Bundle theory for symplectic and contact geometry with applications to Lagrangian and Hamiltonian mechanics

Thesis

BUNDLE THEORY FOR SYMPLECTIC AND CONTACT GEOMETRY WITH APPLICATIONS TO LAGRANGIAN AND HAMILTONIAN MECHANICS

BY

GERARD THOMPSON B.Sc., M.S.

A DISSERTATION OFFERED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN THE SUBJECT OF

MATHEMATICS

THE OPEN UNIVERSITY

WALTON HALL

MILTON KEYNES

JUNE 1984

Date of submission : May 1984

Date of award : 12 July 1984

ABSTRACT

This work is primarily concerned with various aspects of Lagrangian and Hamiltonian mechanics. These different aspects are related in a somewhat complicated way, which is clarified by the introduction of the unifying concept of a jet bundle. Here, bundles are first considered in some generality and then jet bundles are introduced and their bundle structure investigated especially with a view to formulating Lagrangian and Hamiltonian mechanics invariantly.

The second chapter is devoted to a discussion of the two Schouten brackets and it is explained why each is of importance in mechanics. The symmetric Schouten bracket enables the notion of a killing tensor to be defined on a Riemannian or pseudo-Riemannian manifold. A theorem is proved which gives an upper bound on the dimension of the space of killing tensors of a fixed rank.

In chapter 3 Hamilton-Jacobi theory and a version of Noether's theorem are presented from a modern viewpoint. Then conditions are obtained which entail the existence of a constant of motion which is polynomial in momenta. These conditions are used to construct several classical Hamiltonians with "hidden" symmetries - the usage of the latter term is briefly justified. Most importantly, all systems with two degrees of freedom which have a quadratic integral independent of the Hamiltonian are characterized.

In chapter 4 the geometrical properties of TM and the closely related space $J^1(\mathbb{R},M)$ are investigated. It is shown that $J^1(\mathbb{R},M)$

has an intrinsically defined 1-1 tensor in an analogy to TM. The behaviour of these tensors under diffeomorphisms is investigated. The chapter concludes with a discussion of the various notions of symmetry in Lagrangian theory.

Chapter 5 begins with a review of symplectic geometry and continues with the definition and examination of contact manifolds in the same spirit as symplectic geometry. In particular, contact diffeomorphisms are described on the space $J^1(M,\mathbb{R})$. Finally, two theorems are given which endeavour to explain the general principle enabling the Lie algebra of a subalgebra of vector fields to be transferred to a collection of forms.

BUNDLE THEORY FOR SYMPLECTIC AND CONTACT GEOMETRY WITH APPLICATIONS TO LAGRANGIAN AND HAMILTONIAN MECHANICS

CHAPTER

INTRODUCTION

During the past twenty five years there has been renewed interest in the subject of classical mechanics. The coordinate-free techniques of modern differential geometry have provided just the right language for expressing the different approaches to mechanics. Moreover, many topics which have traditionally been hard to state precisely such as Hamilton-Jacobi theory, have been seen to be very natural consequences of an invariant, geometric formulation. At the same time many subtle, new issues have been raised which may have ramifications quite remote from the description of simple mechanical systems.

This work begins by developing geometric machinery which enables Lagrangian and Hamiltonian mechanics to be developed invariantly. In fact much more than this is done. Bundle theory is developed in its own right with a view to defining and exploring the properties of the so-called jet bundles. These provide an appropriate context in which to formulate virtually any system of partial differential equations invariantly.

In the second chapter, Schouten's brackets are defined and their principal algebraic properties obtained. Each of the two brackets is very important in theoretical mechanics. The importance of the symmetric Schouten bracket is explained in section 2.2 and this leads to the concept of a Killing tensor on a Riemannian or pseudo-Riemannian manifold. The skew Schouten bracket is important because it enables Poisson manifolds, which are manifolds having a Poisson bracket structure, to be characterized conveniently. Important examples of Poisson manifolds are cotangent bundles and the dual spaces of finite dimensional Lie algebras. The latter are briefly investigated in section 2.3.

Chapter 3 is concerned with constants of the motion in Hamiltonian mechanics. Noether's theorem and Hamilton-Jacobi theory are presented but the chapter mainly deals with constants of the notion which are polynomial in the momentum variables. Conditions are obtained which ensure the existence of such a constant of motion and these are used to construct several classical systems with hidden symmetries. Exactly why it is reasonable to refer to such systems as "hidden" is explained at the end of section 3.3.

Whereas chapter 3 was concerned with Hamiltonian mechanics, chapter 4 deals with Lagrangian theory. The main geometrical properties of TM - the tangent bundle of some smooth manifold M - are presented and then much the same analysis is presented for $J^1(\mathbb{R},M)$. Each of these spaces have an intrinsically defined 1-1 tensor called S, the one on TM actually being the restriction of the one on $J^1(\mathbb{R},M)$. It is explained why these tensors behave slightly differently under diffeomorphisms. Section 4.3 is concerned with symmetry in Lagrangian mechanics.

Chapter 5 begins with a brief review of symplectic manifolds. In section 5.2 one definition of contact manifolds is presented and the corresponding geometrical properties are developed. These include the construction of a skew bracket analogous to the Poisson bracket of a symplectic manifold and a 2-vector which is geometrically equivalent to the contact form. Section 5.3 is based on some ideas of Lie and gives a way of obtaining all contact transformations on $J^1(M,\mathbb{R})$ - the space of 1-jets of maps from M to $\mathbb{R}$. Section 5.4 endeavours to explain the general principle by which the Lie algebra structure of a subalgebra of vector fields may be transferred to a suitable collection

of forms.

As regards the original element in this work, the author has deliberately tried to give a new twist to whatever "standard" topics which have been presented. Apart from a few novel observations in chapter 1, there is one substantial theorem, namely, 1.8.3. Section 2.2 is the only really original part of chapter 2. In chapter 3, Hamilton-Jacobi theory and what is referred to as the "general Noether theorem" are quite well-known. Apart from that, most of the remainder of chapter 3 is quite original and undoubtedly this constitutes the main original element in this thesis. In chapter 4, section 4.2 and some of the remarks in section 4.3 are new. In chapter 5, the author knows of no reference for proposition 5.1.1 at least in its present form. It is fair to say that most of section 5.2 is known though some of the results were known, at least to the author, only in their classical form. Section 5.3 is, as was mentioned above, essentially a presentation of some original ideas of Lie. Theorems 5.4.4 and 5.4.5 are original. Prior to the submission of this thesis two papers relating to it have been published by the author; these appear as [9] and [39] in the bibliography. [39] is basically the same as example 3 of section 3.5. [9] is a joint work, of which the author's contribution is here covered approximately by sections 1.7 and 4.2. In addition, a letter and a paper based on sections 3.3, 3.4 and 3.5 have been accepted for publication in the Journal of Physics 'A' and the Journal of Mathematical Physics respectively.

I would like to thank Dr. Marek Kossowski for numerous stimulating discussions during the course of this work. I also wish to express my sincere appreciation for the advice and encouragement of my supervisor Dr. Michael Crampin throughout the writing of this dissertation.

NOTATION

Theorems, propositions, lemmas and corollaries are itemized sequentially within each section of each chapter. The differential geometric notation agrees largely with Sternberg's [38]. M,N denote smooth manifolds - M being of dimension m. Occasionally the notation of classical tensor calculus is employed and where appropriate upper and lower indices are used to denote contravariant and covariant geometric objects respectively. However, most of the tensor calculus notation occurs where there is a Euclidean metric δ_{ij} available, in which case whether an index is up or down is of no significance. The summation convention is always in force over repeated indices whether or not they are both up or down or one up and one down. The formula

$$L_X\omega = X\rfloor d\omega + d(X\rfloor\omega)$$

where X is a vector field and ω a p-form and the left-hand side denotes the Lie derivative of ω along X is used repeatedly. The term "Φ is a local map from M to N" means that there is an open subset U of M such that Φ is a map from U to N. I shall summarize the remaining notation below: all geometric objects defined will be assumed to be smooth though of course sometimes smoothness has to be proved to be a consequence of certain assumptions. Recall that N a smooth manifold.

$V^p(N)$ the module of p-vector fields on N (i.e. in classical language totally antisymmetric contravariant p-tensors).

$F^p(N)$ the module of p-forms on N (i.e. in classical language totally antisymmetric covariant p-tensors).

$S^p(N)$	the bundle over N whose fiber at $x \in N$ consists of the p-th symmetric power of the tangent space $T_x N$ to N at x.
$\underline{S^p(N)}$	sections of $S^p(N)$ over N.
TN	the tangent bundle of N.
T*N	the cotangent bundle of N.
$L_X \omega$	the Lie derivative of a p-form ω along a vector field X.
$[\gamma]_x$	a tangent vector to N at x, γ being a map from a neighbourhood of 0 in $\mathbb{R}$ to N such that $\gamma(0) = x$.
$df\|_x$	a covector to N at x, f being a map from a neighbourhood of x in N to $\mathbb{R}$ such $f(x) = 0$.
graph (Φ)	the submanifold of $M \times N$ given by the image of Φ which is a map from M to N.
$X \lrcorner \alpha$	the interior product of a p-form α by a vector X (this notation can also be usefully extended where X is a q-vector).
$X \llcorner \alpha$	the interior product of a q-vector X by a p-vector α.
$\langle X, \alpha \rangle$	the contraction of a p-vector X with a p-covector α : this notation will also be used where α is a 1-p type tensor.
$,i$	derivative in the i-th coordinate vector field $\frac{\partial}{\partial x^i}$ direction.

$;i$	covariant derivative (associated to a metric) in the i-th coordinate vector field $\frac{\partial}{\partial x^i}$ direction.
$A^{(i_1\ldots i_p)}$	the symmetric part of the contravariant p-tensor $A^{i_1\ldots i_p}$ (likewise for covariant tensors).
$A^{[i_1\ldots i_p]}$	the skew-symmetric part of the covariant p-tensor $A^{i_1\ldots i_p}$ (like for covariant tensors).
(E,π,M,F)	a bundle with total space M, projection π and fiber F.
(E,π,M,F,G)	a fiber bundle consisting of a bundle (E,π,M,F) and structure group G.
$J^r(E)$	the bundle of r-jets of local sections of a total space E fibered over M.
$J^r(M,N)$	the bundle of r-jets of maps of M to N.
$\phi^{-1}(E,\pi,M,F)$	the induced bundle of (E,π,M,F) arising from a map $\phi : N \to M$.
VE	the vertical bundle of a bundle (E,π,M,F).
ϕ_*	the pointwise tangent map from T_xM to $T_{\phi(x)}N$ where ϕ is a map $M \to N$ and $x \in M$.
ϕ^*	the pullback of $F^p(N)$ to $F^p(M)$ induced by $\phi : M \to N$
$T\phi$	the tangent bundle map : $TM \to TN$ induced by $\phi : M \to N$

CHAPTER 1

BUNDLE THEORY FOR MECHANICS AND DIFFERENTIAL EQUATIONS

This chapter is very largely a collection of definitions, some of which are quite familiar, others of which are not so easy to find. First of all affine spaces are defined and it is explained how they may be made into a category. Next, in section 4.2, bundles are defined and several ways of defining new bundles from given ones are considered. Then in section 4.4 some particular kinds of bundles are considered. The definition of a Lie group bundle appears to be novel and is a very natural generalization of that of a vector bundle. In fact, the definitions of Lie group and vector bundles are given together so as to emphasize that the latter is a particular case of the former. Affine bundles are also defined and their relationship to vector bundles is explained.

In section 1.5 "fiber bundles" as opposed to bundles are considered, in other words, the structure group of a fiber bundle is introduced. It is then explained how two apparently different definitions of fiber bundle due - originally to Steenrod and Ehresmann respectively - are equivalent. There then follows a variety of examples which give examples of all the various kinds of bundles which have been introduced so far in the chapter. In section 1.6 a new type of bundle is introduced, which provides the appropriate framework for formulating partial differential equations in differential geometric language: these are the jet bundles. It is shown that jet bundles detect when a section of the jet space over the source space is actually the jet of a section of

the target space over the source space. These results are by now fairly standard.

In section 1.7 two fundamental jet bundles - $J^1(\mathbb{R},M)$ and $J^1(M,\mathbb{R})$ are singled out and their various bundle structures investigated. Also the relationship of the spaces $J^1(M,\mathbb{R})$ and T^*M and $J^1(\mathbb{R},M)$ and TM are elucidated. In section 1.8 the canonical dilation vector field of a vector bundle is defined and it is shown that mappings which preserve this vector field are precisely vector bundle morphisms. Finally, in section 1.9 a very brief review of Lagrangian and Hamiltonian mechanics is included.

1.1 Affine Space

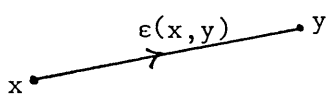
The first topic I deal with is the notion of an affine space. Whilst very important, it is hard to cite a good reference for this concept on the same level as say vector spaces.

Definition 1.1.1 An affine space consists of an ordered triple (V,X,ρ) in which V is a vector space and X is a set, such that V acts on X via ρ (say on the left) freely and transitively.

Sometimes the definition of an affine space is given rather differently. Now as I have defined it there is a map $\rho: V \times X \to X$ with some special properties: if one chooses $x \in X$ there is an induced map $V \to X$ by $v \mapsto \rho(v,x)$ and there is one such for each $x \in X$. Now the assumption of freeness and transitivity in the action ensure that each of these maps is invertible; this is equivalent to saying that there exists a map $\varepsilon: X \times X \to V$ called the *incidence* map. We are thus led to the following equivalent definition of an affine space.

Definition 1.1.2 An affine space consists of an ordered triple (V,X,ε) in which V is a vector space and X is a set and a map $\varepsilon: X \times X \to V$ satisfying (i) $x \in X \Rightarrow \varepsilon(x,x)=0$ and (ii) $x,y,z \in X \Rightarrow \varepsilon(x,y)+\varepsilon(y,z)=\varepsilon(x,z)$.

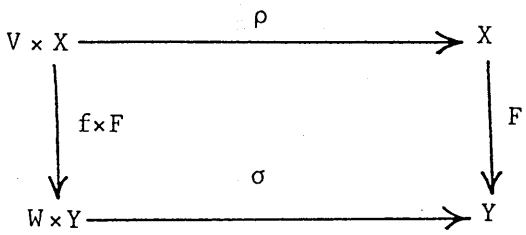
Notice that it is the freeness and transitivity of the action in the original definition which ensure the existence of the map ε in the second definition, whereas the action properties give (i) and (ii). The first definition is more suitable for analytical development whereas the second better for geometric intuition. In fact we may picture a pair of points x,y and $\varepsilon(x,y)$ as the displacement vector which joins them.



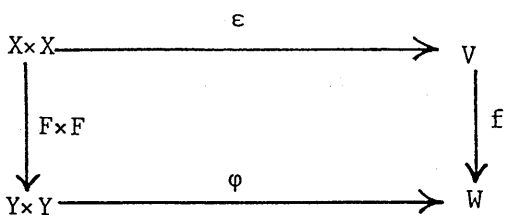
One often says also that the affine space X is *modelled* on the vector space V.

Definition 1.1.3 The dimension of an affine space is the dimension of its underlying vector space.

Definition 1.1.4 Given two affine spaces (V,X,ρ), (W,Y,σ) an affine morphism is a pair of maps $f:V \to W$ and $F:X \to Y$ where f is linear such that the following diagram commutes:



If the affine spaces were defined by the second definition and named (V,X,ε), (W,Y,φ) then instead one would have the following equivalent diagram



Clearly the collection of affine spaces and affine morphisms form a category and the notion of affine isomorphism is clear. This category will be denoted by AFF. The definition of affine morphism just given captures the classical notion of affine transformations as those which "preserve parallelism".

Given the foundations laid here one could proceed to a theory of affine geometry but I will limit myself to a few more remarks. It can be shown that if one selects a particular point x_0 of an affine space (V,X,ρ), then it can be given the structure of a vector space isomorphic to V with x_0 playing the role of origin. Conversely, any vector space V may itself be naturally regarded as an affine space. For, one may interpret the vector addition in V as the action of V on a set X where X is the underlying set

associated to V (technically, one applies a forgetful functor).

For further development of affine geometry I refer to [36]. The main reason for discussing affine spaces is that they are important in making precise the bundle structure of jet bundles. These are defined in section 1.6 here and form the setting for most of what follows later. However, before giving the definition of a jet bundle, it is convenient to consider bundles and fiber bundles in some generality and I turn to this topic next.

1.2 Smooth Bundles

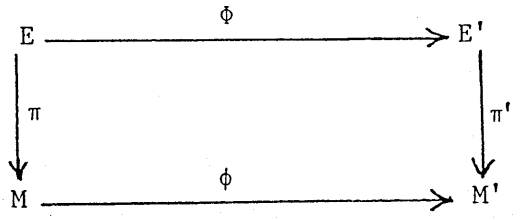
A bundle may be viewed as a purely topological construct or else it may be thought of as a geometric object. Here it is the latter which is of primary interest and so it will tactily be assumed that all maps and geometric objects which are considered are smooth i.e. C^∞. This convention must be excercized with some care because it can happen that objects which are not smooth arise from the interaction of smooth objects. For example, when a Lie group acts on a manifold the quotient space of the manifold by the orbits need not yield a smooth manifold. With this word of caution I now proceed to a fundamental definition.

Definition 1.2.1 A bundle consists of a quadruple (E,π,M,F) in which E,M,F are manifolds and $\pi:E\to M$ is a surjective submersion such that for each point $x\in M$, firstly, $\pi^{-1}(x)$ is diffeomorphic to F; secondly, there is a neighborhood U of x in M and a diffeomorphism φ of $U\times F$ to $\pi^{-1}(U)$ such that for all $z\in U$, $\pi(\varphi(z,y))=z$. E is called the total space (or less accurately the bundle), M the base, F the fiber and π the projection.

Given $x\in M$, $\pi^{-1}(x)$ is the fiber over x; according to the definition $\pi^{-1}(x)$ is diffeomorphic to F which is why F is referred to as *the* fiber. A bundle in where there is a φ which gives a (global) diffeomorphism of $M\times F$ to E is said to be trivial and hence the second condition in the definition is referred to as local triviality. In the definition given above there is no mention of the structure group. This will be introduced in section 1. below in which case I shall speak of "fibre bundle" rather than "bundle". A map p of M to E such that $\pi\circ p=i_M$ is called a section of E over M.

I now explain how bundles may be made into category.

Definition 1.2.2 For two bundles (E,π,M,F), (E',π',M',F') a bundle morphism is a pair of maps Φ,φ such that the following diagram commutes:

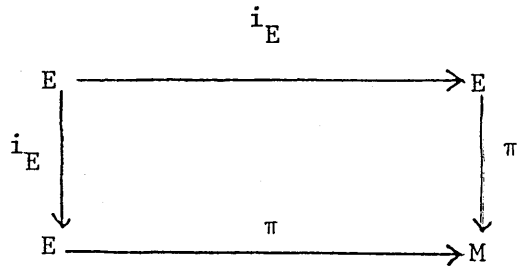


In case Φ and φ are diffeomorphisms one talks of a bundle isomorphism. The following proposition consists simply in checking definitions.

Proposition 1.2.3 The collection of bundles with bundle morphisms forms a category which will be denoted BUN.

The main point of definition 1.2.2 is that it forces Φ to be fiber-preserving: two points in the same fiber in E must map to the same fiber in E'. Now suppose that (Φ,φ) is a bundle morphism as in definition 1.2.2. One may conjecture that φ is a diffeomorphism if Φ is, but the following example shows this to be false.

Example 1.2.4 Suppose that (E,π,M,F) is a bundle. E may be viewed via i_E as a bundle over E with a fiber consisting of a point. The following diagram shows that (i_E,π) defines a bundle morphism; however, since (E,π,M,F) was an arbitrary bundle π need not be a diffeomorphism.



However, one does have the following.

Proposition 1.2.5 Let (Φ,ϕ) be a bundle morphism of (E,π,M,F) and (E',π',M',F') such that Φ is onto. Then ϕ is onto.

Proof: Suppose that $x' \in M'$ and choose $p' \in \pi'^{-1}(x')$. Since Φ is onto there is a $p \in E$ such that $\Phi(p) = p'$. Now let $x = \pi(p)$. Then

$$\phi(x) = \phi \circ \pi(p) = \pi' \circ \Phi(p) = \pi'(p') = x'.$$

Example 1.2.4 shows, using the notation of definition 1.2.2, that if Φ is one-one ϕ need not be. Other examples may be readily constructed in which the surjectivity or injectivity of ϕ does not imply the same property for Φ. However, one result which is useful for global questions is the following.

Proposition 1.2.6 Let (Φ,ϕ) be a bundle morphism of (E,π,M,F) and (E',π',M',F') such that ϕ is one-one and Φ is fiberwise one-one. Then Φ is one-one.

Proof: Suppose that $p,q \in E$ and that $\Phi(p) = \Phi(q)$. Then

$$\pi' \circ \Phi(p) = \pi' \circ \Phi(q) = \phi \circ \pi(p) = \phi \circ \pi(q)$$

and hence $\pi(p) = \pi(q)$ since ϕ is one-one. In other words p and q belong to the same fiber of E and hence by assumption

$$\Phi(p) = \Phi(q) \Rightarrow p = q.$$

An obvious subcategory of BUN consists of all bundles fibred over M and morphisms in which the base map is i_M. I denote this category by BUN(M) and note that this is a category but not full subcategory of BUN. I finish this section with one more basic definition.

Definition 1.2.7 A bundle (E',π',M',F') is a subbundle of (E,π,M,F) if there exists a bundle morphism (i,j) where $i:E' \to E$ and $j:M' \to M$ and both i and j are injective.

Notice that according to the definitions a section of a bundle (E,π,M,F) is precisely the same thing as a BUN(M) morphism of (M,i_M,M,f) to (E,π,M,F), where f just denotes a single point.

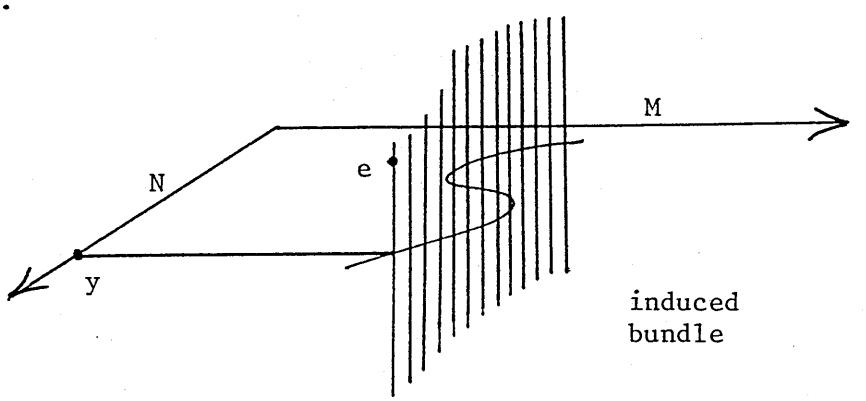
I have purposely avoided all mention of the local description of bundles so far so as to keep the basic definitions clear. However, it should be clear that the local trivialization condition amounts to the introduction of a coordinate chart (x^i,y^A) on E where (x^i) is a chart on M. Suppose that (x'^i,y'^A) is a second such chart on E'. Then the local description of definition 1.2.2 is given by

$$x'^j = \phi^j(x^i) \qquad\qquad y'^B = \Phi^B(x^i,y^A).$$

1.3 Categorical Constructions with Bundles

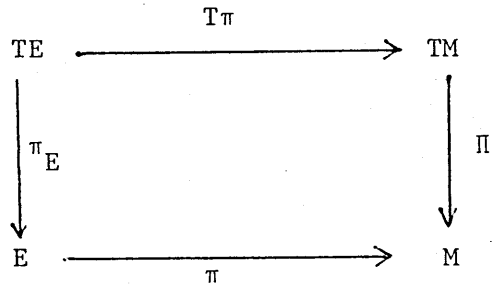
I next describe several important ways of obtaining new bundles from old ones. The first one is simply the categorical product in BUN. Thus, let (E,π,M,F), (E',π',M',F') be two bundles; then a new bundle $(E\times E',\pi\times\pi',M\times M',F\times F')$ may be formed which is called the product of the bundles.

Another important construction is the induced bundle. Suppose that (E,π,M,F) is a bundle and that φ is a map from N to M, N being another manifold. Then one may consider the subset of pairs (y,e) in $N\times E$ such that, for $e\in E$, $\pi(e)=\varphi(y)$ for some $y\in N$. This is a bundle over N and given $y\in N$ the fiber over y is $\pi^{-1}(\varphi(y))$. This generalizes the usual idea of restriction of a bundle and is also referred to as the pullback bundle.



The next construction is the categorical product in BUN(M). Suppose that (E,π,M,F) and (E',π',M,F') are bundles over M. Then the subset of pairs (e,e') of $E\times E'$ such that $\pi(e)=\pi'(e')$ defines a bundle over M whose fiber is $F\times F'$; which is why this construction is also known as the fiber product. Alternatively, the fibre product may be defined as the pullback of the product bundle $(E\times E',\ \pi\times\pi',\ M\times M,\ F\times F')$ via the diagonal map of M into $M\times M$.

Another case in which the induced bundle construction arises is the vertical bundle associated to a bundle. Let (E,π,M,F) be a bundle and let $\Pi:TM \to M$ and $\pi_E : TE \to E$ denote the tangent bundle projections. The following commutative diagram defines a VBUN morphism



π induces a bundle over E which is not TE and which I denote by $\pi^{-1}(TM)$. We now have two bundles over E, namely, TE and $\pi^{-1}(TM)$ and a BUN(E) morphism of TE onto $\pi^{-1}(TM)$. Taking the kernel on each fiber gives a new bundle over E called the vertical bundle of E or sometimes the fiber-tangent bundle denoted by VE. VE is the subbundle of TE consisting of all tangent vectors of E which are tangent to the fiber of E over M.

Another simple but extremely important construction that may be made is the following. Suppose one has two bundles over the same base, say, (E_1,π_1,M,F_1) and (E_2,π_2,M,F_2) and moreover that F_1 and F_2 are vector spaces. A new bundle may be formed with base M and fiber $F_1 \otimes F_2$ by forming $\pi_1^{-1}(x) \otimes \pi_2^{-1}(x)$ for each $x \in M$. I shall denote this bundle by $E_1 \otimes_M E_2$ and refer to it as the tensor product bundle of E_1 and E_2 over M.

1.4 Lie Group, Vector and Affine Bundles

In differential geometry one usually asks that the fiber has more structure than being merely a manifold.

Definition 1.4.1 A bundle is a Lie group (vector bundle) if the fiber is a Lie group (vector space) and if each point in the base has a neighborhood U such that there is a trivializing BUN(U) diffeomorphism Φ and moreover, the restriction of Φ to each fiber is a Lie group (linear) isomorphism.

The next definition goes hand in hand with the previous one.

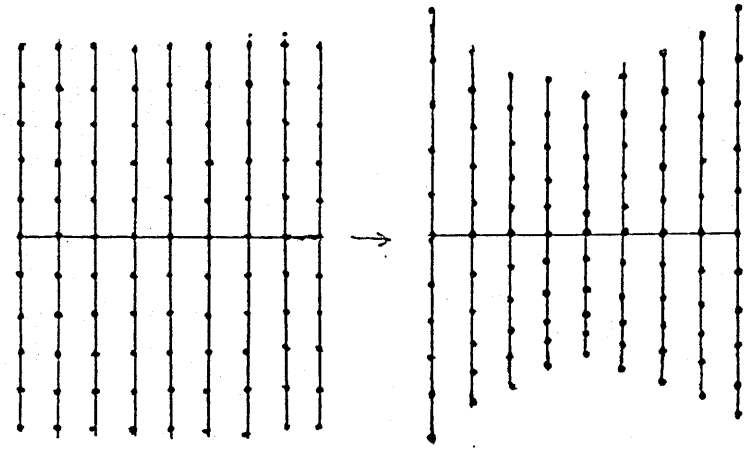
Definition 1.4.2 A morphism of Lie group (vector) bundles is a bundle morphism which is a Lie homomorphism (linear) in each fiber.

In the same way as proposition 1.2.3 one has:

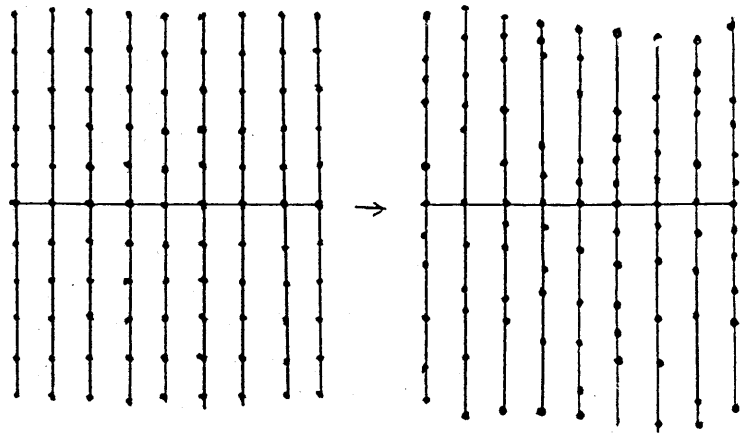
Proposition 1.4.3 The collection of Lie group (vector) bundles and Lie group (vector) bundle morphisms form a category denoted by LGBUN (VBUN). Thus VBUN is a subcategory of LGBUN which is in turn a subcategory of BUN. Moreover, VBUN is a full subcategory of LGBUN though neither VBUN nor LGBUN are full in BUN.

Next, suppose that (E,π,M,G) is a Lie group bundle and that e is the identity in G. If $x \in M$, U is a neighborhood of x and Φ is a local trivialization then there is a distinguished point in the fiber over x i.e. $\Phi^{-1}(x,e)$ since any other trivialization is also a Lie isomorphism on each fiber. Thus (E,π,M,G) has a distinguished section called the identity section which may be shown to be a smooth submanifold of E diffeomorphic to M. In VBUN one refers to "zero section" rather than identity section. The following diagrams are intended to illustrate several different kinds of maps which can occur in VBUN.

