HAMILTONIAN SYSTEMS WITH NILPOTENT STRUCTURES

A thesis presented

by

Malcolm Irving

to

The Department of Engineering

Control Theory Centre

in fulfillment of the requirements

for the degree of

Doctor of Philosophy

University of Warwick

Coventry

July 1983

HAMILTONIAN SYSTEMS WITH NILPOTENT STRUCTURES

A thesis presented

by

Malcolm Irving

to

The Department of Engineering

Control Theory Centre

in fulfillment of the requirements

for the degree of

Doctor of Philosophy

University of Warwick

Coventry

July 1983

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am indebted to my supervisor Dr. P.E. Crouch who has been the source of many productive ideas and stimulating discussions during this period of research and I would like to record here my appreciation for all his help and encouragement. I also wish to thank Hazel Taylor for her excellent typing of a difficult manuscript.

To my parents.

SUMMARY

Symplectic Geometry has proved a powerful method in extending the knowledge of the classical theory of Hamiltonian mechanics without external variables. In this thesis these methods are applied to a class of Hamiltonian systems with controls in order to answer fundamental questions arising from Systems Theory and Classical Mechanics.

The theoretical aspects of this thesis deal with the extension of the Lie algebraic results of Engel and Lie on nilpotent and solvable Lie algebras, respectively, by the introduction of symplectic structures. It provides revealing results on the internal structure of symplectic vector spaces acted on by nilpotent or solvable Lie algebras. Then, using the methods of Kostant and Kirillov, these results are globalized to look at nilpotent transitive actions on simply connected symplectic manifolds and the consequent internal structures.

This theory is then applied to realizations of finite Volterra series with the additional property that the realization is Hamiltonian. These realizations are known to have an underlying nilpotent structure. A canonical realization is found and then shown to be closely linked with the theory of interconnections.

Finally, the concepts of complete integrability on free Hamiltonian systems is put into a feasible framework for Hamiltonian systems with controls which have an associated nilpotent Lie algebra. It is found that it is still possible to integrate these systems by quadratures but the structure is now much more complex.

Chapter 1

INTRODUCTION

Hamiltonian systems have been studied since early in the 19^{th} century starting with the equations of planetary motion. In more recent times there has been a great revival of interest in this area of Classical Mechanics with the advent of Symplectic Geometry, see for example [1], [2].

However, the systems investigated in the above texts are free systems i.e. with no external variables. Hamiltonian systems with controls and their corresponding properties of controllability and observability are a much more recent development, [3], [4]. This thesis aims to develop further the control theoretic aspects of this research to include an extension of the realization results of Crouch [5] to Hamiltonian systems. Also, it introduces a generalization of the classical complete integrability results of Arnold [2] to allow for the greater complexity of structure present in control systems.

Realization Theory

In the linear case it is well known that if the input-output map has the form

$$y(t) = \int_0^t W(t - \sigma_1)\, u(\sigma_1)\, d\sigma_1$$

where $u(\cdot)$ is a $\mathbb{R}^m$ - valued input function and $y(\cdot)$ is a $\mathbb{R}^q$ - valued output function and the Volterra kernel W satisfies

$$W(t - \sigma_1) = H(t)\, G(\sigma_1)$$

with W, H and G continuously differentiable maps, then it has a realization given by the linear system

$$\dot{x} = Ax + Bu \qquad x(o) = x_o$$

$$y = Cx$$

where $x \in \underset{\sim}{R}^n$, $A \in \underset{\sim}{R}^{nxn}$, $B \in \underset{\sim}{R}^{nxm}$, $C \in \underset{\sim}{R}^{qxn}$
and conversely.

In Brockett and Rahimi [6] it is shown that if the Volterra kernel $W(t - \sigma_1) = - W(\sigma_1 - t)$ then the realization is of the form

$$\begin{pmatrix} \dot{q} \\ \dot{p} \end{pmatrix} = \begin{pmatrix} A & B \\ -C & -A' \end{pmatrix} \begin{pmatrix} q \\ p \end{pmatrix} + \begin{pmatrix} \alpha \\ -\beta \end{pmatrix} u \qquad \begin{matrix} B = B' \\ C = C' \end{matrix} \tag{1.1.1}$$

$$y = p'\alpha + q'\beta \qquad (q, p)' \in \underset{\sim}{R}^{2n}$$

$$q(o) = q_o, \quad p(o) = p_o$$

where A, B, C, α, β are appropriately dimensioned matrices.

This is called a Hamiltonian system as it consists of Hamiltonian vector fields on $(\underset{\sim}{R}^{2n}, J)$ where J is the usual symplectic form

$$J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$$

If Hamiltonian functions are given by

$$H = \frac{1}{2} q' Cq + \frac{1}{2} p' Bp + p' Aq$$

$$H_u = p'\alpha + q'\beta$$

using standard notation (1.1.1) can be rewritten as

$$\dot{q} = \left({}^{\partial}/_{\partial p} \, (H + u H_u) \right)', \quad \dot{p} = - \left({}^{\partial}/_{\partial q} \, (H + u H_u) \right)'$$

$$y = H_u$$

Notice that if $u \equiv 0$ then this is exactly the form of Hamilton's equations.

The realizations considered in this thesis are a generalization of this linear case to that given by linear analytic systems of the form

$$\dot{x} = X_H(x) + \sum_{i=1}^{r} u_i X_{H_{u_i}}(x) \qquad x(o) = x_o$$

$$x \in (M, \omega)$$

$$y^i(t) = H_{u_i}(x) \tag{1.1.2}$$

where each vector field is Hamiltonian and all the data is real analytic.

In Brockett [7] conditions are given for finite Volterra Series of the form

$$y(t) = W_o(t) + \int_o^t W_1(t, \sigma_1)\, u(\sigma_1)\, d\sigma_1$$

$$+ \sum_{i=1}^{n} \int_o^t \int_o^{\sigma_1} \cdots \int_o^{\sigma_i} W_{i+1}(t,\sigma_1,\ldots,\sigma_{i+1}) u(\sigma_1)\ldots u(\sigma_{i+1}) d\sigma_{i+1}\ldots d\sigma_1$$

to have linear analytic realizations.

Work done jointly by the author and Dr. P.E. Crouch give the further necessary and sufficient conditions for the realization to be of the form (1.1.2).

<u>Liouville's Theorem</u>

As has been already mentioned the problem of planetary motion or the so called Kepler problem was one of the earliest known examples of a

Hamiltonian system. This is the problem of the motion of a particle in the gravitational field of a fixed point mass. It is modelled by the following system

$$\dot{x} = X_H(x) \qquad x \in (\underset{\sim}{R}^6, J) \tag{1.1.3}$$

where $H(q,p) = \frac{1}{2} p^2 - \alpha/_q$, $\alpha > 0$ some constant

$x = (q_1, q_2, q_3, p_1, p_2, p_3)'$ are the canonical coordinates of position and momentum on $\underset{\sim}{R}^6$ with

$$p^2 = \sum_{i=1}^{3} p_i^2, \quad q = \left(\sum_{i=1}^{3} q_i^2\right)^{\frac{1}{2}}$$

Liouville's theorem, as presented by Arnold [2] gives conditions for the existence of a structure preserving map $\psi : \underset{\sim}{R}^{2n} \to \underset{\sim}{R}^{2n}$ such that the Hamiltonian function H in the new coordinates $(I_1,\ldots,I_n, \phi_1\ldots,\phi_n)$ on $\underset{\sim}{R}^{2n}$ is of the form $H(I, \phi) = H(I)$ so the Hamiltonian vector field is

$$\dot{I}_i = \frac{\partial H}{\partial \phi_i} = 0, \quad \dot{\phi}_i = - \frac{\partial H}{\partial I_i} = \text{constant} \tag{1.1.4}$$

These are the action-angle coordinates and in particular it is the case that the system (1.1.3) satisfies these conditions and thus under ψ (1.1.3) is integrable by quadratures. This high-lights the importance of choice of canonical coordinates for determining the solution of a system.

In Liouville's theorem on free systems the underlying structure is abelian. With the introduction of controls this structure is normally lost. As a first step to find a generalization of Liouville's theorem with controls, nilpotent structures are considered and conditions sought for the existence of canonical coordinates for which the control system is integrable by quadratures.

Notation

The following notations will be used throughout this thesis assuming real analytic differentiability unless otherwise stated.

$\underset{\sim}{R}$	The real numbers
$\underset{\sim}{R}^n$	n - dimensional Euclidean space
$M(M^n)$	n - dimensional connected manifold
TM	Tangent bundle of M
T^*M	Cotangent bundle of M
T_xM	Tangent space to M at x
T_x^*M	Cotangent space to M at x
$V(M)$	Set of all vector fields on M
$\Omega^k(M)$	Set of all k-forms on M
(M^{2n},ω)	2n - dimensional connected symplectic manifold with symplectic form ω
$C(M)$	Ring of real valued functions on M
Diff (M)	Group of diffeomorphisms of M
$\mathcal{M}$ p.c.	Class of piecewise constant functions on $[o, \infty)$ taking values in $\underset{\sim}{R}^m$.

Chapter 2

SYSTEMS THEORY AND MECHANICS

This chapter contains the fundamental concepts and results that are needed in this thesis. Amplification of these can be found in Crouch [8] for systems theory and Abraham and Marsden [1] for the appropriate areas of Symplectic geometry.

2.1 Nonlinear Systems Theory

The class of system considered throughout this thesis is defined by the following equations.

$$\dot{x} = f(x) + \sum_{i=1}^{r} u_i \, g_i(x) \qquad x(o) = x_o, x \in M^n$$

$$y = h(x) \tag{2.1.1}$$

where the associated vector fields $f + \sum_{i=1}^{r} \alpha_i g_i$ for any $(\alpha_1, \ldots, \alpha_r) \in \mathbb{R}^r$, are complete analytic vector fields on M^n, an n-dimensional real analytic connected manifold, and h is a $\mathbb{R}^q$ valued analytic function on M.

Since the associated vector fields are complete, solutions to the above equations are defined on [0,T] for all piecewise constant controls u on [0,T], and all positive times T. Sussmann [9] enables an extension to include measurable controls. However, piecewise constant controls, $u \in \mathcal{U}_{p.c.}$, shall be used throughout.

For the case where the Lie algebra of (2.1.1) is finite dimensional, denoted by

$$\mathcal{L} = \{f, g_1, \ldots, g_r\}_{L.A.}$$

Palais [10] guarantees that $\mathcal{L}$ consists of complete analytic vector fields on M.

All Lie algebras throughout this thesis associated to systems of the form (2.1.1) are assumed to be finite dimensional unless otherwise specified.

Definition 2.1.1 A distribution Δ on a manifold M is the assignment to each $x \in M$ of a subspace $\Delta(x)$ of T_xM, Δ is said to be a C^K ($K = \infty, \omega$) distribution if it is spanned by a family D of C^K vector fields i.e. if $\Delta(x)$ is the linear subspace of T_xM spanned by the family $\{X(x)\}$ with $X \in D$ for every $x \in M$.

Δ is said to be m-dimensional if $\Delta(x)$ has dimension m for each $x \in M$.

Definition 2.1.2 A C^K - distribution Δ is involutive if the vector fields X, Y $\in \Delta$ then $[X, Y] \in \Delta$ i.e. $X(x), Y(x) \in \Delta(x) \Rightarrow [X, Y](x) \in \Delta(x)$, $x \in M$.

Definition 2.1.3 A submanifold N of M is said to be an integrable submanifold of Δ if $T_xN = \Delta(x)$ $\forall x \in N$, and a maximal integral submanifold of Δ is a connected integral submanifold of Δ maximal for the relation of inclusion.

Further, Δ is integrable if through every point of M there passes a maximal integral submanifold of Δ.

Theorem 2.1.4 (Nagano [11]). An analytic involutive distribution is integrable.

Theorem 2.1.5 (Frobenius [12]). A C^∞ constant dimensional involutive distribution is integrable.

