cos(2\beta t)-2xx_0]},\qquad t>0. \quad (10.20)$$

(**2**) The kernel of the operator $\partial_t-\sum\partial_{x_i}^2\pm a|x|^2$ can be found. Consider the operator

$$\varDelta_n-a^2|x|^2=\partial_{x_1}^2+\cdots+\partial_{x_n}^2-a^2(x_1^2+\cdots+x_n^2),\qquad a\geq 0.$$

The associated Hamiltonian is given by

$$H=\frac{1}{2}(\xi_1^2+\cdots+\xi_n^2)-\frac{a^2}{2}(x_1^2+\cdots+x_n^2),$$

with the Hamiltonian system of equations

$$\dot{x}_j=H_{\xi_j}=\xi_j,\qquad \dot{\xi}_j=-H_{x_j}=a^2x_j,\qquad j=1,\cdots,n.$$

The geodesic $x(s)$ starting at $x_0=(x_1^0,\cdots,x_n^0)$ and having final point $x=(x_1,\cdots,x_n)$ satisfies the system,

$$\ddot{x}_j=a^2x_j,\quad x_j(0)=x_j,\quad x_j(t)=x_j,\quad j=1,\cdots,n.$$

As in the one-dimensional case, the law of conservation of energy is

$$\dot{x}_j^2(s) - a^2 x_j^2(s) = 2E_j, \qquad j = 1, \cdots, n,$$

where E_j is the energy constant for the j-th component. The total energy, which is the Hamiltonian, is given by

$$H = \sum_{j=1}^{n} (\frac{1}{2}\dot{x}_j^2 - \frac{1}{2}a^2 x_j^2) = E_1 + \cdots + E_n = E,$$

where E is constant. Proposition 10.3 yields,

$$E_j = \frac{a^2[x_j^2 + (x_j^0)^2 - 2x_j x_j^0 \cosh(at)]}{2\sinh^2(at)},$$

and hence,

$$H = E = \sum_{j=1}^{n} E_j = \frac{a^2[|x|^2 + |x_0|^2 - 2\langle x, x_0\rangle \cosh(at)]}{2\sinh^2(at)}.$$

where $|x|^2 = \sum_{j=1}^{n} x_j^2$ and $\langle x, x_0\rangle = \sum_{j=1}^{n} x_j x_j^0$. The action between x_0 and x in time t satisfies the equation $\partial_t S = -E$ or

$$\frac{\partial S}{\partial t} = -\frac{a^2[|x|^2 + |x_0|^2 - 2\langle x, x_0\rangle \cosh(at)]}{2\sinh^2(at)}$$

$$= \frac{\partial}{\partial t}[\frac{a}{2}(|x|^2 + |x_0|^2)\coth(at) - \frac{a\langle x, x_0\rangle}{\sinh(at)}].$$

Hence, we can choose

$$S = \frac{a}{2\sinh(at)}[(|x|^2 + |x_0|^2)\cosh(at) - 2\langle x, x_0\rangle]. \tag{10.21}$$

Set

$$S_j = \frac{a}{2\sinh(at)}[x_j^2 + (x_j^0)^2)\cosh(at) - 2x_j x_j^0].$$

Then $S = S_1 + \cdots + S_n$ and $\partial_{x_j} S = \partial_{x_j} S_j$ so that Lemma 10.3 yields,

$$\sum_{j=1}^{n} (\partial_{x_j} S)^2 = \sum_{j=1}^{n} (\partial_{x_j} S_j)^2 = \sum_{j=1}^{n} (a^2 x_j^2 + 2E_j) = a^2|x|^2 + 2E,$$

$$\sum_{j=1}^{n} \partial_{x_j}^2 S = \sum_{j=1}^{n} \partial_{x_j}^2 S_j = na\coth(at).$$

A kernel of the form

$$K(x_0, x, t) = V(t)e^{kS(x_0,x,t)}, \qquad k \in \mathbb{R}$$

is to be found. A calculation similar to the one-dimensional case yields,

$$\partial_t K = e^{kS}(V'(t) - kEV(t)), \qquad \partial^2_{x_j} e^{kS} = e^{kS} k[k(\partial_{x_j} S)^2 + \partial^2_{x_j} S].$$