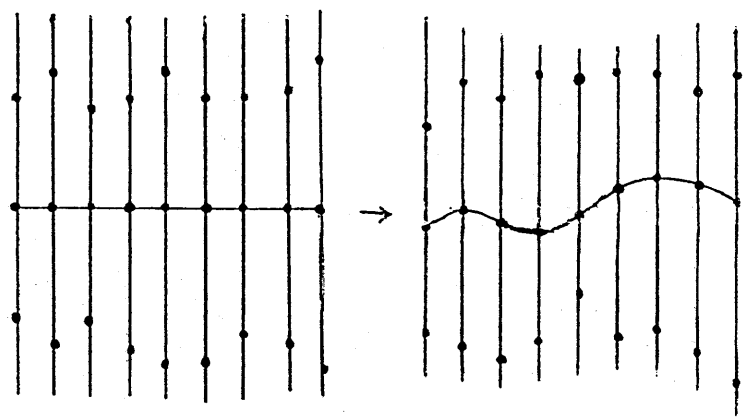
vector bundle morphism



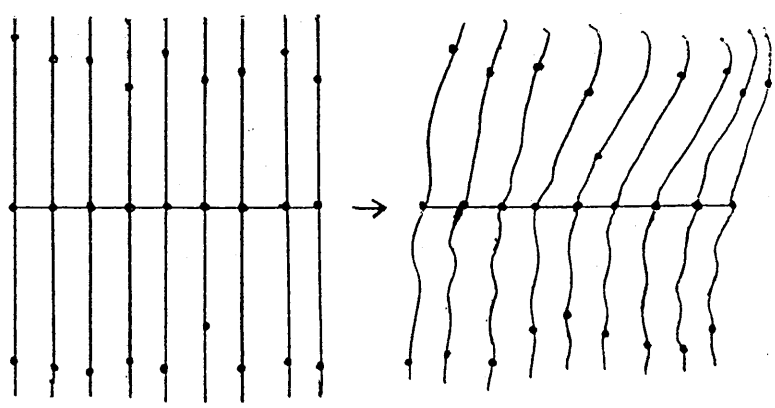
bundle morphism preserving zero section but not linear in fibers



bundle morphism



map preserving zero section but not even a bundle morphism

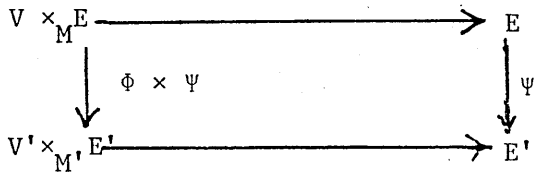


I now turn to the notion of affine bundle . In section 1.1 I showed how affine spaces are very closely related to vector spaces: in fact a choice of point to serve as an origin in an affine space enables all the structure of the associated vector space to be transferred to the affine space. On a bundle level one must choose a section rather than a single point. Thus, one could define an affine bundle as a fibered manifold in which each fiber is an affine space and which is locally trivial the trivialization giving an affine isomorphism on each fiber. However, as Goldschmidt [14] observed there is a rather neater definition. Notice that as a consequence of the provisional definition of affine bundle just given, and the definitions of affine space and vector bundle, an affine bundle must have a vector bundle associated to it. Goldschmidt's definition derives the required structure from this vector bundle.

Definition 1.4.4 An affine bundle (E,π,F,V,ρ,W,M) consists of a vector bundle (V,ρ,M,W), a bundle (E,π,M,F) and a BUN(M) morphism of $V\times_M E$ to E such that for each $x\in M$ $\pi^{-1}(x)$ is an affine space modelled on $\rho^{-1}(x)$ and the BUN(M) morphism is given by the action of $\rho^{-1}(x)$ on $\pi^{-1}(x)$.

The definition of a morphism of affine bundles is now almost obvious.

Definition 1.4.5 A morphism of affine bundles (E,π,F,V,ρ,W,M) $(E',\pi',F',V',\rho',W',M')$ consists of a BUN morphism (Ψ,φ) of (E,π,M,F) and (E',π,M',F'), a VBUN morphism (Φ,ϕ) of (V,ρ,M,W) and (V',ρ',M',W') such that the following diagram commutes



Proposition 1.4.6 The collection of affine bundles and affine bundle morphisms form a category denoted by AFFBUN.

1.5 The Structure Group of a Bundle

In this section I will explain the role of the structure group in a bundle. First of all we should recognize that there are several notions of "bundle". On one level one may simply consider the category of fibered manifolds without any reference to "the fiber". Alternately, one may use the term "bundle" in the sense of definition 1.2.1; however, notice that in the former case if it is assumed that the projection π is a submersion then the fibers will all be submanifolds of the same dimension. Suppose next that one has a trivial bundle $(M \times F, \pi, M, F)$ then clearly any fiber may be unambiguously identified with F. However, if the bundle (E,π,M,F) is not trivial there will be several ways in which to identify $\pi^{-1}(x)$ $(x \in M)$ with F. This ambiguity is measured by the structure group of the bundle. From now on I shall reserve the term "fiber bundle" for a bundle with a given structure group. It would be possible now to give a definition of a fiber bundle and this was indeed the course followed by Steenrod in the first definitive account of the subject [37]. There is an alternative approach, however, which Steenrod attributes to Ehresmann [12] and it is this that I shall outline next.

In the Ehresmann approach to fiber bundles, the fiber bundle itself is not to the forefront. Instead, one considers the appropriately named construct of a principal bundle. The subsequent definition is taken from Sternberg [38] and would also follow from

Husemoller's treatment [21] as applied in the category of smooth manifolds. Actually, Husemoller derives as a consequence of his definition that the fiber is diffeomorphic to the structure group.

Definition 1.5.1 A (right) principal G-bundle (P,π,M,G) is a bundle (in the sense of definition 1.2.1) in which, in addition, G is a Lie group which acts freely (to the right) on P in such a way that the quotient space P/G i.e. the space of orbits is diffeomorphic to M. Moreover, each $x \in M$ has a trivializing neighbourhood U such that $\Phi : \pi^{-1}(U) \cong U \times G$; and also, where $\eta(p) \in G \quad s \in G, \quad p \in \pi^{-1}(U)$

$$\phi(ps) = (\pi(p), \eta(p)s).$$

Thus, a principal bundle is a particular kind of bundle in which the fiber is a group G, which also acts on the total space in such a way that the fibers of π are the orbits of G. One should note that the Lie group bundles defined in section 1.4 are not principal, because the structure group is Aut(G) not G. Also, it is easy to show that a principal bundle is trivial iff it has a section [38]. I now give the definition of morphism for principal bundles.

Definition 1.5.2 A morphism of principal bundles (P,π,M,G) and (P',π',M',G') consists of a bundle morphism (Φ,ϕ) and a homomorphism $f : G \to G'$ such that $\forall\ p \in P$ and $\forall\ s \in G$

$$\Phi(ps) = \Phi(p)f(s).$$

Proposition 1.5.3 The collection of principal bundles and morphisms form a category denoted by PBUN. Also, the collection of principal bundles with fixed base M and morphisms fibered over i_M form a subcategory of PBUN denotes by PBUN(M).

I now consider fiber bundles from the Ehresmann viewpoint and begin with an important proposition which enables the definition of fiber bundle to be given immediately afterwards.

Proposition 1.5.4 Let (P,ρ,M,G) be a right principal bundle and F a smooth manifold upon which G acts effectively on the left. G acts to the right on $P \times F$ by $(p,f)s = (ps, s^{-1}f)$ $(p \in P, f \in F, s \in G)$. We may consider the space of orbits $P \times_G F$ of $P \times F$ under the action of G. Then $P \times_G F$ is a smooth bundle over M with fiber F.

Proof: Firstly, I define a map π from $P \times_G F$ to M by

$$\pi\Big(\bigcup_{s\in G} (ps,s^{-1}f)\Big) = \rho(p) \qquad (p \in P,\ f \in F).$$

π is well-defined because if $\bigcup_{s\in G} (ps,s^{-1}f)$ and $\bigcup_{s\in G} (p's,s^{-1}f')$ are two representatives of the same element of $P \times_G F$ $p = p't$ for some $t \in G$, ρt then $\rho(p) = \rho(p')$ and so $\pi\big(\bigcup_{s\in G} (ps,s^{-1}f)\big) = \pi\big(\bigcup_{s\in G} (p's,s^{-1}f')\big)$. Clearly π is onto because if $x \in M$ and $p \in p^{-1}(x)$, $f \in F$

$$\pi\Big(\bigcup_{s\in G} (ps,s^{-1}f)\Big) = \rho(p).$$

I next show that given $x \in M$, $\pi^{-1}(x)$ is in one-one correspondence with F. In fact we can certainly define a map μ from F to $\pi^{-1}(x)$ by $p\in P$ and $f\in F$.

$$f \longmapsto \bigcup_{s\in G} (ps,s^{-1}f) \quad \text{where}$$

μ is evidently onto and it is also one-one. For, suppose that

$$\bigcup_{s\in G} (ps,s^{-1}f) = \bigcup_{t\in G} (p't,t^{-1}f').$$

Then certainly $f' = uf$ for some $u \in G$ and hence

$$\bigcup_{s\in G} (ps, s^{-1}f) = \bigcup_{t\in G} (p't, t^{-1}f')$$

$$= \bigcup_{t\in G} (p't, t^{-1}uf)$$

$$= \bigcup_{t\in G} (p'u\, u^{-1}t,\ t^{-1}u\, f)$$

$$= \bigcup_{s\in G} (p'us,\ s^{-1}f).$$

From the first and last terms of these equalities above, it follows that $p = p'u^{-1}$, and hence from the freeness of the G action on P that $u = e$. Thus $f = f'$ and F is in one-one correspondence with $\pi^{-1}(x)$.

It is now high time to describe the differentiable structure of $P \times_G F$. Suppose that $\bigcup_{s\in G} (ps, s^{-1}f) \in P \times_G F$. Then $p \in P$ and as such, by the local triviality of P, there is U open in M with $p \in U$ such that $\rho^{-1}(U) \cong U \times G$. Hence, $\rho^{-1}(U) \times F \cong U \times G \times F$. Now since the local trivialization of P commutes with the action of G, we may conclude from the preceding argument, that $\rho^{-1}(U) \times_G F$ is in one-one correspondence with $U \times F$. Taking enough U's to cover M and enough charts for the U's and F will provide an atlas for $P \times_G F$. Moreover, given any point of $P \times_G F$ we can identify π just as the projection from $\mathbb{R}^m \times \mathbb{R}^n$ to $\mathbb{R}^m$ and so it is certainly a smooth submersion. All in all it has been shown that $(p \times_G F, \pi, M, F)$ is a smooth bundle. ∎

If Φ_u, Φ_v are any two local trivializations of $p \times_G F$ i.e.

$$\Phi_u : \pi^{-1}(U) \cong U \times F, \quad \Phi_v : \pi^{-1}(V) \cong V \times F$$

we can define a map $\Phi_{vu} : U \cap V \to G$ by

$$\Phi_{vu}(x)y = \Phi_v(\Phi_u^{-1}(x,y))$$

and the Φ_{vu}'s are a group of automorphisms of F.

Definition 1.5.5 With the notation of the last proposition, $P \times_G F$ is said to be the fiber bundle with structure group G and fiber F associated to the principal bundle (P,ρ,M,G). Conversely, (P,ρ,M,G) is said to be the principal bundle associated to the fiber bundle (E,π,M,F,G) if $E = P \times_G F$.

The maps Φ which appeared above are known as the transition functions of the bundle. They represent restrictions on the possible changes of coordinates which preserve the structure of a fiber bundle. The chief advantage of the Ehresmann definition of fiber bundle is that it sidesteps the introduction of the transition functions. Although it is convenient to avoid mentioning the transition functions in the definition, they are important, and can be used to give an internal description of fiber bundles which is essentially the Steenrod definition. In fact one has the following theorem - versions of which may be found in [21, 32, 38].

Theorem 1.5.6 Let M and F be manifolds, G a Lie group acting effectively on F and A an atlas for M. If $\forall\ U, V \in \forall\ \exists$ maps $\Phi_{vu} : \forall\ U \cap V \to G$ such that $\forall\ U \cap V \in U \cap V \cap W$ $(U, V, W \in A)$ $\Phi_{WV}(x)\ \Phi_{VU}(x) = \Phi_{WU}(x)$, then $\exists$ a fiber bundle (E,π,M,F,G) for which

the Φ_{VU}'s are the transition functions.

In particular, starting from a fiber bundle as it has been defined here, theorem 1.5.6 can be applied to retrieve the associated principal bundle. Specifically, replace the fiber by the structure group, allow it to act on itself by right translation and use the original transition functions. One obtains a principal bundle and if one next forms its associated fiber bundle using the original fiber according to definition 1.5.5, one obtains a bundle isomorphic to the original one. Thus, coming full circle we see why the Steenrod and Ehresmann definition of fiber bundles are equivalent.

I now turn to the notion of morphism for fiber bundles.

Definition 1.5.7 For two fiber bundles (E,π,M,F,G), (E',π',M',F',G') a morphism (Φ,ϕ,f) consists of a bundle morphism (Φ,ϕ) and a homomorphism f of G to G' such that

$$\forall\ e \in E, \forall s \in G \qquad (\Phi(e))f(s) = \Phi(es).$$

Of course, this definition is just the more general version of definition 1.5.2.

Proposition 1.5.8 The collection of fiber bundles and their morphisms form a category denoted by FBUN of which PBUN is a full subcategory. Moreover, PBUN(M) is a full subcategory of FBUN(M) - the collection of all fiber bundles with base space M and morphisms fibered over i_M.

I next give a variety of examples of fiber bundles, some of which will be examined at greater length in subsequent chapters.

Example 1.5.9 Let M be an m-dimensional manifold and let $F = R^m$. If $(x^i_u) : U \to \mathbb{R}^m$ and $(x^i_v) : V \to \mathbb{R}^m$ are two overlapping local charts on M one can define a map $\Phi_{VU} : U \cap V \to GL(m, \mathbb{R})$ by

$$[\Phi_{vu}(x)]^i_j = \left(\frac{\partial x^i_u}{\partial x^j_v}\right) \qquad (x \in U \cap V).$$

From the chain rule, the hypotheses of theorem 1.5.6 are satisfied and the resulting bundle is TM. The principal bundle associated to TM is FM - the frame bundle of M. If instead one takes

$$\left[\Phi_{vu}(x)\right]^i_j = \left[\frac{\partial x^j_v}{\partial x^i_u}\right]^{-1}$$

theorem 1.5.6 is still applicable and yields T*M with associated principal bundle F*M - the coframe bundle of M.

Example 1.5.10 Let S^n denote the unit sphere in $\mathbb{R}^n$ (with Euclidean metric). Then note that S^3 may be identified with the unit sphere in $\mathbb{C}^2$ (with the standard Hermitian metric), S^1 may be identified with the unit circle $\mathbb{C}$. Also, S^2 may be identified with $\mathbb{C}$ P(1), as follows. Firstly, $\mathbb{C}$ P(1) may be viewed as equivalence classes of pairs of complex numbers $[(z_1, z_2)]$ all differing by a non-zero complex multiple and satisfying $z_1 z_2 \neq 0$. Secondly, S^2 may be identified with $\mathbb{C} \cup \{\infty\}$ (the Riemann sphere) by, for example, stereographic projection. I can now define a map $\lambda : \mathbb{C}$ P(1) $\to \mathbb{C} \cup \{\infty\}$ by

$$[(z_1, z_2)] \to \frac{z_1}{z_2} \qquad (z_2 \neq 0)$$

$$[(z_1, 0)] \to \infty.$$

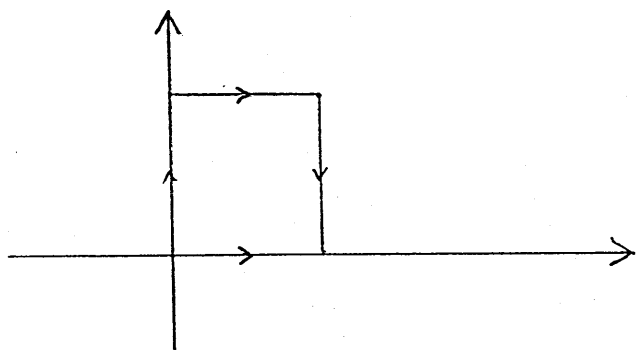
It may be checked that λ is well defined, bijective and smooth. With the above identifications S^3 may be viewed as a bundle over S^2. In fact, if $(z_1, z_2) \in S^3$ define $\pi : S^3 \to S^2$ by $\pi(z_1, z_2) = [(z_1, z_2)]$. It is easy to check that π is a surjective submersion. Next, let $U_1 = S^2 - (1,0)$ and $U_2 = S^2 - (0,1)$ so that $\{U_1, U_2\}$ is an open cover of S^2 and define

$$\Phi_1 : \pi^{-1}(U_1) \to U_1 \times S^1 \text{ by } (z_1,z_2) \to [(\frac{z_1}{z_2}, 1)], (\frac{z_2}{z_2})^{\frac{1}{2}}$$

$$\Phi_2 : \pi^{-1}(U_2) \to U_2 \times S^1 \text{ by } (z_1,z_2) \to [(1, \frac{z_2}{z_1})], (\frac{z_1}{z_1})^{\frac{1}{2}}$$

Then S^3 is a principal bundle over S^2 with fiber S^1 and local trivializations Φ_1, Φ_2; it is known as the Hopf bundle.

Example 1.5.11 Consider the unit square $[0,1] \times [0,1]$ in $\mathbb{R}^2$ and identify the ends as indicated.



This gives the Möbius band in which the base is $[0,1]$, the fiber is $[0,1]$ and the structure group G is $\mathbb{Z}_2$.

Example 1.5.12 In this example I want to consider the generalization of example 1.5.9 to vector bundles in general. Unfortunately, the term "vector bundle" is ambiguous because it does not indicate whether it is to be viewed as an object in BUN or FBUN. Rather than inventing some nonstandard terminology to overcome this problem, I shall just give the extra qualification as needed. The same procedure will be adopted for Lie group and affine bundles. In this example I am working in the category of FBUN.

Suppose that (V,ρ,M,W) is a vector bundle in the sense of definition 1.4.1. Suppose also that $\{U_i\}$ is an open cover of M. Then in order to be able to view (V,ρ,M,W) as an object in FBUN one needs to give transition functions

$$\Phi_{ij} : U_i \cap U_j \to GL(W).$$

If $(x_i^a) : U_i \to \mathbb{R}^m$ and $(x_j^a) : U_j \to \mathbb{R}^m$ are charts and also

$$x_j^a = \phi^a(x_i^b)$$

are the change of coordinates and (x_i^A) and (x_j^A) are the fiber coordinates on $\rho^{-1}(U_i)$ and $\rho^{-1}(U_j)$ respectively, then the Φ_{ij} satisfy

$$v_j^B = [\Phi_{ij}(x)]^{BA} v_i^A$$

(here, there is a summation over A and i,j merely index the cover $\{U_i\}$ on M). In example 1.5.8 the Φ_{ij}'s were defined in terms of the ϕ^a's i.e. in terms of the geometry of M. More generally one may consider the tensor bundles formed by tensoring copies of T_xM and T^*_xM $(x \in M)$.

Example 1.5.13 Following on from the last example I consider affine bundles as objects in FBUN. Let (E,π,F,V,ρ,W,M) be an affine bundle modelled on (V,ρ,W,M) as in definition 1.4.4 and employ the same notation as in example 1.5.11, but also let (e_i^A) denote the fiber coordinates of E over M on U_i. This time the transition functions are maps

$$\Psi_{ij} : U_i \cap U_j \to Aut(F).$$

which satisfy

$$e_j^B = [\Phi_{ij}(x)]^{BA} e_i^A + \Gamma_{ji}^B(x)$$

where the Φ_{ij}'s are the same as in the last example and the Γ_{ij}'s are given functions. This demonstrates in local coordinates the precise relationship between affine and vector bundles. Notice also that any affine bundle gives rises to a principal bundle in which the fiber is a vector space acting on itself by translation : thus, there is a functor from AFFBUN ∩ FBUN to PBUN.

1.6 Jet Bundles

Suppose that (E,π,M,F) is a bundle. If f and g are two sections of E over M then I will say

Definition 1.6.1 f and g are r-equivalent at $x \in M$ if for all curves $\gamma : \mathbb{R} \to M$, such that $\gamma(0) = x$ and all functions $F : E \to \mathbb{R}$ such that $F(x) = 0$

$$\frac{d^s}{dt^s}(F\circ f\circ\gamma(t))\Big|_{t=0} = \frac{d^s}{dt^s}(F\circ g\circ\gamma(t))\Big|_{t=0} \quad (0 \leqq s \leqq r)$$

(t denotes the natural coordinate on $\mathbb{R}$).

This definition gives an equivalence relation on the set of sections of E over M.

Definition 1.6.2 The r-equivalence class of a section f at x is called the r-jet of f at x and is denoted by $j^r_x f$: x is known as the source of $j^r_x f$ and $f(x)$ as the target. The union of all r-jets of sections at all points $x \in M$ is denoted by $J^r(E)$.

I shall outline next the fact that $J^r(E)$ is a manifold : it is moreover a bundle over each of M,E and $J^s(E)$ $(0 \leqq s \leqq r)$ with corresponding projection maps denoted by π^r_M, π^r_E and π^r_s. Suppose that M and E are of dimension m and $m+n$ respectively and that U is a typical open subset of M i.e. part of an atlas. Then $(\pi^r_M)^{-1}(U)$ may be taken as a typical open subset of $J^r(E)$. From a coordinate viewpoint $(x^i) : U \to \mathbb{R}^m$ and $(x^i,z^A) : \pi^{-1}(U) \to \mathbb{R}^{m+n}$ are coordinates on U and $\pi^{-1}(U)$ respectively and $j^r_x f \in (\pi^r_M)^{-1}(U)$ so that f has the local representation as

$$x^i = x^i, \quad z^A = f^A(x^i)$$

and the coordinates on $(\pi_M^r)^{-1}(U)$ are denoted by

$(x^i, z^A, z^A_{j_1}, \ldots, z^A_{j_1 j_2 \cdots j_r})$ then the coordinates of $j^r_x f$ are given by

$$x^i = x^i, \quad z^A = f^A(x^i), \quad z^A_{j_1} = \frac{\partial f^A}{\partial x^{j_1}}, \quad z^A_{j_1 j_2} = \frac{\partial^2 f^A}{\partial x^{j_1} \partial x^{j_2}} \ldots, z^A_{j_1 j_2 \cdots j_r} = \frac{\partial^r f^A}{\partial x^{j_1} \partial x^{j_2} \ldots \partial x^{j_r}}$$

Suppose that V were a second typical open subset of M and that $(\bar{x}^i, \bar{z}^A)$ are coordinates on $\pi^{-1}(V)$ and that $U \cap V \neq \phi$. Then since E is a bundle over M we must have for smooth local functions

$\phi^j : \mathbb{R}^m \to \mathbb{R}^m$ and $\Phi^B : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$

$$\bar{x}^j = \phi^j(x^i) \qquad \bar{z}^B = \Phi^B(x^i, z^A).$$

Using the chain rule and Leibnitz derivation properties these equations give the transition functions of the induced jet coordinates. Indeed, if the section f has the second local representation

$$\bar{x}^i = \bar{x}^i \qquad \bar{z}^A = \bar{f}^A(\bar{x}^i)$$

these transition functions will be given by

$$\bar{x}^j = \phi^j(x^i), \quad \bar{z}^B = \Phi^B(x^i, z^A), \quad \bar{z}^B_j = \frac{\partial(\phi^{-1})^i}{\partial x^j} \frac{\partial \Phi^B}{\partial z^A} z^A_i,$$

$$\bar{z}^B_{j_1 j_2} = \frac{\partial(\phi^{-1})^{i_1}}{\partial \bar{x}^{j_1}} \cdot \frac{\partial(\phi^{-1})^{i_2}}{\partial \bar{x}^{j_2}} \cdot \frac{\partial \Phi^B}{\partial z^A} z^A_{i_1 i_2} + \text{terms with first order jets,}$$

$$\bar{z}^B_{j_1 \ldots j_r} = \frac{\partial(\phi^{-1})^{i_1}}{\partial \bar{x}^{j_1}} \cdot \frac{\partial(\phi^{-1})^{i_2}}{\partial \bar{x}^{j_2}} \cdots \frac{\partial(\phi^{-1})^{i_r}}{\partial \bar{x}^{j_r}} \cdot \frac{\partial \Phi^B}{\partial z^A} z^A_{i_1 \ldots i_r} + \text{terms} \qquad (1.6.1)$$

with $(r-1)^{st}$ order jets.

Theorem 1.5.6 may now be invoked and one concludes that $J^r(E)$ is a manifold which is a bundle over both M and E. Moreover, counting the number of independent coordinates on $J^r(E)$ one also finds that

$$\dim J^r(E) = m + n \binom{m+r}{r}.$$

Jet bundles provide an appropriate context in which to view differential equations geometrically. Central to this approach, which was pioneered by E. Cartan, is the fact that jet bundles come naturally equipped with a canonical module of linear differential forms.

Definition 1.6.3 On $J^r(E)$ the r^{th}-order contact module is the module of 1-forms Ω^r defined by

$$\Omega^r\Big|_{j^r_x f} = (\pi^{r*}_{r-1} - \pi^r_M{}^*(j^{r-1}f)^*)T^*_{j^{r-1}_x f} J^{r-1}(E) \qquad (1.6.2)$$

The importance of the contact module is that it characterizes jets of sections.

Theorem 1.6.4 A section s of $J^r(E)$ over M is the r-jet of a section of E over M iff $s^*\Omega^r = 0$.

Proof: One method of proof would be to show that definition 1.6.3 leads to the usual basis for the contact module. If $(x^i, z^A, z^A_{j_1}, \ldots, z^A_{j_1}, \ldots, z^A_{j_1 \ldots j_r})$ are local coordinates for $J^r(E)$ as given above this basis is

$$(dz^A - z^A_i\, dx^i,\ dz^A_{j_1} - z^A_{j_1 i} dx^i, \ldots, dz^A_{j_1 \ldots j_{r-1}} - z^A_{j_1 \ldots j_{r-1} i} dx^i).$$

However, definition 1.6.3 affords a more elegant argument. Suppose firstly that $s = j^r f$ for some section f of E over M. Then for $x \in M$

$$\begin{aligned}
(s^*\Omega^r)\big|_x &= (j^r f)^*(\pi^{r*}_{r-1} - \pi^{r*}_M (j^{r-1}f)^*)\, T^*_{j^{r-1}_x f} J^{r-1}(E) \\
&= ((\pi^r_{r-1} \circ j^r f)^* - (j^{r-1}f \circ \pi^r_M j^r f)^*) T^*_{j^{r-1}_x f} J^{r-1}(E) \\
&= ((j^{r-1}f)^* - (j^{r-1}f)^*)\, T^*_{j^{r-1}_x f} J^{r-1}(E) \text{ since } \pi^r_M \circ j^r f = i_M \\
&= 0
\end{aligned}$$

The converse will be proved by induction on r. Note firstly that

$$\pi^{r}_{r-1}{}^{*}\,\Omega^{r-1} \subset \Omega^{r}. \tag{1.6.2}$$

This is because

$$\pi^{r*}_{r-1}(\Omega^{r-1}|_{j^{k-1}_x f}) = \pi^{r*}_{r-1}(\pi^{r-1*}_{r-2} - \pi^{r-1*}_{M}(j^{r-2}f)^{*})T^{*}_{j^{r-2}_x f}J^{r-2}(E)$$

$$= (\pi^{r}_{r-1}{}^{*} - (\pi^{r}_{M} \circ j^{r-1}f)^{*})\pi^{r-1*}_{r-2}T^{*}_{j^{r-2}_x f}J^{r-2}(E)$$

and the fact that $\pi^{r-1*}_{r-2}(T^{*}_{j^{r-2}_x f}J^{r-2}(E)) \subset T^{*}_{j^{r-1}_x f}J^{r-1}(E)$ since $J^{r-1}(E)$ is a bundle over $J^{r-2}(E)$.

Now suppose by induction that for t a section of $J^{r-1}(E)$ over M, $t^{*}\,\Omega^{r-1} = 0$ implies that $t = (\pi^{r-1}_{E} \circ t)$. Let s be a section of $J^{r}(E)$ over M such that $s^{*}\Omega^{r} = 0$ and let f be the section of E over M given by $\pi^{r}_{E} \circ s$. The best way to proceed now is to fix a point $x \in M$. Then $s(x) \in J^{r}(E)$ and so by definition there is a section of E over M, say, f_x such that $s(x) = j^{r}_{x}f_x$. Since $s^{*}\Omega^{r} = 0$ (1.6.2) implies that $s^{*}\pi^{r}_{r-1}{}^{*}\Omega^{r-1} = (\pi^{r}_{r-1} \circ s)^{*}\,\Omega^{r-1} = 0$. Hence by the induction hypothesis

$$\pi^{r}_{r-1} \circ s = j^{r-1}(\pi^{r-1}_{E} \circ \pi^{r}_{r-1} \circ s).$$

But $\pi^{r-1}_{E} \circ \pi^{r}_{r-1} \circ s = \pi^{r}_{E} \circ s = f$ and so $\pi^{r}_{r-1} \circ s = j^{r-1}f$. This shows that $j^{r-1}f$ and $j^{r-1}f_x$ agree to $(r-1)^{st}$ order at x.

$$\text{Now } 0 = s^*\,\Omega^r\big|_{j^r_x f_x} = s^*(\pi^{r^*}_{r-1} - \pi^{r^*}_{M}\, j^{r-1}f_x^*)T^*_{j^{r-1}_x f_x} J^{r-1}(E)$$

$$= ((\pi^r_{r-1} \circ s)^* - (j^{r-1}f_x \circ \pi^r_M \circ s)^*)T^*_{j^{r-1}_x f_x} J^{r-1}(E)$$

$$= ((j^{r-1}f)^* - (j^{r-1}f_x)^*)T^*_{j^{r-1}_x f_x} J^{r-1}(E).$$

It follows from the preceding argument that $j^{r-1}f$ and $j^{r-1}f_x$ agree to first order as sections of $J^{r-1}(E)$. Now $J^r(E) \subset J^1(J^{r-1}(E))$ and since $j^r f$ and $j^r f_x$ are sections of $J^r(E)$ it follows that f and f_x agree to order r at x. The induction is completed by noting that the case $r = 1$ gives immediately the condition that f and f_x agree to first order at x. ∎

In the language of jet bundles a system of r^{th} order partial differential equations (P.D.E.r.O.) is simply a local submanifold Σ of the r-jet bundle of sections of a bundle E over M. M is the space of "independent variables" and the fiber of E over M the space of "dependent variables". A solution of Σ is the r-jet of a local section whose image lies in Σ. Since Ω^r detects those sections of $J^r(E)$ over M which are r-jets of sections of E over M, the problem of finding solutions may be rephrased as one of finding m-dimensional integral manifolds of Ω^r restricted to Σ. I now examine the bundle structure of the jet spaces. Arbitrary order jet bundles are not in general vector bundles. Instead one has

Proposition 1.6.5 Let E be a bundle over M. Then $J^r(E)$ via π^r_{r-1} has the structure of an affine bundle over $J^{r-1}(E)$.