The associated vector fields for system (2.1.1) are the vector fields

$$x \longmapsto f(x) + \sum_{i=1}^{r} u_i g_i(x) \text{ for each fixed } u \in \mathcal{U}_{p.c.}, \text{ assumed}$$

throughout to be complete, and denote the set of associated vector fields by F.

The group G_F (semi-group S_F) of F is defined to be the sub (semi -) group of diffeomorphisms of M of the form

$$G_F = \{g \in \mathrm{Diff}(M) : g = \gamma_{X_1}(t_1) \circ \ldots \circ \gamma_{X_n}(t_n),$$

$$X_i \in F,\ t_i \in \underset{\sim}{R},\ n \in \underset{\sim}{Z}^+\}$$

$$S_F = \{g \in \mathrm{Diff}(M) : g = \gamma_{X_1}(t_1) \circ \ldots \circ \gamma_{X_n}(t_n),$$

$$X_i \in F,\ t_i \in \underset{\sim}{R}^+,\ n \in \underset{\sim}{Z}^+\},$$

where $\gamma_{X_i}(t_i)$ is the flow of $X_i \in F$.

Definition 2.1.6 A linear analytic system is said to be accessible if $S_F(x)$ has nonempty interior in M, $\forall x \in M$.

The following theorem, Krener and Hermann [13], gives algebraic conditions for accessibility as well as a proof of importance in a later chapter.

Theorem 2.1.7 A linear analytic system is accessible if and only if $\dim \mathcal{L}(x) = n = \dim M$, $\forall x \in M$.

Proof Let $x \in M$ and U a neighbourhood of x. An open set V of M is constructed such that $V \subset U$ and each point $v \in V \subset S_F(x)$ - S_F system semi-group- is joined to x by a trajectory of the system remaining in U. In fact a sequence of submanifolds V_j of dimension j are constructed with these properties as follows:

For $j = 1$ choose $X_1 \in F$ such that $X_1(x) \neq 0$ and let $(t,x) \to \gamma_1(t)x$ be the flow of X_1. Choose $\varepsilon > 0$ such that $\gamma_1(t_1)x \in U$ for $0 \leq t_1 < \varepsilon$ then $V_1 = \{\gamma_1(t_1)x : 0 < t_1 < \varepsilon\}$ is a submanifold of dimension one which satisfies the requirements. V_j is now constructed by induction.

Suppose V_{j-1} satisfies the above requirements and is defined by

$$V_{j-1} = \{\gamma_{j-1}(t_{j-1}) \circ \ldots \circ \gamma_1(t_1)x : (t_{j-1}, \ldots, t_1) \in$$

an open subset of the positive orthant in $\underset{\sim}{R}^{j-1}\}$ where $(t,x) \to \gamma_k(t_k)x$ is the flow of a vector field $X_k \in F$

Now if $j < m$ find $X_j \in F$ and $v \in V_{j-1}$ such that $X_j(v) \notin T_v V_{j-1}$. If this was not possible $F_v \in T_v V_{j-1}$, $\forall v \in V_{j-1}$ and so $\mathcal{L}(v) \subset T_v V_{j-1}$, $\forall v \in V_{j-1}$ which implies that the distribution defined by $\mathcal{L}$ has dimension $j-1$ on $V_{j-1} \subset U$ which is a contradiction. It follows that the map $(t_j, \ldots, t_1) \to \gamma_j(t_j) \circ \ldots \circ \gamma_1(t_1)x$ is of full rank on some open subset of the positive orthant of $\mathbb{R}^j$. Thus by restriction if necessary it can be assumed that the range V_j of this map is a submanifold of M of dimension j with the required properties. The desired set V can now be taken to be V^m.

Conversely, if for $x \in M$ it is assumed that $r = \dim \mathcal{L}(x) < \dim T_xM = n$ then there exists a maximal integral submanifold N of dimension r passing through x. Using Nagano's theorem 2.1.4 clearly the reachable set $S_F(x)$ of the system from x is contained in N and so cannot have non-empty interior in M with respect to the topology on M. Therefore the system cannot be accessible, which is a contradiction.

Q.E.D.

Denote the reachable set from x_o at time $T \geq 0$ by $R(T, x_o)$, it is defined to be the set of all points $x_1 \in M$ such that there exists a piecewise constant control $u(\cdot)$ on $[0, T]$ and corresponding solution $x(\cdot)$ on $[0, T]$ such that $x(T) = x_1$.

The reachable set from x_o, denoted by $R(x_o)$ is defined by

$$R(x_o) = \bigcup_{T \geq 0} R(T, x_o)$$

Notice that $S_F(x_o) = R(x_o)$.

Definition 2.1.8 A linear analytic system is said to be strongly accessible if for each $x \in M$ there exists a $T > 0$ such that $R(x, T)$ has a nonempty interior in M.

Let S denote the ideal of $\mathcal{L}$ generated by $\{g_i\}^i_{L.A.}$, S has codimension one or zero in $\mathcal{L}$.

Corollary 3.4 Sussmann [14], shows that S(x) has constant dimension on M if the system is accessible and the following theorem, Sussmann [14], is a characterization of strong accessibility.

<u>Theorem 2.1.9</u> A linear analytic system is strongly accessible if and only if $\dim S(x) = n = \dim M$, $\forall x \in M$. In this case R(x, T) has non-empty interior in M $\forall T > 0$.

An immediate corollary from the proof of theorem 4.3 in Crouch [5] for systems which are strongly accessible and such that $S = \mathcal{L}$ is the following

<u>Theorem 2.1.10</u> Given a strongly accessible linear analytic system of the form

$$\dot{x} = f(x) + \sum_{i=1}^{r} u_i g_i(x) \qquad x(0) = x_o, \quad x \in M^n \tag{2.1.2}$$

such that S is a finite dimensional nilpotent Lie algebra and $S = \mathcal{L}$. Then there exists an analytic diffeomorphic change of coordinates such that (2.1.2) takes the form

$$\begin{array}{lll}
\dot{z}_1 = f_1 & + \sum_{i=1}^{r} u_i g^i{}_1 & z_1(0) = 0 \\
\dot{z}_2 = f_2(z_1) & + \sum_{i=1}^{r} u_i g^i{}_2(z_1) & z_2(0) = 0 \\
\quad\vdots & & \\
\dot{z}_j = f_j(z_1,\dots,z_{j-1}) & + \sum_{i=1}^{r} u_i g^i{}_j(z_1,\dots,z_{j-1}) & z_j(0) = 0 \\
\quad\vdots & & \\
\dot{z}_m = f_m(z_1,\dots,z_{m-1}) & + \sum_{i=1}^{r} u_i g^i{}_m(z_1,\dots,z_{m-1}) & z_m(0) = 0
\end{array}$$

where f_j, g^i_j are vector valued polynomials of the components of the vectors $z_i \in \underset{\sim}{R}^{n_i}$, $\sum_{i=1}^{m} n_i = n$, being the dimension of the state space.

If a linear analytic system is accessible then G, the connected Lie group corresponding to $\mathcal{L}$ the Lie algebra generated by the system, acts transitively on M, the state space. By a standard result, Helgason [15], M is therefore analytically diffeomorphic to the homogeneous space G/G_{x_o} where G_{x_o} is the isotropy group

$$G_{x_o} = \{g \in G : g.x_o = x_o\}$$

If the system is strongly accessible the connected Lie subgroup N of G, corresponding to S also acts transitively on M, and so M can be expressed as a homogeneous space N/N_{x_o}.

A number of concepts of observability will also be required throughout this thesis.

Definition 2.1.11 Two points x_o, $x_1 \in M$ are said to be indistinguishable if the input-output maps for the system initialized at x_o and x_1 are equal for all $u(\cdot) \in \mathcal{U}_{p.c.}$.

Definition 2.1.12 A system is said to be observable if x_o and $x_1 \in M$ being indistinguishable implies $x_o = x_1$.

Definition 2.1.13 A system is said to be weakly observable if for all states $x_o \in M$ there exists a neighbourhood U of x_o, such that if $x_1 \in U$ is indistinguishable from x_o then $x_o = x_1$.

Let $\mathcal{H}$ denote the smallest linear subspace of C(M) containing the functions h_i, $i=1,\ldots,q$ and closed under Lie differentiation by elements of $\mathcal{L}$. Thus $\mathcal{H}$ consists of all linear combinations of the functions.

$$L_{X_1}(L_{X_2}(\ldots(L_{X_n}(h_i)\ldots)\quad , X_i \in \mathcal{L}$$

The following theorem is found in Krener and Hermann [13],

<u>Theorem 2.1.14</u> A linear analytic accessible system is weakly observable if and only if

$T_x^*M = d\mathcal{H}(x) = \{d\tau(x) : \tau \in \mathcal{H}\}$, $\forall x \in M$ where T_x^*M is the cotangent space to M at x.

Given two systems $\sum_1$ and $\sum_2$, with the same control class and output space, $x_1 \in M_1$ and $x_2 \in M_2$ are said to be indistinguishable if $h_1 \circ g_1(x_1) = h_2 \circ g_2(x_2)$ where $g_1 \in G_{F_1}$ and $g_2 \in G_{F_2}$ is obtained using the same piecewise constant control as the one involved in g_1.

<u>Definition 2.1.15</u> $\sum_1$ and $\sum_2$ are strongly equivalent if every state in $\sum_1$ is indistinguishable from some state in $\sum_2$ and conversely.

For further results on strong equivalence refer to Goncalves [16].

2.2 Basic theory of Classical Mechanics

Most of the work in this thesis depends on a knowledge of Symplectic Geometry. This section presents the basic concepts. Good references for a more extensive survey are Abraham and Marsden [1] and Arnold [2].

Definition 2.2.1 Let V be a vector space and ω a non-degenerate two form on V. Then the pair (V, ω) is said to be a symplectic vector space.

A globalization of this gives the following definition.

Definition 2.2.2 Let M be a manifold with a nondegenerate closed two form ω defined on M. The pair (M, ω) is said to be a symplectic manifold. ω is often called a symplectic form.

The non-degeneracy of ω guarantees that M is even dimensional.

One of the most fundamental examples of a symplectic manifold is the cotangent bundle of a manifold i.e. $M = T^*Q$. This has a naturally defined symplectic structure as follows:

Let $\pi : T^*Q \longmapsto Q$ be the cotangent bundle projection. Define the 1 - form θ in T^*Q by

$$\theta(\alpha)(x) = \alpha_o \pi_*(x)$$

where $x \in T_\alpha T^*Q$ and π_* denotes the push forward.

Then the natural symplectic form ω is given by $\omega = -d\theta$.

The symplectic manifolds throughout this thesis are assumed to be finite dimensional unless otherwise specified.

Locally, the above forms can be written as

$$\theta = \sum_{i=1}^{n} p_i \, dq_i$$

$$\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$$

where $q_1,\ldots,q_n$ are coordinates on Q and $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ are canonical coordinates on $M = T^*Q$.

<u>Definition 2.2.3</u> Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. A C^K - mapping $f : M_1 \longmapsto M_2$ is called either symplectic or a symplectomorphism if $f^*\omega_2 = \omega_1$.

The definition of locally symplectic follows obviously from the above.

<u>Definition 2.2.4</u> Let (M, ω) be a symplectic manifold and $H \in C^K(M)$. The Hamiltonian vector field X_H on M is defined by

$$i\,(X_H)\omega = dH$$

where $i(X_H)\omega$ is the inner product of X_H and ω.

Locally, this corresponds to

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad i = 1,\ldots,n$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad i = 1,\ldots,n$$

where $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ are canonical coordinates on M and $\omega = \sum_{i=1}^{n} dq_i \wedge dp_i$.

These are Hamilton's equations.

Note that H is called the Hamiltonian function of the Hamiltonian vector field X_H.

Hamiltonian vector fields yield an important example of a symplectomorphism. For if X_H is a Hamiltonian vector field on (M, ω) then its flow $\gamma(t)$ is a symplectomorphism.