Therefore, we have that

$$\Delta_n e^{kS} = ke^{kS}[k(a^2|x|^2 + 2E) + na\coth(at)].$$

In order to obtain the kernel for the heat operator, a multiplier method is employed. Consider the parabolic operator

$$P_n = \partial_t - \Delta_n + \alpha a^2 |x|^2,$$

where α is a multiplier to be determined later. Then,

$$P_n K$$

$$= e^{kS}[V'(t) - kEV(t)] - ke^{kS}[k(a^2|x|^2 + 2E) + na\coth(at)]V(t) + \alpha a^2 |x|^2 V(t) e^{kS}$$

$$= e^{kS} V(t)[\frac{V'(t)}{V(t)} - kE(1+2k) + (\alpha - k^2)a^2|x|^2 - kna\coth(at)].$$

Now choose $k = -1/2$ and $\alpha = 1/4$ and let $b = a/2 \geq 0$, then we have

$$P_n K = e^{kS} V(t)[\frac{V'(t)}{V(t)} + \frac{na}{2}\coth(at)].$$

For this to vanish, we choose $V(t)$ to satisfy,

$$\frac{V'(t)}{V(t)} = -nb\coth(2bt), \qquad t > 0.$$

This can be integrated to yield,

$$V(t) = \frac{C}{\sinh^{n/2}(2bt)}.$$

Hence, it is concluded that the fundamental solution for the operator $P_n = \partial_t - \Delta_n + b^2|x|^2$ expressed in the form $K(x_0, x, t)$ is given by

$$K(x_0, x, t) = \frac{C}{(2bt)^{n/2}} \frac{(2bt)^{n/2}}{\sinh^{n/2}(2bt)} e^{-\frac{b}{2\sinh(2bt)}((|x|^2 + |x_0|^2)\cosh(2bt) - 2\langle x, x_0\rangle)}.$$

When $b \to 0$, the kernel of the heat operator $\partial_t - \Delta_n$ should be obtained. This is

$$\frac{1}{(4\pi t)^{n/2}} e^{-\frac{1}{4t}|x - x_0|^2}, \quad t > 0.$$

By comparison, the following value is obtained for the constant C,

$$C = \frac{b^{n/2}}{(2\pi)^{n/2}}.$$

Proposition 10.6. Let $b \geq 0$ and $\Delta_n = \sum_{j=1}^n \partial_{x_j}^2$. The fundamental solution for the operator $P_n = \partial_t - \Delta_n + b^2|x|^2$ is for $t > 0$,

$$K(x_0, x, t) = \frac{1}{(4\pi t)^{n/2}}(\frac{2bt}{\sinh(2bt)})^{n/2} e^{-\frac{b}{2t\sinh(2bt)}[(|x|^2+|x_0|^2)\cosh(2bt)-2\langle x,x_0\rangle]}. \tag{10.22}$$

In a similar way as with the one-dimensional case, choosing $b = -i\beta$ gives rise to the following result.

Proposition 10.7. Let $\beta \geq 0$ and $\Delta_n = \sum_{j=1}^n \partial_{x_j}^2$. The fundamental solution for the operator $P = \partial_t - \Delta_n - \beta^2|x|^2$ when $t > 0$ is given by

$$K(x_0, x, t) = \frac{1}{(4\pi t)^{n/2}}(\frac{2\beta t}{\sin(2\beta t)})^{n/2} e^{-\frac{\beta}{2\sin(2\beta t)}[(|x|^2+|x_0|^2)\cos(2\beta t)-2\langle x,x_0\rangle]}. \tag{10.23}$$

10.5 The Laplacian and Some Geometric Implications.

The Laplacian has many remarkable properties on its own and can be used in conjunction with other information to discuss properties of differentiable manifolds. The results presented here overlap with many other areas, such as Hodge theory and cohomolgy theory. To this end, the Laplace operator will be examined from a geometric point of view **[60]**. Most of the short theorems which will be developed here will involve vector fields and one-forms. Some background information will be required first.