Proof: Two parts $j^r_x f$, $j^r_{\bar{x}} \bar{f}$ of $J^r(E)$ are in the same fiber over $J^{r-1}(E)$ iff $x = \bar{x}$ and f and $\bar{f}$ are $(r-1)$-equivalent at x.

Considering now equation (1.6.1) we see that the difference of two such points depends only on their r-jet and transforms as tensor. If we use two sets of standard coordinates $(x^i, z^A, z^A_i, \ldots, z^A_{i_1 \ldots i_r})$, $(\bar{x}^i, \bar{z}^A, \bar{z}^A_i, \ldots, \bar{z}^A_{i \ldots i_r})$ and make a transformation $\bar{x}^i = \phi^i(x^j)$, $\bar{z}^A = \phi^A(x^j, z^B)$ we see that this tensor transforms with

$$\frac{\partial(\phi^{-1})^{j_1}}{\partial \bar{x}^{i_1}} \frac{\partial(\phi^{-1})^{j_2}}{\partial \bar{x}^{i_2}} \cdots \frac{\partial(\phi^{-1})^{j_r}}{\partial \bar{x}^{i_r}} \cdot \frac{\partial \Phi^A}{\partial z^B}.$$ This means that it transforms as an element of $S^{r*}(M) \otimes V(E)$. Hence by theorem 1.5.6 it follows that $J^r(E)$ is an affine bundle modelled on $S^r(M) \otimes V(E)$. ∎

1.7 $J^1(\mathbb{R},M)$ and $J^1(M,\mathbb{R})$

Let M be a smooth m-dimensional manifold. We may think of $\mathbb{R} \times M$ and $M \times \mathbb{R}$ as trivial fiber bundles over $\mathbb{R}$ and M respectively, and then sections are simply curves and functions on M respectively. Thus, we may consider the bundles $J^1(\mathbb{R} \times M)$ and $J^1(M \times \mathbb{R})$. In the case where one deals with jets of maps as opposed to jets of sections, it is convenient to write $J^r(M,N)$ rather than $J^r(M \times N)$, where $M \times N$ is regarded as a trivial fiber bundle over M. Hence, I write $J^1(\mathbb{R}, M)$ and $J^1(M, \mathbb{R})$ rather than $J^1(\mathbb{R} \times M)$ and $J^1(M \times \mathbb{R})$ respectively. It should be clear from the definition of $J^1(\mathbb{R}, M)$ and $J^1(M, \mathbb{R})$ and the "usual" definition of TM and T*M (see for example [38]), that $J^1(\mathbb{R},M)$ is globally diffeormorphic to $\mathbb{R} \times TM$ and $J^1(M, \mathbb{R})$ globally diffeomorphic to $\mathbb{R} \times T^*M$. Indeed, TM and T*M may actually be defined as subbundles of $J^1(\mathbb{R},M)$ and $J^1(M,\mathbb{R})$ respectively.

All of the spaces $J^1(\mathbb{R},M)$, TM, $J^1(M,\mathbb{R})$ and T*M are very important for what is to follow in this work. $J^1(\mathbb{R},M)$ is the setting for Lagrangian mechanics and chapter 4 examines some of the intrinsic, geometrical properties of $J^1(\mathbb{R},M)$ and TM. T*M occurs in chapters 2 and 3 in connection with Hamiltonian theory. The first three sections of chapter 5 are essentially concerned with imitating the geometric constructions of T*M on $J^1(M,\mathbb{R})$, in other words, passing from symplectic to contact geometry. Clearly then, it is important to know the various bundle structures of $J^1(\mathbb{R},M)$ and $J^1(M,\mathbb{R})$. These will now be investigated and I start with $J^1(M,\mathbb{R})$.

$J^1(M,\mathbb{R})$ may be viewed as a bundle over $\mathbb{R}$, M, T*M or $M \times \mathbb{R}$ and the associated projection maps will be denoted by α, β, γ and $\alpha \times \beta$ respectively. α, β, γ may be defined by, for $j^1_x f \in J^1(M,\mathbb{R})$

$$\alpha(j^1_x f) = x, \quad \beta(j^1_x f) = f(x), \quad \gamma(j^1_x f) = df|_x.$$

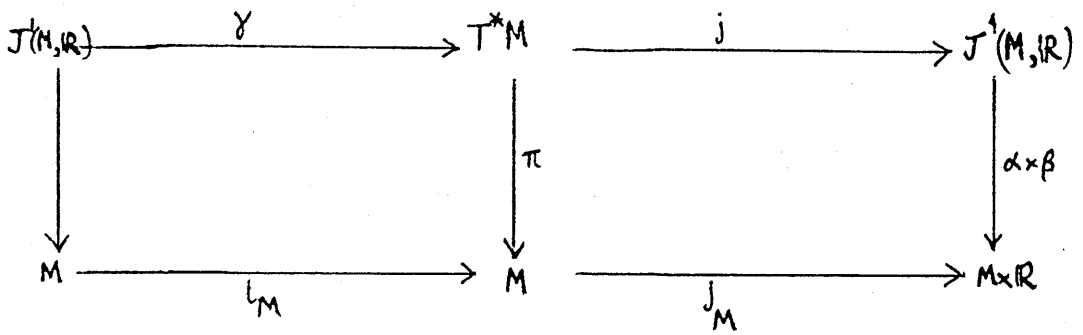
Indeed, $J^1(M,\mathbb{R})$ is naturally a vector bundle over M or T*M by defining, for $\lambda \in \mathbb{R}$ $j^1_x f$, $j^1_x g \in J^1(M,\mathbb{R})$

$$j^1_x f + j^1_x g = j^1_x (f+g), \quad \lambda(j^1_x f) = j^1_x (\lambda f).$$

$J^1(M,\mathbb{R})$ may also be viewed as a vector bundle over $M \times \mathbb{R}$ but one must define instead

$$j^1_x f + j^1_x g = j^1_x (f+g - f(x)), \quad \lambda(j^1_x f) = j^1_x (\lambda f - (\lambda-1)f(x))$$

where f(x) denotes the identically constant function with value f(x). The precise relationship of the spaces T*M and $J^1(M,\mathbb{R})$ may be clarified by the following diagram:



Here, π is the cotangent bundle projection and j_M denotes the inclusion $x \longmapsto (x,0)$. Also, j is the map given by $df|x \longmapsto j^1_x (f-f(x))$, where again $f(x)$ denotes the constant function $f(x)$. Both γ and j are VBUN morphisms - indeed T^*M is the pullback of $J^1(M,\mathbb{R})$ via j_M.

Each of $J^1(M,\mathbb{R})$ and T^*M carries an intrinsically defined 1-form. In the case of $J^1(M,\mathbb{R})$ this 1-form, say, Θ is defined by, for $j^1_x f \in J^1(M,\mathbb{R})$ where $i_{\mathbb{R}}$ is the identity function on $\mathbb{R}$,

$$\Theta \Big|_{j^1_x f} = d\beta - \alpha^* f^* \, d(i_{\mathbb{R}}) \ .$$

In any standard coordinate system (x^i, z, p_i), Θ has the local expression $dz - p_i dz^i$. Θ is an example of a contact structure whose properties will be examined in chapter 5. On T^*M the canonical 1-form θ is defined by, where $p \in T^*M$ with $\pi(p) = x$

$$\theta(p) = \pi^*(p) .$$

Here, $\pi^* : T_x{}^*M \to T_p{}^*T^*M$ and so θ defines a section of T^*T^*M over T^*M i.e. is a 1-form on T^*M. θ is also characterized by the universal property that $\sigma^*\theta = \sigma$ for any $\sigma \in F^1(M)$. If (x^i, p_i) are standard coordinates on T^*M, θ has the local form $p_i dx^i$ and hence $j^*\theta = -\Theta$. The exterior derivative of θ is the canonical example of a symplectic structure, whose properties are briefly reviewed in section 5.1. More generally, $d\theta$ may be dualized to give a 2-vector and this is an example of a cosymplectic structure, which are discussed in chapter 2.

There is one further property of $J^1(M,\mathbb{R})$ that is needed in chapter 5 and which it is appropriate to give here. $J^1(M,\mathbb{R})$ may be embedded in $T^*(M \times \mathbb{R})$: in fact if $j^1_x f \in J^1(M,\mathbb{R})$ define a map $i : J^1(M,\mathbb{R}) \to T^*(M \times \mathbb{R})$ by

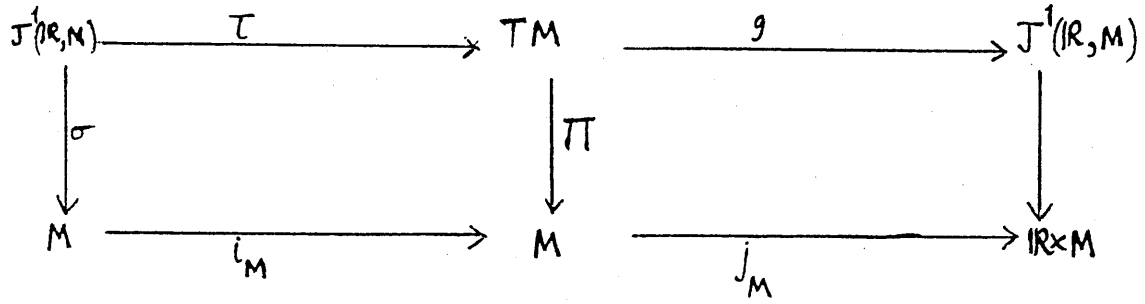
$$j^1 x f \longmapsto dF|_{x,f(x)} \quad \text{where } F(x,z) = f(x) - z.$$

It is easy to see that if $\theta_{M\times\mathbb{R}}$ denotes the canonical 1-form on $T^*(M \times \mathbb{R})$ then $i^*(\theta_{M\times\mathbb{R}}) = -\theta$.

The decision to examine $J^1(M,\mathbb{R})$ before $J^1(\mathbb{R},M)$ was not made without reason. $J^1(M,\mathbb{R})$ enjoys its many linear properties because it consists of functions into a linear space . Now $J^1(\mathbb{R},M)$ also has linear properties which it derives in virtue of its duality with $J^1(M,\mathbb{R})$. $J^1(\mathbb{R},M)$ may be viewed as a bundle over $\mathbb{R}$, M, TM or $\mathbb{R} \times M$ and the corresponding projection maps will be denoted by ρ, σ, τ and $\rho \times \sigma$. ρ,σ and τ are defined by, for $j^1_t \gamma \in J^1(\mathbb{R},M)$

$$\rho(j^1_t \gamma) = t, \quad \sigma(j^1_t \gamma) = \gamma(t), \quad \tau(j^1_t \gamma) = \gamma(t).$$

The relationship between the spaces TM and $J^1(\mathbb{R},M)$ may be summarized in the following diagram.



Here, Π is the tangent bundle projection and j_M denotes the inclusion $x \to (0,x)$. Also g is the map given by $[\gamma]_x \to j_0^1\gamma$.

I showed above that $J^1(M,\mathbb{R})$ is naturally a vector bundle over both M and $M \times \mathbb{R}$ via β and $\alpha \times \beta$ (although $(i_{J^1(M,\mathbb{R})}, j_M)$ does not define a VBUN morphism). There is a pairing of elements $j_t^1\gamma$ in $J^1(\mathbb{R},M)$ with those $j_x^1 f$ in $J^1(M,\mathbb{R})$ which satisfy $\gamma(t) = x$ by

$$\langle j_t^1\gamma,\ j_x^1 f\rangle = \frac{d}{dt}\,(f \circ \gamma(t))\Big|_{t=0} .$$

It is clear that this pairing is non-degenerate and demanding that it is bilinear, endows $J^1(\mathbb{R},M)$ with a vector bundle structure both over M via σ, and over $\mathbb{R} \times M$ via $\rho \times \sigma$. TM may be made into equivalent vector bundles using either of these two vector bundles structures; then, both (τ, i_M) and (g, j_M) are VBUN morphisms.

1.8 Canonical dilation vector field on a vector bundle

Suppose that (V,π,M,W) is a vector bundle in the sense of definition (1.4.1). The scalar multiplication by $\mathbb{R}$ on each fiber gives rise to an action a of the multiplicative group of $\mathbb{R}^*$ on V. In fact if $v_x \in V$ define $a : \mathbb{R}^* \times V \to V$ by

$$(\lambda, v_x) \to (\lambda v)_x .$$

Since for $\lambda,\mu \in \mathbb{R}^*$ $a(\lambda\mu, v_x) = a(\lambda,(\mu v)_x) = a(\lambda,a(\mu, v_x))$ a does indeed define an action. It follows from Lie's second fundamental theorem [32] that there is an induced vector field Δ on V. Δ is called the canonical dilation vector field and is defined by,

$\forall f \in F(V)$, $\forall v_x \in V$

$$\Delta f(v_x) = \frac{d}{d\lambda} (f(a(\lambda, v_x)))\Big|_{\lambda=0}.$$

Clearly Δ is tangent to the fibration $\pi : V \to M$: in fact if coordinates (x_i) are introduced on M and (v_A) are fiber coordinates then Δ has the local expression $v_A \frac{\partial}{\partial v_A}$. The existence of Δ renders it possible to speak of a function being homogeneous in the fiber : $f \in F(V)$ will be said to be homogeneous of degree s in the fiber if $\Delta f = sf$. The following observation is important.

Proposition 1.8.1 A function $f : V \to \mathbb{R}$ is homogeneous of degree zero iff it is (the pullback of) a function on M.

Sketch of proof $\Delta f = 0$ means geometrically that the levels of f are tangent to Δ: but given any point $v_x \in V$ there is an integral curve through v_x "ending" on the zero section and so f is entirely

determined by its value on the zero section. This means that f is the pullback of a function on M. Conversely, if $f = \pi^*\phi$ for some function ϕ on M then $\Delta f = \Delta(\pi^*\phi) = (\pi_*\Delta)\phi = 0$ since Δ is tangent to the fiber. ∎

The vector field Δ may also be considered on a vector space which corresponds to the preceding case when M degenerates to a point. On a vector space V any vector field may be identified as an endomorphism of V. From this point of view, Δ corresponds to the identity automorphism i_V. Next, a point of notation. If $\Phi : M \to M'$ is a map of manifolds then I use $T\Phi$ to denote the induced bundle map of TM to TM'. In general Φ does not induce a map of $V^1(M)$ to $V^1(M')$, indeed this only happens when Φ is a diffeomorphism. Again, if Φ is not a diffeomorphism there is no bundle morphism of T^*M' to T^*M. There is, however, always a map of sections i.e. of $F^1(M')$ to $F^1(M)$ and this will be denoted by Φ^*. Finally, Φ_* will be used to denote the pointwise map of tangent spaces; hence in the case where Φ is a diffeomorphism and $X \in V^1$

$$\Phi_* X \circ \Phi = (T\Phi) \circ X.$$

Lemma 1.8.2 Let V, V' be vector spaces with canonical dilation vector fields Δ and Δ' respectively and Φ a smooth map of V to V'. Then Δ and Δ' are Φ-related i.e. $\Delta' \circ \Phi = (T\Phi) \circ \Delta$ iff Φ is linear.

Proof: Make the identification of Δ and Δ' with i_V and $i_{V'}$ respectively. The condition of Φ-relatedness then becomes

$$i_{V'} \circ \Phi = T\Phi \circ i_V$$

where Φ_* is regarded as a map from V to V'. But the last equation says of course that $\Phi = T\Phi$ which is precisely the condition that Φ be linear. ∎

The last result prepares the way for the next on the bundle level.

Theorem 1.8.3 Let (V,π,M,W), (V',π',M',W') be vector bundles and Δ and Δ' their respective canonical dilation vector fields. Then a map Φ of V to V' is a VBUN morphism iff Δ and Δ' are Φ-related.

Proof: If Φ is a VBUN morphism then, since Δ and Δ' are tangent to the fiber it follows that Δ and Δ' are $\Phi|V_x$-related on each fiber V_x by lemma 1.8.2. This gives the necessity and the sufficiency also follows, provided only that one knows that Φ is a BUN morphism. Next suppose that $v_1, v_2 \in V$ satisfy $\pi(v_1) = \pi(v_2)$. Then there are integral curves of Δ passing through v_1 and v_2 respectively and "meeting" on 0_v (the zero section of V). Since Δ is Φ-related to Δ' there must be integral curves of Δ' passing through $\Phi(v_1)$ and $\Phi(v_2)$ respectively which "meet". However, two integral curves of Δ^1 "meet" only if they lie in the same fiber of V^1; hence $\pi' \circ \Phi(v_1) = \pi' \circ \Phi(v_2)$ i.e. the fibration is preserved. ∎

Corollary 1.8.4 With the notation of theorem 1.8.3, Φ is a VBUN isomorphism iff $\Delta' \circ \Phi = (T\Phi) \circ \Delta$. ∎

1.9 A brief review of Analytical Mechanics

In this section I give a brief summary of mechanics as it is needed in Chapters 2, 3 and 4. A great deal of apparatus has been developed in this chapter and some of the manifolds defined provide the kinematics. It remains to add the dynamics.

There are essentially two distinct approaches to mechanics which bear the names of Lagrange and Hamilton respectively. Traditionally it has usually been assumed that the difference between Lagrangian and Hamiltonian mechanics is rather formal. As Goldstein [16] says "...the Hamiltonian formulation usually does not materially decrease the difficulty of solving any given problem in mechanics. The advantages of the Hamiltonian formulation lie not in its use as a calculational tool, but rather in the deeper insight it affords into the formal structure of mechanics." Indeed, the development of intrinsic geometric methods has led to a deeper appreciation of the difference inherent in the two approaches; these can be quite significant especially when one comes to formulate field and quantum theories for example.

Lagrangian dynamics is usually formulated on $\mathbb{R} \times TM \cong J^1(\mathbb{R},M)$ and Hamiltonian dynamics on T*M where, as always, M denotes a smooth m-dimensional manifold. In both cases M is called the configuration space and is the space in which the actual motions occur. TM(T*M) is called phase space and $\mathbb{R} \times TM$ evolution space. The fact that Lagrangian theory contains the extra $\mathbb{R}$ factor enables

it to deal with systems which are explicitly time dependent. There are subtle differences between the spaces TM and $\mathbb{R} \times TM$ and these are investigated in chapter 4. In the rest of this section it will be assumed that Lagrangian theory is formulated on $\mathbb{R} \times TM$ and Hamiltonian theory on T*M. However, Crampin [8] has recently explained how Lagrangian theory may be formulated directly on TM. Also, time dependent Hamiltonian systems are easily accommodated on $\mathbb{R} \times T^*M$.

Lagrange's equations of motion are second order equations constructed from a function L called the Lagrangian which depends on first order quantities. As follows from section 1.6, a system of second order ordinary differential equations (O.D.E. 2.0) is locally, from the geometric viewpoint, a codimension m submanifold Σ in $J^2(\mathbb{R},M)$. Adopting standard coordinates $(t, x_i, \dot{x}_i, \ddot{x}_i)$ for $J^2(\mathbb{R},M)$ the problem of solving an O.D.E. 2.0 may be restated as that of finding m-dimensional, integral manifolds of the contact module $\{dx_i - \dot{x}_i dt, d\dot{x}_i - \ddot{x}_i dt\}$ restricted to Σ for which also $dt \neq 0$. It may be possible to write the system of O.D.E. 2.0 in the form $\ddot{x}_i = \Gamma_i(t, x_i, \dot{x}_i)$ (where the Γ_i's are smooth functions); geometrically, this means that Σ is transverse to the fibration $\pi_1^2 : J^2(\mathbb{R},M) \to J^1(\mathbb{R},M)$. If this happens the restricted contact module is $\{dx_i - x_i dt, d\dot{x}_i - \Gamma_i dt\}$ and one sees that the whole problem has been pulled down to $J^1(\mathbb{R},M)$. In other words Σ has been identified with the space $J^1(\mathbb{R},M)$ which explains how the second order theory case be formulated entirely on $J^1(\mathbb{R},M)$. On the other hand, it should be appreciated that not every system of O.D.E. 2.0 can be treated in this way; I shall call

O.D.E. 2.0. with this property regular. In the case of the Euler-Lagrange equations the submanifold Σ is determined by the vanishing of the m functions (see Crampin [7])

$$\frac{\partial^2 L}{\partial t \partial \dot{x}^i} + \dot{x}^j \frac{\partial^2 L}{\partial x^j \partial \dot{x}^i} + \ddot{x}^j \frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j} - \frac{\partial L}{\partial x^i} .$$

Clearly then, the condition that allows the equations to be pulled down to $J^1(\mathbb{R},M)$ is $\det \left(\frac{\partial^2 L}{\partial \dot{x}^i \partial \dot{x}^j}\right) \neq 0$. A Lagrangian which satisfies this condition at all points of $J^1(\mathbb{R},M)$ is said to be regular. Hence, a regular Lagrangian gives regular Euler-Lagrange equations. Geometrically, the Euler-Lagrange equations arise as the condition obeyed by the extrremals of the variational problem $\int L dt$. This derivation is standard and may be found in [16] for example.

I now turn to the formulation of Hamilton's equations which, in contrast to the Euler-Lagrange equations are not derived from a variational problem. Instead, one simply chooses some function h on T*M. Letting θ denote the canonical 1-form on T*M, $d\theta$ is a symplectic structure and hence there is some vector field X_h on T*M defined by $X_h \rfloor d\theta = -dh$. X_h is the global Hamiltonian vector field associated to h and, unlike the Lagrangian, <u>any</u> smooth h can be chosen. In terms of standard coordinates (x^i, p_j) for T*M Hamilton's equations are obtained:

$$\dot{x}^i = \frac{\partial h}{\partial p_i} , \quad \dot{p}_i = - \frac{\partial h}{\partial x_i} .$$

However, from the geometrical point of view they are not, nor necessarily equivalent to, second order equations. Indeed, the construction of X_h depends only on the symplectic structure of

T^*M i.e. $d\theta$ rather than θ. Hence, Hamiltonian dynamics may be formulated on any symplectic manifold: all one need do is choose any smooth function to obtain the dynamics.

A vector field X which satisfies the condition $d(X \lrcorner d\theta) = 0$ is said to be locally Hamiltonian. By the Poincaré lemma, it follows that locally there is a function h such that $X \lrcorner d\theta = -dh$, though h need not be globally defined. However, it is a standard fact that the Lie bracket of two locally Hamiltonian vector fields is always a global Hamiltonian vector field (see Arnold [1]).

CHAPTER 2

SCHOUTEN'S BRACKETS WITH APPLICATIONS

Chapter 2 is concerned with Schouten's brackets. It is shown that there are actually two quite distinct brackets on $S(M)$ and $\Lambda(M)$ - the space of symmetric and skew symmetric, contravariant tensors associated to a smooth manifold M. The definitions and essential properties are proved rigorously for the bracket on $S(M)$ and sketched in the case of the bracket on $\Lambda(M)$. Each of the two brackets has an important role to play in theoretical mechanics. The bracket on $S(M)$ allows the concept of a Killing tensor to be defined and this is done in section 2.2. A theorem is proved which gives an upper bound for the maximum possible dimension of the space of Killing tensors on Riemannian or pseudo-Riemannian manifold. As a corollary, one can also obtain the dimension of the homogeneous Killing tensors of rank n and degree r on a flat manifold.

The bracket $\Lambda(M)$ allows the notion of a Poisson manifold to be defined. A sketch using local coordinates is given of why Poisson manifolds are precisely those which have Poisson bracket structures i.e. commutative and Lie algebra structures related by a derivation type property. One of the two natural examples of a Poisson manifold is T*M: this has the stronger property of actually being symplectic. The other natural example is the dual space of a finite dimensional Lie algebra g. This is examined in section 2.3. It is shown that g* is indeed a Poisson manifold and that $S(g)$ has a natural (Schouten-type)

bracket which enables it to be identified with a subalgebra of $F^0(g^*)$. The results in section 2.3 are more or less standard.

2.1 Schouten's brackets

Schouten's brackets were introduced quite some time ago but it is only comparatively recently that their important in mechanics has been recognized [29, 43]. In the first place, it should be emphasized that there are actually two distinct Schouten brackets. These are defined on $\underline{S(M)}$ and $\underline{\Lambda(M)}$ the space of symmetric, contravariant and skew-symmetric, contravariant tensors respectively, associated to a smooth manifold M, which, as usual, will be assumed to be m-dimensional. $\underline{S(M)}$ and $\underline{\Lambda(M)}$ are each graded $\mathbb{R}$-algebras under $\odot$ and $\wedge$ respectively; these products arise from the associated bundles S(M) and Λ(M) and hence $\underline{S(M)}$ is commutative whereas $\underline{\Lambda(M)}$ satisfies $[A,B] = (-1)^{pq}[B,A]$ $(A \in \overset{p}{\underline{\Lambda}}(M),\ B \in \overset{q}{\underline{\Lambda}}(M))$. The symmetric Schouten bracket on $\underline{S(M)}$ makes it into a Lie algebra in which the two algebraic structures are related by a derivation formula. The situation on $\underline{\Lambda(M)}$ is rather similar except it is a graded Lie algebra and the derivation formula respects the grading.

The bracket on $\underline{S(M)}$ is defined as follows: firstly, if $A,B \in \underline{S(M)}$

$[A,B] = 0.$

Secondly, if $A \in \underline{S^1(M)}$ and $B \in \underline{S^0M}$

$[A,B] = AB$ i.e. the directional derivative of the function B.

Thirdly, if $A,B \in \underline{S^1(M)}$

[A,B] is the Lie bracket of vector fields.

This defines the bracket on elements of degree zero and one. Since $\underline{S(M)}$ is a commutative, graded algebra the bracket is now determined

automatically on the other degrees by insisting that it be $\mathbb{R}$-bilinear, and that it be a derivation in both factors i.e. for A, B, C $\in$ S(M)

$$[A,B \odot C] = [A,B] \odot C + B \odot [A,C] \qquad (2.1.1)$$

$$[A \odot B, C] = [A,C] \odot B + A \odot [B,C] \qquad (2.1.2)$$

The definition of the symmetric Schouten bracket just given is essentially the same as that given by Woodhouse [43]. If A, B $\in$ S(M) and local coordinates are chosen so that A and B have the components $A^{i_1 \cdots i_p}$ and $B^{j_1 \cdots j_q}$ then $[A,B] \in S^{p+q-1}(M)$ and it has the components (see [29])

$$[A,B]^{j_1 \cdots j_{p+q-1}} = A^{i(j_1 \cdots j_{p-1}} B_{,i}^{\ \ j_p \cdots j_{p+q-1})} - B^{i(j_1 \cdots j_{q-1}} A_{,i}^{\ \ j_q \cdots j_{p+q-1})} . \qquad (2.1.3)$$

(2.1.3) may be used to give a proof that the symmetric Schouten bracket satisfies the Jacobi identity. Alternatively, one may proceed as follows.

Proposition 2.1.1 For all A, B, C $\in$ S(M)

$$[A,[B,C]] = [[A,B], C] + [B, [A,C]] \qquad (2.1.4)$$

Proof: In view of the definition of the symmetric Schouten bracket (2.1.4) is certainly true when the A, B, C's are of degree zero and one (or any mixture of zeros and ones). Since any element of $S^p(M)$ may be written as a sum of decomposable (simple) symmetric tensors, and in view of the $\mathbb{R}$-bilinearity of the bracket, (2.1.4) in general results by induction, from the following two part argument. Firstly, if $A \in S^p(M)$, $B \in S^q(M)$, $C \in S^{r'}(M)$, and if the Jacobi identity holds for all $p \leqq p'$, $q' \leqq q$, $r \leqq r'$ then it holds also for all $p \leqq p' + 1$, $q' \leqq q$, $r \leqq r'$.

Proof: Let $D \in S^1(M)$. Then I shall show that

$$[A\Theta D, [B,C]] = [[A\Theta D,B], C] + [B, [A\Theta D, C]]. \qquad (2.1.5)$$

Now reducing the left hand side gives, using (2.1.1) and (2.1.2),

$$[A\Theta D, [B,C]] = A\Theta[D,[B,C]] + [A,[B,C]] \odot D$$

$$= A\Theta[[D,B],C] + A\Theta[B,[D,C]] + [[A,B],C]\Theta D + [B,[A,C]]\Theta D$$

by Jacobi.

Next, considering the right hand side of (2.1.5) and using (2.1.1) and (2.1.2)

$$[[A\Theta D,B],C] + [B,[A\Theta D,C]]$$

$$= [[A,B]\Theta D,C] + [A\Theta[D,B],C] + [B,[A,C]\Theta D] + [B,A\Theta[D,C]]$$

$$= [[A,B],C]\Theta D + [A,B]\Theta[D,C] + [A,C]\Theta[D,B] + A\Theta[[D,B],C]$$

$$+ [B, [A,C]]\Theta D + [A,C]\Theta[B,D] + [B,A]\Theta[D,C] + A\Theta[B,[D,C]]$$

(again by (2.1.1) and (2.1.2))

$$= [[A,B],C]\Theta D + A\Theta[[D,B],C] + [B,[A,C]]\Theta D + A\Theta[B,[D,C]]$$

(by skew-symmetry).