A weaker version of Hamiltonian vector field is a locally Hamiltonian vector field on (M, ω) which is a vector field X on (M, ω) such that for every $x \in M$ there exists a neighbourhood U of x such that X restricted to U is Hamiltonian. An equivalent form of this is that $L_X\omega = 0$, where $L_X\omega$ is the Lie derivative of ω with respect to X.

Definition 2.2.5 Let (M, ω) be a symplectic manifold and $f, g \in C^K(M)$ then the Poisson bracket of f and g is given by

$$\{f, g\} = \omega(X_f, X_g)$$

In canonical coordinates $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ on (M, ω) this is given by

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i}$$

Proposition 2.2.6 The real vector space C(M), together with the Poisson bracket { , } forms a Lie algebra.

A number of mathematical identities will be required repeatedly. These are listed here,

(i) $d\,d\,\alpha = 0$, $d(\alpha \wedge \beta) = d\,\alpha \wedge \beta + (-1)^k\,\alpha \wedge d\,\beta$, α - k-form

(ii) $f^* d\,\alpha = d\,f^* \alpha$, $f : M \longmapsto M$.

(iii) $L_X\,\alpha = d\,i(X)\,\alpha + i(X)d\,\alpha$, $X \in V(M)$

(iv) f a diffeomorphism, $f^* L_X\,\alpha = L_{f_*^{-1}X}\, f^*\alpha$

(v) $[X_f, X_g] = X_{\{f,\, g\}}$ on (M, ω).

Proofs and further details may be found in chapters 2 and 3 of Abraham and Marsden [1].

One of the main theorems in Classical Mechanics, and certainly one which has attracted much attention in recent years, is one associated with Liouville and Arnold. The form it is stated in here follows Arnold [2].

<u>Theorem 2.2.7</u> Let (M, ω) be a $2n$ - dimensional symplectic manifold, and consider the system

$$\dot{x} = X_H(x) \qquad x(0) = x_0,\ x \in M \tag{2.2.1}$$

Suppose there exists n - functions $f_1, \ldots, f_n$ on M in involution i.e. $\{f_i, f_j\} = 0$, $1 \leq i, j \leq n$ with $f_1 = H$ and let

$$M_{f(c)} = \{x : f_i(x) = c_i\ ,\quad i = 1, \ldots, n\}$$

with the n - one - forms df_i linearly independent at each point of

$M_{f(c)}$. Then

(i) $M_{f(c)}$ is a n - dimensional submanifold of M invariant under the flow of the Hamiltonian vector field X_H on (M, ω).

(ii) If $M_{f(c)}$ is compact and connected, then it is diffeomorphic to the n - dimensional torus T^n.

(iii) There exists canonical coordinates $(I_1,\dots,I_n, \phi_1,\dots,\phi_n)$ on (M, ω) such that $H = H(I_1,\dots,I_n)$ called the action - angle coordinates.

(iv) The canonical equations with Hamiltonian function H can be integrated by quadratures.

An excellent proof of this theorem may be found in Vinogradov and Kuperschmidt [17].

Throughout the field of Classical Mechanics and Symplectic Geometry much work has been done on finding systems of the form (2.2.1) which have n - functions in involution or perhaps fewer functions in involution but n - functions which commute with H under Poisson bracket see Nehorosev [18]. Much work has also been done in this area using the methodology of Kostant [19], Souriau [20] and Kirillov [21], see for instance Guillemin and Sternberg [22] or Marsden and Weinstein [23].

The basic difference between this and the work presented in later chapters is that commuting Hamiltonian functions under Poisson bracket are no - longer considered. The following example shows that for control systems the underlying structure is much more complex.

<u>Example 2.2.8</u> Consider the linear analytic dynamical system on $(\underset{\sim}{R}^4, \sum_{i=1}^{2} dq_i \wedge dp_i)$ with Hamiltonian functions defined by

$$H = p_1 + \frac{1}{2} p_2 q_1^2$$

$$H_u = p_1$$

so $\dot{x} = X_H(x) + u X_{H_u}(x) \qquad x(0) = x_o \qquad x \in (\underset{\sim}{R}^4, \omega)$

$$x = (q_1, q_2, p_1, p_2)'$$

That is
$$\begin{aligned} \dot{q}_1 &= 1 + u & q_1(0) &= q_1^o \\ \dot{q}_2 &= \frac{1}{2} q_1^2 & q_2(0) &= q_2^o \qquad (2.2.2) \\ \dot{p}_1 &= -p_2 q_1 & p_1(0) &= p_1^o \\ \dot{p}_2 &= 0 & p_2(0) &= p_2^o \end{aligned}$$

the Poisson brackets of the Hamiltonian functions are found as follows,

$$\{H_u, H\} = \sum_{i=1}^{2} \frac{\partial H_u}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial H_u}{\partial p_i} \frac{\partial H}{\partial q_i} = - p_2 q_1$$

$$\{H_u \{H_u, H\}\} = \quad p_2 \quad = \{H \{H_u, H\}\}$$

All other Poisson brackets vanish.

Notice that although H and H_u do not commute under Poisson brackets the system (2.2.2) can still be integrated by quadratures and further the level surface

$N = \{x \in \underset{\sim}{R}^4$: $f(x) = 0$ for all f such that $X_f \in \mathcal{L}$, $\mathcal{L} = \{X_H, X_{H_u}\}_{L.A.}\}$ is a 2 - dimensional submanifold of the symplectic manifold $(\underset{\sim}{R}^4, \omega)$ described by $p_1 = p_2 = 0$.

Also, on calculation of the Lie brackets of $\mathcal{L}$, $\mathcal{L}$ spans the tangent space of N at each point of N since

$$[X_{H_u}, X_H] = [{}^{\partial}/_{\partial q_1} \;, \; {}^{\partial}/_{\partial q_1} + \tfrac{1}{2} q_1^2 \, {}^{\partial}/_{\partial q_2} - p_2 q_1 \, {}^{\partial}/_{\partial p_1}]$$

$$= q_1 \, {}^{\partial}/_{\partial q_2} - p_2 \, {}^{\partial}/_{\partial p_1}$$

$$[X_{H_u} [X_{H_u} X_H]] = [{}^{\partial}/_{\partial q_1} \;, \; q_1 \, {}^{\partial}/_{\partial q_2} - p_2 \, {}^{\partial}/_{\partial p_1}]$$

$$= {}^{\partial}/_{\partial q_2} = [X_H [X_{H_u}, X_H]]$$

So $X_{H_u}(x)$ and $[X_{H_u} [X_{H_u} X_H]](x)$ span T_xN, $\forall\, x \in N$, all other Lie brackets vanish, so by theorem 2.1.7 this system is not accessible on $\underset{\sim}{R}^4$ but it is on the level surface N.

It is this type of generalization of Liouville's theorem that will come under investigation in Chapter 5.

Chapter 3

GENERAL THEORY OF NILPOTENT STRUCTURES ON SYMPLECTIC MANIFOLDS

This chapter begins with a section investigating the presence of nilpotent and solvable structures on a symplectic vector space extending the classical Lie algebraic results of Engel and Lie. The second section then considers a globalization to symplectic manifolds of the above. The appropriate conditions are found on a sequence of involutive distributions for the existence of canonical coordinates which have the same properties as found in section one. Finally, in the third section the existence of such sequences of involutive distributions is investigated using the techniques of Kostant [19] and Kirillov [21].

3.1 Hamiltonian Endomorphisms

This section investigates the new properties which arise when a symplectic structure is placed on the classical theorems of Engel and Lie, see Humphreys [24].

Let $\mathcal{L}$ be a Lie algebra. Define a sequence of ideals by

$$\mathcal{L}^1 = \mathcal{L} \ , \ \mathcal{L}^2 = [\mathcal{L},\mathcal{L}],\ldots,\mathcal{L}^k = [\mathcal{L},\mathcal{L}^{k-1}]$$

Recall that $\mathcal{L}$ is said to be nilpotent if $\mathcal{L}^n = 0$ for some n.

If $\mathcal{L}$ is any Lie algebra $X \in \mathcal{L}$ is ad-nilpotent if ad_X is a nilpotent endomorphism, i.e. $(ad_X)^n = 0$ for some n. If each element of a Lie algebra $\mathcal{L}$ is ad-nilpotent it is said to consist of nilpotent endomorphisms.

Proposition 3.1.1 Let $\mathcal{L}$ be a Lie algebra.

(i) If $\mathcal{L}$ is nilpotent then so are all sub and quotient algebras

(ii) If $\mathcal{L}/Z(\mathcal{L})$ is nilpotent, so is $\mathcal{L}$, where $Z(\mathcal{L})$ is the centre of $\mathcal{L}$.

(iii) $\mathcal{L}$ nilpotent, $\mathcal{L} \neq 0 \Rightarrow Z(\mathcal{L}) \neq 0$.

Theorem 3.1.2 (Engel) Let V be a non-zero finite dimensional vector space over $\underset{\sim}{R}$, and let $\mathcal{L}$ be a subalgebra of gl(V) consisting of nilpotent endomorphisms. Then

(i) $\mathcal{L}$ is nilpotent

(ii) There exists a vector $v \neq 0$ in V such that $X.\ v = 0$ for all $X \in \mathcal{L}$.

(iii) There exists a basis $e_{1,\ldots,}e_n$ in V such that all the endomorphisms $X \in \mathcal{L}$ are expressed by matrices with zeros on and above the diagonal.

It is part (iii) of Engel's theorem with vector space replaced by symplectic vector space that will be of particular interest.

First some notation,

Definition 3.1.3 If V is a vector space then a flag in V is a sequence of subspaces

$$0 = V_o \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V$$

such that $\dim V_i = i$.

If $X \in \mathrm{End}(V)$ and $X.\ V_i \subset V_i$ then X is said to stabilize the flag.

In terms of flags theorem 3.1.2(iii) could be rewritten as, there exists a flag in V such that $\mathcal{L}.V_i \subset V_{i-1}$. Since $V_{i-1} \subset V_i$, $\mathcal{L}$ stabilizes the flag.

Definition 3.1.4 Let (V, ω) be a symplectic vector space, then a linear mapping $X \in L(V, V)$ is infinitesimally symplectic with respect to the symplectic form ω if

$$\omega(Xv, v') + \omega(v, Xv') = 0$$

for all $v, v' \in V$.

Denote the set of all linear mappings in $L(V,V)$ that are infinitesimally symplectic with respect to ω by $sp(V, \omega)$.

Infinitesimal symplectic linear mappings are often referred to as linear Hamiltonian mappings.

An ordered basis can be chosen for V such that the matrix of ω is

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

where, if $\dim V = 2n$, I is the $n \times n$ identity matrix. Thus definition 3.1.4 says that $X \in sp(V, J)$ if

$$X'J + JX = 0$$

if and only if $X = \begin{array}{c c} & \begin{array}{cc} n & n \end{array} \\ \begin{array}{c} n \\ n \end{array} & \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \end{array}$ $\quad x_1 = -x_4'$, $x_2 = x_2'$, $x_3 = x_3'$

The matrix $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ is said to be a Hamiltonian matrix.

If $X_{1,\ldots,}X_n$ are infinitesimally symplectic their Lie algebra $\mathcal{L} = \{X_{1,\ldots,}X_n\}_{L.A.}$ is said to be a Lie algebra of Hamiltonian endomorphisms i.e. it is clear that each element of $\mathcal{L}$ is infinitesimally symplectic. In fact $sp(V, \omega)$ is a Lie subalgebra of $gl(V)$.

<u>Definition 3.1.5</u> Let (V, ω) be a symplectic vector space and $U \subset V$ a subspace. The ω-orthogonal complement of U is the subspace defined by

$$U^{\perp} = \{v \in V : \omega(v, u) = 0 \quad \forall\, u \in U\}$$

U is said to be

(i) isotropic if $U \subset U^{\perp}$ i.e. $\omega(u, u') = 0 \;\; \forall u, u' \in U$

(ii) co-isotropic if $U^{\perp} \subset U$ i.e. $\omega(u, u') = 0 \; \forall\, u' \in U \Rightarrow u \in U$

(iii) Lagrangian if U is isotropic and has an isotropic complement, i.e. $V = U \oplus U'$ where U' is isotropic.