Given a vector field X which has local components X^i, there is associated with it a one form η defined by writing

$$\eta = g_{ij} X^i dx^j = X_i \, dx^i.$$

The codifferential of η is given by

$$\delta\eta = -\nabla_i X^i = -g^{ji}\nabla_j X_i. \tag{10.24}$$

This is also denoted as δX. Frequent use will be made of the following form of Green's theorem, which has been adapted to the case of vector fields.

Theorem 10.4 (Green) In a compact, orientable Riemannian manifold M without boundary,

$$\int_M (\delta X)\, dv = 0, \tag{10.25}$$

for any vector field X. In terms of components, (10.25) is given as

$$\int_M (\nabla_i X^i)\, dv = \int_M g^{ij}\nabla_i X_j \, dv = 0.$$

Here dv is the volume form on M.

Given a function f on M, the codifferential δdf of df can be formed. This is called the Laplacian of f and in terms of coordinates, it reads,

$$\Delta f = -g^{ji} \nabla_j \nabla_i f. \tag{10.26}$$

In fact, it can be shown that (10.26) is identical to (10.1). Regarding $X_i = \nabla_i f$ as a vector field, Theorem 10.4 leads to Theorem 10.5.

Theorem 10.5 In a compact and orientable Riemannian manifold M without boundary,

$$\int_M \Delta f \, dv = 0, \tag{10.27}$$

for any function f on M.

Definition 10.3. A p-form ω is said to be harmonic if it satisfies

$$d\omega = 0, \qquad \delta\omega = 0. \tag{10.28}$$

Definition 10.3 implies that the Laplacian, which can be written intrinsically as well in the form

$$\Delta = \delta d + d\delta$$

vanishes on a harmonic p-form, $\Delta\omega = 0$. The existence of harmonic p-forms in M is closely related to the topology of M.

Theorem 10.6. (Hodge) In a compact and orientable Riemannian manifold, the number of linearly independent, with constant real coefficients, harmonic p-forms is equal to the p-th dimensional Betti number b_p of the manifold.

The integral formula (10.27) holds for any function f in M, with Δf defined by (10.26). Thus, if $\Delta f \geq 0$ or $\Delta f \leq 0$ in M, it must be that $\Delta f = 0$. Since

$$\frac{1}{2} \Delta f^2 = (\Delta f) f - |\nabla f|^2,$$

invoking (10.27) again implies that

$$\int_M [(\Delta f) f - |\nabla f|^2] \, dv = 0. \tag{10.29}$$

Thus, if $\Delta f \geq 0$ or $\Delta f \leq 0$ in M, since it has been concluded that $\Delta f = 0$ in M by the remark above, then from (10.29) it follows that,

$$\int_M |\nabla f|^2 \, dv = 0.$$

This of course implies that $\nabla f = 0$, and therefore $f = C$ in M, where C is a constant.

Lemma 10.3. If $\Delta f \geq 0$ or $\Delta f \leq 0$ in M, then f is constant in M.

Suppose for a real nonzero constant c the function f in M satisfies

$$\Delta f = cf. \tag{10.30}$$

Substituting (10.30) into (10.29), it is found that

$$\int_M [cf^2 - |\nabla f|^2] \, dv = 0.$$

Clearly, for this to hold, the constant c must be positive.

Lemma 10.4. If $\Delta f = cf$ for a nonconstant function f in M, the constant c must be positive.

Let X be a vector field in M and define the function $f = |X|^2 = g(X, X)$. Taking the covariant derivative of f, we obtain

$$\nabla_j \nabla_i f = 2\nabla_j X_s \nabla_i X^s + 2X_s \nabla_j \nabla_i X^s.$$

Contracting this with the metric gives

$$\Delta f = g^{ji} \, \nabla_j \nabla_i f = 2(|\nabla X|^2 + X_s g^{ji} \nabla_j \nabla_i X^s). \tag{10.31}$$

Integrating (10.31) over M and applying (10.27) proves Theorem 10.7.