This shows the equality of the left and right hand sides of (2.1.5) and the first part of the induction argument is complete. The second part asserts that with A, B or C as above, the Jacobi identity also holds for $p \leq p'$, $q' \leq q$, $r \leq r' + 1$.

Proof: Again let $D \in S^1(M)$. Then I shall show that

$$[A,[B,C\Theta D]] = [[A,B],C\Theta D] + [B,[A,C\Theta D]] \qquad (2.1.6)$$

Firstly, reducing the left hand side gives using (2.1.1) and (2.1.2)

$$[A,[B,C\odot D]] = [A,[B,C]]\odot D + [B,C]\odot[A,D] + [A,C]\odot[B,D] + C\odot[A,[B,D]]$$

$$= [[A,B],C]\odot D + [B,A,C]\odot D + [B,C]\odot[A,D] + [A,C]\odot[B,D]$$

$$+ C\odot[[A,B],D] + C\odot[B,[A,D]] \text{ by Jacobi.}$$

Next reducing the right hand side of (2.1.6) with (2.1.1) and (2.1.2) gives

$$[[A,B],\ C\odot D] + [B,[A,\ C\odot D]]$$

$$= [[A,B],C]\odot D + C\odot[[A,B],D] + [B,[A,C]\odot D] + [B,[A,D]\odot C]$$

$$= [[A,B],C]\odot D + C\odot[[A,B],D] + [B,[A,C]]\odot D + [A,C]\odot[B,D]$$

$$+ [B,[A,D]]\odot C + [B,C]\odot[A,D]$$

showing that the left hand and right hand sides of (2.1.6) are equal. The proof of proposition 2.1.1 is now complete because although the Jacobi identity involves three arguments, the induction follows from just (2.1.5) and (2.1.6) in view of the skew-symmetry of Schouten's bracket. ∎

In chapter 3 I shall be concerned with constants of motion in Hamiltonian mechanics. The symmetric Schouten bracket is of fundemental importance in this context and the following theorem will be used frequently. It depends on the fact that a symmetric, contravariant tensor M may be identified with a function on T^*M which is polynomial in the fiber. If $A \in \underline{S^n(M)}$ then a is defined by $a(p) = A(x)(p,\ldots,p)$ where $p \in T^*M$ and $\pi(p) = x$, there being n arguments on the right hand side.

Theorem 2.1.2 The map defined by $A \to a$ just described, defines a Lie algebra isomorphism from $S(M)$ to $F^0(T^*M)$. The proof can be given from a coordinate calculation using (2.1.3) or, more elegantly, by establishing the result for $S^0(M)$ and $S^1(M)$ and then using $\mathbb{R}$-linearity and (2.1.1) and (2.1.2) to deduce the general case by induction.∎ In fact the Hamiltonian formalism on T*M briefly described in section 1.9 provides yet another way to view the symmetric Schouten bracket. Any $a \in F^0(T^*M)$ defines a global Hamiltonian vector field X_a on T*M as was explained in section 1.9. Hence, any symmetric, contravariant tensor field A on M may be lifted to a vector field X_a on T*M. Thus, starting from $A,B \in S(M)$ one may construct X_a and X_b and thence $[X_a, X_b]$. Since $[X_a, X_b]$ is globally Hamiltonian, it determines a function on T*M which may be normalized so as to give a function which is polynomial in the fiber (it must be made to vanish on the zero section of T*M) and which thus can be identified with an element of $S(M)$.

In section 2.2 the Schouten bracket on S(M) is used to isolate the notion of a Killing tensor. This concept, which is a generalization of the familiar notion of a Killing vector, applies in the category of Riemannian or pseudo-Riemannian manifolds. Specifically, a Killing tensor is a symmetric, (contravariant) tensor which Schouten commutes with the metric. As we shall see, Killing tensors are first integrals of the geodesic flow, but play a role in the problem of finding first integrals of more general Hamiltonians.

I turn next to the skew Schouten bracket on $\underline{\Lambda(M)}$. It should be emphasized that it is not possible to define a bracket on $\underline{C(M)}$. If one tries to define it along the lines of the symmetric Schouten bracket it will not be well-defined and one obtains different answers according to the order in which various operations are performed. Of course general elements of $\underline{C(M)}$ cannot be identified with elements of $F^0(T^*M)$ which is another manifestation of the same phenomenon. However, in the case of $\underline{\Lambda(M)}$ something analogous can be done by way of constructing a bracket. In fact Schouten's skew bracket is defined in the same way as the symmetric bracket for elements of degree zero and one. Again it will be extended to higher degrees by using $\mathbb{R}$-linearity and the following rule instead of (2.1.1) and (2.1.2): for $A \in \underline{\Lambda^p(M)}$ and $B \in \underline{\Lambda^q(M)}$ with A and B decomposable, say,

$$A = A_1 \wedge \ldots \wedge A_p, \quad B = B_1 \wedge \ldots \wedge B_q$$

where each A_i and B_j are vector fields on M define

$$[A,B] = \sum_{i,j} A_1 \wedge \ldots \wedge A_{i-1} \wedge A_i \wedge \ldots \wedge A_p \wedge [A_i, B_j] \wedge B_1 \wedge \ldots \wedge B_{j-1} \wedge B_j \wedge \ldots \wedge B_q$$

Technically, the Lie derivative operation has been extended to $\underline{\Lambda(M)}$ as a "biderivation of degree one". This means that one has instead of (2.1.1) and (2.1.2), for A, $C \in \underline{\Lambda(M)}$, $\underline{B \in \Lambda^q(M)}$.

$$[A, B \wedge C] = [A,B] \wedge C + (-1)^q B \wedge [A,C] \tag{2.1.7}$$

$$[A \wedge B, C] = (-1)^q [A,C] \wedge B + A \wedge [B,C] \tag{2.1.8}$$

Again if local coordinates are chosen so that A and B have the components $A^{i_1 \ldots i_p}$ and $B^{j_1 \ldots j_q}$, $[A,B] \in \Lambda^{p+q-1}(M)$ and its components are given by

$$[A,B]^{j_1\cdots j_{p+q-1}} = A^{i[j_1\cdots j_{p-1}}{}_{,i}B^{j_p\ldots j_{p+q-1}]} + (-1)^{pq}\, B^{i[j_1\cdots j_{q-1}}{}_{,i}A^{j_q\cdots j_{p+q-1}]} \tag{2.1.9}$$

Hence, rather than being skew-symmetric the bracket on $\underline{\Lambda(M)}$ satisfies

$$[A,B] = (-1)^{pq}[B,A] \tag{2.1.10}$$

Similarly, one also has a graded Jacobi identity which may be proved from (2.1.9) (see [29]), or much in the same way as proposition 2.1.1 but this time respecting the grading

$$[A,[B,C]] = [[A,B],C] + (-1)^{pq}[B,[A,C]] \tag{2.1.11}$$

where $A \in \underline{\Lambda^p(M)}$, $B \in \underline{\Lambda^q(M)}$, $C \in \underline{\Lambda(M)}$

Schouten's bracket on $\Lambda(M)$ is important because it entirely characterizes menifolds which have a Poisson bracket structure. Suppose Ω is an element of $\Lambda^2(M)$. Then we may certainly define a map from $F^0(M) \times F^0(M)$ to $F^0(M)$ by

$$(f,g) \to \Omega(df, dg) = \{f,g\}.$$

The map is evidently skew-symmetric and bilinear but it does not make $F^0(M)$ into a Lie algebra in general. In fact this happens, as Lichnerewicz showed, iff $[\Omega,\Omega] = 0$ where the bracket is Schouten's bracket on $F^0(M)$. One can easily see this from (2.1.9) because, using coordinates (x_i) on M

$$\{\{f,g\},h\} + \{\{g,h\},F\} + \{\{h,f\},g\} = \Omega^{i[j}\,\Omega^{kl]}_{,i}\frac{\partial f}{\partial x_j}\,\frac{\partial g}{\partial x_k}\,\frac{\partial h}{\partial x_l} \tag{2.1.12}$$

and the right hand side of (2.1.12) is just $[\Omega,\Omega]$ (df, dg, dh). Sometimes the tensor Ω is referred to as a cosymplectic structure [19]

and a manifold equipped with a cosymplectic structure is known as a Poisson manifold. Of course, a symplectic structure is a familiar, special case of a Poisson manifold. There is also another context in which Poisson manifolds arise naturally, namely, on the dual space of a finite dimensional Lie algebra and I consider these in section 2.3 after presenting some properties of Killing tensors in the next section.

2.2 Killing Tensors

Let (M,g) be a Riemannian or pseudo-Riemannian manifold. Let ∇ denote the metric connection and G denote the contravariant metric corresponding to g. A symmetric covariant n-tensor field K on M is said to be a (symmetric) Killing tensor if the symmetric part of the $n+1$ tensor ∇K vanishes. As is well known Nijenhuis [29], Woodhouse [43] this is equivalent to the condition that the symmetric Schouten bracket of K with G vanishes; in terms of local co-ordinates the index condition

$$K_{(i_1 \ldots i_n ; i_{n+1})} = 0. \qquad (2.2.1)$$

The space of all symmetric Killing tensors forms a subring under symmetrized tensor product, and a Lie subalgebra under the Schouten bracket of the collection of all contravariant tensor fields on M.

(2.2.1) is obviously a very natural generalization of the conditions for the existence of Killing vector fields whose importance has long been recognized in the context of general relativity. It is well-known that in flat spaces i.e. those for which the curvature tensor of g vanishes the number of Killing vectors is maximal. One is thus led to conjecture that something similar holds for Killing tensors and indeed this is the case. In the sequel K_n^m denotes the dimension of the vector space of analytic, symmetric rank n Killing tensors on (M,g) and K will denote a Killing tensor of rank n.

Proposition 2.1: If $K_{b_1\cdots b_n}$ is a Killing tensor in a flat space then $K_{b_1\cdots b_n, a_1\cdots a_{n+1}} \equiv 0$ i.e. the Killing tensors of rank n are polynomials of degree less than or equal to n.

Proof: Suppose as induction hypothesis that for $0 \leq k < n$

$$K_{b_1\cdots b_k}(a_1\cdots a_{n-k}, a_{n-k+1}\cdots a_{n+1}) = 0.$$

Then

$$\begin{aligned} 0 &= K_{b_1\cdots b_k}((a_1\cdots a_{n-k}, a_{n-k+1}\cdots a_{n+1})b_{k+1}) \\ &= \frac{n-k}{n+2} K_{b_1\cdots b_{k+1}}(a_1\cdots a_{n-k-1}, a_{n-k}\cdots a_{n+1}) \\ &\quad + \frac{k+2}{n+2} K_{b_1\cdots b_k}(a_1\cdots a_{n-k}, a_{n-k+1}\cdots a_{n+1})b_{k+1} \end{aligned}$$

From the induction hypothesis, it follows that the second term on the right is zero whence so is the first. The result now follows by induction which begins successfully because $K_{b_1\cdots b_n}$ is a Killing tensor. ∎

The next result gives an upper bound on K_n^m.

Theorem 2.2.2 $K_n^m \leq \frac{(m+n-1)!(m+n)!}{(m-1)!m!n!(n+1)!}$

Proof: Recall that the Killing tensors are assumed to be real analytic. The argument is pointwise at any point p of M so I shall refer to all derivatives and functions as being evaluated at p. It follows from (2.2.1) that

$$K_{(i_1\cdots i_n, i_{n+1}\ j_1\cdots j_r)} = B_{i_1\cdots i_{n+1}j_1\cdots j_r} \quad (2.2.2)$$

where the B's are functions which are linear combinations of derivatives of K of order less than or equal to r with coefficients which depend only on the metric and its derivatives. Before proceeding I show

that the functions $K_{i_1\ldots i_n, j_1\ldots j_n k}$ may be isolated and expressed as a linear combination of lower order derivatives. This follows from (2.2.2) with $r=n$ and $r=n-1$ and because

$$K_{(i_1\ldots i_n, j_1\ldots j_n k)} = K_{i_1\ldots(i_n, j_1\ldots j_n k)}$$

$$= \frac{1}{n+2} K_{i_1\ldots i_{n-1} k, i_n j_1\ldots j_n}$$

$$= \frac{n+1}{n+2} K_{(i_1\ldots i_{n-1} i_n, j_1\ldots j_n)k} \quad .$$

Thus all derivatives of order *greater* than n maybe recursively computed from lower derivatives. Now for $r = 0,1,\ldots,n$ (2.2.2) maybe regarded as a system of linear equations whose dependent variables are all the derivatives of orders between 1 and n+1. Starting from $r = n$ we may recursively compute all *lower* order derivatives. Linear algebra tells us that at each stage we pick up a number of free parameters which is precisely the number of $K_{i_1\ldots i_n, i_{n+1}\, j_1\ldots j_r}$ i.e. $\binom{m+r}{n+1}\binom{m+n-1}{n}$ less the number of independent equations contained in (2.2.2) which is $\binom{m+r-1}{r}\binom{m+n}{n+1}$. The recursion continues to $r = 0$; there remain, however, a further $\binom{m+n-1}{n}$ free parameters, the constant Killing tensors, which trivially satisfy (2.2.1) If the metric is not flat there will be further integrability conditions imposed by the expressions for $K_{i_1\ldots i_n;\ i_{n+1} j_1\ldots j_n}$ in terms of g and its derivatives. In flat space in view of proposition 2.2.1 these hold identically. Hence

$$K_n^m \leq \binom{m+n-1}{n} + \sum_{r=0}^{n} \binom{m+r}{r+1}\binom{m+n-1}{n} - \sum_{r=0}^{n}\binom{m+r-1}{r}\binom{m+n}{n+1}$$

$$= \frac{(m+n-1)!\,(m+n)!}{(m-1)!\,m!\,n!\,(n+1)!}$$ ∎

The preceding argument is a generalization of that used by Kalnins and Miller [23] to compute a bound for K_2^m. It also follows from the argument:

Corollary 2.2.3 If M is flat the dimension of the space of homogeneous, rank n Killing tensors of degree r denoted by $K_{n,r}^m$ is for $m \geq 2$

$$\binom{m+r-1}{r}\binom{m+n-1}{n} - \binom{m+r-2}{r-1}\binom{m+n}{n-1} = (n-r+1)\frac{(m+r-2)!\,(m+n-1)!}{r!\,(n+1)!\,(m-1)!\,(m-2)!}$$ ∎

The corollary only makes intrinsic sense for flat space. Again in [23] it is shown that spaces of non-zero constant curvature possess the maximum number of independent Killing vectors. For these spaces and flat spaces it is natural to conjecture that the higher rank Killing tensors are simply symmetrized products of Killing vectors. Unfortunately, to actually to compute the dimension of the space of symmetrized products of Killing vectors seems, at least to the present author, a rather forbidding combinatorial prbblem. However, I can confirm the conjecture in spaces of constant curvature in several cases: $n = 1$, $n = 2$ m arbitrary; $m = 2$, n arbitrary (in which case a basis for the Killing tensors is easily written down); for m,n arbitrary and $r = 0$ $K_{n,0}^m = \binom{m+n-1}{n}$ etc. For spaces other than those of constant curvature the analogous conjecture is false, and it was to investigate just this phenomenon that Hauser and Malhiot [17] initiated their program.

I end this section with a result which is sometimes useful.

Proposition 2.2.4 Let (M,g) be a flat pseudo-Riemannian manifold. Let $A_{a_1\dots a_n}$, $A_{a_1\dots a_{n-1}}$ be covariant, symmetric tensors of rank $n,n-1$ respectively. Suppose also that $A_{a_1\dots a_n}$ is Killing. Then $K_{a_1\dots a_{n-1}} = A_{a_1\dots a_{n-1}i,i} - A_{a_1\dots a_{n-1}}$ is Killing iff

$$nA_{(a_1\dots a_{n-1},a_n)} + A_{a_1\dots a_n,ii} = 0.$$

Proof: $K_{a_1\dots a_{n-1},a_n} = A_{a_1\dots a_{n-1}i,i\,a_n} - A_{a_1\dots a_{n-1},a_n}$

$$= K_{(a_1\dots a_{n-1},a_n)} = A_{i(a_1\dots a_{n-1},a_n)i} - A_{(a_1\dots a_{n-1},a_n)}.$$

Thus $K_{(a_1\dots a_{n-1},a_n)} = 0$ iff $A_{i(a_1\dots a_{n-1},a_n)i} - A_{(a_1\dots a_{n-1},a_n)} = 0.$

Since $A_{a_1\dots a_n}$ is Killing, it follows that

$$0 = \frac{n+1}{n} A_{(a_1\dots a_n,\,i)i} = A_{i(a_1\dots a_{n-1},a_n)i} + \frac{1}{n} A_{a_1\dots a_n,ii}. \qquad \blacksquare$$

2.3 The cosymplectic structure on the dual of a Lie algebra

In this section I shall show that if g is a finite dimensional Lie algebra then g^* has a natural cosymplectic structure. If $p \in g^*$ and $X \in T^*_p g^*$ then since g^* is, in particular, a vector space one has that $T^*_p g^* \cong g^{**}$. Since g is assumed to be finite dimensional $g^{**} \cong g$. Hence, a covector on g^* may be identified with an element of g. This enables the cosymplectic structure Ω of g^* to be defined by

$$\Omega(X,Y)_p = \langle [X,Y], p \rangle \qquad (2.3.1)$$

Here, on the left hand side X and Y are to be thought of as elements of $T^*_p g^*$ and on the right hand side as elements of g.

Theorem 2.3.1 Ω is a cosymplectic structure.

Proof: Let $F, G \in F^0(g^*)$ and let (x^i) and (p_j) be dual coordinate systems for g and g^* respectively. Then at each point of g^*, $(\frac{\partial}{\partial p_i})$ may be thought of as an element of g. Using (2.3.1) one can define a "Poisson bracket" as follows:

$$\{F,G\} = C^k_{ij}\, p_k \, \frac{\partial F}{\partial p_i} \cdot \frac{\partial G}{\partial p_j} \,. \qquad (2.3.2)$$

Here C^k_{ij} are the structure constants of g. Now if indeed (2.3.2) does yield a Poisson bracket structure, then by Lichnerowicz's characterization of Poisson manifolds [24], Ω must be a cosymplectic structure. Thus, it must be shown that Ω is a 2-vector and that the Jacobi identity holds. It is clear from (2.3.1) that Ω is alternating and $F^0(g^*)$ - linear and hence Ω is a 2-vector. Finally, if $H \in F^0(g^*)$, then

$$\{\{F,G\},H\} = C^n_{km} C^k_{ij} \frac{\partial F}{\partial p_i} \cdot \frac{\partial G}{\partial p_j} \cdot \frac{\partial H}{\partial p_m} + C^n_{lm} C^k_{ij} p_k p_n \left(\frac{\partial^2 F}{\partial p_i \partial p_l} \cdot \frac{\partial G}{\partial p_j} + \frac{\partial F}{\partial p_i} \cdot \frac{\partial^2 G}{\partial p_j \partial p_l} \right) \frac{\partial H}{\partial p_m} .$$

Hence $\{\{F,G\},H\} + \{\{G,H\},F\} + \{\{H,F\},G\}$

$$= (C^n_{km} C^k_{ij} + C^n_{ki} C^k_{jm} + C^n_{kj} C^k_{mi}) \frac{\partial F}{\partial p_i} \cdot \frac{\partial G}{\partial p_j} \cdot \frac{\partial H}{\partial p_m}$$

$$+ C^n_{lm} C^k_{kj} p_k p_n \left(\frac{\partial^2 F}{\partial p_i \partial p_l} \cdot \frac{\partial G}{\partial p_j} + \frac{\partial F}{\partial p_i} \cdot \frac{\partial^2 G}{\partial p_j \partial p_l} \right) \frac{\partial H}{\partial p_m}$$

$$+ C^n_{lm} C^k_{ij} p_k p_n \left(\frac{\partial^2 G}{\partial p_i \partial p_l} \cdot \frac{\partial H}{\partial p_j} + \frac{\partial G}{\partial p_i} \cdot \frac{\partial^2 H}{\partial p_j \partial p_l} \right) \frac{\partial F}{\partial p_m}$$

$$+ C^n_{lm} C^k_{ij} p_k p_n \left(\frac{\partial^2 H}{\partial p_i \partial p_l} \cdot \frac{\partial F}{\partial p_j} + \frac{\partial H}{\partial p_i} \cdot \frac{\partial^2 F}{\partial p_j \partial p_l} \right) \frac{\partial G}{\partial p_m} .$$

The first term of the right hand side above vanishes because of the Jacobi identity: the other terms may be grouped into three pairs of which one is

$$C^n_{lm} C^k_{ij} p_k p_n \frac{\partial^2 F}{\partial p_i \partial p_l} \cdot \frac{\partial G}{\partial p_j} \cdot \frac{\partial H}{\partial p_m} + C^n_{lm} C^k_{ij} p_k p_n \frac{\partial^2 F}{\partial p_j \partial p_l} \cdot \frac{\partial H}{\partial p_i} \cdot \frac{\partial G}{\partial p_m}$$

$$= (C^n_{lm} C^k_{ij} + C^n_{lj} C^k_{mi}) p_k p_n \frac{\partial^2 F}{\partial p_i \partial p_l} \cdot \frac{\partial G}{\partial p_j} \cdot \frac{\partial H}{\partial p_m} \qquad \text{(re-indexing)}$$

$$= (C^k_{im} C^n_{lj} + C^k_{ij} C^n_{ml}) p_k p_n \frac{\partial^2 F}{\partial p_l \partial p_i} \cdot \frac{\partial G}{\partial p_j} \cdot \frac{\partial H}{\partial p_m} .$$

Because of the skew-symmetry of the C^i_{jk}'s it follows that the preceding terms are each zero and so Jacobi is satisfied. ∎

In section 2.1 it was shown how the symmetric Schouten bracket enables $\underline{S(M)}$ to be viewed as a subalgebra of $F^0(T^*M)$ both with respect to the commutative and Lie algebra structures. A rather similar construction can

be made with S(g). Firstly, the bracket may be extended from g to S(g) by linearity, defining the bracket of any element of g with a scalar to be zero and the derivation rule

$$[A \otimes B, C] = [A,C] \otimes B + A \otimes B + A \otimes [B,C] \quad (A,B,C \in S(g)) \qquad (2.3.3)$$

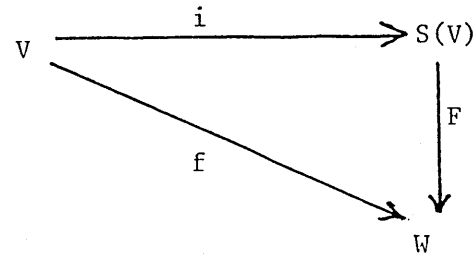
Notice that in this purely algebraic construction, one must define the bracket of elements of S(g) with scalars so that the extended bracket, being defined on elements of degree 0 and 1, is completely determined by the derivation property. The precise relationship of S(g) to $F^0(g^*)$ is given by the following result.

Proposition 2.3.2 S(g) is a subalgebra of both the commutative and Lie algebra structures of $F(g^*)$ - the two being related by (2.3.3).

Proof: As usual an element A of S(g) may be identified with a polynomial on g^* and with this identification, S(g) with symmetrized tensor product, is a subalgebra of $F^0(g^*)$. Moreover, in the identification just mentioned, scalars correspond to constant polynomials; this, together with (2.3.1) shows that the homomorphism property holds for elements of degree zero and one i.e. $[A,B] \to \{a,b\}$. But (2.3.3) and the analogous property for $F(g^*)$ imply now that the homomorphism property holds for elements of arbitrary degree. ∎

In his discussion of these topics, Hermann suggests another way of viewing S(g) as a Lie algebra [19]. To present this it is necessary to recall two standard algebraic ideas.

To begin with let V be a vector space neither necessarily finite dimensional nor of zero characteristic. Then the symmetric tensor algebra associated to V is defined as follows: it is a pair $(i,S(V))$ such that $i:V\to S(V)$ and if f is any linear map from V to W an associative algebra satisfying $f(x)f(y)=f(y)f(x)(x,y\in V)$ there exists an associative algebra homomorphism F from $S(V)$ to W such that



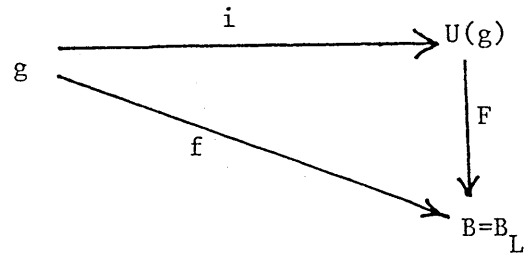
As is well known $S(V)$ is unique and may be realized concretely as the quotient of the contravariant tensor algebra $T(V)$ by the ideal generated by elements of the form $x\otimes y - y\otimes x \quad (x,y\in V)$.

I next turn to the universal enveloping algebra associated to a Lie algebra g. Firstly, however, it should be noted that whenever A is an associative algebra A may naturally be regarded as a Lie algebra by defining

$$[a,b] = ab-ba \qquad (a,b\in A).$$

In fact every Lie algebra arises as a subalgebra from a Lie algebra constructed in this way [22]. Whenever an associative algebra A is thus regarded as a Lie algebra I shall use the notation A_L. With this preliminary the universal enveloping associated to g may be defined as follows: Suppose that B is any associative algebra and f any Lie algebra homomorphism from g into B_L, the universal enveloping algebra associated to

g consists of a map i from g to an associative algebra U(g) such that there is an associative algebra homomorphism F from U(g) to B such that:



Here $B=B_L$ regarded as sets. Again the universal enveloping is unique and may be considered as the quotient of the tensor algebra C(g) by the ideal generated by all elements of the form $a\otimes b-b\otimes a-[a,b]$ $(a,b\in g)$.

Next, recall that an algebra A is said to be *graded* if $A=\sum_{i=0}^{\infty} A_i$ where each A_i is a subspace and moreover $A_iA_j\subset A_{i+j}$. Given such a graded algebra A, one can define $B_i=\sum_{j=0}^{i} A_i$ and then A may be viewed as a filtered algebra. An algebra B is said to be *filtered* if for any non-negative integer i there exists a subspace B_i such that (i) $B_i\subset B_j$ whenever $i\leq j$ (ii) $B=\bigcup_{i=0}^{\infty} B_i$ (iii) $B_iB_j\subset B_{i+j}$. Thus, it is possible to pass from a graded algebra to a filtered algebra. Conversely, starting from a filtered algebra B as above, one may define $A_i=B_i/B_{i-1}$, set $A=\sum_{i=0}^{\infty} A_i$ and define a multiplication in A by

$$(b_i+B_{i-1})(b_j+B_{j-1})=b_ib_j+B_{i+j-1}.$$

In view of (iii) above this multiplication is well defined and makes A into a graded algebra.

The considerations of the last paragraph may be applied to the universal enveloping algebra U(g) of a Lie algebra g. I denote the graded algebra corresponding to U(g) by $U^{GR}(g)$. The following important theorem holds:

THEOREM 2.5.1 S(g) and $U^{GR}(g)$ are isomorphic as graded algebras. ∎

For the proof of this result I refer to [20]. The result is not hard to believe when one considers the concrete characterizations of $S(g)$ and $U(g)$ given earlier. This result is essentially the Poincaré-Birkhoff-Witt theorem and its proof is a tour de force in the indexing of monomials [20, 22].

In addition to their Lie structures, $S(g)$ and $U^{GR}(g)$ also have commutative algebra structures and by the very construction of $U^{GR}(g)$ these are isomorphic.

CHAPTER 3

CONSTANTS OF MOTION IN HAMILTONIAN MECHANICS

This chapter is devoted to the Hamiltonian approach to mechanics. I begin in section 3.1 with several new results which ensure that Hamiltonians of various types have particular integrals of motion. In section 3.2 I give a modern treatment of two familiar, classical topics - Noether's theorem and Hamilton-Jacobi theory; whilst the results given in this section are not new, they are of obvious importance for the existence of first integrals in the Hamiltonian context. In section 3.3 conditions are obtained which entail the existence of an integral of motion which is polynomial in momenta for a Hamiltonian of standard mechanical type. These conditions are formulated, firstly invariantly, and then with coordinates using the machinery of chapter 2. In section 3.4 an explicit example is worked out which demonstrates how the conditions obtained in the preceding section may be used in practice to obtain a system with an integral of motion of degree four. In section 3.5 the results of several calculations akin to those of section 3.4 are given. Most importantly, it is shown that two of the results from section 3.1 entirely characterize all those Hamiltonians with two degrees of freedom which have an integral of motion quadratic in the momenta - a problem considered and partially resolved by Darboux.

3.1 Some general results on integrals of motion

The results which I give here belong properly to the realm of symplectic geometry. N thus throughout denotes a symplectic manifold and $\{\ ,\ \}$ denotes the Poisson bracket on $F^o(N)$ the ring of functions on N.

Proposition 3.1.1 Suppose that H is the Hamiltonian of a system and that

$$H = f(A,B_1,\ldots,B_r)$$

where $A,B_1,\ldots,B_r \in F^o(N)$ and f is a function of the r+1 arguments indicated. Suppose also that $\{A,B_i\} = 0 \quad (1 \leq i \leq r)$. Then $\{H,A\} = 0$.

The proof of proposition 3.1.1 is trivial from the properties of $\{\ ,\ \}$, and though it may seem innocuous it can sometimes yield useful results. The next result has been given before [39] but I shall now expand upon it considerably. Again the proof is straightforward using the derivation properties of $\{\ ,\ \}$.

Proposition 3.1.2 Suppose that $H,A,B,P,Q \in F^o(N)$ and that

$$H = \frac{A+B}{P+Q}$$

where $\{A,B\} = \{P,Q\} = \{A,Q\} = \{P,B\} = 0$. Then $\{H, \frac{AQ-BP}{P+Q}\} = 0$.

The last result leads immediately to the following, the proof being similar.

Proposition 3.1.3 Suppose that $H,A_i,P_j \in F^o(N) \quad (1 \leq i,j \leq r)$ and that

$$H = \frac{A_1+A_2+\ldots+A_r}{P_1+P_2+\ldots+P_r} \qquad \text{where}$$

$\{A_i,A_j\} = \{P_i,P_j\} = \{A_i,P_j\} = 0 \quad (i \neq j)$.