The following two propositions the proofs of which appear for example, in Abraham and Marsden [1], are of continual use.

<u>Proposition 3.1.6</u> Suppose $U_1, U_2 \subset V$ are subspaces

(i) $U_1 \subset U_2 \Rightarrow U_2^{\perp} \subset U_1^{\perp}$

(ii) $U_1^{\perp} \cap U_2^{\perp} = (U_1 + U_2)^{\perp}$

(iii) $\dim V = \dim U_1 + \dim U_1^{\perp}$

(iv) $U_1 = U_1^{\perp\perp}$

(v) $(U_1 \cap U_2)^{\perp} = U_1^{\perp} + U_2^{\perp}$

Proposition 3.1.7 Let (V, ω) be a symplectic space and $U \subset V$ a subspace. Then the following statements are equivalent:

(i) U is Lagrangian

(ii) $U = U^{\perp}$

(iii) U is isotropic and $\dim U = \frac{1}{2} \dim V$.

Note that (iii) can be rephrased by saying that Lagrangian subspaces are maximal isotropic subspaces.

Lemma 3.1.8 Let $\mathcal{L}$ be a Lie algebra of Hamiltonian endomorphisms on a symplectic vector space (V, ω), and let U, W be subspaces of V such that $\mathcal{L} : U \longmapsto W$. Then,

$$\mathcal{L} : W^{\perp} \longmapsto U^{\perp} .$$

Proof. Let $X \in \mathcal{L}$. Since $\mathcal{L}$ consists of Hamiltonian endomorphisms X is infinitesimally symplectic, and therefore satisfies

$$\omega(X v_1, v_2) + \omega(v_1, X v_2) = 0 \quad \forall\, v_1, v_2 \in V$$

Let $v \in W^{\perp}$, $u \in U$

$$\omega(X v, u) = -\omega(v, Xu) = 0$$

since $X.u \in W$ which implies $X.v \in U^{\perp}$.

Q.E.D.

Theorem 3.1.9 Let $\mathcal{L}$ be a nilpotent Lie algebra of Hamiltonian endomorphisms on a 2n - dimensional symplectic vector space (V, ω). Then there exists a chain of isotropic subspaces V_i of dimension i

$$\{0\} = V_o \subset V_1 \subset V_2 \subset \ldots \subset V_{n-1} \subset V_n$$

where V_n is a Lagrangian subspace, and

$$\mathcal{L} . V_i \subset V_{i-1}, \qquad 1 \leq i \leq n$$

Moreover, $\mathcal{L} . V_{i-1}^{\perp} \subset V_i^{\perp}$, and so there exists a flag of the form

$$\{0\} = V_o \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \subset \ldots \subset V_1^{\perp} \subset V_o^{\perp} = V$$

with $\mathcal{L}$ stabilizing the flag.

Proof. By Engel's theorem there exists an $\alpha_1 \neq 0$ in V such that $\mathcal{L} . \alpha_1 = 0$. Set $V_1 = \mathbb{R}\, \alpha_1$.

Thus $\mathcal{L} : V_1 \hookrightarrow 0 \subset V_1$. Since V_1 is one dimensional it is obviously isotropic with respect to (V, ω) i.e. $V_1 \subset V_1^{\perp}$, and from lemma 3.1.8

$$\mathcal{L} : V_1^{\perp} \longmapsto V_1^{\perp}$$

Proceed by induction. Assume that V_i is an i-dimensional isotropic subspace of V with $\mathcal{L}: V_i \longmapsto V_{i-1} \subset V_i$ then $\mathcal{L}: V_i^{\perp} \longmapsto V_i^{\perp}$.

Consider the reduced space $V_i^{\perp}/V_i = W_i$. Now, since $\mathcal{L} : V_i \longmapsto V_i$ and $\mathcal{L} : V_i^{\perp} \longmapsto V_i^{\perp}$, induces an endomorphism $\mathcal{L}^{\dagger}$ on W_i. Moreover, since $\mathcal{L}$ is nilpotent, by proposition 3.1.1., $\mathcal{L}^{\dagger}$ is nilpotent on W_i.

Thus, there exists $0 \neq \beta_{i+1} \in W_i$ such that $\mathcal{L}^{\dagger} \cdot \beta_{i+1} = 0$. Let α_{i+1} be a representative for β_{i+1} in $V_i^{\perp} \subset V$. Then

$$\mathcal{L} \cdot \alpha_{i+1} \subset V_i$$

Define $V_{i+1} = V_i + \underset{\sim}{R}\, \alpha_{i+1}$, so $\mathcal{L} : V_{i+1} \longmapsto V_i \subset V_{i+1}$ and $\mathcal{L} : V_{i+1}^{\perp} \longrightarrow V_{i+1}^{\perp}$ by lemma. Claim V_{i+1} is isotropic i.e.

$$V_i + \underset{\sim}{R}\alpha_{i+1} \subset (V_i + \underset{\sim}{R}\, \alpha_{i+1})^{\perp}$$

But $(V_i + \underset{\sim}{R}\, \alpha_{i+1})^{\perp} = V_i^{\perp} \cap (\underset{\sim}{R}\, \alpha_{i+1})^{\perp}$ by proposition 3.1.6 (ii). By construction $\underset{\sim}{R}\, \alpha_{i+1} \subset V_i^{\perp}$ so $V_i \subset (\underset{\sim}{R}\, \alpha_{i+1})^{\perp}$

In particular, since $V_i \subset V_i^{\perp}$, $V_i \subset V_i^{\perp} \cap (\underset{\sim}{R}\, \alpha_{i+1})^{\perp}$ and since $\underset{\sim}{R}\, \alpha_{i+1} \subset (\underset{\sim}{R}\, \alpha_{i+1})^{\perp}$, $\underset{\sim}{R}\, \alpha_{i+1} \subset V_i^{\perp} \cap (\underset{\sim}{R}\, \alpha_{i+1})^{\perp}$

Thus,

$$\begin{aligned} V_{i+1} = V_i + \underset{\sim}{R}\, \alpha_{i+1} &\subset V_i^{\perp} \cap (\underset{\sim}{R}\, \alpha_{i+1})^{\perp} \\ &= (V_i + \underset{\sim}{R}\, \alpha_{i+1})^{\perp} \\ &= V_{i+1}^{\perp} \end{aligned}$$

Hence the claim is proved and V_{i+1} is isotropic and i+1 - dimensional. By lemma 3.1.8 since $\mathcal{L} : V_i \longmapsto V_{i-1}$ then $\mathcal{L} : V_{i-1}^{\perp} \longmapsto V_i^{\perp}$.

Proceed until V_n is reached, which will be maximal isotropic since $\dim V_n = \frac{1}{2} \dim V$, and therefore by proposition 3.1.7 is a Lagrangrian subspace. Also since $V_i \subset V_n \Rightarrow V_n^{\perp} \subset V_i^{\perp}$ there exists a flag

$$\{0\} = V_o \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \subset \ldots \subset V_1^{\perp} \subset V_o^{\perp} = V$$

with $\dim V_i = i \qquad 0 \leq i \leq n$

$\dim V_i^{\perp} = 2n-i \qquad 0 \leq i \leq n$

such that $\mathcal{L}$ stabilizes the flag.

Q.E.D.

This theorem shows that there exists an ordered basis on

V, $e_{q_1}, \ldots e_{q_n}$, $e_{p_1}, \ldots, e_{p_n}$ say, such that

$e_{p_1}, \ldots, e_{p_i}$ spans $V_i \qquad 1 \leq i \leq n$

and

$e_{p_1}, \ldots, e_{p_n}, e_{q_n}, \ldots, e_{q_{i+1}}$ spans $V_i^{\perp}$, $\quad 1 \leq i \leq n-1$

In coordinates relative to this basis theorem 3.1.9 says that each $X \in \mathcal{L}$ can be expressed by

$$X = \begin{pmatrix} N & 0 \\ -Q & -N' \end{pmatrix}$$

where $Q = Q'$ and N is a strictly lower triangular matrix, $N,Q \in \underset{\sim}{R}^{n\times n}$.

For example, consider a free system of the form

$$\dot{x} = X(x) \qquad x(0) = x_o , \qquad x \in (\underset{\sim}{R}^{2n}, J)$$

with $x = \begin{pmatrix} N & 0 \\ -Q & -N' \end{pmatrix}$, $x = (q_1,\dots,q_n, p_1,\dots,p_n)'$

$(q_1,\dots,q_n, p_1,\dots,p_n)$ canonical coordinates with respect to the above basis.

It is a Hamiltonian system with Hamiltonian H given by,

$$H = p' Nq + \frac{1}{2} q' Qq$$

$$q = (q_1,\dots,q_n)' , \qquad p = (p_1,\dots,p_n)'$$

So

$$\dot{x} = X_H (x) \qquad x(o) = x_o, \quad x \in (\underset{\sim}{R}^{2n}, J)$$

Equally, it is possible to choose an ordered basis $e_{q_1},\dots,e_{q_n}$, $e_{p_1},\dots,e_{p_n}$ on V such that $e_{q_1},\dots,e_{q_i}$ spans V_i

$e_{q_1},\dots,e_{q_n}, e_{p_n},\dots,e_{p_{i+1}}$ spans $V_i^{\perp}$ and express each element X in coordinates relative to this basis as,

$$X = \begin{pmatrix} N & P \\ 0 & -N' \end{pmatrix}$$

where N is strictly upper triangular matrix, $P = P'$, N, $P \in \underset{\sim}{R}^{n\times n}$. then for a typical free system the Hamiltonian function is

$$H = \frac{1}{2}\ p'\ Pp + p'\ Nq$$

where $(q,p)' = (q_1,\ldots,q_n,\ p_1,\ldots,p_n)'$ are the canonical coordinates relative to the symplectic basis $e_{q_1},\ldots,e_{q_n},\ e_{p_1},\ldots,e_{p_n}$.

Example 3.1.10 Suppose the state space is the symplectic vector space $(\underset{\sim}{R}^4,\ \omega)$ with canonical coordinates (q_1,q_2,p_1,p_2) and $\omega = \sum_{i=1}^{2} dq_i \wedge dp_i$. Consider the following bilinear system

$$
\begin{aligned}
\dot{q}_1 &= u & q_1(0) &= 0 \\
\dot{q}_2 &= q_1 + uq_1 & q_2(0) &= 0 \\
\dot{p}_1 &= -\ p_2 - up_2 & p_1(0) &= 0 \\
\dot{p}_2 &= -\ q_2 & p_2(0) &= 0
\end{aligned}
\qquad (3.1.1)
$$

This system is obviously Hamiltonian, if

$$\dot{x} = X_H + u\ X_{H_u} \qquad x(0) = 0$$

$$x = (q_1,\ q_2,\ p_1,\ p_2)'$$

then $H = p_2q_1 + \frac{1}{2}\ q_2^2$, $H_u = p_1 + p_2q_1$

So $X_H = q_1\ {}^{\partial}/_{\partial\, q_2} - p_2\ {}^{\partial}/_{\partial\, p_1} - q_2\ {}^{\partial}/_{\partial\, p_2}$

$X_{H_u} = {}^{\partial}/_{\partial q_1} + q_1\ {}^{\partial}/_{\partial q_2} - p_2\ {}^{\partial}/_{\partial p_1}$

(3.1.1) can be rewritten as

$$\dot{x} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & -1 & 0 & 0 \end{pmatrix} x + u \left[\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right]$$

Introduce new coordinates

$$\dot{q}_o = 0 \qquad \dot{p}_o = -u\, p_1$$

$$q_o(0) = 1 \qquad p_o(0) = 1$$

let $\bar{x} = (q_o, q_1, q_2, p_o, p_1, p_2)'$ then (3.1.1) becomes

$$\dot{\bar{x}} = \underset{A}{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}} \bar{x} + u \underset{B}{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}} \bar{x}$$

The Lie algebra $\mathcal{L} = \{A, B\}_{L.A.}$ is nilpotent and the subspaces in theorem 3.1.9 correspond to

V_1 is spanned by $\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix} = (e_{p_1})$

V_2 is spanned by $\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix} = (e_{p_1}, e_{p_2})$

V_3 is spanned by $\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix} = (e_{p_1}, e_{p_2}, e_{p_3})$

$V_2^{\perp}$ is spanned by $\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}$

$= (e_{p_1}, e_{p_2}, e_{p_3}, e_{q_3})$

$V_1^{\perp}$ is spanned by $\begin{bmatrix}0\\0\\0\\1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\0\\0\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\\0\\0\\0\end{bmatrix}$

$= (e_{p_1}, e_{p_2}, e_{p_3}, e_{q_3}, e_{q_2})$

In theorem 3.1.9 the dimension of each subspace was one greater than the previous subspace. The question now addressed is to whether a similar result is obtained when this is not necessarily so.