Theorem 10.7.

$$\int_M [(g^{ji} \nabla_j \nabla_i X^s) X_s + |\nabla X|^2] \, dv = 0. \tag{10.32}$$

From (10.32), if the second covariant derivative $\nabla\nabla X$ of a vector field X vanishes, then it must be that the first covariant derivative ∇X vanishes.

As introduced at the start, with a vector field X on M is associated a one-form $\eta = X_i \, dx^i$. For this one-form, the Laplacian is determined to be

$$(\Delta \eta)_i = -g^{kj} \nabla_k \nabla_j X_i + g^{kj} (\nabla_k \nabla_i - \nabla_i \nabla_k) X_j. \tag{10.33}$$

If R is the Riemann curvature tensor for the manifold M, then for a one-form ω with components ω_i,

$$(\nabla_k \nabla_i - \nabla_i \nabla_k) \omega_j = -g^{st} \, R_{kijt} \omega_s.$$

Using η as the one-form,

$$(\nabla_k \nabla_i - \nabla_i \nabla_k) X_j = -g^{st} R_{kijt} X_s = g^{st} R_{kitj} \, X_s. \tag{10.34}$$

Contracting both sides with the metric, (10.34) becomes,

$$g^{kj} (\nabla_k \nabla_i - \nabla_i \nabla_k) X_j = g^{st} R_{it} X_s. \tag{10.35}$$

Substituting (10.35) into (10.33), the Laplacian of the one-form η is given by

$$(\Delta \eta)_i = -g^{kj} \nabla_k \nabla_j X_i + R_i^{\,t} X_t. \tag{10.36}$$

A remarkable series of results can be obtained by starting with a vector field $X = X^i \partial_i$ in M and forming a related vector field

$$X^j(\nabla_j X^i) - (\nabla_j X^j) X^i. \tag{10.37}$$

Applying ∇_i to (10.37), it is found that

$$\nabla_i [X^j (\nabla_j X^i) - (\nabla_j X^j) X^i]$$
$$= X^i (\nabla_j \nabla_i X^j - \nabla_i \nabla_j X^j) + (\nabla_i X^j)(\nabla_j X^i) - \nabla_j X^j \, \nabla_i X^i.$$

For a vector field W^i, it is the case that $\nabla_k \nabla_j W^s - \nabla_j \nabla_k W^s = g^{st} R_{kjit} W^i$. Substituting this into the result above, the following equation comes out.

$$\nabla_i [X^j (\nabla_j X^i) - (\nabla_j X^j) X^i] = R_{is} X^i X^s + (\nabla^j X^i)(\nabla_i X_j) - (\nabla_j X^j)^2. \tag{10.38}$$

Applying the result in (10.25) to (10.38), the following theorem is proved.

Theorem 10.8. Let X be a vector field in M, then

$$\int_M [R_{ij} X^i X^j + (\nabla^j X^i)(\nabla_i X_j) - (\nabla_j X^j)^2] \, dv = 0. \tag{10.39}$$

Two other equivalent forms of (10.39) can also be written:

$$\int_M [R(X, X) - \frac{1}{2} |d\eta|^2 + |\nabla X|^2 - (\delta \eta)^2] \, dv = 0. \tag{10.40}$$

If L_X represents Lie differentiation with respect to X, then

$$\int_M [R(X, X) + \frac{1}{2} |L_X g|^2 - |\nabla X|^2 - (\delta X)^2] \, dv = 0. \tag{10.41}$$

Proof: It remains to prove (10.40) and (10.41), as (10.39) has been shown. Since

$$|d\eta|^2 = (\nabla_j X_i - \nabla_i X_j)(\nabla^j X^i - \nabla^i X^j) = 2(\nabla_j X_i)^2 - 2(\nabla_i X_j)(\nabla^j X^i).$$

Therefore, this implies that

$$(\nabla_i X_j)(\nabla^j X^i) = (\nabla_j X_i)^2 - \frac{1}{2} |d\eta|^2. \tag{10.42}$$

Therefore, substituting (10.42) into integral formula (10.39) using $\delta \eta = -\nabla^j X_j$, (10.40) is obtained.