Then $\dfrac{A_1(P_2+\ldots+P_r)-P_1(A_2+\ldots A_r)}{P_1+P_2+\ldots+P_r}$, $\dfrac{A_2(P_3+\ldots+P_r+P_1)-P_2(A_3+\ldots+A_r+A_1)}{P_1+P_2+\ldots+P_r}$

- $\ldots$, $\dfrac{A_r(P_1+\ldots+P_{r-1})-P_r(A_1+\ldots+A_{r-1})}{P_1+P_2+\ldots+P_r}$

are r integrals of motion for H which themselves mutually commute. In particular, if $r=\frac{1}{2}\dim(N)$ and these integrals are independent, the system determined by H is completely integrable in the sense of Louiville's theorem.

The last two results seem very closely related to some classical results of Louiville (see Whittaker [42]). Also, one could write down more integrals by using proposition 3.1.3 and choosing, for example, $A=A_1+A_2$, $B=A_3+\ldots+A_r$ $P=P_1+P_2$, $Q=P_3+\ldots+P_r$ etc. The preceding results are valid for arbitrary symplectic manifolds. By contrast, the next result and its corollary hold in the case that $N=T^*M$ or, more generally, on an exact symplectic manifold. In the former case, denoting the canonical 1-form by θ and the canonical radial vector field by Δ, it is easily shown [40] that for any $f\in F^0(T^*M)$ $\Delta f=\langle X_f,\theta\rangle$; here X_f is the Hamiltonian vector field associated to f. Hence, f is homogeneous of degree s in the fibers iff $\langle X_f,\theta\rangle = sf$. Recall that the next propostion and corollary apply to exact symplectic manifolds.

Proposition 3.1.4. For two Hamiltonians H,K the conditions $\{H,K\} = 0$ and $\{\langle X_H,\theta\rangle,K\} = 0$ imply that $\{H,\langle X_K,\theta\rangle\} = 0$.

Proof: $\{H,\langle X_K,\theta\rangle\} = X_H\langle X_K,\theta\rangle$

$= \langle[X_H,X_K],\theta\rangle + X_K\langle X_H,\theta\rangle + d\theta(X_H,X_K)$

$= \langle X_{\{H,K\}},\theta\rangle - \{\langle X_H,\theta\rangle,K\} + \{H,K\}$

$= 0 + 0 + 0.$ ■

Corollary 3.1.5 For two Hamiltonians H,K the conditions $\{H,K\}=0$ and $\langle X_H,\theta\rangle=sH$ imply that $\{H,\langle X_K,\theta\rangle\}=0$. ∎ The use of this result is that starting from a given homogeneous Hamiltonian and one first integral for the corresponding Hamiltonian vector field one can generate new integrals. A version of it is to be found in Lie [25].

I give one final result which is valid only for $\mathbb{R}^{2m}$.

Proposition 3.1.6 Let $H:T^*(\mathbb{R}^m)\simeq\mathbb{R}^{2m}\to\mathbb{R}$ be a Hamiltonian given by

$$H=\tfrac{1}{2}p\cdot p+e(\underline{x})+f(x)$$

where f is an arbitrary function of $x=(\underline{x}\cdot\underline{x})^{\frac{1}{2}}$ and e satisfies

$$\underline{x}\cdot\mathrm{grad}(e)+2e=0$$

i.e. e is homogeneous of degree minus two. Then the function E is a constant of motion where

$$E=x^2p^2-(\underline{x}\cdot\underline{p})^2+2\underline{x}\cdot\underline{x}e \quad .$$ ∎

Again I forgo the proof of this result which is an unenlightening computation. However, notice that the term which is quadratic in the momenta in E corresponds to a rank two Killing tensor of the (co-) metric δ_{ij} - an observation which considerably eases the proof. Moreover the result makes essential use of the linear and metrical properties of $\mathbb{R}^m$ and would not make sense on a general T^*M.

3.2 Hamilton-Jacobi theory and Noether's theorem

In this section I consider two topics which are important as far as constants of motion are concerned. Presently, I shall discuss Noether's theorem but begin with an examination of Hamilton-Jacobi theory. As the name implies this has had a long history, being the most powerful tool available for the explicit integration of systems in classical mechanics. Today it also has considerable theoretical importance, and the equation known as the Hamilton-Jacobi equation, plays an analagous role in classical mechanics to that played by the Schrödinger equation in quantum theory, at least so folklore has it.

To formulate Hamilton-Jacobi theory in modern language we need one preliminary concept.

Definition 3.2.1 Let (N,ω) be a $2m$-dimensional symplectic manifold. An m-dimensional integral manifold of ω (i.e. a submanifold of N such that ω pulled back to it vanishes) is called a Lagrangian submanifold.

It is a standard result that Lagrangian submanifolds are maximal in the sense that the pullback of ω cannot vanish on any submanifold of dimension greater than m [41]. To describe all the Lagrangian submanifolds of a symplectic manifold is a delicate business. Certainly on T^*M any fiber is a Lagrangian submanifold as too is the zero section. Moreover, one has

Propostion 3.2.2 On T^*M the graph of a 1-form ϕ (i.e. the m-dimensional submanifold of T^*M its image describes) is a Lagrangian submanifold iff $d\phi = 0$.

Proof: It follows from the universal property of θ that

$$\phi^*\theta = \phi.$$

Hence

$$\phi^*d\theta = d\phi \qquad \text{and the result follows.} \quad \blacksquare$$

I explained how, in section 1.9 a choice of a function h on T^*M gives a vector field X_h on T^*M. Using the map γ of section 1.7, h may be pulled back to give a function on $J^1(M,\mathbb{R})$. If one now chooses a level of h and hence a level of γ^*h one obtains, at least locally, codimension one submanifolds of T^*M and $J^1(M,\mathbb{R})$ respectively. However, the latter is precisely what one means by a first order partial differential equation (P.D.E. 1.0) with one dependent variable. In fact there is no need to work on $J^1(M,\mathbb{R})$. Codimension one submanifolds of T^*M are P.D.E. 1.0's in which the dependent variable does not occur explicitly. Since h is a function on T^*M, the different levels of h form, at least locally, a codimension one foliation of T^*M, and each leaf may be interpreted as a P.D.E. 1.0. The main idea behind Hamilton-Jacobi theory is that this foliation and the vector field X_h interact nicely; for example, the conservation of energy law $X_h h = 0$ is a consequence of the fact that X_h is tangent to the levels of h.

In the classical literature one often runs across the phrase "complete solution of a P.D.E. 1.0" (see for example [6, 25]). Roughly speaking, a complete solution is one which depends on m arbitrary constants. More precisely, if the P.D.E. is described

locally by $h = 0$, a complete solution is a foliation of T^*M by Lagrangian submanifolds, each one of which is tangent to some level of h. Actually, this definition allows a more generous notion of solution since it allows the possibility of many-valued solutions; also, although stated globally the idea is mainly a local one. From the definition of complete solution there are m (local) functions on T^*M which parametrize the leaves of the foliation. I call these a_i $(1 \leq i \leq m)$ so that if (x_i) are coordinates on M, (a_i, x_i) are coordinates for T^*M. I now give the statement of the Hamilton-Jacobi theorem.

THEOREM 3.2.3 Under the circumstances just described i.e. given a function h on T^*M and a complete solution in which the leaves are parametrized by the a_i's then

(i) the a_i's are constants of motion for X_h

(ii) the a_i's are in involution i.e. $\{a_i, a_j\} = 0$. ∎

If the conditions of the theorem are met, then the system determined by h is, according to the modern usage "completely integrable" . Unfortunately this is not the same thing as in the Frobenius theorem and the classical authors said simply "integrable".

I now turn to the proof of the theorem which depends on a relatively simple lemma from linear symplectic geometry.

Lemma 3.2.4 Let (V, Ω) be a symplectic vector space and W a hyperplane. Then there is a unique 1-dimensional subspace A of V characterized by the property that it is the collection of vectors ξ such that $\Omega(\xi, W) = 0$. Moreover, if Λ is a Lagrangian plane i.e. a linear subspace

which is a Lagrangian submanifold then $A \subset \Lambda$.

Proof: To show existence, let α be a covector whose kernel is W. Define ξ by $\xi \rfloor \Omega = \alpha$. Then

$$\Omega(\xi,W) = \langle W, \xi \rfloor \Omega \rangle = \langle W, \alpha \rangle = 0$$

since W is the kernel of α. To show uniqueness, suppose that $\xi \in V$ satisfies $\Omega(\xi,W) = 0$. Then $\langle W, \xi \rfloor \Omega \rangle = 0$ and hence $W = \ker(\xi \rfloor \Omega)$. Since Ω is nondegenerate, one must have that $W = \ker(\xi \rfloor \Omega) = \ker \alpha$. This means that α and $\xi \rfloor \Omega$ differ only by a non-zero factor; the uniqueness of A follows. To show that $A \subset \Lambda$ consider the subspace spanned by A and Λ. Clearly $\Omega(A,A) = 0$, and since Λ is Lagrangian $\Omega(\Lambda,\Lambda) = 0$; also $\Omega(\Lambda,A) = 0$ by the characterization of A and the fact that $\Lambda \subset W$. Hence Ω vanishes on the subspace spanned by A and Λ but by the maximality of Λ we have $A \subset \Lambda$. ∎

Now we can prove theorem 3.2.3.

Proof of theorem 3.2.3:

(i) It follows from the definitions and lemma 3.2.4 applied in the tangent space that X_h is tangent to all the submanifolds $a_1 = c_1$, $a_2 = c_2 \ldots$, $a_m = c_m$. Hence X_h is also tangent to the submanifold $a_1 = c_1$ i.e. $X_h a_1 = 0$ and likewise for the remaining a_i's.

(ii) Again, from lemma 3.2.2 it follows that X_{a_i} is tangent to all the Lagrangian submanifolds. Hence X_{a_i} is also tangent to the bigger submanifolds $a_j = c_j$ i.e. $X_{a_i} a_j = 0$ or $\{a_i, a_j\} = 0$ as claimed. ∎

One may consider Hamilton-Jacobi theory for simultaneous P.D.E's $h_1 = 0,\ldots,h_q = 0$ in which a complete solution is a foliation of T*M by Lagrangian submanifolds each one of which lies in some level of h_i for each $1 \leq i \leq q$.

Proposition 3.2.5 The simultaneous P.D.E's $h_1 = 0,\ldots,h_q = 0$ have a complete solution only if $\{h_i,h_j\} = 0$.

Proof: Picking a level $h_i = c_i$ we know from lemma 3.2.4 that X_{h_j} is tangent to the Lagrangian submanifold lying in $h_i = c_i$ hence tangent to $h_i = c$ itself. ∎

It is possible that non-involutive functions may possess isolated solutions even though they cannot possess a complete solution. Also, this idea of involution is the simplest case of a far more general notion in P.D.E. theory; for example see Cartan [4].

The usefulness of these ideas is that when confronted with some Hamiltonian vector field which had to be integrated, the classical authors found that in practice they could often obtain complete solutions to the H-J equation and hence they had a maximal collection of independent pairwise involutive first integrals, namely, the a_i's of theorem 3.2.3. These a_i's also give rise to the vector fields $\{a_i, \}$ and since $\{a_i,a_j\} = 0$ these Lie-commute and so form a completely integrable system of vector fields in the Frobenius sense! This system can thus be integrated, and we

may suppose that $\{a_i, \} = \frac{\partial}{\partial b_i}$ $(1 \leq i \leq m)$ where the b_i's are functions on T*M. Hence, at least locally $d\theta = da_i \wedge db_i$. Another point to note in theorem 3.2.3 is that h is a function of the a_i's i.e. $dh \wedge da_1 \wedge \ldots \wedge da_m = 0$; this follows from the hypothesis that the leaves of the complete solution all lie in the levels of h. This then leads to a more familiar looking version of Hamilton-Jacobi theory (see Goldstein [16] or Arnold [1]). On p. 260 of the latter reference the Jacobi theorem is stated. Essentially, this just comes down to the fact that since H is a function of the a_i's, Hamilton's equations

$$\frac{da_i}{dt} = \frac{\partial h}{\partial b_i} = 0, \qquad \frac{db_i}{dt} = \frac{\partial h}{\partial a_i}$$

are trivial to integrate.

I next turn to Noether's theorem. This theorem was originally presented in terms of the invariance of the Lagrangian of a mechanical system [30]. There is a very closely related result based not on the action of a system but rather on its Cartan 1-form and the relationship between this result and Noether's theorem will be investigated in some detail in section 4.5. The former result has a direct analogue in Hamiltonian mechanics, and this is the version of Noether's theorem considered here. One of the main points that I wish to make here is that both this version of Noether's theorem and Hamilton-Jacobi theory appear as very natural consequences of the geometric formulation of Hamiltonian mechanics.

THEOREM 3.2.6 Let (N,ω) be a symplectic manifold and h be a function on it. Then

(1) If f is constant on the flow of the vector field X_h defined by $X_h \lrcorner \omega = -dh$, the vector field X_f defined by $X_f \lrcorner \omega = -df$ satisfies

$$X_f h = 0 \quad \text{and} \quad L_{X_f}\omega = 0 \ .$$

(2) Conversely, if a vector field X on T^*M satisfies

$$Xh = 0 \quad \text{and} \quad L_X \omega = 0$$

then, locally, a function f determined by $X \lrcorner \omega = -df$ is constant on the trajectories of X_h.

Proof: I shall prove only (2), since (1) is the reverse argument which, however, works regardless of the topology of N since any exact form is closed but not necessarily vice-versa. Thus

$$0 = L_X\omega = X \lrcorner d\omega + d(X \lrcorner \omega)$$

$$\Rightarrow d(X \lrcorner \omega) = 0$$

and so, at least locally, there is a function f on N such that

$$X \lrcorner \omega = -df$$

Now

$$\begin{aligned} X_h f &= X_h \lrcorner df \\ &= -X_h \lrcorner X_f \lrcorner \omega \\ &= X_f \lrcorner X_h \lrcorner \omega \\ &= -X_f \lrcorner dh \\ &= -X_f h \\ &= 0. \end{aligned}$$

■

3.3 Constants of the motion polynomial in momenta

I shall refer to theorem 3.2.6 as the "general" Noether theorem in the Hamiltonian context; in particular, it applies in the case where N is the cotangent bundle of some manifold, say M. In this case, any symmetric tensor A of rank n on M gives rise to a function a on T^*M which is a homogeneous polynomial of rank n in the fibers: one simply defines, for $p_x \in T^*M$ with $\pi(p) = x$.

$$a(p) = A(x)(p,\ldots,p)$$

there being n arguments on the right hand side. More generally, a collection of symmetric tensors of various ranks may be used to define a function on T^*M which is an inhomogeneous polynomial. I shall write

$$H = A_0 + A_1 + A_2 + \ldots$$

where the A_i's are tensors of rank i respectively. According to the convention used above

$$h = a_0 + a_1 + a_2 + \ldots$$

is the corresponding function on T^*M.

Suppose that Y is a vector field on M. Then it is not hard to show that there is a unique vector field X on T^*M with the properties that

(i) $\pi_* X = Y$

(ii) $L_X \theta = 0$.

In fact, if (x^i) is a coordinate system on M and (x^i, p_i) the induced coordinates on T^*M and $Y = Y^i \frac{\partial}{\partial x^i}$ then $X = Y^i \frac{\partial}{\partial x^i} - p_i \frac{\partial Y^j}{\partial x^i} \frac{\partial}{\partial p_j}$.

The foregoing considerations lead to the following result which I refer to as the "special" Noether theorem.

Theorem 3.3.1 Let X be a vector field on T^*M which satisfies (i) and (ii) above with Y a vector field on M. Suppose also that $h:T^*M \to \mathbb{R}$ and that $Xh=0$. Then Y regarded as a function on T^*M is a first integral for X.

Proof: The result is merely a specialization of theorem 3.2.6 (2). For, clearly X satisfies $Xh=0$ and $L_X\omega=0$ where $\omega=d\theta$. Thus, by the theorem there is a first integral associated to h, and since X has the local expression $Y^i\frac{\partial}{\partial x^i} - p_i\frac{\partial Y^j}{\partial x^i}\frac{\partial}{\partial p_j}$, it follows that this integral is $Y^i p_i$ i.e. Y regarded as a function on T^*M. ∎

In case h is of the form considered previously i.e. $h=a_0+a_1+a_2+\ldots$ the condition $Xh=0$ is equivalent to the conditions $[Y,A_i]=0$ $(0\leq i\leq n)$. This will be investigated more closely presently but for now I will further specialize the last result to obtain what I shall refer to as the "classical" Noether theorem.

Theorem 3.3.2 Let G be a rank two symmetric contravariant tensor field i.e. a cometric on M and V be a function on M. Define $H=\frac{1}{2}G+V$ and let h denote the corresponding function on T^*M. Suppose that Y is a vector field on M satisfying

(i) $L_Y G = 0$

(ii) $YV=0$

Then Y regarded as a function on T^*M is a first integral of X_h. ∎

These last two results are particular instances of an even more general result. In fact notice that since a not necessarily homogeneous, contravariant, symmetric tensor A on M determines a function a on T^*M, it also determines a vector field X_a on T^*M. Since X_a is Hamiltonian it obviously satisfies $L_{X_a}\, d\theta = 0$, so one obtains immediately

the following result which is a direct generalization of theorem 3.3.1 and whose proof is similar.

THEOREM 3.3.3 Let $h : T^*M \to \mathbb{R}$ and suppose that A is a not necessarily homogeneous contravariant, symmetric tensor on M such that $X_a h = 0$ where a is the function on T^*M associated to A. Then a is a first integral of X_h. ∎

This is an appropriate point to give some attention to the question of "obvious" versus "hidden" symmetries. It would seem natural to call a first integral arising from theorem 3.3.1 as "obvious". Indeed, in that case the symmetry vector field associated via the general Noether theorem is the natural lift of a vector field on M. In all other cases, the symmetry vector fields on T^*M will not be projectable to vector fields on M and these may therefore be called "hidden". Furthermore, when h itself has the standard form of a classical mechanical system, the conditions satisfied by these obvious symmetries i.e. (i) and (ii) of theorem 3.3.2 are entirely determined by the geometry of M.

As we have seen, a function on T^*M which is polynomial in the fibers may be identified with contravariant, inhomogeneous, symmetric tensor fields on M. Thus, one may consider a Hamiltonian h where

$$h=h_0+h_1+h_2+\ldots+h_s$$

and $h_0,h_1,\ldots,h_s$ are the functions of degree $0,1,\ldots,s$ in the fiber corresponding to some tensors $H_0,H_1,\ldots,H_s$ of rank $0,1,\ldots,s$ respectively. Similarly, one may consider a function a where

$$a = a_0+a_1+a_2+\ldots+a_n$$

and $a_0, a_1, \ldots, a_n$ are the functions of degree $0,1,\ldots,n$ respectively. The conditions that h and a Poisson commute are obtained from theorem 3.2.6 and equating to zero terms of each grade i.e.

$$\begin{aligned} [H_0, A_0] &= 0 \\ [H_0, A_1] + [H_1, A_0] &= 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ [H_{s-1}, A_n] + [H_s, A_{n-1}] &= 0 \\ [H_s, A_n] &= 0 \end{aligned} \qquad (3.3.1)$$

Conditions (3.3.1) may be obtained in coordinates in the special case where h is a classical Hamiltonian.

Consider a standard Hamiltonian of classical mechanics

$$h = \tfrac{1}{2} p_j p_j + v(x_i)$$

where (x_i, p_j) is a coordinate system. Suppose that f is a constant of motion for the system determined by h and that

$$f = A_{a_1 \ldots a_n} p_{a_1} \ldots p_{a_n} + A_{a_1 \ldots a_{n-1}} p_{a_1} \ldots p_{a_{n-1}} + \ldots A_{a_1} p_{a_1} + A$$

where $A_{a_1 \ldots a_n}$, $A_{a_1 \ldots a_{n-1}}, \ldots, A_{a_1}$, A are symmetric tensors of rank $n, n-1, \ldots, 1, 0$ respectively. Conditions (3.3.1) in that case are easily seen to be

$$\begin{aligned} A_{(a_1 \ldots a_n, a_{n+1})} &= 0 \\ A_{(a_1 \ldots a_{n-1}, a_n)} &= 0 \\ A_{(a_1 \ldots a_{n-2}, a_{n-1})} &= n V_{,i} A_{a_1 \ldots a_{n-1} i} \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ A_{,a_1} &= 2 V_{,i} A_{a_1 i} \\ 0 &= V_{,i} A_i \end{aligned} \qquad (3.3.2)$$

Several remarks can be made about (3.3.2). Firstly, if V is itself a polynomial in co-ordinates then so too is f. Secondly, the alternate equations of (3.3.2) decouple into two sets, and so it suffices to look for constants of purely odd and purely even degrees. Thirdly, the first two equations of (3.3.2) define $A_{a_1 \ldots a_m}$ and $A_{a_1 \ldots a_{n-1}}$ as Killing tensors of the metric δ_{ij}. Recall from proposition 2.2.1 that $A_{a_1 \ldots a_n}$ must then be polynomials of degree at most n.

Returning to 3.3.2 it is clear that linear integrals are just the obvious symmetries of theorems 3.3.1 and 3.3.2. I shall be interested here in hidden symmetries rather than obvious ones. Quadratic integrals, which from the preceding comments may be taken in the form $A_{ij}p_ip_j + A$, correspond to rank two Killing tensors which also satisfy the conditions

$$A_{,i} = 2V_{,j}A_{ij}. \qquad (3.3.3)$$

A may be eliminated from these conditions leaving several *linear* second order partial differential equations to be satisfied by V. For the case of two degrees of freedom conditions (3.3.3) reduce to a single independent condition which is analyzed later in example 3 of section 3.5. More generally, when using (3.3.2) to detect polynomial integrals of odd or even degree the second highest degree term is always subject to some linear equations which also involve the components of the Killing tensors; these may be obtained, at least in theory, by differentiating the conditions

$$A_{(a_1 \ldots a_{n-2}, a_{n-1})} = nV_{,i}A_{a_1 \ldots a_{n-1}i}$$

enough times so as to be able to eliminate the $A_{a_1 \ldots a_{n-2}}$ components. However, for degree three constants or higher there will also enter

nonlinear equations due to the other conditions in (3.3.2), which makes the problem of finding such integrals much more complicated. Example 6 in section 5 gives an example of such a complication.

3.4 A specific example

I shall next give a rather detailed example of how equations (3.3.2) may be actually be used in practice. Referring to section 3.3, I suppose that $m=2$ and write $x=x_1, y=x_2, p_x=p_1, p_y=p_2$. I also assume that $V=V(x-y)$ and so the quantity $M=\frac{1}{2}(p_x+p_y)$ is a constant of motion by Noether's theorem. Besides M and the Hamiltonian H there must be one more functionally independent integral depending on (x,y,p_x,p_y), and one may ask whether this third integral is polynomial in momenta. It is quite straightforward to show that if the polynomial has degree two or three, then up to various inessential additive and multiplicative constants

$$V = x-y \qquad \text{or} \qquad V = \frac{1}{(x-y)^2} \tag{3.4.1}$$

Now suppose that f is a degree four integral. By considering the sequence $f,\{f,M\}, \{\{f,M\},M\}$ .... it is clear that it is sufficient to look for an f whose Killing components are of degree less than or equal to one in x and y. Now $\{f,M\} \neq 0$ otherwise there would be three mutually commmuting integrals $\{H,M,f\}$ which would force f to be dependent on H and M. More generally,one may refer to a polynomial integral as trivial if it can be obtained from polynomial combinations of constants of lower degree. In two dimensions it is certainly true that the Killing tensors are generated by the Killing vectors as was explained in section 2.2. Hence, it is sufficient to take the degree four term of the integral as

$$\begin{aligned} A_{ijkl}p_ip_jp_kp_l &= 4(yp_x - xp_y)(D_1p_x^3 - 3D_3p_x^2p_y + 3D_4p_xp_y^2 - D_2p_y^3) \\ &+ E_1p_x^4 + 4E_3p_x^3p_y + 6E_5p_x^2p_y^2 + 4E_4p_xp_y^3 + E_2p_y^4 \end{aligned} \tag{3.4.2}$$

where the D's and E's are constants.

Now applying conditions (3.3.2) one obtains the following system of equations where ' denotes differentiation with respect to the variable x-y.

$$\begin{aligned} A_{11,1} &= 4V'(A_{1111} - A_{1112}) \\ A_{11,2} + 2A_{12,1} &= 12V'(A_{1112} - A_{1122}) \\ A_{22,1} + 2A_{12,2} &= 12V'(A_{1122} - A_{1222}) \\ A_{22,2} &= 4V'(A_{1222} - A_{2222}) \end{aligned} \tag{3.4.3}$$

and

$$\begin{aligned} A_{,1} &= 2V'(A_{11} - A_{12}) \\ A_{,2} &= 2V'(A_{12} - A_{22}) \end{aligned} \tag{3.4.4}$$

(3.4.3) yields partial integrability conditions on the A_{ij}: use the first pair and last pair to obtain, after differentation, expressions for $A_{12,11}$ and $A_{12,22}$. Then demand that $A_{12,1122} = A_{12,2211}$. One obtains

$$15(D_1+D_2-D_3-D_4)V''' + (2D_1(x-2y)+2D_2(2x-y)+6D_3y-6D_4x-E_1+E_2-2E_3+2E_4)V'''' = 0. \tag{3.4.5}$$

It follows since $V = V(x-y)$ that

$$15(D_1+D_2-D_3-D_4)V''' + (3(D_1+D_2-D_3-D_4)(x-y)-E_1+E_2-2E_3+2E_4)V''''' = 0 \tag{3.4.6}$$

and $$(D_1-D_2-D_3+D_4)V''''' = 0 \tag{3.4.7}$$

or else $$D_1+D_2-D_3-D_4=D_1-D_2-D_3+D_4=E_1-E_2+2E_3-2E_4 = 0 \tag{3.4.8}$$

If (3.4.6) and (3.4.7) hold, V is of the form

$$V = K(x-y) \qquad \text{or} \qquad V = Q(x-y) + \frac{q}{(x-y)^2} \tag{3.4.9}$$

where K is a cubic polynomial, Q a quadratic polynomial and q some constant. Next, using (3.4.4) one obtains the further condition:

$$(A_{11}-A_{22})V''' = (A_{11,2}-A_{12,1}-A_{12,2}+A_{22,1})V' \qquad (3.4.10)$$

It follows from either of conditions (3.4.9) together with (3.4.10) that V has one of the forms given by (3.4.1) and that f is necessarily trivial.

I next turn to the other alternative i.e. that (3.4.8) holds. Using the fact that M^4, HM^2, H^2 are all polynomial integrals one may further suppose that either

$$E_1 = E_2 = E_3 = E_4 = E_5 \qquad (3.4.11)$$

or

$$D_1 = -D_2 = 1, \quad E_1 = E_2 = E_3 = E_4 = E_5 = 0. \qquad (3.4.12)$$

(3.4.11) leads once again to $V = x-y$. (3.4.12) in conjunction with (3.4.10) leads to the following condition where $W' = V$

$$WW''' + 3(x-y)W'W'' + 3(x-y)W''W'' + 12W'W'' = 0. \qquad (3.4.13).$$

Besides the solutions equivalent to (3.4.1), (3.4.13) gives a third possibility i.e. that $V = \dfrac{1}{(x-y)^{2/3}}$. Thus the Hamiltonian given by

$$H = \tfrac{1}{2}(p_x^2 + p_y^2) + \frac{1}{(x-y)^{2/3}}$$

has the quartic integral f given by

$$f = 4(yp_x - xp_y)(p_x - p_y)^3 + \frac{8(p_x-p_y)((yp_x-xp_y)-(x-y)(p_x+p_y))}{(x-y)^{2/3}} + \frac{32(x+y)}{(x-y)^{4/3}}$$

Moreover, this is essentially the only system which admits a non-trivial quartic polynomial.

The result may be generalized as follows. Define the Hamiltonians

H_k $k=0,1,2,\ldots$ by

$$H_k = \tfrac{1}{2}(p_x^2+p_y^2) + \frac{1}{(x-y)^{\frac{2}{2k+1}}} .$$

Then f_k is an integral of degree $2(k+1)$ where

$$\frac{1}{16} f_k = (p_x+p_y)\left(\frac{1}{2^{2k+1}}\binom{k}{0}\frac{2k+1}{2k+1}(x-y)^{\frac{2k+1}{2k+1}}(p_x-p_y)^{2k+1} + \frac{1}{2^{2k-1}}\binom{k}{1}\frac{2k+1}{2k-1}(x-y)^{\frac{2k-1}{2k+1}}(p_x-p_y)^{2k-1} + \cdots + \tfrac{1}{2}\binom{k}{k}\frac{(2k+1)}{1}(x-y)^{\frac{1}{2k+1}}(p_x-p_y)\right) - 2(x+y)\left(\tfrac{1}{4}(p_x-p_y)^2 + \frac{1}{(x-y)^{\frac{2}{2k+1}}}\right)^{k+1} .$$

3.5 Other Examples

1. Consider the system with m degrees of freedom whose Hamiltonian is given by

$$h = \tfrac{1}{2}p_i p_i + e$$

where e is a homogeneous function of degree minus two and also $\{\sum_{j=1}^{m} p_j, e\} = 0$. This system is a variation of the Calogero system [3,28,29,44] The following integrals were found using (3.3.2)

$$E = x_i x_i p_j p_j - x_i x_j p_i p_j + 2x_i x_i e$$

$$F = (\sum_{i=1}^{m} x_i) p_j p_j - x_j p_j (\sum_{i=1}^{m} p_i) + 2(\sum_{i=1}^{m} x_i) e \quad .$$

2. As another variation on the Calogero system consider the system with m degrees of freedom whose Hamiltonian is given by

$$h = \tfrac{1}{2}p_i p_i + e + f(x_i x_i)$$

where e is a homogeneous function of degree minus two, $\{\sum_{j=1}^{m} p_j, e\} = 0$ and f is any function of $x = (\underline{x}.\underline{x})^{\frac{1}{2}}$. This time one has the following integrals

$$E = x_i x_i p_j p_j - x_i x_j p_i p_j + 2x_i x_i e$$

$$D = (\Sigma_i(x_i)p_j - x_j \Sigma_i(p_i))(\Sigma_i(x_i)p_j - x_j \Sigma_i(p_i)) + 2(\Sigma x_i)^2(e+f).$$

so that for $m = 3$ this system is completely integrable, Since $\{D,E\} = 0$.