Theorem 3.1.11 Let (V, ω) be a symplectic vector space and

$$\{0\} = V_o \subset V_1 \subset \ldots \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \ldots \subset V_o^{\perp} = V$$

a sequence of subspaces contained in V. Then there exists symplectic basis of the form,

$$e_{q_1^1}, \ldots, e_{q_{r_1}^1}, \; e_{q_1^2}, \ldots, e_{q_{r_2}^2}, \ldots, e_{q_1^n}, \ldots, e_{q_{r_n}^n}$$

$$e_{p_1^1}, \ldots, e_{p_{r_1}^1}, \; e_{p_{r_1}^2}, \ldots, e_{p_{r_2}^2}, \ldots, e_{p_1^n}, \ldots, e_{p_{r_n}^n}$$

such that

$$e_{p_1^1}, \ldots, e_{p_{r_i}^i} \text{ spans } V_i \qquad 1 \leqslant i \leqslant n$$

and

$$e_{p_1^1}, \ldots, e_{p_{r_n}^n}, e_{q_{r_n}^n}, \ldots, e_{q_1^n}, \ldots, e_{q_{r_{i+1}}^{i+1}}, \ldots, e_{q_1^{i+1}} \text{ spans } V_i \quad 1 \leqslant i \leqslant n-1$$

Proof. By definition 3.1.5(iii), since V_n is Lagrangian there exists a V_n' such that $V = V_n \oplus V_n'$ and V_n' is a Lagrangian subspace. Furthermore let $v' \in V_n'$ then the map $i(v')\omega|_{V_n}$ defined by $\hat{\omega}$, $\hat{\omega} : V_n' \longmapsto V_n^*$, is obviously an isomorphism, since $\ker \hat{\omega} = \{0\}$ and dimension of V_n' is equal to that of V_n^*. Thus V can be rewritten as $V = V_n \oplus V_n^*$.

If $e_{q_1},\ldots,e_{q_n}$ is any basis for V_n, take the dual basis for V_n^* and let $e_{p_1},\ldots,e_{p_n}$ be the isomorphic basis of V_n', then $e_{q_1},\ldots,e_{q_n}, e_{p_1},\ldots,e_{p_n}$ is a symplectic bases for (V, ω). In this basis ω can be represented by the matrix

$$J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

Now proceed by induction. Find a basis for V_1, say $e_{\hat{p}_1^1},\ldots,e_{\hat{p}_{r_1}^1}$, which are of course, linear combinations of the e_{p_i}'s. Orthonormalise using the Gram-Schmidt process using the Euclidean metric, and then denote the resultant basis for V_1 by $e_{p_1^1},\ldots,e_{p_{r_1}^1}$.

Now find a vector $e_{\hat{p}_1^2}$ such that $e_{\hat{p}_1^2} \in V_2$ but $e_{\hat{p}_1^2} \notin V_1$. Again use the Gram-Schmidt process to ensure $e_{p_1^1},\ldots,e_{p_{r_1}^1}, e_{p_1^2}$ is Euclidean orthonormal.

Repeat this for a vector $e_{\hat{p}_2^2} \in$ span $\{V_1, e_{p_1^2}\}$. Continue this until an orthonormal basis for V_2 is found. The same procedure is repeated for $V_3,\ldots,V_n$.

Now let

$$e_{q_k^j} = J\, e_{p_k^j} \qquad 1 \leq j \leq n, \quad 1 \leq k \leq r_i, \; 1 \leq i \leq n$$

Clearly,

$$e_{p_j^i}'\ J e_{p_\ell^k} = 0$$

since $e_{p^i_j}$'s are linear combinations of the e_{p_i}'s which have this property.

Similarly,

$$e_{q^i_j}{}' \, J \, e_{q^k_\ell} = 0, \text{ since}$$

$$e_{q^i_j} = J \, e_{p^i_j} \quad \text{and } J' \, J \, J = J$$

Now $e_{q^i_j}{}' \; J \, e_{p^k_\ell} = e_{p^i_j}{}' \; e_{p^k_\ell}$, these are zero unless $i = k$ and $j = \ell$ by construction.

Q.E.D.

If the Lie subalgebra $\mathcal{L}$ of gl(V) over an algebraically closed field F is solvable then the classical theorem in the literature is Lie's theorem.

Recall that a Lie algebra $\mathcal{L}$ is solvable if it satisfies $\mathcal{L}^{(1)} = \mathcal{L}$, $\mathcal{L}^{(k)} = [\mathcal{L}^{(k-1)}, \mathcal{L}^{(k-1)}]$, $k > 1$ and there exists some integer n such that $\mathcal{L}^{(n)} = 0$.

<u>Proposition 3.1.12</u> Let $\mathcal{L}$ be a Lie algebra.

(i) If $\mathcal{L}$ is solvable then so are all sub and quotient algebras.

(ii) If I is an ideal of $\mathcal{L}$ and $\mathcal{L}$/I and I are solvable then so is $\mathcal{L}$.

(iii) If I and J are solvable ideals of $\mathcal{L}$ then I + J is a solvable ideal.

<u>Lemma 3.1.13</u> Let V be a non-zero, finite dimensional vector space and $\mathcal{L}$ a solvable Lie subalgebra over an algebraically closed field F of

gl(V). Then V contains a common eigenvector, v, for each element of $\mathcal{L}$.

i.e. $X.v = \lambda(X)v, \ \forall X \in \mathcal{L}$ where $\lambda : \mathcal{L} \longmapsto F$.

<u>Theorem 3.1.14</u> (Lie) If $\mathcal{L}$ is a solvable subalgebra of gl(V), $\{0\} \neq V$ a finite dimensional vector space, then $\mathcal{L}$ stabilizes some flag in V.

Replace the finite dimensional vector space by a finite dimensional symplectic vector space to obtain.

<u>Theorem 3.1.15</u> Let (V, ω) be a symplectic vector space of dimension 2n. If $\mathcal{L}$ is a solvable Lie subalgebra of Hamiltonian endomorphisms contained in gl(V) over an algebraically closed field F, then there exists a sequence of isotropic subspaces V_i of dimension i

$$\{0\} = V_o \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n$$

with V_n a Lagrangian subspace, and

$$\mathcal{L} : V_i \longmapsto V_i, \qquad 1 \leq i \leq n.$$

Moreover, $\mathcal{L} : V_i^{\perp} \longmapsto V_i^{\perp}$, and so there exists a flag of the form

$$\{0\} = V_o \subset V_1 \subset \ldots \subset V_{n-1} \subset V_n = V_n^{\perp} \subset V_{n-1}^{\perp} \subset \ldots \ V_1^{\perp} \subset V_o^{\perp} = V$$

and $\mathcal{L}$ stabilizes the flag.

Proof. Identical to the proof of theorem 3.1.9. Let $\alpha_1 \neq 0$ be a common eigenvector of all elements of $\mathcal{L}$, which exists by lemma 3.1.13., and consider the subspace $V_1 = F\alpha_1$. V_1 is isotropic since it is one-dimensional.

Again proceed by induction. Assume V_i is an i-dimensional isotropic subspace of V with $\mathcal{L} : V_i \longmapsto V_i$ then by lemma 3.1.8. $\mathcal{L} : V_i^{\perp} \longmapsto V_i^{\perp}$. Consider the reduced space $W_i = V_i^{\perp}/V_i$ and the induced Lie subalgebra of endomorphisms, $\mathcal{L}^{\dagger}$, on W_i. By proposition 3.1.12. $\mathcal{L}^{\dagger}$ is solvable and therefore there exists a common eigenvector $0 \neq \beta_{i+1} \in W_i$ such that $X^{\dagger}\beta_{i+1} = \lambda(X^{\dagger})\beta_{i+1}$, $\forall\ X^{\dagger} \in \mathcal{L}^{\dagger}$. Let α_{i+1} be a representative for β_{i+1} in $V_i^{\perp} \subset V$ then define

$$V_{i+1} = V_i + F\alpha_{i+1}$$

and then $\mathcal{L} : V_{i+1} \longmapsto V_{i+1}$ by construction.

That V_{i+1} is isotropic follows exactly as in theorem 3.1.9. Again, stop the procedure when V_n is reached. V_n is isotropic and $\dim V_n = \frac{1}{2}\dim V = n$ and therefore V_n is a Lagrangian subspace of (V, ω).

Use of proposition 3.1.6(i) and lemma 3.1.8 completes the flag which is stabilized under $\mathcal{L}$.

Q.E.D.

Essentially, this theorem says that there exists an ordered basis for (V, ω) such that each element of $\mathcal{L}$ can be expressed by a matrix of the form relative to the basis,

$$X = \begin{pmatrix} S & 0 \\ -Q & -S' \end{pmatrix} \qquad \text{where } S, Q \in F^{n\times n},\ Q = Q' \text{ and } S \text{ is lower triangular}$$

3.2 Canonical coordinates

In a globalization of theorem 3.1.11 one would assume the existence of a sequence of involutive distributions of the form

$$\{0\} = \Delta_o \subset \Delta_1 \subset \dots \subset \Delta_{n-1} \subset \Delta_n$$

on a symplectic manifold (M, ω) with

$$\Delta_n(x) = (\Delta_n(x))^{\perp}, \ \forall\, x \in M.$$

then find conditions for the existence of canonical coordinates $(q_1^1,\dots,q_{r_1}^1,\ q_1^2,\dots,q_{r_n}^n,\ p_1^1,\dots,p_{r_1}^1,\ p_1^2,\dots,p_{r_n}^n)$ on M such that locally

Δ_1 is spanned by $\partial/\partial p_1^1,\dots,\partial/\partial p_{r_1}^1$

$\vdots$

Δ_i is spanned by $\partial/\partial p_1^1,\dots,\partial/\partial p_{r_1}^1,\dots,\partial/\partial p_1^i,\dots,\partial/\partial p_{r_i}^i$

$\vdots$

Δ_n is spanned by $\partial/\partial p_1^1,\dots,\partial/\partial p_{r_1}^1,\dots,\partial/\partial p_1^n,\dots,\partial/\partial p_{r_n}^n$

This is the goal of this section.

Definition 3.2.1 Let (M, ω) be a symplectic manifold and $i : L \hookrightarrow M$ be an immersion. We say L is an isotropic (co-isotropic) immersed submanifold of (M, ω) if

$(T_x i)\ (T_x L) \subset T_{i(x)}M$ is an isotropic (co-isotropic) subspace for each $x \in L$.

The same terminology is used for submanifolds of M and for subbundles of TM over submanifolds of M.

A submanifold $L \subset M$ is called Lagrangian if it is isotropic and there is an isotropic subbundle $E \subset TM|_L$ such that $TM|_L = TL \oplus E$.

Note that $i : L \hookrightarrow M$ is isotropic if and only if $i^*\omega = 0$, and also if $L \subset M$ is Lagrangian then $\dim L = \frac{1}{2} \dim M$ and $(T_xL)^{\perp} = T_xL$, $\forall x \in L$.