To show (10.41), the Lie derivative with respect to X of g_{ij} is given by

$$L_X \, g_{ij} = \nabla_j X_i + \nabla_i X_j. \tag{10.43}$$

Therefore, from (10.43), it is found that,

$$|L_X g|^2 = (\nabla_j X_i + \nabla_i X_j)(\nabla^j X^i + \nabla^i X^j) = 2 \nabla_i X_j \nabla^i X^j + 2 \nabla_i X_j \nabla^j X^i.$$

Solving this, the following alternate expression is found,

$$\nabla_i X_j \nabla^j X^i = \frac{1}{2}|L_X g|^2 - |\nabla X|^2. \tag{10.44}$$

Substituting (10.44) into integral formula (10.38), the result in (10.41) is obtained.

Corollary 10.1.

$$\int_M [g(\Delta X, X) - \frac{1}{2}|d\eta|^2 - (\delta X)^2]\, dv = 0. \tag{10.45}$$

Proof: This follows by subtracting (10.26) from (10.40).

Theorem 10.8 has very important consequences in the case in which the vector field is harmonic or if it is Killing. Each case will be examined in turn. The definition of a harmonic form appears in Definition 10.3.

1. Suppose that a harmonic one-form η is the differential of a function f, so that $\eta = df$. Then it follows that $\delta\eta = \Delta f = 0$, since η is harmonic. Lemma 3 then implies that f is constant. Consequently, η vanishes identically.

Theorem 10.9. If a harmonic one-form in M is the differential of a function, then it is identically zero.

Theorem 10.10. For a vector field X in M,

$$\int_M [R(X,X) + |\nabla X|^2]\, dv \geq 0, \tag{10.46}$$

with equality if and only if X is a harmonic vector.

This is an obvious consequence of (10.40). In the case of equality in (10.46), the following proposition can be stated.

Theorem 10.11. If the Ricci curvature in M satisfies $R(X,X) \geq 0$, then a harmonic vector field X in M has a vanishing covariant derivative. If the Ricci curvature in M is positive definite, then a harmonic vector field other than zero does not exist in M.

Based on the statement of Hodge Theorem 10.6, this theorem combined with Theorem 10.11 gives the following result.

Theorem 10.12. In a compact and orientable manifold with positive Ricci curvature, the first Betti number vanishes.

Suppose that X is a harmonic vector field which means that (10.28) holds, where η is the one-form associated with X. Then it follows that

$$\Delta\eta = 0, \qquad g^{kj}\nabla_k \nabla_j X^s - R_i^{\ s} X^i = 0. \tag{10.47}$$

Conversely, if $\Delta\eta = 0$, then from integral formula (10.45),

$$\int_M [\frac{1}{2}|d\eta|^2 + |\delta\eta|^2]\, dv = 0. \tag{10.48}$$

From (10.48), it follows that $d\eta = 0$, $\delta\eta = 0$. Therefore, X is a harmonic vector field, and this is summarized below.

Theorem 10.13. In order that a vector field X be harmonic, it is necessary and sufficient that

$$\Delta\eta = 0, \qquad g^{kj}\nabla_k\nabla_j X^s - R_i^{\,s} X^i = 0, \tag{10.49}$$

where η is the one-form associated with X.

2. A vector field X is Killing if it satisfies

$$L_X g_{ij} = \nabla_j X_i + \nabla_i X_j = 0. \tag{10.50}$$

Contracting (10.50) with g^{ij} gives

$$\nabla_i X^i = 0. \tag{10.51}$$

This is an important result, since it implies that if X is Killing, then $\delta X = 0$. Suppose that a one-form associated with a Killing vector field X is the differential of a function $X_i = \nabla_i f$, then $\Delta f = 0$, which implies that f is constant so $X = 0$.

Lemma 10.5. If the one-form associated with a Killing vector field is the differential of a function, then it is identically zero.

Proof: $|L_X g|^2 \geq 0$ so (10.41) implies Lemma 10.5.

Lemma 10.6. For a vector field X in M,

$$\int_M [R(X,X) - |\nabla X|^2 - (\delta X)^2]\, dv \leq 0, \tag{10.52}$$

with equality if and only if X is Killing.