3. In this example I consider a system whose Hamiltonian h is given by

$$h = \tfrac{1}{2}p_i p_i + V(x_j)$$

and enquire for which V there exists a constant of motion which is quadratic in the momenta. The case of two degrees of freedom has been

;cussed several times but never completely resolved [10,26,42]. Resuming from
3.3) one has

$$A_{,i} = 2V_{,j}A_{ij}. \qquad (3.5.1)$$

:egrability conditions on A yield

$$A_{ij}V_{,jk} - A_{kj}V_{,ji} + V_{,j}(A_{ij,k} - A_{kj,i}) = 0 \quad . \qquad (3.5.2)$$

/ assume $m = 2$. Then from section 2.2 above for some constants $a, b_1, b_2, c_1, c_2, c_3$
: obtains from (3.5.2)

$$(axy + b_1x + b_2y - c_3)(V_{xx} - V_{yy}) - (a(x^2 - y^2) + 2b_2x - 2b_1y - c_1 + c_2)V_{xy}$$

$$+ 3(ay + b_1)V_x - 3(ax + b_2)V_y = 0 \quad . \qquad (3.5.3)$$

By performing canonical transformations which leave invariant $p_x^2 + p_y^2$
.5.3) may be reduced to four different cases which are

$$xy(V_{xx} - V_{yy}) - (x^2 - y^2 - c_1 + c_2)V_{xy} + 3yV_x - 3xV_y = 0 \qquad (3.5.4)$$

$$xy(V_{xx} - V_{yy}) - (x^2 - y^2)V_{xy} + 3yV_x - 3xV_y = 0 \qquad (3.5.5)$$

$$(x+y)(V_{xx} - V_{yy}) - (x-y)V_{xy} + 3V_x - 3V_y = 0 \qquad (3.5.6)$$

$$c_3(V_{xx} - V_{yy}) + (c_2 - c_1)V_{xy} = 0 \quad . \qquad (3.5.7)$$

(3.5.4)it is assumed that $c^2 = c_1 - c_2 \neq 0$ and in (3.5.7) that not both
, $c_2 - c_1$ are zero as the corresponding constant of motion in that case is
rely the Hamiltonian itself.

.5.5) may be solved directly to give

$$V = e(x,y; -2) + f(x^2 + y^2)$$

ere f is an arbitrary function and $e(x,y;-2)$ indicates a function homogeneous
degree minus two. Proposition(3.1.6) gives the corresponding integral as

$$(yp_x - xp_y)^2 + 2(x^2 + y^2)e \quad . \qquad (3.5.8)$$

For (3.5.4) define the canonical transformation

$$x = \frac{uv}{c},\quad y = \frac{1}{c}((u^2-c^2)(c^2-v^2))^{1/2},\quad p_x = \frac{cv\left(\frac{u^2-c^2}{c^2-v^2}\right)^{1/2} p_u + cu\left(\frac{c^2-v^2}{u^2-c^2}\right)^{1/2} p_v}{v^2\left(\frac{u^2-c^2}{c^2-v^2}\right)^{1/2} + u^2\left(\frac{c^2-v^2}{u^2-c^2}\right)^{1/2}}$$

$$p_y = \frac{cup_u - cvp_v}{u^2\left(\frac{c^2-v^2}{u^2-c^2}\right)^{1/2} + v^2\left(\frac{u^2-c^2}{c^2-v^2}\right)^{1/2}} .$$

In these coordinates the solution to (3.5.4) is

$$V = \frac{f(u) - g(v)}{u^2 - v^2} \quad \text{for arbitrary functions } f \text{ and } g.$$

The Hamiltonian is

$$2H = \frac{1}{(u^2-v^2)}((u^2-c^2)p_u^2 + 2f(u)) + \frac{1}{u^2-v^2}((c^2-v^2)p_v^2 - 2g(v)) \qquad (3.5.9)$$

From theorem (2.2) the corresponding constant of motion is

$$\frac{v^2}{u^2-v^2}((u^2-c^2)p_u^2 + 2f(u)) + \frac{u^2}{u^2-v^2}((c^2-v^2)p_v^2 - 2g(v)) . \qquad (3.5.10)$$

Likewise for (3.5.6) define the canonical transformation

$$u = (2(x^2+y^2))^{1/2} + x + y \qquad v = (2(x^2+y^2))^{1/2} - (x+y)$$

$$p_u = \frac{((2(x^2+y^2))^{1/2} - 2y)p_x - ((2(x^2+y^2))^{1/2} - 2x)p_y}{4(x-y)}$$

$$p_v = \frac{((2(x^2+y^2))^{1/2} + 2y)p_x + ((2(x^2+y^2))^{1/2} + 2x)p_y}{4(x-y)} .$$

The solution to (3.5.6) is

$$V = \frac{g(u) + h(v)}{u + v} \quad \text{for arbitrary functions } g,h .$$

The Hamiltonian may be written as

$$H = \frac{1}{u+v}(4up_u^2 + g(u)) + \frac{1}{u+v}(4vp_v^2 + h(v)) \qquad (3.5.11)$$

and again by theorem 3.1.2 the corresponding constant of motion is

$$-\frac{v}{u+v}(4up_u^2 + g(u)) + \frac{u}{u+v}(4vp_v^2 + h(v)) \quad .$$

It remains only to discuss (3.5.8). There is a variety of cases depending on the values of c_1, c_2, c_3. However, in each case there is a canonical coordinate system (x', y', p_x', p_y') so that the Hamiltonian may be written as

$$H = \frac{1}{2}p_x'^2 + V_1(x') + \frac{1}{2}p_y'^2 + V_2(y') \qquad (3.5.12)$$

for some functions V_1, V_2. Thus the Hamiltonian is additively separable and the constants of motion are obvious. It is interesting to observe that proposition 3.1.2 also applies to (3.5.12) and so I have established the result that together propositions 3.1.2 and 3.1.6 completely characterize all Hamiltonians with two degrees of freedom which admit quadratic integrals in addition to the Hamiltonian itself.

Some special cases of the above results have been given recently by several authors [11,13,44]; they consider systems with potentials of the form $V = \frac{1}{2}(c_1x^2 + c_2y^2) + ax^2y - \frac{1}{3}by^3$ (where a, b, c_1, c_2 are constants). These systems are variations on the infamous Henon-Heiles potential which corresponds to the case $a = b = c_1 = c_2 = 1$ and which was introduced twenty years ago as a model in celestial mechanics [18]. The system was conjectured to be integrable; however, one may easily see directly from

(3.5.3) that it certainly has no quadratic integral besides the Hamiltonian. Fordy [13] distinguishes three cases among the more general potentials such that there is a quadratic integral and shows that these are the only ones which do have such an integral.

4. In this example I consider a system with two degrees of freedom whose Hamiltonian h is given by

$$h = \tfrac{1}{2}(p_x^2 + p_y^2) + V(x^2 + y^2).$$

This is the angular momentum analogue of the example treated in section 3.4 so that $yp_x - xp_y$ is an integral. If one asks for those V which have non-trivial quadratic or cubic integrals it turns out that in the former case

$$V = x^2 + y^2 \qquad \text{or} \qquad V = \frac{1}{(x^2 + y^2)^{\frac{1}{2}}} \quad .$$

In the latter case one finds there are no V's which have non-trivial cubic integrals. This underlines the importance of the harmonic oscillator and Kepler potentials which is indeed what these two are.

5. In the next example I consider a class of systems which includes the type considered in special relativity. Let h be the Hamiltonian where g_{ij} is a flat metric of any signature and

$$h = (1 + g^{ij}p_i p_j)^{\frac{1}{2}} + V \quad .$$

The analog of (3.3.2) gives, where ; denotes the covariant derivative of the metric connection obtained from g_{ij},

$$A^{a_1 \cdots a_{n-1}} V_{;i} = 0$$

$$A^{a_1 \cdots a_{n-2}} V_{;i} = 0$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$A^{a \uparrow i} V_{;i} = 0$$

$$A^{i} V_{;i} = 0$$

It follows that in looking for polynomial integrals it is sufficient to consider just homogeneous polynomials. Also, this system has precisely the same *linear* integrals as the classical Hamiltonian h' given by

$$h' = \tfrac{1}{2}g^{ij}p_ip_j + V.$$

6. In this sixth example I will first of all consider the problem of trying to obtain cubic integrals for systems in general, then specialize to a particular kind of system with two degrees of freedom. The cubic integral may be assumed to have the form

$$A_{ijk}p_ip_jp_k + A_ip_i \qquad (3.5.13)$$

where A_{ijk} is a Killing tensor. One of the two remaining conditions from (3.3.2) is

$$A_{(i,j)} = 3V_{,k}A_{ijk} \quad . \qquad (3.5.14)$$

Now differentiate (3.5.14) twice to obtain

$$A_{(i,j)lm} = 3V_{,klm}A_{ijk} + 3V_{,kl}A_{ijk,m} + 3V_{,km}A_{ijk,l} + 3V_{,k}A_{ijk,lm} \qquad (3.5.15)$$

The left-hand side of (3.5.15) is symmetric in all four indices and so insisting that the right-hand side be symmetric too, gives the following system of linear, third order, partial differential equations for V

$$\begin{aligned} &A_{ijk}V_{,klm} + A_{ijk,m}V_{,kl} + A_{ijk,l}V_{,km} + A_{ijk,lm}V_{,k} \\ = \; &A_{jlk}V_{,kim} + A_{ljk,m}V_{,ki} + A_{ljk,i}V_{,km} + A_{ljk,im}V_{,k} \end{aligned} \qquad (3.5.16)$$

Clearly one also obtains analogous linear conditions on V starting with any polynomial constant whose degree is bigger than one.

Now I specialize to the case $m = 2$ and consider, on grounds of tractability, a cubic integral of the form

$$(yp_x - xp_y)^3 + Ap_x + Bp_y$$

where again I write $x = x_1$ $y = x_2$, $p_x = p_1$, $p_y = p_2$ and $A = A_1$, $B = A_2$.
The remaining conditions of (3.3.2) may be written as

$$\frac{\partial A}{\partial x} = 3y^2\left(y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y}\right)$$

$$\frac{\partial A}{\partial y} + \frac{\partial B}{\partial x} = -6xy\left(y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y}\right) \qquad (3.5.17)$$

$$\frac{\partial B}{\partial y} = 3x^2\left(y\frac{\partial V}{\partial x} - x\frac{\partial V}{\partial y}\right)$$

and

$$A\frac{\partial V}{\partial x} + B\frac{\partial V}{\partial y} = 0 \qquad (3.5.18)$$

Condition (3.5.16) is most easily obtained from (3.5.17) directly, much as (3.4.5) was obtained in section 3.4. There is a single equation which is

$$x^2yV_{xxx} - (x^3 - 2xy^2)V_{xxy} + (y^3 - 2x^2y)V_{xyy} - xy^2V_{yyy}$$

$$+ 8xyV_{xx} + 8(y^2 - x^2)V_{xy} - 8xyV_{yy} \qquad (3.5.19)$$

$$+ 12yV_x - 12xV_y = 0$$

The solution of (3.5.19) is

$$V = f(x^2 + y^2) + g + h \qquad (3.5.20)$$

where f is an arbitrary function and g and h are homogeneous functions of degree minus two and minus three respectively. Thus, applying the second condition in (3.3.2) imposes strong conditions on the form of V.

In view of (3.5.20) it is convenient to change coordinates so that

$$\zeta = x^2 + y^2, \qquad \eta = \frac{y}{x} \quad .$$

In order that the transformation should be canonical, one must also have

$$p_x = 2xp_\zeta - \frac{y}{x^2} p_\eta \quad , \quad p_y = 2yp_\zeta + \frac{1}{x} p_\eta$$

and the Hamiltonian, in view of (3.5.20) may be written as

$$H = 2\zeta p_\zeta^2 + f(\zeta) + \frac{(1+\eta^2)p_\eta^2 + G(\eta)}{2\zeta} + \frac{H(\eta)}{\zeta^{3/2}} \tag{3.5.21}$$

for some functions G and H. The cubic integral now assumes the form

$$(1+\eta^2)^3 p_\eta^3 + ap_\xi + bp_\eta \tag{3.5.22}$$

for some functions a and b. It remains to satisfy the last condition in (3.3.2) as well as to relate a and b to f,G and H. One finds that

$$2\zeta \frac{\partial a}{\partial \zeta} = a \tag{3.5.23}$$

$$4\zeta^2 \frac{\partial b}{\partial \zeta} + (1+\eta^2)^2 \frac{\partial a}{\partial \eta} = 0 \tag{3.5.24}$$

$$\frac{a(1+\eta^2)}{2\zeta^2} + \frac{(1+\eta^2)}{\zeta} \frac{\partial b}{\partial \eta} - \frac{2b\eta}{\zeta} = 3(1+\eta^2)^2 \left(\frac{G'(\eta)}{\zeta} + \frac{H'(\eta)}{\zeta^{3/2}}\right) \tag{3.5.25}$$

but the last condition (3.5.18) has still not been applied. Still without applying it, it follows that the cubic integral has the form

$$(1+\eta^2)^3 p_\eta^3 + \theta(\eta)\zeta^{\frac{1}{2}} p_\zeta + \left(\frac{(1+\eta^2)^2\theta'(\eta)}{2\zeta^{\frac{1}{2}}} + 3(1+\eta^2)G(\eta)\right)p_\eta \tag{3.5.26}$$

where θ is a function of η satisfying

$$\theta + (1+\eta^2)^2 \theta'' = 6(1+\eta^2)H' \tag{3.5.27}$$

When (3.5.18) is applied one finds that the function f is a sum of three functions homogeneous of degrees -1, -2, -3 respectively. Then one may argue in several stages that there is no loss of generality in supposing that $f \equiv 0$ and $G \equiv 0$. The cubic integral then has the form

$$(1+\eta^2)^3 p_\eta^3 + \theta\zeta^{\frac{1}{2}} p + \frac{(1+\eta^2)^2\theta'}{2\zeta^{\frac{1}{2}}} p_\eta \tag{3.5.28}$$

where in addition to (3,5.27), θ and H are also bound by the relation

$$3\theta H = (1 + \eta^2)^2 \theta' H'. \tag{3.5.28}$$

Next, setting $K\sec z = \theta$ and $\eta = \tan z$ (3.5.28) and (3.5.29) and are transformed respectively to

$$K''+2K = 6H' \cos z \tag{3.5.30}$$

$$3HK = H'K'+H'K \tan z \tag{3.5.31}$$

where $'$ denotes differentiation with respect to z. One can obtain a single, albeit rather complex, third order equation for K by differentiating each of (3.5.30) and (3.5.31) with respect to z and then using all four equations to eliminate H, H' and H''.

Finally, I shall summarize the results of this example using the original notation. It has been shown that the only systems which have a potential of the type given by (3.5.20), which have an integral of the form $(yp_x - xp_y)^3 + Ap_x + Bp_y$ are those which have a Hamiltonian given by

$$\tfrac{1}{2}(p_x^2 + p_y^2) + \frac{H(\frac{y}{x})}{(x^2+y^2)^{3/2}} .$$

In this case the integral is given by

$$-(yp_x - xp_y)^3 + \frac{1}{2(x^2+y^2)^{\frac{1}{2}}} [Kx(xp_x+yp_y) - (K\frac{\sqrt{x^2+y^2}}{x})' (yp_x - xp_y)].$$

where $'$ denotes differentiation with respect to $z = \arctan(\frac{y}{x})$. Moreover H and K are related by (3.5.30) and (3.5.31) - conditions which imply that H depends on three arbitrary constants.

7. In this seventh and final example I investigate the question of when a Hamiltonian of the type considered in section 3.3 and giving rise to conditions (3.3.1) has an integral of motion which is a *homogeneous* polynomial in the momenta. In fact (3.3.2) shows that the existence of integrals which are homogeneous polynomials in the velocities is a good deal more restrictive than Xanthopoulos [46] has claimed. For in that case clearly the polynomial is a Killing tensor, say $A_{a_1 \ldots a_n}$, related also to the potential V by the conditions

$$A_{a_1 \ldots a_{n-1} i} V^{,i} = 0 \qquad (3.5.32)$$

What has been said so far applies equally to systems with any number of degrees of freedom. I now specialize to the case of two degrees of freedom and use coordinates x and y. In this case, the Killing tensors consist precisely of symmetrized products of the three Killing vector fields $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$, $y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}$. Suppose now that one asks for all systems with two degrees of freedom which possess integrals which are homogeneous of degree one. Then it is not hard to show that these integrals are necessarily of the momentum type in the sense that there exists a *canonical* coordinate system (x, y, p_x, p_y) so that the potential is either given by $V = V(x-y)$, or $V = V(x^2+y^2)$. Likewise, if one asks for systems which have homogeneous quadratic integrals independent of the Hamiltonian, one finds again that V has one of the two forms given above and that the integrals are merely the squares of momentum integrals. The latter results can be obtained directly or by specializing example 3 of this section.

I shall now investigate systems which have integrals which are homogeneous cubic polynomials. By what has been remarked earlier, this cubic integral may be assumed to be of the form

$$A(yp_x - xp_y)^3 + 3(yp_x - xp_y)^2(B_1p_x - B_2p_y) + 3(yp_x - xp_y)(C_1p_x^2 + 2C_3p_xp_y + C_2p_y^2)$$
$$+ D_1p_x^3 - 3D_3p_x^2p_y^2 + 3D_4p_xp_y^2 - D_2p_y^3$$

where $A, B_1, B_2, C_1, C_2, C_3, D_1, D_2, D_3, D_4$ are constants. This integral could be written in classical tensor notation as $A_{ijk}p_ip_jp_k$. (3.3.2) now gives three conditions relating the A_{ijk} and V. In order that these be compatible to yield non-constant V it is necessary that

$$A_{111}A_{122} - A_{112}^2 = 0 \qquad (3.5.33)$$
$$A_{222}A_{112} - A_{122}^2 = 0 \qquad (3.5.34)$$
$$A_{111}A_{222} - A_{112}A_{122} = 0 \qquad (3.5.35)$$

When these conditions are imposed one finds that once more they can be satisfied essentially only in the two ways mentioned above i.e. after a suitable change of coordinates

$$V = V(x-y), \qquad V = V(x^2+y^2) \qquad (3.5.36)$$

Indeed, this is almost obvious from just (3.5.35). The preceding argument may be generalized to the case of arbitrary n, but this shall not be done because the calculations become rather cumbersome. Thus, one is led to conclude that the only standard classical systems with two degrees of freedom which have integrals which are homogeneous polynomials in the momenta are of the type given by (3.5.36) a result which obviously conflicts with those given by Xanthopoulos [46].

CHAPTER 4

SECOND ORDER EQUATIONS AND LAGRANGIAN MECHANICS

This chapter is principally concerned with Lagrangian mechanics. Section 4.1 describes some of the geometrical features of TM. This is taken from Crampin's paper [8] but contains a new result about the kinds of diffeomorphisms which preserve the tensor field S which is also introduced in this section. Section 4.2 is much the same for $J^1(\mathbb{R},M)$ as section 4.1 was for TM. Notation is chosen so that geometric objects, when suitably restricted, give the corresponding object on TM. However, the constructions on $J^1(\mathbb{R},M)$ are not trivial modifications of those for TM. Also the Cartan 1-form is introduced and its role in formulating Lagrangian theory briefly explained. Finally, section 4.3 is devoted to a discussion of several types of symmetry in Lagrangian mechanics. The way in which several different notions of symmetry interact is investigated. Also, the precise relationship between Noether and Cartan symmetries is explained. This last section is based on papers of Crampin [7] and Prince [33] but contains some new observations.

4.1 Geometrical Structures on TM

In section 1.8 the relationship of TM to $J^1(\mathbb{R},M)$ was explored (here M is a smooth manifold of dimension m) and in section 1.9 it was explained that TM provides a suitable setting for time-independent Lagrangian mechanics whereas $J^1(\mathbb{R},M)$ is appropriate for dealing with time dependent systems. Both TM and $J^1(\mathbb{R},M)$ have special geometrical features which are needed to set up Lagrangian theory. These have been examined in some detail by Crampin [8] in the case of TM and used to shed light on such questions as the inverse problem of Lagrangian mechanics i.e. when is an O.D.E. 2.0. equivalent to a Euler-Lagrange equation? In this section I review some of the constructions on TM with a view to carrying out analogous constructions on $J^1(\mathbb{R},M)$. I assume that coordinates (t, x_i) have been chosen on $\mathbb{R} \times M$ and that (x_i, u_i) and (t, x_i, u_i) are the induced coordinates on TM and $J^1(\mathbb{R},M)$ respectively. Since TM and $J^1(\mathbb{R},M)$ are vector bundles over M and $\mathbb{R} \times M$ respectively, they both have canonical dilation vector fields which have the local form $u_i \frac{\partial}{\partial u_i}$.

In [47] it is shown that TM has a canonical 1-1 tensor field S which in coordinates has the form $\frac{\partial}{\partial u_i} \otimes dx_i$; S is the key ingredient in Crampin's formulation of Lagrangian dynamics [8]. S may be defined in several different ways, perhaps the simplest of which follows from the vertical lift construction. If $X \in T_x M$ then its vertical lift to $T_{(x,u)}TM$ where u is any point of $\Pi^{-1}(x)$ is given by $X^v(x,u) = \Delta(x,X)$. If X is a vector field then the process may obviously be done fiber by fiber to obtain a vector field X^v on TM.

X^v is clearly vertical and "constant on fibers". If X has the local expression $X^i \frac{\partial}{\partial x^i}$ then X^v has the form $X^i \frac{\partial}{\partial u_i}$. The tensor S may now be defined by

$$S_{(x,u)}(v) = (\Pi_* v)^V_{(x,u)} \qquad (v \in T_{(x,u)}\, TM).$$

In other words S may be viewed as a field of endomorphisms on each tangent space to TM, and I have given the effect of S on an arbitrary tangent vector.

A second important lifting procedure on TM is the so-called complete lift [47]. In this one prolongs the flow of a vector field by using the tangent map. This yields a flow on TM and hence a vector field X^c on TM. If X has the local expression $X^i \frac{\partial}{\partial x^i}$ then X^c is given by $X^i \frac{\partial}{\partial x_i} + u_i \frac{\partial X^j}{\partial x_i} \cdot \frac{\partial}{\partial u_j}$. One should note that the complete lift construction has no pointwise equivalent. Now let Γ be a (regular) autonomous O.D.E. 2.0. I will single out seven basic formulas which relate the objects above:

$$[X^v, Y^v] = 0 \tag{4.1.1}$$

$$[X^v, Y^c] = [X,Y]^v \tag{4.1.2}$$

$$[X^c, Y^c] = [X,Y]^c \tag{4.1.3}$$

$$S(X^c) = X^v \tag{4.1.4}$$

$$< \Gamma, S > = \Delta \tag{4.1.5}$$

$$[X^v, \Delta\ ] = X^v \tag{4.1.6}$$

$$[X^c, \Delta\ ] = 0 \tag{4.1.7}$$

These may be verified by local calculations or, with more or less difficulty, by using the intrinsic definitions.

Almost all the other results in [8] may be derived from these formulas. For example, one has

$$L_{X^V} S = 0 \qquad (4.1.8)$$

$$L_{X^C} S = 0 \qquad (4.1.9).$$

I next give a proof of a result which is fundamental to Crampin's analysis in [8].

Proposition 4.1.1 Let Γ be a regular autonomous second order equation. Then, $L_\Gamma S$ acts as the identity on vertical vector fields i.e.

$$X \in V^1(TM) \text{ and } \pi_* X = 0 \Rightarrow \langle X, L_\Gamma S\rangle = X.$$

Proof: Note first that by the derivation property of Lie derivative

$$\langle X, L_\Gamma S\rangle = L_\Gamma \langle X, S\rangle - \langle [\Gamma, X], S\rangle$$

and as $\langle X, S\rangle = 0$ it is sufficient to show that

$$\langle [X,\Gamma], S\rangle = X.$$

Secondly, note that if the last equation holds for $X = Y^V$ ($Y \in V^1(M)$) then if $f \in F(TM)$ one has

$$\begin{aligned} \langle [fY^V,\Gamma], S\rangle &= \langle f[Y^V,\Gamma] - (\Gamma f)Y^V, S\rangle \\ &= f[Y^V,\Gamma] \\ &= fY^V. \end{aligned}$$

Hence, it is sufficient to show that $\langle [Y^V,\Gamma], S\rangle = Y^V$, since the vertical lifts form a basis for the vertical vector fields. But now

$$\begin{aligned} \langle [Y^V,\Gamma], S\rangle &= L_Y v\langle \Gamma, s\rangle = -\langle \Gamma, L_Y v S\rangle \\ &= L_Y v\Delta \qquad \text{(from (4.1.5) and (4.1.8))} \\ &= Y^V \qquad \text{(from (4.1.6))}. \end{aligned}$$

∎

Another point which is worth adding about complete and vertical lifts is that if (X_i) are a basis of vector fields on M then (X_i^v, X_i^c) are a basis of vector fields on TM.

One may enquire about the kind of diffeomorphisms of TM which leave S invariant. The following proposition resolves this question entirely.

Theorem 4.1.3 A diffeomorphism Φ of TM preserves S iff it is the lift of a diffeomorphism ϕ of M i.e. $\Phi = T\phi$, followed by a translation of the fibers.

Proof: I begin with the necessity. The first point to note is that Φ must be a BUN morphism because given any $(x,u) \in TM$ the vertical vectors in $T_{(x,u)}TM$ are a subspace distinguished by S viz., either the kernel or image of S regarded as an endomorphism of $T_{(x,u)}TM$. Now introduce standard coordinate systems (x_i,u_i), (x_i',u_i') on TM so that Φ is locally described by $u_k' = \Phi_k(x_i,u_j)$ $x_l' = \phi_l(x_i)$. Then one finds that such a coordinate transformation changes S as follows:

$$\frac{\partial}{\partial u_i} \otimes dx_i = \frac{\partial \Phi_j}{\partial u_i} \cdot \frac{\partial (\phi_i^{-1})}{\partial x'_k} \cdot \frac{\partial}{\partial u'_j} \otimes dx'_k$$

Hence $\frac{\partial \Phi_j}{\partial u_i} \cdot \frac{\partial (\phi_i^{-1})}{\partial x'_k} = \delta_j k$, which implies that Φ consists of T_ϕ followed by a fiber translation. On the other hand, it is clear that if ϕ is a diffeomorphism of M then $T\phi$ followed by a fiber translation changes coordinates in just the same way and so preserves S. ∎

4.2 Geometrical Structures on $J^1(\mathbb{R},M)$

I now consider some geometrical features of $J^1(\mathbb{R},M)$. As was pointed out above $J^1(\mathbb{R},M)$ is a vector bundle over $\mathbb{R} \times M$ and has a canonical dilation vector field Δ, which has the local expression $u_i \frac{\partial}{\partial u_i}$. Of course, (x_i,u_i) may be used simultaneously as coordinates for TM. In what follows we shall frequently be in a position where corresponding to some geometric object on TM there is an analogous object on $J^1(\mathbb{R},M)$. The notation and terminology will be chosen as far as possible so that by suitably restricting the object on $J^1(\mathbb{R},M)$ one obtains the corresponding object on TM. However, I shall not distinguish between the object and its restriction. The context will make it clear when I am dealing with $J^1(\mathbb{R},M)$ rather than TM.

As was mentioned in section 1.8, $J^1(\mathbb{R},M) \cong \mathbb{R} \times (TM)$. Hence the tensor $\frac{\partial}{\partial u_i} \otimes dx^i$ TM may be regarded as a tensor on $J^1(\mathbb{R},M)$. Letting t denote the canonical coordinate on $\mathbb{R}$ and subtracting $\Delta \otimes dt$ from the tensor just mentioned, gives a 1-1 tensor on $J^1(\mathbb{R},M)$ which will also be called S. Clearly S restricted to TM makes sense and coincides with the tensor which was called S on TM. In standard coordinates (t,x_i,u_i) on $J^1(\mathbb{R},M)$, S has the form $\frac{\partial}{\partial u_i} \otimes (dx_i - u_i dt)$. This last expression renders evident the following proposition.

Proposition 4.2.1 S is the unique 1-1 tensor on $J^1(\mathbb{R},M)$ which

(i) vanishes on vertical vectors and regular second order equations.

(ii) if $X \in V^1(J^1(\mathbb{R},M))$, S(X) is vertical.

(iii) satisfies $S(\frac{\partial}{\partial t}) = -\Delta$.

The analog of the complete lift construction on TM proceeds as follows. Let X be a vector field on $\mathbb{R} \times M$. Then there is a unique vector field $X^{(1)}$ or $J^1(\mathbb{R},M)$ determined by the conditions that it be $(\rho \times \sigma)$-related to X and that $L_{X^{(1)}} \Omega^{(1)} \subset \Omega^{(1)}$ where $\Omega^{(1)}$ denotes the module of contact forms. $X^{(1)}$ is called the first prolongation of X and if $X = T \frac{\partial}{\partial t} + X^i \frac{\partial}{\partial x^i}$

$$X^{(1)} = T \frac{\partial}{\partial t} + \dot{X}^i \frac{\partial}{\partial x_i} + (\frac{\partial}{\partial t} + u_j \frac{\partial}{\partial x_j})(X^i - Tu_i) \frac{\partial}{\partial u_i} .$$

Again if $X \in V^1(\mathbb{R} \times M)$ its vertical lift X^v will be defined by $X^v = S(X^1)$ and it has the local expression $(X^i - Tu_i) \frac{\partial}{\partial u_i}$. The following are the analog of formulas (4.1.1 - 7) where $X,Y \in V^1(\mathbb{R} \times M)$ and $Xt = T$ and $Yt = V$.

$$[X^v,Y^v] = TY^v - VX^v \qquad (4.2.1)$$

$$[X^v,Y^{(1)}] = [X,Y]^v + \dot{T}Y^v - \dot{V}X^v \qquad (4.2.2)$$

$$[X^{(1)},Y^{(1)}] = [X,Y]^{(1)} \qquad (4.2.3)$$

$$S(X^{(1)}] = X^v \qquad (4.2.4)$$

$$<\Gamma,S> = 0 \qquad (4.2.5)$$

$$[X^v, \Delta] = X^v - T\Delta \qquad (4.2.6)$$

$$[X^{(1)}, \Delta] = [\frac{\partial}{\partial t}, X^v] + \dot{T}\Delta \qquad (4.2.7)$$

Here the symbol $\dot{T}$ denotes the function obtained by differentiating T (regarded as a function on $J^1(\mathbb{R},M)$) along any regular second order equation.