<u>Proposition 3.2.2</u> Let (M, ω) be a symplectic manifold and $L \subset M$ a submanifold. Then L is Lagrangian if and only if L is isotropic and $\dim L = \frac{1}{2} \dim M$.

A number of preparatory results are required before the proof of the main theorem in this section can be given. The first says that symplectomorphisms preserve Hamiltonian vector fields.

<u>Theorem 3.2.3</u> (Jacobi) Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds and $\psi : M_1 \longmapsto M_2$ a diffeomorphism. Then ψ is symplectic if and only if for all $H \in C(M_2)$, $\psi_*^{-1} X_H = X_{H_o \psi}$.

The following interpretation of this theorem will be used repeatedly. Suppose X_f and X_g are Hamiltonian vector fields with Hamiltonian functions f and g respectively, on a symplectic manifold (M, ω) Let $(t, x) \longmapsto \gamma(t)x$ be the flow of $X_f(x)$, then since the flows of Hamiltonian vector fields are symplectic Jacobi's theorem says,

$$\gamma(-t)_* X_g (\gamma(t)\, x) = X_{g(\gamma(t))} (x)$$

The following results are standard to Symplectic Geometry and can be found for example in Abraham and Marsden [1] and Guillemin and Sternberg [22].

Theorem 3.2.4 (Darboux-Weinstein) Let N be a submanifold of the symplectic manifold M, and let ω_o and ω_1 be two non-degenerate closed two forms on M such that $\omega_o|_N = \omega_1|_N$. Then there exists a neighbourhood, U, of N and a diffeomorphism $f : U \longmapsto M$ such that $f|_N = \text{id}$ and $f^* \omega_1 = \omega_o$.

This extends the classical Darboux Theorem which is the above theorem when N is a point. It then states that if two symplectic forms agree on the tangent space at a point, then, up to a symplectomorphism, they agree in a neighbourhood of a point.

The case where N is a Lagrangian submanifold is of particular interest.

Corollary 3.2.5 (Kostant) Let L be a Lagrangian submanifold of a symplectic manifold (M, ω). Let L also be regarded as the zero section of T^*L and let ω_o be the natural symplectic form on T^*L. Then there exists a neighbourhood, U, of L in M and a diffeomorphism h of U into T^*L such that $h|_L = \text{id}$ and $h^* \omega_o = \omega_1$.

Definition 3.2.6 Let (M, ω) be a symplectic manifold. A polarization of M is an involutive distribution Δ of M such that $\Delta(x) \subset T_xM$ is a Lagrangian subspace for all $x \in M$.

The above type of polarization is usually referred to as a real polarization in the literature, it will be the kind most used in this thesis.

The fundamental example of a polarization is the foliation of a cotangent bundle by its fibres. In fact,

<u>Proposition 3.2.7</u> (Kostant-Weinstein) Suppose Δ is a polarization of a symplectic manifold (M, ω) which is transverse to a Lagrangian submanifold L. Then there is a symplectomorphism, f, of some neighbourhood, U, of L in M onto some neighbourhood, V, of the zero section of T^*L carrying the leaves of Δ onto the fibres of T^*L.

On T^*L there is a naturally defined one form, θ_L, such that $\omega_L = - d\theta_L$ is a symplectic form on T^*L. This form θ_L can be characterised as (i) $- d\theta_L = \omega_L$, (ii) $i(X)\, \theta_L = 0$ if X is tangent to the cotangent foliation and (iii) $\theta_L\big|_L = 0$ i.e. θ_L restricted to the zero section vanishes. By proposition 3.2.7. if Δ is any polarization of a symplectic manifold (M, ω) which is transversal to a Lagrangian submanifold L then locally, about L, there is a uniquely defined one form β satisfying (i) $d\beta = \omega$ (ii) $i(X)\beta = 0$ for X lying in Δ and (iii) $\beta\big|_L = 0$.

Let (M, ω) be a 2m - dimensional symplectic manifold and suppose there exists a sequence of involutive distributions on M of the form

$$\{0\} = \Delta_o \subset \Delta_1 \subset \ldots \subset \Delta_{n-1} \subset \Delta_n$$

such that they satisfy:

(1) Δ_n is a polarization of M which is transversal to a Lagrangian submanifold L in M.

This immediately implies that each Δ_i, $1 \leq i \leq n$ is isotropic in the sense that

$$\Delta_i(x) \subset \Delta_n(x) = (\Delta_n(x))^{\perp} \subset (\Delta_i(x))^{\perp}, \ \forall\, x \in M.$$

(2) there exists Hamiltonian vector fields on M,

$X^1_1,\ldots,X^1_{r_1},\ldots,X^n_1,\ldots,X^n_{r_n}$ such that $X^1_1,\ldots,X^1_{r_1},\ldots,X^i_1,\ldots,X^i_{r_i}$

span Δ_i for $1 \leq i \leq n$.

(3) $[X^i_j, X^k_\ell] \in \Delta_k$ for all $k \geq i$

(4) there exists m Hamiltonian vector fields $X_1,\ldots, X_m$ on M such that

$X_1(x),\ldots,X_m(x)$ spans T_xL $\forall x \in L$

and $[X_i, X^k_\ell] \in \Delta_k$, $1 \leq i \leq m$, $X^k_\ell \in \Delta_k$

$1 \leq k \leq n$

<u>Theorem 3.2.8</u> Let (M^{2m}, ω) be a symplectic manifold in which there exists a sequence of involutive distributions satisfying (1) - (4) above. Then there exists canonical coordinates $(q^1_1,\ldots,q^1_{r_1},\ldots,q^n_1,\ldots,q^n_{r_n}, p^1_1,\ldots,p^n_1,\ldots,p^n_{r_n})$ on a neighbourhood, U', of L in M such that $\partial/\partial p^1_1,\ldots,\partial/\partial p^1_{r_1},\ldots,\partial/\partial p^i_1,\ldots,\partial/\partial p^i_{r_i}$ spans Δ_i.

<u>Proof</u>. As the analysis performed here will be local, M can be identified with T^*L on application of corollary 3.2.5. Fix $x \in L \subset M = T^*L$ and let (U, ϕ) be a chart in T^*L, $x \in U$ on which there exist canonical coordinates $(\hat{q}_1,\ldots,\hat{q}_m, \hat{p}_1,\ldots,\hat{p}_m)$. Thus

$$\phi^* \omega = \sum_{i=1}^{m} d\hat{q}_i \wedge d\hat{p}_i$$

The necessary p coordinates are first constructed. Introduce the map

$\psi : \underset{\sim}{R}^m \longmapsto I(\Delta_n, x)$ ($I(\Delta_n, x)$ the integral submanifold of Δ_n through x) defined by

$$\psi : (p_1^1,\ldots,p_{r_1}^1,\ldots,p_1^n,\ldots,p_{r_n}^n) \longmapsto \gamma_1^1(p_1^1)\circ\ldots\circ\gamma_{r_1}^1(p_{r_1}^1)\circ\ldots\circ$$
$$\gamma_1^n(p_1^n)\circ\ldots\circ\gamma_{r_n}^n(p_{r_n}^n)x$$

where γ_j^i is the flow of X_j^i.

Let $p_i = (p_1^1,\ldots,p_{r_1}^1,\ldots,p_1^i,\ldots,p_{r_i}^i,o,\ldots,o)$

Now using exactly the same argument as in the proof of Krener's theorem 2.1.7 along with condition (2) a neighbourhood $\bar{U} \subset U$ with $x \in \bar{U}$ is obtained such that for $1 \leq i \leq n$,

$p_i \longmapsto \psi(p_i)$ is a diffeomorphism onto $\bar{U} \cap I(\Delta_i, x)$, where $I(\Delta_i, x)$ is the integral submanifold of Δ_i through x. These are the required p coordinates.

Let γ_i (q_i) be the flow of X_i, $1 \leq i \leq m$ and define the map $\Psi : \underset{\sim}{R}^{2m} \longmapsto \bar{U} \subset M$ by $\Psi(q, p) = \gamma(q)\circ\psi(p)$ where

$\gamma(q) = \gamma_1(q_1)\circ\ldots\circ\gamma_m(q_m)$

Claim. (a) Ψ is a diffeomorphism of some neighbourhood of the origin in $\underset{\sim}{R}^{2m}$ onto some neighbourhood $U' \subset \bar{U} \subset M$.

(b) $\Psi_* \,\partial/\partial p_j^i$ is a Hamiltonian vector on U'

(c) $\Psi_* \,\partial/\partial p_j^i$ spans Δ_i on U' for $1 \leq j \leq r_i$

$1 \leq i \leq n$

Suppose these claims are valid for the moment and proceed to prove the theorem.

Define the coordinates q^i_j, $1 \leq i \leq n$, $1 \leq j \leq r_i$ by

$$dq^i_j = - i(\partial/_{\partial p^i_j}) \quad (\Psi^* \omega) \tag{3.2.1}$$

for then if this were true

$$\Psi^* \omega = \sum dq^i_j \wedge dp^i_j.$$

For (3.2.1) to be valid the following must be checked.

(i) $dq^i_j \, (\partial/_{\partial p^\ell_m}) = 0$. That is $(\Psi^* \omega) \, (\partial/_{\partial p^\ell_m}, \partial/_{\partial p^i_j}) = 0$. But $\omega \, (\Psi_* \partial/_{\partial p^\ell_m}, \Psi_* \partial/_{\partial p^i_j}) = 0$ since by (c) $\Psi_* \partial/_{\partial p^i_j} \subset \Delta_i$ which by assumption is isotropic for $1 \leq i \leq n$.

(ii) $d \left(i(\partial/_{\partial p^i_j}) \right) (\Psi^* \omega) \quad = 0$

Since then $i(\partial/_{\partial p^i_j}) \, (\Psi^* \omega)$ can be expressed locally, by the derived derivative of some function.

For this Cartan's identity is required i.e.

$$L_X \, \omega = di(X)\omega + i(X)d\omega, \; X \in V(M), \; \omega \in \Omega^k(M)$$

so $d \left(i(\partial/_{\partial p^i_j}) \quad (\Psi^* \omega) \right)$

$= L_{\partial/_{\partial p^i_j}} \quad \Psi^* \omega \; - i(\partial/_{\partial p^i_j}) \quad d \left((\Psi^* \omega) \right)$

but $d\,\Psi^*\omega = \Psi^* d\omega = 0$ since ω is closed. Also,

$$L_{\partial/\partial p^i_j}\ \Psi^*\omega = \Psi^* L_{\Psi_*\partial/\partial p^i_j}\ \omega$$

But from (b) $\Psi_* \partial/\partial p^i_j = X$ is a Hamiltonian vector field on U' thus $L_X\ \omega = 0$ and $d\left(i(\partial/\partial p^i_j)\ (\Psi^*\omega)\right) = 0$ as required.

(iii) $\Psi^*\omega = \bar{\omega}$, clearly $d\bar{\omega} = 0$ and non-degeneracy follows from (a).

Therefore (3.2.1) is a valid choice of q^i_j. This is the required coordinate system satisfying the conditions of the theorem, since they are now seen to be equivalent to (c).

Now it just remains to prove claims (a), (b) and (c).

(a) Proof. Differentiating the expression for Ψ and evaluating at $q = 0$, $p = 0$ gives

$$\partial/\partial q_i\ \ \Psi(q, p)\Big|_{(o,o)} = X_i\ (x)$$

$$\partial/\partial p^i_j\ \ \Psi(q, p)\Big|_{(o,o)} = X^i_j\ (x)$$

These vectors span T_xM by construction, now on application of the inverse function theorem the result is obtained.