When X is Killing, (10.51) holds hence

$$\int_M [R(X,X) - |\nabla X|^2]\, dv = 0. \tag{10.53}$$

From equation (10.53), the following Theorem follows.

Theorem 10.14. If the Ricci curvature in M satisfies $R(X,X) \leq 0$, then a Killing vector field X in M has vanishing covariant derivative. If the Ricci curvature in M is negative definite, then a Killing vector field other than zero does not exist in M.

Another interesting integral formula can be obtained by adding (10.32) and (10.41). It will be used in the following Theorem and is given by

$$\int_M [(g^{ji}\nabla_j\nabla_i X_t + R_t^{\,h} X_h)X^t - \frac{1}{2}|L_X g|^2 + (\delta X)^2]\, dv = 0. \tag{10.54}$$

Theorem 10.15. A vector field will be Killing if and only if

$$g^{ij}\nabla_i\nabla_j X_k + R_k^{\,t} X_t = 0, \qquad \delta X = 0. \tag{10.55}$$

Proof: Suppose that X is a Killing vector field on M. Applying definition (10.50) to the second right-hand term of (10.34), we have,

$$\nabla_k \nabla_i X_j + \nabla_i \nabla_j X_k = -g^{st} R_{kijt} X_s.$$

Contract both sides of this with g^{ij} to obtain

$$\nabla_k \nabla^i X_i + g^{ij} \nabla_i \nabla_j X_k = -R_k^t X_t.$$

Using (10.51), the first term is gone, and so putting all terms on one side yields the first equation of (10.55). The second equation in (10.55) need not be assumed when X is Killing on account of (10.51).

Conversely, substituting the first equation of (10.55) and $\delta\eta = 0$, where η is the one-form associated with X, into (10.54), it then follows that $L_X g = 0$. This is the condition (10.50) that X must satisfy to be Killing.

Theorem 10.16. For a harmonic one-form ω and a Killing vector field X, the inner product $\omega(X)$ is constant over the manifold.

Proof: In terms of components,

$$-g^{ji} \nabla_j \nabla_i (\omega_s X^s)$$

$$= -(g^{ji} \nabla_j \nabla_i \omega_s) X^s - g^{ji} \nabla_i \omega_s \nabla_j X^s - g^{ji} \nabla_j \omega_s \nabla_i X^s - \omega_s g^{ji} (\nabla_j \nabla_i X^s)$$

$$= -R_s^i \omega_i X^s - \nabla^i \omega_s \nabla_i X^s - \nabla^i \omega_s \nabla_i X^s + \omega_s R_i^s X^i$$

$$= -\nabla^i \omega_s \nabla_i X^s + \nabla^i \omega_s \nabla_s X^i = -\nabla^i \omega_s \nabla_i X^s + \nabla^s \omega_i \nabla_s X^i = 0.$$

To simplify this, (10.55), (10.50) and $\nabla_i \omega_s = \nabla_s \omega_i$ have been substituted. Therefore, by Lemma 10.3, it follows that $\omega_i X^i$ is constant over M.