As is well-known, any diffeomorphism ϕ of M lifts uniquely to a diffeomorphism $T\phi$ of TM. However, the situation on $J^1(\mathbb{R},M)$ is not quite so simple. Now certainly any diffeomorphism ϕ of M may be lifted to a contact transformation of $J^1(\mathbb{R},M)$; in fact since $J^1(\mathbb{R},M) \cong \mathbb{R} \times TM$ this map is just $i_{\mathbb{R}} \times T\phi$. However, an arbitrary diffeomorphism ϕ of $\mathbb{R} \times M$ may not always be lifted so as to give a global contact transformation of $J^1(\mathbb{R},M)$. A further condition is necessary. I shall examine the situation with the help of two sets of standard local coordinates (t,x_i,u_i) and (t',x'_i,u'_i). We are given a diffeomorphism of $\mathbb{R} \times M$ and this can be described locally by the equations

$$t' = \phi_0(t,x_i), \quad x'_j = \phi_j(t,x_i)$$

In order to be able to lift ϕ to a contact transformation, one must impose the condition that the contact module $\{dx'_i - u'_i dt'\}$ be pulled back to $\{dx_i - u_i dt\}$. I shall present the calculation, since it is almost the same as that which appears in the next theorem; however, applying the last condition one finds that

$$u'_j = \frac{\dfrac{\partial \phi_j}{\partial t} + u_i \dfrac{\partial \phi_j}{\partial x_i}}{\dfrac{\partial \phi_0}{\partial t} + u_i \dfrac{\partial \phi_0}{\partial x_i}}$$

and this determines the transformation Φ. We want $\dfrac{\partial \phi_0}{\partial t} + u_i \dfrac{\partial \phi}{\partial x_i}$ to be always non-zero, which will be the case iff

$$\frac{\partial \phi_0}{\partial t} \neq 0, \qquad \frac{\partial \phi_0}{\partial x_i} = 0 \qquad (1 \leq i \leq m).$$

But this says precisely that ϕ must preserve the fibration determined by the map $\mathbb{R} \times M \longrightarrow \mathbb{R}$ which projects onto the first factor. Thus

the diffeomorphisms of $\mathbb{R} \times M$ which can be lifted to contact transformations of $J^1(\mathbb{R},M)$ are precisely those which preserve the fibration of $\mathbb{R} \times M$ over $\mathbb{R}$.

Theorem 4.2.2 A diffeomorphism Φ of $J^1(\mathbb{R},M)$ preserves S iff it is the lift of a transverse (in the sense defined above) diffeomorphism ϕ of $\mathbb{R} \times M$.

Proof: On $J^1(\mathbb{R},M)$ the vertical vectors form $\mathrm{im}(S)$ at each point. Hence Φ must be fibered over ϕ say on $\mathbb{R} \times M$. Let (t,x_i,u_i), (t',x'_i,u'_i) be standard coordinate systems on $J^1(\mathbb{R},M)$ such that Φ is locally described by

$$t' = \phi_0(t,x_i),\ x'_j = \phi_j(t,x_i),\ u'_j = \Phi_j(t,x_i,u_i).$$

I shall now show that if Φ preserves S then u'_j is determined by ϕ_0 and the ϕ_j's. Since Φ is fibered over ϕ, $\frac{\partial}{\partial u_i} = \frac{\partial \Phi_j}{\partial u_i} \frac{\partial}{\partial u'_j}$ and hence $\frac{\partial}{\partial u'_j} = (\frac{\partial \Phi_j}{\partial u_i})^{-1} \frac{\partial}{\partial u_i}$. Thus, pulling back S by Φ gives

$$(\frac{\partial \Phi_j}{\partial u_i})^{-1} \frac{\partial}{\partial u_i} \otimes (\frac{\partial \phi_j}{\partial x_i} dx_i + \frac{\partial \phi_j}{\partial t} dt - \Phi_j(\frac{\partial \phi_0}{\partial t} dt + \frac{\partial \phi_0}{\partial x_i} dx_i)).$$

This will be the same as S iff

$$\frac{\partial \Phi_j}{\partial u_i} = \frac{\partial \phi_j}{\partial x_i} - \Phi_j \frac{\partial \phi_0}{\partial x_i}$$

and $$u_i \frac{\partial \Phi_j}{\partial u_i} = \Phi_j \frac{\partial \phi_0}{\partial t} - \frac{\partial \phi_j}{\partial t}$$

which are in turn equivalent to

$$\Phi_j = \frac{\frac{\partial \phi_j}{\partial t} + u_i \frac{\partial \phi_j}{\partial x_i}}{\frac{\partial \phi_0}{\partial t} + u_i \frac{\partial \phi_0}{\partial x_i}} \qquad (4.2.8)$$

Now (4.2.8) is precisely the condition that Φ be the prolongation to $J^1(\mathbb{R},M)$ of ϕ from $\mathbb{R} \times M$. Hence, in order that Φ be globally defined, it is necessary that ϕ be transverse in the sense described above. Conversely, the prolongation of a transverse diffeomorphism of $\mathbb{R} \times M$ changes coordinates in just the same way and so preserves S. ∎

To end this section I shall explain how the tensor S figures in the geometric formulation of Lagrangian theory.

Definition 4.2.3 Given a Lagrangian $L : J^1(\mathbb{R},M) \to \mathbb{R}$, its associated Cartan 1-form Θ_L is given by $Ldt + \langle S, dL \rangle$ (t is the canonical coordinate on $\mathbb{R}$). $d\Theta_L$ is known as the Cartan 2-form. Clearly, in standard coordinates (t,x_i,u_i) Θ_L is given by

$$\Theta_L = Ldt + \frac{\partial L}{\partial u_i}(dx_i - u_i dt) \qquad (4.2.9)$$

Now $d\Theta_L = \frac{\partial L}{\partial x_i} dx_i \wedge dt + d(\frac{\partial L}{\partial u_i}) \wedge (dx_i - u_i dt) = (d(\frac{\partial L}{\partial u_i}) - \frac{\partial L}{\partial x_i} dt) \wedge (dx_i - u_i dt)$.

$$(d\Theta_L)^m = (d(\frac{\partial L}{\partial u_1}) - \frac{\partial L}{\partial x_1} dt) \wedge \ldots \wedge (d(\frac{\partial L}{\partial u_m}) - \frac{\partial L}{\partial x_m} dt) \wedge (dx_1 - u_1 dt) \wedge \ldots \wedge (dx_m - u_m dt).$$

Thus

$$\Theta_L \wedge (d\Theta_L)^m = L \det(\frac{\partial^2 L}{\partial j_i \partial u_j}) dt \wedge dx_1 \wedge \ldots \wedge dx_m \wedge du_1 \wedge \ldots \wedge du_m.$$

Since L is assumed to be regular i.e. $\det(\frac{\partial^2 L}{\partial u_i \partial u_j}) \neq 0$, then provided L is nowhere zero, $\Theta_L \wedge (d\Theta_L)^m$ is a volume form and so Θ_L is an example of a contact structure. Alternatively, Θ_L may be thought of as a contact structure on the open subset for which L is non-zero. Contact structures will be studied in some detail in Chapter 5. Anticipating developments there, one can assert the existence of a canonically defined one-dimensional distribution Γ which is characteristic to $d\Theta_L$ i.e. satisfies $\Gamma \rfloor d\Theta = 0$. Γ may be further normalized by insisting that $\Gamma t = 1$. In chapter 5 a different normalization will be given and the canonical characteristic vector field Z will be defined. Z satisfies $Z \rfloor d\Theta = 0$ but $\langle Z, \Theta \rangle = 1$. Clearly, $Z = \frac{1}{L} \Gamma$. This has the rather interesting consequence that Γ is always defined provided only that L is regular, whereas Z will not be defined at points where L is zero.

Γ is of course the Euler-Lagrange vector field associated to a regular Lagrangian as may readily be seen from using (4.2.9). The preceding remarks illustrate what might be considered one of the drawbacks of Lagrangian theory. Θ_L is the key geometrical object in the formalism and yet its connection with the variational problem for Ldt is not a priori evident.

4.3 Symmetry in Lagrangian mechanics

This section is intended to supplement works by Crampin [7] and Prince [33] which deal with the notion of symmetry in Lagrangian mechanics. Briefly a symmetry of a differential equation is a one-parameter group of diffeomorphisms which permutes solutions of the equation. There are several kinds of symmetry which may be usefully distinguished and to which first integrals of the equations may be associated either directly or indirectly. I shall begin by considering two kinds of symmetry which apply to regular, second order equations rather than those just of Euler-Lagrange type.

Definition 4.3.1 Suppose that Σ is a regular, second order equation so that Σ can be identified with $J^1(\mathbb{R},M)$ and that Γ is the associated vector field. Then a vector field Y on $J^1(\mathbb{R},M)$ such that

$$[Y,\Gamma] = \lambda\Gamma \qquad (4.3.1)$$

for some $\lambda \in F^0(J^1(\mathbb{R},M))$ is called a dynamical symmetry. If also $Y = X^{(1)}$ for some $X \in V^1(\mathbb{R} \times M)$ then X is called a Lie symmetry. ∎

I shall assume that a specific, regular second order equation has been chosen and use A and C respectively to denote the collection of Lie and dynamical symmetries.

Suppose that $X, Y \in C$ so that

$$[X,\Gamma] = \lambda\Gamma \qquad [Y,\Gamma] = \mu\Gamma$$

where $\lambda,\mu \in F(J^1(\mathbb{R},M))$. Then

$$[[X,Y],\Gamma] = (X\mu - Y\lambda)\Gamma. \qquad (4.3.2)$$

It follows from (4.2.3) and (4.3.2) that A and C are $\mathbb{R}$-Lie algebras and that A is a subalgebra of C. In fact more is true in that A is finite dimensional though C is not.

C may also be viewed as a Lie algebra in a rather different way. Firstly, note that the collection of all first integrals of Γ form a subgroup (indeed subalgebra) of $F^0(J^1(\mathbb{R},M))$. Secondly, suppose that f is such a first integral and that $X \in C$. Then since

$$[fX,\Gamma] = f[X,\Gamma] - (\Gamma f)X$$

and $\Gamma f = 0$, fX is another dynamical symmetry. Thus, C may be considered as a finite rank module/algebra over the ring of first integrals of Γ.

It is worth emphasizing that the preceding remarks held for arbitrary regular second order equations. I now suppose that in addition Γ is the Euler-Lagrange vector field associcated to a regular second order equation. This opens up the possibility of other kinds of symmetries. As in section 4.2 Θ_L denotes the Cartan 1-form associated to L.

Definition 4.3.2 A vector field Y on $J^1(\mathbb{R},M)$ is called a Cartan symmetry if $L_Y\Theta$ is exact. If $Y = X^{(1)}$ for $X \in V^1(\mathbb{R} \times M)$ then Y will be said to be a special Cartan symmetry.

Definition 4.3.3 A vector field Y in $J^1(\mathbb{R},M)$ is called a Noether symmetry if Y is an infinitesimal symmetry of $\Omega^{(1)}$ and $L_Y Ldt$ is exact mod $\Omega^{(1)}$. If $Y = X^{(1)}$ for $X \in V^1(\mathbb{R} \times M)$, in which case the first condition holds automatically, then Y will be said to be a special Noether symmetry. I shall denote the collection of Cartan (special Cartan) and Noether (special Noether) symmetries by D(B) and E(F) respectively. The lettering used above for A, B, C, D agrees with that used by Prince [33] for identical concepts though the terminology is slightly different. It turns out that one of the categories above is redundant because of the following result whose proof is evident.

Proposition 4.3.4 Suppose that $X \in V^1(J^1(\mathbb{R},M))$ and $L_X\Omega^{(1)} \subset \Omega^{(1)}$. Then $L_X\Theta \equiv L_X Ldt \bmod \Omega^{(1)}$. ∎ I shall show presently that proposition 4.3.4 implies that essentially B = F. First I quote without proof some results from Crampin [7] which elucidate the relationship between symmetries based on the Cartan form and symmetries based on a version of Noether's theorem.

Theorem 4.3.5

(i) If X is a Cartan symmetry, say $L_X\Theta = df$, then $f - \langle X,\Theta\rangle$ is a constant of motion.

(ii) If F is a constant of motion, there is a vector field X on $J^1(\mathbb{R},M)$ such that $L_X\Theta = d(F + \langle X,\Theta\rangle)$ and any vector field $X + f\Gamma$ with $f \in F^o(J^1(\mathbb{R},M))$ has the same property.

(iii) Of all vector fields Y differing by a multiples of Γ which generate a given constant of motion there is a unique one which

satisfies Yt = 0: this establishes a one-one correspondence between constants of motion and a Lie subalgebra of vector fields on $J^1(\mathbb{R},M)$. ∎

Theorem 4.3.6 If X is a Noether symmetry, say $L_X Ldt - df \in \Omega^{(1)}$, then $f - \langle X,\Theta\rangle$ is a constant of motion. ∎

These results do not quite explain the relationship of Noether and Cartan symmetries. In fact the situation is this: one may weaken the definition of Cartan symmetry by insisting only that $L_Y\Theta$ be exact mod $\Omega^{(1)}$, say, $L_Y\Theta \equiv df \bmod \Omega^{(1)}$. Then, since Γ is a second order equation, it follows that $f - \langle Y,\Theta\rangle$ is a constant of motion. Suppose now that one starts with a Noether symmetry X. Then, from proposition 4.3.4 X is a Cartan symmetry in the extended sense just mentioned. On the other hand, not every Cartan symmetry is an infinitesimal symmetry of $\Omega^{(1)}$, and it is this which is responsible for the fact that theorem 4.3.6 has no converse.

Suppose now that D' and B' denote Cartan and special Cartan symmetries in the extended sense just mentioned i.e.
$D' = \{Y \in V^1(J^1(\mathbb{R},M)) \,|\, L_Y\Theta \text{ is exact mod } \Omega^{(1)}\}$ and
$B' = \{Y \in V^1(\mathbb{R} \times M) \,|\, L_X(1)\Theta \text{ is exact mod } \Omega^{(1)}\}$. Then one has the following relations, where < denotes "is a Lie subalgebra of".

Theorem 4.3.7 (i) $B < A < C$

(ii) $B < B' = F < E < D'$

(iii) $D < C$

Proof: (i) Suppose that $L_{X^{(1)}}\Theta = df$ so that $X \in B$. Then

$$L_{X^{(1)}} d\Theta = 0 \text{ and so}$$

$$[X^{(1)}, \Gamma]\, d\Theta = L_{X^{(1)}}(\Gamma \rfloor d\Theta) - \Gamma \rfloor L_{X^{(1)}} d\Theta = 0.$$

Hence $[X^{(1)}, \Gamma]$ is characteristic to $d\Theta$ and so $X \in A$. $A < C$ is clear.

(ii) The fact that $B' = F$ and $E < D'$ is immediate from proposition 4.3.4. $B < B'$ and $F < E$ are clear.

(iii) $D < C$ follows from Crampin's proof of theorem 4.3.5 (iii) given in [7]. ∎

Notice that since $B < A$ the special Cartan symmetries form a finite dimensional $\mathbb{R}$-Lie algebra.

CHAPTER 5

SYMPLECTIC AND CONTACT GEOMETRY

This chapter begins by reviewing some of the basic properties of symplectic manifolds. Then linear symplectic spaces are considered and it is explained why, on the purely algebraic level, the Jacobi identity is equivalent to a derivation of degree one on the exterior algebra of the dual space. In section 5.2 the notion of a contact structure is introduced and it is shown how a program for contact manifolds somewhat analogous to that of symplectic geometry may be carried out. The geometry of contact manifolds is rather more complex than that of symplectic manifolds and several of the formulas appearing in section 5.2 are new. For example, it is shown how the canonical characteristic vector field behaves when a contact transformation is applied. As a corollary those contact transformations which are fibered over T*M are characterized. Recall that symplectic manifolds can be dualized to give cosymplectic structures. Although there is a bracket on the ring of functions on a contact manifold it cannot lead to a cosymplectic structure. It does, however, give a 2-vector with Cartan properties and this is defined. This 2-vector allows an invariant proof of a classical theorem in contact geometry to be given: the theorem says that contact transformations are precisely those which scale the bracket of functions by a factor.

Section 5.2 was valid generally for arbitrary contact manifolds. Section 5.3 by contrast, focuses on the space $J^1(M,\mathbb{R})$. A modern version of some results originally due to Lie is presented;

specifically, it is explained how all contact transformations may be obtained from transformations or even suitable submanifolds of $J^1(M,\mathbb{R}) \times J^1(M,\mathbb{R})$.

Finally, section 5.4 is an attempt to unify the process by which the Lie algebra structure on a collection of vector fields is transferred to a collection of forms. Two particular theorems are proved and it is shown that they explain several different cases in which Lie algebras of forms occur.

5.1 A Symplectic Review

Before getting to the concept of contact structure in the next section I review some symplectic preliminaries. A symplectic structure ω on a manifold N, necessarily of even dimension, is simply a closed non-degenerate 2-form. According to the Darboux theorem, coordinates (x_i, p_i) may be introduced so that locally $\omega = dp_i \wedge dx_i$. The canonical example of a symplectic structure occurs when $N = T^*M$ (the cotangent bundle of some underlying manifold M) as was mentioned in section 1.7. In this case the symplectic form is the exterior derivative of the canonical 1-form θ.

Returning to the general symplectic structure it is well known that because of the nondegeneracy of ω a pointwise isomorphism is defined between $T_{(x,p)}T^*M$ and $T^*_{(x,p)}T^*M$ and hence an identification of vector fields with 1-forms. In particular if $f \in F^o(T^*M)$ its associated Hamiltonian vector field X_f satisfies

$$X_f \lrcorner \omega = - df.$$

As was mentioned in Chapter 2 every symplectic form gives rise to a Poisson manifold. For, if ω is simply a non-degenerate 2-form (not necessarily closed) there is a corresponding 2-vector Ω. One may form the 3-form $d\omega$ and then "raise the indices" with ω to obtain a 3-vector. It turns out that the resulting 3-vector is just $[\Omega,\Omega]$ i.e. the skew Schouten bracket of Ω with itself.

If $f,g \in F^o(T^*M)$ the condition $d\omega = 0$ ensures that at least locally, that there is a function, denoted by $\{f,g\}$, such that

$$[X_f, X_g] \lrcorner \omega = - d\{f,g\} \quad .$$

Equivalently, $d\omega = 0$ implies that the Hamiltonian vector fields form a Lie algebra.

It is instructive to trace the purely linear analog of the latter remarks. I begin by assuming V is a vector space equipped simply with a skew-symmetric bilinear product denoted by $[\ ,\ ]$. As such, $[\ ,\ ]$ determines a unique linear map from $\Lambda^2(V)$ to V which will also be denoted by $[\ ,\ ]$. One may also consider the trilinear map from $V \times V \times V$ to V given by, for $X,Y,Z \in V$

$$(X,Y,Z) \longmapsto [X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]$$

and I denote its associated linear map from $\Lambda^3(V)$ to V by γ. One may also consider the dual maps of $[\ ,\ ]$ and γ. I denote by d the dual map of $[\ ,\ ]$ so that d is a linear map from V^* to $\Lambda^2(V^*)$ and by γ^* the dual of γ so that γ^* is a linear map from V^* to $\Lambda^3(V^*)$. It is not hard to show that any linear map from $\Lambda^p(V)$ to V (hence one from $\Lambda^p(V^*)$ to V^*) may be uniquely extended to a derivation of the entire exterior algebra. Thus, d and γ may be extended to derivations of degree 1 and 2 respectively. One would expect d and γ to be intimately related and indeed they are.

Proposition 5.1.1 $d^2 = \gamma^*$

Proof: It is easy to check that d^2 is a derivation and by the way both it and γ are extended as derivations, they agree on, in fact annihilate, scalars i.e. elements of $\Lambda^0(V^*)$. Hence it is sufficient to check that d^2 and γ agree on $V^* = \Lambda^1(V^*)$. Moreover, picking $X,Y,Z \in V$ and $\alpha \in V^*$, because of linearity the following calculation

is sufficient to obtain the result.

$$\begin{aligned} d^2\alpha(X,Y,Z) &= d\alpha([X,Y],Z) + d\alpha([Y,Z],X) + d\alpha([Z,X],Y) \\ &= \alpha([[X,Y],Z) + \alpha([[Y,Z],X]) + \alpha([[Z,X],Y]) \\ &= \alpha([[X,Y],Z] + [[Y,Z],X] + [[Z,X],Y]]) \\ &= \gamma^*\alpha(X,Y,Z) \end{aligned}$$

The first equality above follows from the fact d is a derivation. □

Corollary 5.1.2 Using the notation of the last proposition, V is a Lie algebra under $[\ ,\ ]$ iff $d^2 = 0$.

In this section the definition of a contact structure and several familiar examples are given. Thereafter, I set up contact geometry, much as one sets up symplectic geometry, and give coordinate free proofs of some classical results. Here is the principal definition.

Definition 5.2.1

A contact structure θ on a manifold N of dimension $2m+1$ is a 1-form θ such that $\theta\wedge(d\theta)^m$ is a volume form. The pair (N,θ) is together referred to as a contact manifold. ∎

Suppose that θ is a contact structure on N, then it is well known, imitating the Darboux theorem, that co-ordinates (x_i,z,p_i) may be introduced on N so that $\theta = dz - p_i dx_i$. I next give some examples which suggest that contact manifolds are worth studying.

Example 1.

The canonical 1-form θ on $J^1(M,\mathbb{R})$ is a contact structure.

Example 2.

Let $H: T^*M \to \mathbb{R}$. Then $H = 0$ gives a codimension one submanifold, at least locally, in T^*M. Restricting the canonical 1-form θ to this submanifold gives a contact structure. This situation occurs in Hamilton-Jacobi theory for example, as was shown in chapter 3.

Example 3.

Let $L: \mathbb{R} \times TM \to \mathbb{R}$. Choosing standard coordinates (t,x_i,u_i) for $\mathbb{R}\times TM$, the 1-form $\theta_L = Ldt + \frac{\partial L}{\partial u_i}(dx_i - u_i dt)$ is the Cartan 1-form, as was defined in section 4.2. It was shown there that away from the open set where $L = 0$, provided that the determinant

of the matrix $\frac{\partial^2 L}{\partial u_i \partial u_j}$ is nowhere zero, then θ_L is a contact structure.

Example 4.

Let $H: \mathbf{R} \times T^*M \to \mathbf{R}$. Denoting the canonical coordinate on $\mathbf{R}$ as t, the 1-form $\theta_H = \theta - Hdt$ is called the Poincaré-Cartan 1-form [1]. If H is nowhere zero then θ_H is a contact structure and clearly this is the appropriate context for studying time-dependent Hamiltonian mechanics.

In the remainder of this section I will always use (N,θ) to denote a contact manifold.

Definition 5.2.2

On a contact manifold (N, θ) the canonical characteristic vector field Z is the unique vector field such that

$$Z \lrcorner d\theta = 0 , \quad \langle Z, \theta \rangle = 1. \qquad \blacksquare$$

I next develop the contact analog of Hamiltonian theory which needs to be emended since $d\theta$ is degenerate.

Definition 5.2.3

If τ is a 1-form on (N, θ) then its associated vector field X_τ is uniquely that which satisfies

$$X_\tau \lrcorner d\theta = \langle Z, \tau \rangle \theta - \tau \quad \text{and} \quad \langle X_\tau, \theta \rangle = 0. \qquad \blacksquare$$

The last definition is mainly of interest in the case that $\tau = df$ for some $f \in F^0(N)$ in which case I write X_f rather than X_{df}. Notice also that from definition 5.2.2 $X_\theta \equiv 0$ and so θ spans the kernel of the map

from $F^1(N)$ to $V^1(N)$ by $\tau \mapsto X_\tau$. Using the definition one may construct the analog of the Poisson bracket on a symplectic manifold.

Definition 5.2.4

Given $f,g \in F^o(N)$ the Jacobi bracket $[f,g]$ is defined as $X_f g$. If one adopts Darboux coordinates for θ such that $\theta = dz - p_i dx_i$ then

$$X_f = \frac{\partial f}{\partial x_i}\frac{\partial}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial}{\partial x_i} + p_i \frac{\partial f}{\partial z}\frac{\partial}{\partial p_i} - p_i \frac{\partial f}{\partial p_i}\cdot\frac{\partial}{\partial z} \quad . \qquad \blacksquare$$

This coordinate representation yields one way to prove the next result; I shall not prove it since it is given with slight modifications in [5] and in any case, it is very similar to the analogous symplectic result.

Proposition 5.2.5 Let $f,g,h \in F^o(N)$ and $\lambda, \mu \in \mathbb{R}$. Then

(i) $[f,g] = -[g,f]$

(ii) $[\lambda f + \mu g, h] = \lambda[f,h] + \mu[g,h]$

(iii) $[f,[g,h]] + [g,[h,f]] + [h,[f,g]] = Zf[g,h] + Zg[h,f] + Zh[f,g]$.

(iv) $[fg,h] = f[g,h] + [f,h]g$. $\blacksquare$

Actually, Cartan used exterior calculus to give an alternative definition of $[\ ,\ ]$.

Proposition 5.2.6

Given $f,g \in F^o(N)$ $m df \wedge dg \wedge \theta \wedge (d\theta)^{m-1} = [f,g]\theta \wedge (d\theta)^m$. $\blacksquare$

Recalling the definition of the map γ in section 1.8 the following result justifies calling contact geometry a generalization of symplectic geometry.

Proposition 5.2.7

Given $f,g \in F^o(T^*M)$ $[\gamma^* f, \gamma^* g] = \gamma^*\{f,g\}$. $\blacksquare$

Proposition 5.2.8

On (N,θ) let (x_i,z,p_i) be Darboux co-ordinates for θ. Then

$$[x_i,x_j] = [p_i,p_j] = 0 \qquad [x_i,p_j] = \delta_{ij}$$

$$[z,x_i] = 0 \qquad [z,p_i] = p_i . \qquad \blacksquare$$

The last proposition is easy to show from Cartan's definition of $[\ ,\ ]$. Recall that any vector bundle has an intrinsically defined vector field Δ corresponding to the action of the multiplicative group of $\mathbb{R}^*$ on the fibers. If (p_i) are fiber coordinates this vector field, say Δ, is just $p_i \frac{\partial}{\partial p_i}$ as we saw in section 1.7.

Proposition 5.2.9

For $J^1(M,\mathbb{R})$ as a vector bundle over $M \times \mathbb{R}$, $\Delta = X_\beta$. (Recall that β: $J^1(M,\mathbb{R}) \to \mathbb{R}$ so β is in particular a function on $J^1(M,\mathbb{R})$). ∎

Again, the coordinate verification of the last proposition is straightforward. In proposition 5.2.5 we saw that $[\ ,\ ]$ does not satisfy the Jacobi identity. Thus, there can be no question of defining a Lie algebra map from $F^o(N)$ to $V^1(N)$ as in the symplectic case. However, we do have

Proposition 5.2.10

Let $f,g \in F^o(N)$. Then

$$[X_f,X_g] = X_{[f,g]} + (Zg)X_f - (Zf)X_g$$

where $[X_f,X_g]$ is the Lie bracket of vector fields and $[f,g]$ the Jacobi bracket of functions.

Proof: One may show that

$$([X_f,X_g] - X_{[f,g]}) \lrcorner d\theta = (Zf)dg - (Zg)df.$$

The result then follows from definition 5.2.3 and the observations that

$$((Zg)X_f - (Zf)X_g) \lrcorner d\theta = (Zf)dg - (Zg)df$$

and

$$\langle (Zg)X_f - (Zf)X_g, \theta \rangle = 0.$$ ∎

It follows from the last proposition that if one has a collection of functions $(f_i) \in F^o(N)$ such that $[f_i,f_j] = 0$ for all i,j, then the associated vector fields X_{f_i} form a completely integrable distribution.

The next definition is motivated by consideration of the space $J^1(M,\mathbb{R})$ and θ.

Definition 5.2.11

If (N, θ) is a contact manifold, a contact transformation is a map of N to itself such that there is a nowhere zero $\rho \in F^0(N)$ which satisfies $\Phi^* \theta = \rho\theta$. ∎ It follows automatically that if Φ is a contact transformation then Φ is a diffeomorphism since $\Phi^*\theta \wedge (d\theta)^m = \rho^{m+1}\theta \wedge d\theta^m$ and $\theta \wedge (d\theta)^m$ is a volume form.

Proposition 5.2.12

If Φ is a contact transformation $\Phi_* Z = X_\rho + \rho Z$.

Proof: $\Phi_* Z \lrcorner d\theta = Z \lrcorner d(\Phi^*\theta)$

$= Z \lrcorner d(\rho\theta)$

$= Z \lrcorner (d\rho \wedge \theta + \rho\, d\theta)$

$= (Z\rho)\theta - d\rho$ from the definition of Z. It follows from definition 5.2.3 that $\Phi_* Z = X_\rho + \lambda Z$. However, since $\Phi^*\theta = \rho\theta$

$$\rho = \langle Z, \Phi^*\theta \rangle = \langle \Phi_* Z, \theta \rangle = \langle X_\rho, \theta \rangle + \lambda \langle Z, \theta \rangle = \lambda.$$ ∎

Corollary 5.2.13

A contact transformation on $J^1(M,\mathbb{R})$ is fibred over T^*M iff ρ is a non-zero constant.