(b) Proof. Since

$$\Psi_*\ \partial/\partial p^i_j\Big|_{q,\ p} = \gamma(q)_* \frac{\partial\psi}{\partial p^i_j}(p) \qquad (3.2.2)$$

it is necessary to evaluate $\dfrac{\partial\psi}{\partial p^i_j}$

This is equal to

$$\gamma^1_1(p^1_1)_*\ldots_*\gamma^i_{j-1}\,(p^i_{j-1})_*X^i_j(\gamma^i_j(p^i_j)_o\ldots_o\gamma^n_{r_n}(p^n_{r_n}).x) \qquad (3.2.3)$$

where X^i_j is the Hamiltonian vector field on M with Hamiltonian function H^i_j. Then by Jacobi's theorem

$$\gamma^i_{j-1}\,(p^i_{j-1})_*\,X_{H^i_j}\left(\gamma^i_{j-1}\,(-\,p^i_{j-1})\,\cdot\right) \text{ can be expressed by}$$

$$X_{H^i_j}\circ\gamma^i_{j-1}(-\,p^i_j)\,(\,\cdot\,)$$

Thus (3.2.3) can be written as,

$$X_{H^i_j}\circ\gamma^i_{j-1}(\,-\,p^i_{j-1})_o\ldots\gamma^1_1(\,-\,p^1_1)^{(y)}$$

where $y = \gamma^1_1(p^1_1)_o\ldots_o\gamma^n_{r_n}\,(p^n_{r_n})\;x = \psi(p)$ and (3.2.2) becomes, since $\gamma_i(q_i)$ $1 \leq i \leq m$ are flows of Hamiltonian vector fields and thus symplectic,

$$X_{H^i_j}\,{}_o\gamma^i_{j-1}(-p^i_{j-1})_o\ldots_o\gamma^1_1(-p^1_1)_o\gamma_m(-q_m)_o\ldots_o\gamma_1(-q_1)\;\left(\Psi\,(q,\,p)\right)$$

and this can be re-expressed by considering

$$H^i_{j_o}\;\gamma^i_{j-1}(-p^i_{j-1})_o\ldots_o\gamma_1(\,-q_1)\left(\Psi\,(q,p)\right)$$

$$=H^i_{jo}\;\gamma^i_j\;(p^i_j)_o\ldots_o\gamma^n_{r_n}(p^n_{r_n})(x)$$

let $k^i_j : \underset{\sim}{R}^{2m} \longmapsto \underset{\sim}{R}^{2m}$ be defined by

$$k^i_j(q_1,\ldots,q_m, p^1_1,\ldots, p^1_{r_n}) = (o,\ldots,o, p^i_j,\ldots,p^n_{r_n})$$

therefore (3.2.4) can be written as

$$X_{H^i_j o\, \Psi o\, k^i_j o\, \Psi^{-1}}\left(\Psi\,(q,p)\right)$$

which is a Hamiltonian vector field on U'.

(c) <u>Proof</u>. Again proved by differentiation and repeated use of the Campbell-Baker-Hausdorff formula.

$$\partial/_{\partial p^i_j}\ \Psi(q,\, p) = \gamma(q)_*\ \partial/_{\partial p^i_j}\left(\psi(p)\right)$$

But

$$\partial/_{\partial p^i_j}\ \psi(p) = \partial/_{\partial p^i_j}\left(\gamma^1_1(p^1_1)o\ldots o\gamma^n_{r_n}\ (p^n_{r_n})\ .x\right)$$

$$= \gamma^1_1(p^1_1)_{*\ldots *}\gamma^i_{j-1}(p^i_{j-1})_* X^i_j\ (\gamma^i_{j-1}\ (-p^i_{j-1})o\ldots o\gamma^1_1\ (-p^1_1)y)$$

where $y = \gamma^1_1\ (p^1_1)o\ldots o\gamma^n_{r_n}\ (p^n_{r_n}).x$

But by the Campbell-Baker-Hausdorff formula and condition (3) this is seen to be contained in Δ_i. However, by condition (4) it is clear that on further application of the Campbell-Baker-Hausdorff formula that $\gamma(q)_*\ \Delta_i \subset \Delta_i$. Further, since Ψ is a diffeomorphism on U' it follows that

$$\partial/\partial p^i_j\, \Psi\,(q,\ p) \;=\; \Psi_*\ \partial/\partial p^i_j \Big|_{(q,\ p)}$$

spans $\Delta_i\,\big(\Psi\,(q,\ p)\big) \quad 1 \le j \le r_i,\ 1 \le i \le n$

in a neighbourhood U' of $x \in L$ in M.

Q.E.D.

A generalization of theorem 3.1.9 to symplectic manifolds would involve searching for a sequence of isotropic distributions and the existence of a Lie algebra consisting of Hamiltonian vector fields on a symplectic manifold $(M,\ \omega)$ which in some sense stabilizes the sequence.

Let $\mathcal{L}$ be a Lie Algebra of Hamiltonian vector fields on a symplectic manifold $(M,\ \omega)$ i.e. $\mathcal{L} = \{X_{H_1},\dots,X_{H_m}\}_{L.A.}$ Suppose there exists a sequence of involutive distributions satisfying the conditions of theorem 3.2.8 and let

$$q_1 = (q^1_1,\dots,q^1_{r_1}),\dots,q_n = (q^n_1,\dots,q^n_{r_n})$$

$$p_1 = (p^1_1,\dots,p^1_{r_1}),\dots,p_n = (p^n_1,\dots,p^n_{r_n})$$

and also let $f(q^1_1,\dots,p^n_{r_n})\ \partial/\partial p_j$ represent

$$f^1(q^1_1,\dots,p^n_{r_n})\ \partial/\partial p^j_1 + \dots + f^{r_j}\ (q^1_1,\dots,p^n_{r_n})\,\partial/\partial p^j_{r_j}\ , \text{ and}$$

e.g. $\dfrac{\partial H}{\partial p_j} = \begin{pmatrix} \dfrac{\partial H_j}{\partial p^j_1} \\ \vdots \\ \dfrac{\partial H_j}{\partial p^j_{r_j}} \end{pmatrix}$, for $1 \le j \le n$.

Then, locally

Δ_1 is spanned by $\partial/\partial p_1$

Δ_2 is spanned by $\partial/\partial p_1$, $\partial/\partial p_2$

$\vdots$

Δ_n is spanned by $\partial/\partial p_1$, $\partial/\partial p_2, \ldots, \partial/\partial p_n$.

This can be used to give the following generalization of theorem 3.1.9.

Theorem 3.2.9 Let (M, ω) be a symplectic manifold. Let $\mathcal{L}$ be a Lie algebra consisting of Hamiltonian vector fields on M and suppose

$$\{0\} = \Delta_o \subset \Delta_1 \subset \ldots \subset \Delta_{n-1} \subset \Delta_n$$

is a sequence of involtive distribution on M satisfying the conditions of theorem 3.2.8. with local coordinates as described above.

If each $X_H \in \mathcal{L}$ with Hamiltonian function H satisfies

$$[X_H, \Delta_i] \subset \Delta_i \qquad 1 \leq i \leq n \tag{3.2.5}$$

then locally each H has the form

$$H = p_1 f_1(q_1) + \ldots + p_n f_n(q_1, \ldots, q_n) + Q(q_1, \ldots, q_n)$$

Proof. Locally $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ are canonical coordinates on M and each Hamiltonian vector field X_H on M can be expressed by

$$X_H = \sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \, {}^{\partial}/_{\partial q_i} - \frac{\partial H}{\partial q_i} \, {}^{\partial}/_{\partial p_i}$$

Consider the case i = n in (3.2.5) and, without loss of generality, calculate

$$[X_H, f \, {}^{\partial}/_{\partial p_j}] \;, \quad f : \underset{\sim}{R}^{2N} \longmapsto \underset{\sim}{R}^{r_j} \;, \quad j \in \{1,\dots,n\}$$

an analytic map, $N = \sum_{k=1}^{n} r_k$.

$$[X_H, f \, {}^{\partial}/_{\partial p_j}] = [\sum_{i=1}^{n} \frac{\partial H}{\partial p_i} \, {}^{\partial}/_{\partial q_i} - \frac{\partial H}{\partial q_i} \, {}^{\partial}/_{\partial p_i} \,, \, f \, {}^{\partial}/_{\partial p_j}]$$

$$= \sum_{i=1}^{n} \left(\frac{\partial H'}{\partial p_i} \frac{\partial f}{\partial q_i} \, {}^{\partial}/_{\partial p_j} - \frac{\partial H'}{\partial q_i} \frac{\partial f}{\partial p_i} \, {}^{\partial}/_{\partial p_j} \right.$$

$$\left. - f' \frac{\partial^2 H}{\partial p_i \partial p_j} \, {}^{\partial}/_{\partial q_i} + \frac{f' \partial^2 H}{\partial q_i \partial p_j} \, {}^{\partial}/_{\partial p_i} \right) \quad (3.2.6)$$

$$(3.2.5) \Rightarrow \frac{\partial^2 H}{\partial p_i \partial p_j} = 0, \qquad 1 \leq i, j \leq n$$

$\Rightarrow$ H is linear in p

thus $H = \sum_{i=1}^{n} p_i' \, f_i \, (q_1,\dots,q_n) + Q \, (q_1,\dots,q_n)$.

For the case i = n -1 it is obvious from (3.2.6) that (3.2.5) implies

$$\frac{\partial^2 H}{\partial q_n \partial p_j} = 0 \qquad 1 \leq j \leq n-1$$

thus $H = \sum_{i=1}^{n-1} p'_i f_i\ (q_1,\ldots,q_{n-1}) + p'_n f_n\ (q_1,\ldots,q_n) + Q(q_1,\ldots,q_n)$

Proceed by induction. Assume $[X_H, \Delta_{k+1}] \subset \Delta_{k+1}$ implies

$$H = \sum_{i=1}^{k+1} p'_i f_i\ (q_1,\ldots,q_{k+1}) + \sum_{i=k+2}^{n} p'_i f_i (q_1,\ldots,q_i)$$

$$+ Q(q_1,\ldots,q_n) \qquad (3.2.7)$$

then consider $[X_H, \Delta_k] \subset \Delta_k$. By calculation (3.2.6)

$$\frac{\partial^2 H}{\partial q_i \partial p_j} = 0, \qquad 1 \leq j \leq k < i \leq n$$

which by (3.2.7) implies

$$H = \sum_{i=1}^{k} p'_i\ f_i\ (q_1,\ldots,q_k) + \sum_{i=k+1}^{n} p'_i f_i (q_1,\ldots,q_i) + Q(q_1,\ldots,q_n)$$

as required.

Continue this procedure until i=1 in (3.2.5) which then gives the required result. Q.E.D.

Essentially this theorem says that systems of the form

$$\dot{x} = X_{H_1}(x) + \sum_{i=2}^{m} u_i X_{H_i}(x), \quad x \in (M, \omega)$$
$$x(o) = x_o$$

with $\mathcal{L} = \{X_{H_1},\ldots,X_{H_m}\}_{L.A.}$

and satisfying the above conditions, can be integrated by quadratures.

Example 3.2.10 Consider the following system on the symplectic manifold $(\underset{\sim}{R}^2 \setminus \{0\}, \quad dq \wedge dp)$

$$\dot{x} = X_H(x) + u\, X_{H_u}(x) \qquad x(0) = x_o, \quad x \in \underset{\sim}{R}^2 \setminus \{0\}$$

where $x = (q, p)'$

where $H = \frac{1}{2} p^2 + \frac{1}{2} q^2$

$$H_u = q \Big/ (p^2 + q^2)^{\frac{1}{2}}$$

Therefore,

$$X_H = p\, {}^{\partial}/_{\partial q} - q\, {}^{\partial}/_{\partial p}$$

$$X_{H_u} = - \frac{p}{(p^2+q^2)^{3/2}} \left(q\, {}^{\partial}/_{\partial q} + p\, {}^{\partial}/_{\partial p} \right)$$

Let $L = \{x \in \underset{\sim}{R}^2 \setminus \{0\} : \frac{1}{2} p^2 + \frac{1}{2} q^2 = \text{constant}\}$

Obviously L is a Lagrangian submanifold with tangent space spanned by $p\, {}^{\partial}/_{\partial q} - q\, {}^{\partial}/_{\partial p}$.