11 References.

[1] M. E. Taylor, Partial Differential Equations, Springer-Verlag, 1996.
[2] A. Das, Integrable Models, World Scientific Lecture Notes, vol. 30, NJ, 1989.
[3] C. Rogers and W. Schief, Bäcklund and Darboux Transformations, Cambridge Texts in Applied Mathematics, 2002.
[4] L. Debnath, Nonlinear Partial Differential Equations, 2nd Ed, Birkhäuser, Boston, 2005.
[5] V. I. Arnold, Mathematical Methods of Classical Mechanics, Springer-Verlag, 1978.
[6] J. E. Marsden and T. S. Ratiu, Introduction to Mechanics and Symmetry, Springer-Verlag, NY, 1994.
[7] M. Ablowitz and H. Segur, Solitons and the Inverse Scattering Transform, SIAM Publications, 1981.
[8] M. Ablowitz and P. A. Clarkson, Solitons, Nonlinear Evolution Equations and Inverse Scattering, Cambridge University Press, Cambridge, 1991.
[9] A. C. Newell, Solitons in Mathematics and Physics, SIAM, Philadelphia, 1985.
[10] M. J. Ablowitz, Nonlinear Dispersive Waves, Asymptotic Analysis and Solitons, Cambridge University Press, Cambridge, 2011.
[11] B. G. Konopelchenko, Solitons in Multidimensions, World Scientific, Singapore, 1993.
[12] P. Bracken, The interrelationship of integrable equations differential geometry and the geometry of their associated surfaces, in "Solitary Waves in Fluid Media", ed. C. David and Z. Feng, Bentham Science Publishers, pgs. 249-295, 2010.
[13] S. Sternberg, Lectures on Differential Geometry, Prentice-Hall Inc, NJ, 1964.
[14] J. M. Lee, Manifolds and Differential Geometry, American Mathematical Society, vol. 107, Providence, RI, 2009.
[15] P. Peterson, Riemannian Geometry, Springer-Verlag, NY, 1998.
[16] R. K. Miller and A. N. Michel, Ordinary Differential Equations, Dover, NY, 1982.
[17] A. M. Wazwaz, Partial Differential Equations Methods and Applications, A. A. Balkana Publishers, Tokyo, 2002.
[18] A. M. Kamchatnov, Nonlinear Periodic Waves and their Modulations, World Scientific, Singapore, 2000.
[19] J. Jost, Partial Differential Equations, 2nd Ed, Springer-Verlag, NY, 2007.
[20] A. C. Newell, The general structure of integrable evolution equations, Proc. Roy. Soc. London, **A 365**, 283-311, 1978.
[21] P. Bracken, Symmetry Properties of a Generalized Korteweg-de Vries Equation and Some Explicit Solutions, Int. J. Mathematics and Mathematical Sciences, **2005, 13**, 2159-2173, 2005.
[22] N. J. Zabusky and M. D. Kruskal, Interactions of "solitons" in a collision-

less plasma and the recurrence of initial states, Phys. Rev. Lett., **15**, 240-243, 1965.

[23] C. S. Gardner, J. Greene, M. Kruskal and R. M. Miura, Method for solving the Korteweg-de Vries equation, Phys. Rev. Lett., **19**, 1095-1097, 1967.

[24] C. S. Gardner, J. M. Greene, M. D. Kruskal and R. M. Miura, Korteweg-de Vries and generalizations. VI. Methods for exact solution, Commun. Pure Appl. Math., **27**, 97-133, 1974.

[25] P. Bracken, "The Generalized Weierstrass System Inducing Surfaces in Euclidean Three Space and Higher Dimensional Spaces", in Partial Differential Equations: Theory, Analysis and Applications, Nova Science Publishers, Hauppauge, NY, 223-264, 2011.

[26] H. Cartan, Differential Forms, Dover, NY, 2006.

[27] H. Flanders, Differential Forms with Applications to the Physical Sciences, Dover, NY, 1989.

[28] R. L. Bishop and S. I. Goldberg, Tensor Analysis on Manifolds, Dover, NY, 1980.

[29] I. Agricola and T. Friedrich, Global Analysis, American Mathematical Society, Providence, RI, 2002.

[30] R. Bryant, S. S. Chern and P. A. Griffiths, Exterior differential systems, Proceedings of the Beijing Symposium on Partial Differential Equations and Geometry, China Scientific Press, 1982.

[31] R. Bryant, S. S. Chern, R. Gardner, H. Goldschmidt and P. A. Griffiths, Exterior Differential Systems, Springer-Verlag, NY, 1991.

[32] R. B. Gardner and N. Kamran, Normal Forms and Focal Systems for Determined Systems of Two First-Order Partial Differential Equations in the Plane, Indiana University Mathematics Journal, **44**, 1127-1161, 1995.

[33] É. Cartan, Les systéms différentials extérieurs et leurs applications géometriques, Hermann, Paris, 1971.

[34] P. E. Hydon, Symmetry Methods for Differential Equations, Cambridge University Press, Cambridge, 2000.