Proof: Note that Z is tangent to the fibration $\gamma : J^1(M,\mathbb{R}) \to T^*M$, so Φ preserves the fibration iff it preserves Z. Examining the form of X_ρ (see definition 5.2.3) this is only possible if $X_\rho \equiv 0$, which in turn holds precisely when $\rho \equiv$ constant. ∎

Proposition 5.2.14

If $\sigma, \tau \in F^1(N)$ then defining $W(\sigma,\tau) = \langle X_\sigma, \tau \rangle$ defines W as a 2-vector i.e. skew-symmetric (2,0) tensor field.

Proof: This is immediate from properties (i), (ii) and (iv) of proposition 5.2.5. ∎

In Darboux coordinates W has the form $(\frac{\partial}{\partial x_i} + p_i \frac{\partial}{\partial z}) \wedge \frac{\partial}{\partial p_i}$. It is clearly the analog of the cosymplectic structure on a symplectic manifold. However, since $[\ ,\ ]$ does not make $F^o(N)$ into a Lie algebra its Poisson-Jacobi tensor does not vanish.

Proposition 5.2.15

$W \llcorner \tau = 0$ where $\tau \in F'(N)$ iff $\tau = \rho\theta$ for some $\rho \in F(N)$.

Proof: Since $X_\theta \equiv 0$, $W \llcorner \rho\theta = \rho X_\theta = 0$. Conversely, $W \llcorner \tau = 0$ implies that $X_\tau = 0$ and so from $X_\tau \lrcorner d\theta = \langle Z, \tau \rangle \theta - \tau$ it follows that

$$\tau = \langle Z, \tau \rangle \theta \ . \qquad \blacksquare$$

The foregoing results enable one to prove a classical theorem (see [6]).

THEOREM 5.2.16

On (N, θ) the following conditions on a map Φ of N to itself are equivalent where in each case $\rho \in F^o(N)$ is never zero.

(i) Φ is a contact transformation i.e. $\Phi^* \theta = \rho\theta$.

(ii) $\Phi_* W = \rho W$

(iii) For all $f, g \in F^o(J^1(M, \mathbb{R}))$ $\quad \rho\Phi^*[f,g] = [\Phi^* f, \Phi^* g]$.

Proof: The equivalence of (ii) and (iii) is immediate from the definitions of W and $[\ ,\ ]$, so I shall show that (i)$\Rightarrow$(iii) and then that (ii)$\Rightarrow$(i).

(i)$\Rightarrow$(iii) Using the Cartan definition of $[f,g]$ and applying Φ^* to it gives $\Phi^*[f,g]\rho^{m+1}\theta \wedge (d\theta)^m = \rho^m \Phi^* df \wedge \Phi^* dg \wedge \theta \wedge (d\theta)^{m-1}$ and the result follows since $\rho \neq 0$.

(ii)$\Rightarrow$(iii) It follows from proposition 5.2.15 that W is a maximally non-degenerate 2-vector whose one-dimensional characteristic subspace is spanned by θ . Now (ii) ensures that $\Phi^* \theta = \lambda\theta$ for some $\lambda \in F(N)$ but since also

$\langle W, d\theta \rangle = 1$ it follows that $\lambda = \rho$. ∎

In his original paper (op. cit.) Lie gives two ways of "determining all contact transformations" of $J^1(M,\mathbb{R})$. The first way specifically requires the full structure of $J^1(M,\mathbb{R})$ and will be discussed in section 5.3. The second way, which shall be briefly examined now, holds for an arbitrary contact manifold. The underlying idea is that one may view the problem of finding a contact transformation as that of finding Darboux coordinates for θ or, to be rather more precise, θ up to a non-zero factor. Suppose then that

$$\theta = \rho(dz - p_i dx_i)$$

where $\rho, x_i, z, p_i \in F(N)$ with ρ nowhere zero. It is readily apparent from the Cartan definition of $[\ ,\ ]$ that $[z, x_i] = 0$ and $[x_i, x_j] = 0$; this observation yields the necessity of theorem 1 in Lie [25]. The converse of the same theorem argues that if one knows $m+1$ independent functions $\{x_i, z\}$, satisfying the above bracket relations then a local contact transformation is obtained; i.e. there exist functions $\{\rho, p_i\}$ ρ being nowhere zero and the p_i's independent such that θ has the canonical form given above. Moreover, ρ and the p_i's are uniquely determined simply by solving some linear equations. To see this result notice that immediately from Cartan's definition of $[\ ,\ ]$ one has $dz \wedge dx_1 \wedge \theta \wedge (d\theta)^{m-1} = 0$. Now one proceeds by induction, working firstly on the submanifolds $z = \text{const.}$, then $z = \text{const.}$, $x_1 = \text{const.}$, etc. This makes sense since the condition $[z, x_i] = 0$ has the geometric consequence that the vector fields X_{x_i} are tangent to the levels of z. Eventually one reaches the conclusion that

$$dz \wedge dx_1 \wedge \ldots \wedge dx_m \wedge \theta = 0$$

which essentially finishes the result.

5.3 Contact Transformations on $J^1(M,\mathbb{R})$

At the outset of this section it is worthwhile remarking that almost all of the results in section 5.2 held for an arbitrary contact manifold. By contrast, this section reaches back to the very origin of contact manifolds themselves and is concerned with $J^1(M,\mathbb{R})$ and the second description of contact transformations given by Lie. Since contact transformations are necessarily diffeomorphisms suppose that Φ is a diffeomorphism of $J^1(M,\mathbb{R})$ which does not necessarily respect any of the fibrations determined by α, β, γ or $\alpha\times\beta$ - actually, for reasons which will become apparent, we wish to consider $J^1(M,\mathbb{R})$ as being fibred over $M\times\mathbb{R}$. Even though Φ need not be fibered over $M\times\mathbb{R}$ we can associate to it an integer invariant k which measures to what extent Φ is fibered i.e. $\dim((\alpha\times\beta)\circ\Phi)_*$; in particular, if $k = m$ then Φ is fibered over $M\times\mathbb{R}$.

The key to understanding Lie's first description of contact transformations is the observation that such a transformation Φ may be interpreted geometrically as a $(2m+1)$-dimensional submanifold of $J^1(M,\mathbb{R})\times J^1(M,\mathbb{R})$ $\text{graph}(\Phi)$ which is transverse to both fibrations determined by projection onto either factor. As such, the following proposition, whose proof is a straightforward use of the theorem of implicit functions and which I therefore shall not supply, gives another geometric interpretation of k.

Proposition 5.3.1

With Φ as above, $\text{graph}(\Phi)$ lies in a $(4m-k+1)$-dimensional submanifold of $J^1(M,\mathbb{R})\times J^1(M,\mathbb{R})$ which is a union of fibers with respect to the projection map $(\alpha\times\beta)^2$ onto $(M\times\mathbb{R})\times(M\times\mathbb{R})$.

Continuing with the notation used above, it is best now to go to the opposite extreme from the case that Φ is fibered i.e. when $k=0$. For notational convenience I now index two copies of M as M_1 and M_2 and

likewise their respective geometric objects and ignore projection maps onto either factor of $J^1(M_1,\mathbb{R}_1) \times J^1(M_2,\mathbb{R}_2)$. Also, objects on $J^1(M_1,\mathbb{R}_1)$ and $J^1(M_2,\mathbb{R}_2)$ will be identified with objects on the product.

Propostion 5.3.2

Let $F: M_1 \times \mathbb{R}_1 \times M_2 \times \mathbb{R}_2 \to \mathbb{R}$ satisfy $Z_1F \neq 0$, $Z_2F \neq 0$ (cf. definition 5.2.5). These hypotheses ensure that the maps

$$F(\ ,\ ,x_2,z_2): M_2 \times \mathbb{R}_2 \to \mathbb{R} \ ((x_2,z_2) \text{ fixed})$$

and

$$F(x_1,z_1,\ ,\ ): M_2 \times \mathbb{R}_2 \to \mathbb{R} \ ((x_1,z_1) \text{ fixed})$$

can be solved to give maps $f_1: M_1 \to \mathbb{R}$, and $f_2: M_2 \to \mathbb{R}_2$ respectively. Now define a map $e: M_1 \times \mathbb{R}_1 \times M_2 \times \mathbb{R}_2 \to J^1(M_1,\mathbb{R}_1) \times J^1(M_2,\mathbb{R}_2)$ by

$$e(x_1,z_1,x_2,z_2) = (j^1_{x_1} f_1, j^1_{x_2} f_2).$$

Then $e^*(Z_1F\theta + Z_2F\theta_2) = dF$.

Proof: Compose e with $i_1 \times i_2$ (for the definition of i see section 1.8). This defines a 1-form on $M_1 \times \mathbb{R}_1 \times M_2 \times \mathbb{R}_2$ which is "not quite" dF. Denoting the canonical 1-forms on $M_1 \times \mathbb{R}_1$ and $M_2 \times \mathbb{R}_2$ by $\theta_{M_1 \times \mathbb{R}_1}$ and $\theta_{M_2 \times \mathbb{R}_2}$ respectively, pull back via $((i_1 \times i_2) \circ e)^*$ the 1-form $Z_1F\,\theta_{M_1 \times \mathbb{R}_1} + Z_2F\theta_{M_2 \times \mathbb{R}_2}$. It follows from the universal property that

$$(i_1 \times i_2) \circ e^*(Z_1F\,\theta_{M_1 \times \mathbb{R}_1} + Z_2F\theta_{M_2 \times \mathbb{R}_2}) = dF.$$

However, from the definition of θ

$$(i_1 \times i_2)^*(Z_1F\theta_{M_1 \times \mathbb{R}_1} + Z_2F\theta_{M_2 \times \mathbb{R}_2}) = (Z_1F)\theta_1 + (Z_2F)\theta_2$$

and the two equalities give the result. □

Corollary 5.3.3

Using the notation of the last proposition the map from the hyper-

surface $F=0$ in $M_1 \times \mathbb{R}_1 \times M_2 \times \mathbb{R}_2$ by

$$(x_1,z_1,x_2,z_2) \longmapsto (j^1_{x_1} f_1, j^1_{x_2} f_2)$$

pulls back $(Z_1F)\theta_1 + (Z_2F)\theta_2$ to zero. Since the image under this map is $(m+1)$-dimensional it determines a local contact transformation with $\rho = -Z_1F/Z_2F$ which, by hypothesis, is never zero. It is by no means clear that corollary 5.3.3 will yield global contact transformations, but one may wish to consider a generalized notion of contact transformation much as one does with Lagrangian submanifolds.

Recall that corollary 5.3.3 deals with the case $k=0$: I now show how to extend these ideas to $k>0$. Incidentally, it may be helpful to point out that in Lie's original paper $q = k-1$ and so $0 \leq q \leq m$. Suppose then that k independent functions $F_1,F_2,\ldots,F_k$ on $M_1 \times \mathbb{R}_1 \times M_2 \times \mathbb{R}_2$ are given. Then assuming that the levels $F_1=0$, $F_2=0,\ldots,F_k=0$ define a codimension k submanifold, which will certainly be true locally, one may consider $F = F_1 + \lambda_2F_2 + \lambda_3F_3 + \ldots + \lambda_kF_k$. As the λ's vary, $F=0$ defines a $(k-1)$-parameter family of submanifolds all intersecting in Σ. The classical word used to describe such a situation was "pencil". Suppose also that all the functions satisfy $Z_1F_i \neq 0$, $Z_2F_i \neq 0 \leq i \leq k$. Now one may consider the map from $\mathbb{R}^{k-1} \times \Sigma$ to $J^1(M_1,\mathbb{R}_1) \times J^1(M_2,\mathbb{R}_2)$ obtained simply by imitating the construction in proposition 5.3.2 for fixed $(\lambda_2,\ldots,\lambda_k) \in \mathbb{R}^{k-1}$. What is obtained is a $(2m+1)$-dimensional submanifold in $J^1(M_1,\mathbb{R}_1) \times J^1(M_2,\mathbb{R}_2)$ on which the 1-form $Z_1(F_1 + \lambda_2F_2 + \ldots + \lambda_kF_k)\theta_1 + Z_2(F_1 + \lambda_2F_2 + \ldots + \lambda_kF_k)\theta_2$ vanishes which generalizes corollary 5.3.3. Thus one obtains the converse to proposition 5.3.1 so that a suitably transverse $2m-k+1$ dimensional submanifold in $(M \times \mathbb{R})^2$ determines a unique contact transformation on $J^1(M,\mathbb{R})$. This statement corresponds to proposition 1 in Lie's paper.

5.4 One generalization of Symplectic Structure

In this section I want to investigate one direction in which symplectic structure may be generalized. In section 5.1 it was emphasized that the role of the closure of a symplectic form is important in that it ensures that the set of Hamiltonian vector fields forms a Lie subalgebra. With this remark in mind, I consider what carries over for a geometry based on a closed p-form. Suppose that Ω is a closed form on a manifold N of dimension n. Using Ω one may define a map $\lambda: V^1(N) \to F^{p-1}(N)$ by

$$X \mapsto -X \lrcorner \Omega \qquad X \in V^1(N).$$

Of course λ may also be viewed pointwise as a linear map from $T_x N$ to $\Lambda^{p-1}(T_x^* N)$ and as such its rank may vary with x. Now the dimension of $T_x M$ is n and the dimension of $\Lambda^{p-1}(T_x^* N)$ is $\binom{n}{p-1}$: thus λ has a cokernel of dimension at least $\binom{n}{p-1}-n$. On the other hand, the kernel of λ could have any dimension between 0 and n inclusive. I assume to begin with that λ has a trivial kernel at each $x \in N$. Now revent to the first view of λ, i.e. as a map of sections. Then clearly not every element of $F^{p-1}(N)$ lies in the image of λ. However, because pointwise λ has a trivial kernel, the image of λ will be a module of rank n. If $d\theta$ is in the image of λ then I will use the notation X_θ to denote its unique preimage.

I shall consider next equivalence classes of (p-2)-forms whose exterior derivative lies in the image of λ: two such p-2 forms θ and ψ will be called equivalent if $d(\theta - \Psi)=0$ and I will denote the set of such equivalence classes by $B(\Psi)$. I define a bracket on these equivalence classes as follows: if $\overline{\theta}$ and $\overline{\varphi}$ are two such equivalence classes then since

λ is injective a unique X_θ is determined and I may set

$$[\overline{\theta},\overline{\varphi}] = \overline{X_\theta \lrcorner d\varphi}. \qquad (5.4.1)$$

This bracket is clearly skew-symmetric, and so to show that it is well defined it suffices to check just the second argument. So suppose that for some $\alpha \in F^{p-3}(N)$ $\varphi = \Psi + d\alpha$. Then $d\varphi = d\Psi$ and so

$$[\overline{\theta},\overline{\Psi}] = \overline{X_\theta \lrcorner d\Psi}.$$

which shows that the bracket is well defined.

Next, suppose that $X_\theta \lrcorner \Omega = -d\theta$ and $X_\varphi \lrcorner \Omega = -d\varphi$. Then

$$\begin{aligned}
[X_\theta, X_\varphi] \lrcorner \Omega &= (L_{X_\theta} X_\varphi) \lrcorner \Omega \\
&= L_{X_\theta}(X_\varphi \lrcorner \Omega) - X\varphi \lrcorner (L_{X_\theta} \Omega) \\
&= L_{X_\theta}(-d\varphi) - X_\varphi \lrcorner (X_\theta \lrcorner d\Omega + d(X_\theta \lrcorner \Omega)) \\
&= -d(X_\theta \lrcorner d\varphi + d(X_\theta \lrcorner \varphi)).
\end{aligned}$$

Hence

$$[X_\theta, X_\varphi] \lrcorner \Omega = -d[\theta, \varphi] \qquad (5.4.2)$$

An important consequence of (5.4.2) is that $X_{[\theta,\varphi]} \lrcorner \Omega = -d[\theta,\varphi]$ and so

$$([X_\theta, X_\varphi] - X_{[\theta,\varphi]}) \lrcorner \Omega = 0 .$$

Since the map λ was assumed to have trivial kernel it follows that

$$([X_\theta, X_\varphi] = X_{[\theta,\varphi]} \qquad (5.4.3)$$

in other words the map $\overline{\theta} \to X_\theta$ defines a Lie isomorphism from B to, what in view of (5.4.2), is a Lie subalgebra of $V^1(N)$.

I examine now how the preceding arguments must be modified if λ is no longer assumed to be injective. Of course the notation X_θ does not make sense but (5.4.1) is still well defined. For suppose that $d\theta$ and $d\varphi$ lie in the image of λ and also that $X \lrcorner \Omega = -d\theta$, $Y \lrcorner \Omega = -d\theta$ and $Z \lrcorner \Omega = -d\phi$.

Then

$$\begin{aligned}(X-Y)\lrcorner d\varphi &= X\lrcorner d\varphi - Y\lrcorner d\varphi \\ &= -X\lrcorner Z\lrcorner\Omega + Y\lrcorner Z\lrcorner\Omega \\ &= Z\lrcorner(X-Y)\lrcorner\Omega \\ &= 0 \qquad \text{since } (X-Y)\lrcorner\Omega=0.\end{aligned}$$

Thus, $X\lrcorner d\varphi = Y\lrcorner d\varphi$ and so defining $[\overline{\theta},\overline{\varphi}]=X\lrcorner d\varphi$, where $d\theta$ and $d\varphi$ belong to the image of λ and X is any element of $V^1(N)$ such that $X\lrcorner\Omega=-d\theta$, makes sense.

Suppose next that $d\theta$ and $d\varphi$ belong to the image of λ and also that $X,Y\in V^1(N)$ satisfy $X\lrcorner\Omega=-d\theta$ and $Y\lrcorner\Omega=-d\varphi$ respectively. Then in analogy with (5.4.2) one has

$$[X,Y]\lrcorner\Omega=-d[\theta,\varphi] \qquad (5.4.4)$$

and again one may conclude that the X's which satisfy $X\lrcorner\Omega=-d\theta$ for some θ in $F^{p-2}(N)$ form a Lie subalgebra, say, $A(\Omega)$ of $V^1(N)$. It is convenient now to isolate the kernel of λ which I denote by $\chi(\Omega)$ and also call the characteristic distribution associated to Ω. With this terminology I can now state the following proposition.

Proposition 5.4.1 Suppose that Ω is a closed p-form on a manifold N and that $Z\in\chi(\Omega)$ and $X\in V^1(N)$ satisfies $d(X\lrcorner\Omega)=0$. Then $[X,Z]\lrcorner\Omega=0$.

Proof: $0=L_X(Z\lrcorner\Omega)=[X,Z]\lrcorner\Omega+Z\lrcorner L_X\Omega$. However, $L_X\Omega=X\lrcorner d\Omega+d(X\lrcorner\Omega)=0$ and so as required

$$[X,Z]\lrcorner\Omega=0. \qquad \square$$

Corollary 5.4.2 $\chi(\Omega)$ is a Lie subalgebra of $V^1(N)$ and hence, at least locally, defines a foliation of N.

Of course versions of this last corollary hold in greater generality, but it is important to note that $\chi(\Omega)$ is not usually an *ideal* in $V^1(N)$. However, one does have

Corollary 5.4.3 $\chi(\Omega)$ is an ideal in $A(\Omega)$.

Proof: $\chi \in A(\Omega) \Rightarrow \Omega = -d\theta$ for some θ in $F^{p-2}(N)$

$d(X \lrcorner \Omega) = 0$

and the result follows from proposition 5.4.1 immediately. □

Most of the preceding remarks may be subsumed in the following theorem the proof of which is a consequence of (5.4.4).

THEOREM 5.4.4 $0 \to \chi(\Omega) \to A(\Omega) \to B(\Omega) \to 0$ is a short exact sequence of Lie algebras.

I shall give some applications of this theorem presently, but before that I consider a slight variation on it. Continuing with the form Ω and the map λ etc. as before, define $A'(\Omega) = \{X \in V^1(N) \mid d(\lambda(X)) = 0\}$ Suppose that X and Y $\in A'(\Omega)$. Then, in a way almost identical to that used to derive (5.4.2), one finds

$$[X,Y] \lrcorner \Omega = d(X \lrcorner Y \lrcorner \Omega) \qquad (5.4.5).$$

Hence $[X,Y] \in A(\Omega)$ which is the analog of the well-known result [1] in symplectic geometry that the bracket of two locally Hamilonian vector fields is globally Hamiltonian. It follows from this that $A(\Omega)$ is actually an ideal in $A'(\Omega)$. Moreover, it is immediate from proposition 5.4.1 that $\chi(\Omega)$ is an ideal in $A'(\Omega)$.

Next, define B' to be the subspace of $F^{p-1}(N)$ consisting of those closed (p-1)-forms which belong to the image of λ. Then in a manner wholly analogous to theorem 5.4.4 one has

Theorem 5.4.5 $0 \to \chi(\Omega) \to A'(\Omega) \to B'(\Omega) \to 0$ is a short exact sequence of Lie algebras.

I shall try to show next that theorems 5.4.4 and 5.4.5 are the general principle which underlies the transference of an $\mathbb{R}$-Lie algebra structure on an algebra of vector fields to a subspace of forms of some fixed degree.

Applications of theorems 5.4.4 and 5.4.5

1. The simplest application is of course where $p = 2$ and Ω is a symplectic form. In this case the map λ is injective indeed bijective. If $p = 2$ but Ω is now longer symplectic theorem 5.4.4 still guarantees the existence of a Lie algebra on equivalence classes of a subspace of $F^o(N)$. Moreover, it is clear from (5.4.1) that this Lie algebra will satisfy the familiar derivation rule i.e. for $f,g,h \in$ im (λ) and denoting the corresponding elements of $B(\Omega)$ by $\bar{f}$, $\bar{g}$, $\bar{h}$

$$[\bar{f}\ \bar{g},\ \bar{h}] = [\bar{f},\bar{h}]\bar{g} + \bar{f}[\bar{g},\bar{h}] \qquad (5.4.6)$$

Hence for $p = 2$ $[,]$ is actually a Poisson bracket. Using corollary 5.4.2 one may say that, at least locally, N is fibered over a symplectic manifold and that $\chi(\Omega)$ spans the tangent space to the fibers. Notice that it is only in the case when $p = 2$ that (5.4.6) can apply, for only the pointwise product of functions makes sense, not of forms in general.

2. A second case in which λ is injective is when $p = n$ i.e. Ω is a volume form, in which case the condition $d\Omega = 0$ is empty. Indeed it follows from the algebraic fact any $(n-1)$-form on a manifold of dimension n is decomposable that λ is actually

surjective. Hence one has a Lie algebra structure on (equivalence classes of) a subspace of $F^{n-2}(N)$ and on the closed $(n-1)$-forms. This structure has been given before in the context of Louiville dynamics which is an approach to mechanics based on a volume form rather than a symplectic form [27].

3. A special case of $p = 2$ is worth investigating more closely i.e. when (N,Θ) is a contact manifold and $\Omega = d\Theta$. Recalling the definition of Z from section 5.2, and assuming it to be complete, N may be viewed as a principal bundle with structure group $\mathbb{R}$ over some base B, which need not in general be a smooth manifold [31]. Assuming that it is, then as in example (1) it will be a symplectic manifold. To specialize still further, suppose that $N = J^1(\mathbb{R},M)$ and that Θ_L is the Cartan form associated to a regular Lagrangian L. One then obtains a Poisson bracket structure on the first integrals of the Euler-Lagrange vector field Γ.

4. The last example may be generalized in a rather different way which leads to the Hamilton-Cartan formalism [15,35]. Let (E,π,M,F) be a bundle with M of dimension m and E of dimension $m + n$. As in section 1.6 $J^1(E)$ will denote the bundle of 1-jets of local sections of E over M. Suppose one is given a function L on $J^1(E)$ and a volume form ω on M. Then relative to these two choices there is an m-form Θ called the Cartan form. If (X^a, Z^A) are coordinates on E with (X^a) coordinates on M and (Z^A_a) denote the induced coordinates for the fibers of $J^1(E)$ over E. Θ is given locally by $\Theta = L\omega + \frac{\partial L}{\partial Z^A_a}(dz^A - z^A_b\, dx^b) \wedge \frac{\partial}{\partial X^a} \lrcorner\, \omega$. This of course is a

generalization of what was done in Chapter 4 because $\mathbb{R} \times M$ may be thought of as a trivial bundle over $\mathbb{R}$ and then $J^1(\mathbb{R},M)$ may be identified with $J^1(\mathbb{R} \times M)$.

The main point of the Hamilton-Cartan formalism is that provided the function L satisfies the regularity condition $\det\left(\frac{\partial^2 L}{\partial z_a^\Lambda \partial z_b^B}\right) \neq 0$ (this is mn by mn matrix) the solutions of the Euler-Lagrange equations for the extremals of the m-form $L\omega$ are the same as the m-dimensional submanifolds of the module of m-forms $\{V \rfloor d\Theta : V \in V^1(J^1(E))\}$, say Σ, which also satisfy the independence condition $\omega \neq 0$ [15]. In this context it is of interest to consider conserved currents which are equivalence classes of $(m-1)$-forms $\bar{\xi}$ on $J^1(E)$ such that $d\xi \in \Sigma$: two such $(m-1)$-forms are considered equivalent if $d\xi = d\eta$. Of course this is just a particular instance of theorem 5.4.4 which thus explains how the collection of conserved currents derives its Lie algebra structure.

REFERENCES

[1] Arnold, V. I. "Mathematical Methods of Classical Mechanics", Springer, New York, 1978.

[2] Abraham, R., Marsden, J. E., "Foundations of Mechanics", Benjamin, Reading, Mass., 1978.

[3] Calogero, F., J. Math. Phys. 12, 419, 1971.

[4] Cartan, E., Ann. Ec. Normale, t. 27, 109-192, 1910.

[5] Cartan, E., "Lecons sur les Invariants Integraux", Hermann, Paris, 1922.

[6] Corathedory, C. "Calculus of Variations and Partial Differential Equations of first order", Holden-Day, 1956.

[7] Crampin, M., Int. J. Th. Phys. 16, 1977, 741.

[8] Crampin, M., J. Phys. A. Math. Gen. 16, 3755-3772, 1983.

[9] Crampin, M., Prince, G. E., Thompson, G., J. Phys. A. Math. Gen. 17, 1437-47, 1984.

[10] Darboux, G., Archives Neerlandaises (ii) 6, 371, 1901.

[11] Dorizzi, B., Grammaticas, B., Ramani, A., J. Math. Phys. 24, 9, 1983.

[12] Ehresmann, C., "Les connexions infinitesimales dans un espace fibre"., Colloque de Topologie, Brussels, 1950.

[13] Fordy, A. P., Phys. Lett. 97A, 1, 2, 21, 1983.

[14] Goldschmidt, H., J. Diff. Geom. 1, 269-307, 1967.

[15] Goldschmidt, H., Sternberg, S., Ann. Inst. Fourier (Grenoble), 23, 1 203-67, 1973.

[16] Goldstein, H., "Classical Mechanics", Addison-Wesley, 1950.

[17] Hauser, I., Malhiot, R. J., J. Math. Phys. 15, 6, 816, 1974.

[18] Henon, M., Heites, C., Astron. J., 69, 73, 1964.

[19] Hermann, R., Interdisciplinary Math. V., Math. Sci. Press, Boston, 1977.

[20] Humphreys, J. E. "Lie Algebras and Representation Theory", Springer, New York, 1972.

[21] Husemoller, D., "Fiber Bundles"., Springer, New York, 1974.

[22] Jacobson, N., "Lie Algebras", Interscience, New York, London, 1962.

[23] Kalnins, E. G., Miller, W., SIAM J. M. A. 11, 1011-26, 1980.

[24] Lichnerowicz, A., J. Diff. Geom., 12, 253-300, 1977.

[25] Lie, S., Math. Ann. VIII, 2, 215-88.

[26] Makarov, A. A., Smorodinsky, J. A., Valiev, Kh., Winternitz, P., Nuovo Cim, 52A, 1061, 1967.

[27] Marmo, G., Saletan, E. J., Simoni, A., Zaccaria, F., J. Math. Phys. 224 835-43, 1981.

[28] Moser, J., Adv. in Math. 16, 197, 1975.

[29] Nijenhuis, A., Nederl. Akad. Wetensch. Proc. Ser. A58, 390, 1955.

[30] Noether, E., Nachr. kgl. Ges. Wiss. Gottingen, 235, 1918.

[31] Pirani, F. A. E., Crampin, M., "Geometrical Mechanics" (in preparation).

[32] Pommaret, J. F., "Systems of Partial Differential Equations and Lie Pseudogroups, Gordon and Breech, London, Paris, 1978.

[33] Prince, G. E., Bull. Austral. Math. Soc., 27, 53-71, 1983.

[34] Sarlet, W., Cantrijn, F., SIAM Rev. 23, 467-494, 1981.

[35] Shadwick, W. F., Lett. Math. Phys., 4, 241, 1980.

[36] Snapper, E., Troyer, R. J., "Metric Affine Geometry", Academic Press, New York, London, 1971.

[37] Steenrod, N., "The Topology of Fiber Bundles", Princeton, 1951.

[38] Sternberg, S., "Lectures on Differential Geometry", Prentice-Hall, Englewood Cliffs, New Jersey, 1964.

[39] Thompson, G., J. Phys. A: Math. Gen. 17, 958-8, 1984.

[40] Thompson, G., "Commentary on Lie's 1874-5 Contact Transformation Paper". Master's Thesis, Chapel Hill, North Carolina, 1983.

[41] - Vinogradov, A. M., Kuperschmidt, B. A., "The Structure of Hamiltonian Mechanics" in "Integrable Systems", London Math. Soc., Lecture Notes. Vol. 60.

[42] Whittaker, E., "A Treatise on the Analytical Dynamics of Particles and Rigid Bodies." Cambridge 1937.

[43] Woodhouse, N. M. J., Comm. Math. Phys. 44, 9, 1975.

[44] Wociechowski, S., Phys. Lett. 66A, 4, 265, 1978.

[45] Wociechowski, S., Phys. Lett. 100A, 6, 277, 1984.

[46] Xanthopoulos, B., J. Phys. A: Math. Gen. 17, 87-94, 1984.

[47] Yano, K., Ishihara, S., "Tangent and Cotangent Bundles", Marcel Dekker, New York, 19.