[35] B. J. Cantwell, Introduction to Symmetry Analysis, Cambridge University Press, Camnbridge, 2002.

[36] G. Duffing, Erzwungene Schwingungen bei Veränderlicher Eigenfrequenz, F. Vieweg u. Sohn, Braunschweig, 1918.

[37] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer Verlag, NY, 1983.

[38] M. Lakshmanan and S. Rajasekar, Nonlinear Dynamics: Integrability, Chaos and Patterns, Springer Verlag, NY, 2003.

[39] B. van der Pol and J. van der Mark, Frequency demultiplication, Nature, **120**, 363-364, 1927.

[40] R. Hermann, The Geometry of Non-Linear Differential Equations, Bäcklund Transformations and Solitons, Part A, Math Sci Press, MA, 1976.

[41] R. Hermann, Geometric Theory of Nonlinear Differential Equations, Bäcklund Transformations and Solitons, Part B, Math Sci Press, MA, 1977.

[42] E. van Groesen and E. M. de Jager, 'Mathematical Structure in Con-

tinuous Dynamical Systems', Studies in Mathematical Physics, vol. 6, North Holland, 1994.
[43] H. D. Wahlquist and F. B. Estabrook, Prolongation Structures of Nonlinear Evolution Equations, J. Math. Phys., **16**, 1-7, 1975.
[44] H. D. Wahlquist and F. B. Estabrook, Prolongation Structures of Nonlinear Evolution Equations, II, J. Math. Phys., **17**, 1293-1297, 1976.
[45] F. B. Estabrook and H. D. Wahlquist, Classical geometries defined by exterior differential systems on higher frame bundles, Classical and Quantum Gravity, **6**, 263-274, 1989.
[46] F. B. Estabrook, 'Covariance, frame bundle, and Ricci flat solutions', Gravitational Collapse and Relativity, eds. H. Sato and T. Nakamura, World Scientific, Singapore, 185-196, 1986.
[47] P. Bracken, An Exterior Differential System for a Generalized Korteweg-de Vries Equation and its Associated Integrability, Acta Applicandae Mathematicae, **95**, 223-231, 2007.
[48] P. Bracken, Exterior Differential Systems for Higher Order Partial Differential Equations, Journal of Mathematics and Statistics, **6**, 52-55, 2010.
[49] P. Bracken, A Geometric Interpretation of Prolongation By Means of Connections, J. Math. Phys., **51**, 113502, 2010.
[50] P. Bracken, Exterior Differential Systems and Prolongations for Three Important Nonlinear Partial Differential Equations, Communications on Pure and Applied Analysis, **10**, 1345-1360, 2011.
[51] P. Bracken, Integrable Systems of Partial Differential Equations Determined by Structure Equations and Lax Pair, Physics Letters **A 374**, 501-503, 2010.
[52] R. Dodd and A. Fordy, The prolongation structures of quasi-polynomial flows, Proc. R. Soc. London, **A 385**, 389-429, 1983.
[53] M. Crampin, Solitons and $SL(2,\mathbb{R})$, Physics Letters **A 66**, 170-172, 1978.
[54] F. B. Estabrook, Moving Frames and Prolongation Algebras, J. Math. Phys., **23**, 2071-2076, 1982.
[55] Z. Qin, A finite-dimensional integrable system related to a new coupled KdV hierarchy, Physics Letters **A 355**, 452-459, 2006.
[56] G-Z. Tu, A New Hierarchy of Integrable Systems and its Hamiltonian Structure, Science in China, Series A, **32, 2**, 142-153, 1989.
[57] P. Bracken, The Hodge-de Rham Decomposition Theorem and an Application to a Patial Differential Equation, Acta Mathematica Hungarica, **133**, 332-341, 2011.
[58] O. Calin and D-C. Chang, Geometric Mechanics on Riemannian Manifolds, Birkhäuser, Boston, 2005.
[59] I. Chavel, Eigenvalues in Riemannian Geometry, Academic Press, Orlando, Florida, 1984.
[60] K. Yano, Integral Formulas in Riemannian Geometry, Marcel Dekker, Inc., NY, 1970.