Let Δ_1 be spanned by $\frac{q}{p^2+q^2}\, {}^{\partial}/_{\partial q} + \frac{p}{p^2+q^2}\, {}^{\partial}/_{\partial p}$ a Hamiltonian vector field on $\underset{\sim}{R}^2 \setminus \{0\}$ with Hamiltonian function $\tan^{-1} p/q$. It is clear that Δ_1 is transversal to L.

Changing to polar coordinates gives

$$p = r\cos\theta$$

$$q = r\sin\theta$$

$$\begin{aligned} dq \wedge dp &= (\sin\theta\, dr + r\cos\theta\, d\theta) \wedge (\cos\theta\, dr - r\sin\theta\, d\theta) \\ &= r\cos^2\theta\, d\theta \wedge dr - r\sin^2\theta\, dr \wedge d\theta \\ &= d\theta \wedge d(r^2/_2) \end{aligned}$$

Let $\phi = r^2/_2$, $I = \theta$ then

$$H(I, \phi) = \phi$$

$$H_u(I, \phi) = \sin I$$

a symplectic change of coordinates. Thus

$$\dot{I} = 1 \qquad\qquad I(o) = I_o$$

$$\dot{\phi} = -u\cos I \qquad\qquad \phi(o) = \phi_o$$

Obviously integrable by quadratures and satisfies

$$[X, \Delta_1] \subset \Delta_1 \quad \forall X \in \mathcal{L}$$

where $\mathcal{L} = \{X_H, X_{H_u}\}_{L.A.}$

Notice $\mathcal{L}$ is solvable and S the ideal in $\mathcal{L}$ generated by X_{H_u} is nilpotent i.e.

$$S = \{\cos I\ \partial/_{\partial\phi},\ \sin I\ \partial/_{\partial\phi}\}_{L.A.}$$

In fact S is abelian.

This leads into the investigation of the existence of such distributions as described in this section.

3.3 Existence of Sequences of Involutive Distributions

In the previous section the existence of a sequence of involutive distributions was assumed and conditions found for a particularly interesting set of canonical coordinates. Here, it is shown by the use of concepts developed by Kostant [19], Kirillov [21] and Souriau [20], how by the introduction of a nilpotent structure that the above objects arise naturally.

The main references used in this section will be Abraham and Marsden [1], Guillemin and Sternberg [22] and Wallach [25].

Definition 3.3.1 Let (M, ω) be a symplectic manifold. It is called a symplectic G - space if G is a Lie group which acts on M by diffeomorphisms such that if

$$\phi : G \times M \longmapsto M$$

$$\phi\ (g, x) = \phi_g\ (x) = g.x$$

then $\phi_g^* \omega = \omega, \ \forall\ g \in G$ (3.3.1)

Furthermore (M, ω) is called a homogeneous symplectic G - Space if G acts transitively on M.

Throughout the rest of this section G will be assumed to be a connected Lie group unless otherwise stated.

Define

$$d\phi\ (X)\ (f)\ (x) = \frac{d}{dt}\bigg|_{t=o} \Big(f\ (\exp - tX.x)\Big),\ f \in C(M)$$

which gives a map from $\mathcal{L}$ to $V(M)$, where $X \in \mathcal{L}$, the Lie algebra associated to G.

It can be shown that $d\phi$ is a Lie algebra homomorphism i.e.

$$[d\phi\ (X),\ d\phi(Y)] = d\phi\ ([X,\ Y]),\quad \forall X,\ Y \in \mathcal{L}$$

By (3.3.1) $d\phi(X)$ is a locally Hamiltonian vector field for all $X \in \mathcal{L}$. However, if

$$d\phi(X) \subset \text{Ham}\ (M,\ \omega)$$

where Ham (M, ω) is the set of all Hamiltonian vector fields on (M, ω), then with this further condition (M, ω) is called a strongly symplectic G - Space.

<u>Definition 3.3.2</u> Let (M, ω) be a strongly symplectic G - space. A lift, λ, of $d\phi : \mathcal{L} \longmapsto \text{Ham}\ (M, \omega)$ is a Lie algebra homomorphism

$$\lambda : \mathcal{L} \longmapsto C(M)$$

such that

$$\begin{tikzcd}
0 \arrow[r] & \underset{\sim}{R} \arrow[r] & C(M) \arrow[r] & \text{Ham}(M,\ \omega) \arrow[r] & 0 \\
 & & & \mathcal{L} \arrow[ul, "\lambda"] \arrow[u, "d\phi"'] &
\end{tikzcd}$$

commutes

Homogeneous simply connected symplectic manifolds can be completely classified, as in Wallach [25], with the aid of co-adjoint actions.

The adjoint representation is a map from the Lie group G to the isomorphisms of $\mathcal{L}$, the Lie algebra of G, i.e.

$$Ad : G \longmapsto Aut\ (\mathcal{L}) \text{ defined by}$$

$$Ad_g\ X = \frac{d}{dt}\Big|_{t=o}\ g \exp t\ X\ g^{-1}$$

Ad_g is called the adjoint operator. More importantly, the co-adjoint is defined by

$$Ad_g^* : \mathcal{L}^* \longmapsto \mathcal{L}^* \qquad \forall g \in G$$

$$Ad_g^*\ f\ (X) = \langle f, Ad_{g^{-1}}\ X\rangle,\ f \in \mathcal{L}^*,\ X \in \mathcal{L}$$

where < , > denotes the duality pairing.

Also, note that Ad_g is a Lie algebra homomorphism for all $g \in G$. Let 0 be an orbit of G acting on $\mathcal{L}^*$ i.e. $0 = Ad_G^*\ f,\ \ f \in \mathcal{L}^*$.
0 has the manifold structure of G/G_f where $G_f = \{g \in G : Ad_g^*\ f = f\}$
Define an action on 0 by G as

$$\sigma : G \times 0 \longmapsto 0$$
$$\sigma(g, f) = Ad_g^*\ f, \quad \forall f \in 0, \forall g \in G.$$

Also note that $T_f\ 0 \cong \mathcal{L}/\mathcal{L}_f$ where $\mathcal{L}_f = \{X \in \mathcal{L} : d\sigma\ (X)\ (f) = 0\}$

Lemma 3.3.3 Define a skew form $B_f(X, Y) = -f([X, Y])$, $X, Y \in \mathcal{L}$, $f \in \mathcal{L}^*$, then B_f is a bilinear form with kernel $\mathcal{L}_f$.

Thus B_f induces a non-degenerate 2-form $\omega(f)$ on $\mathcal{L}/\mathcal{L}_f$ and hence gives a non-degenerate 2 - form on $T_f\, O$ defined by

$$w(f)\big(d\sigma(X)(f),\ d\sigma(Y)(f)\big) = -f([X, Y])$$

Hence, define $\omega_o \in \Omega^2(O)$ by $\omega_o\big((Ad_g^* f)\big)\big(d\sigma(X)\ (Ad_g^* f),\ d\sigma(Y)\ (Ad_g^* f)\big)$ $= -Ad_g^* f\ ([X, Y])$, for $g \in G$, $f \in O$.

Suppose a lift exists on a symplectic G-space (M, ω) then the action $\phi : G \times M \longmapsto M$ is said to be Poisson. (M, ω) may then be referred to as a Poisson G-space.

Theorem 3.3.4 (O, ω_o) is a symplectic manifold and the action $\sigma : G \times O \longrightarrow O$ is Poisson with lift given by

$$\lambda(X)f = f(X), \qquad \forall f \in O.$$

Let (M, ω) be a Poisson G-space. Fix $x \in M$, then there exists a linear mapping

$$\tau(x) : X \longrightarrow \lambda(X)(x)$$

so $\tau(x) : \mathcal{L} \longmapsto \mathbb{R}$

and thus $\tau : M \longmapsto \mathcal{L}^*$

This is usually referred to as the moment map. It acts as a map which intertwines the action on M with the coadjoint action on $\mathcal{L}^*$.

In the situation of theorem 3.3.4 $\tau : O \longmapsto \mathcal{L}^*$ is the inclusion map.

Lemma 3.3.5 If $\tau : M \longmapsto \mathcal{L}^*$ is the moment map of a Poisson action ϕ then

$$\tau_o \phi_g = Ad^*_g {}_o \tau.$$

This is equivalent to saying that τ is Ad^* - equivariant.

Definition 3.3.6 If (M_1, ω_1) and (M_2, ω_2) are Poisson G-spaces with moment maps τ_1 and τ_2 respectively then a morphism is a smooth map $f : M_1 \longmapsto M_2$ such that $f^* \omega_2 = \omega_1$ and $\tau_2 {}_o f = \tau_1$.

Theorem 3.3.7 (Kostant-Souriau) Let (M, ω) be a homogeneous Poisson G-space with moment map τ. Then $\tau(M) = O \subset \mathcal{L}^*$ is an orbit, τ is a morphism from (M, ω) to (O, ω_o) and a covering map.

Corollary 3.3.8 There is a bijection between simply connected homogeneous Poisson G-spaces and co-adjoint orbits.

It is not always the case that a lift exists for a given Lie Algebra $\mathcal{L}$. Normally Lie algebra cohomology provides necessary conditions for existence and uniqueness. Here, however, an alternative method is considered for the specific case of interest. Further, it it always possible to extend the Lie algebra $\mathcal{L}$ to gain a lift, as will be shown.

Suppose $\tau : M \longmapsto \mathcal{L}^*$ is a moment map on a symplectic G-space (M, ω). Let O be an orbit of G acting on $\mathcal{L}^*$ and set

$$N_o = \tau^{-1}(0)$$

Let $x \in N_o$ and $f = \tau(x)$. Define

$$V_x = (d\tau)^{-1}(x)(T_f 0).$$

Assume N_o is a smooth manifold at x, so V_x is the tangent space at x. By definition

$$\tau(x)(X) = \lambda(X)(x), \quad \forall X \in \mathcal{L}.$$

Let $\xi \in V_x$ then the value of

$d\tau(\xi) \in \mathcal{L}^*$ on $X \in \mathcal{L}$ is $d\lambda(X)(x)(\xi) = \omega(x)(X_{\lambda(X)}, \xi)$.

Let $\sigma : G \times 0 \longmapsto 0$ be the action on 0 defined by

$$\sigma(g, f) = Ad_g^* f$$

$d\sigma(X)(f)$ is the tangent vector to 0 at f, $X \in \mathcal{L}$, i.e.

$$d\sigma(X)(f) \in T_f 0.$$

Thus $\xi \in V_x$ if there exists a $Y \in \mathcal{L}$ such that

$$\omega(x)(X_{\lambda(X)}, \xi) = d\sigma(Y)(f)(X), \quad \forall X \in \mathcal{L}. \tag{3.3.2}$$

Now

$$V_x^{\perp} = \{X_{\lambda(X)}(x) \in T_xM : \omega(x)\big(X_{\lambda(X)}(x), \xi(x)\big) = 0 \quad \forall\, \xi \in V_x\}$$

From (3.3.2.) if $X_{\lambda(X)} \in V_x^{\perp}$ then $d\sigma(Y)(f)(X) = 0$, $\forall\, X \in \mathcal{L}$.

By definition of $\mathcal{L}_f = \{X \in \mathcal{L} : d\sigma(X)(f) = 0\}$ thus $V_x^{\perp} = \mathcal{L}_f(x)$ where $\mathcal{L}_f(x)$ is the corresponding subspace of the tangent space of M at x.

Theorem 3.3.9 If $\phi : G \times M \longmapsto M$ is a nilpotent transitive action on a symplectic G-space (M, ω) then this action is not Poisson.

Proof. Assume the action is Poisson, so there exists a moment map $\tau : M \longmapsto \mathcal{L}^*$. Since the action is transitive $V_x = T_xM$ $\forall\, x \in M$ which implies that $V_x^{\perp} = \{0\}$.

Let $Z(\mathcal{L})$ denote the centre of $\mathcal{L}$ i.e.

$$Z(\mathcal{L}) = \{X \in \mathcal{L} : [X, Y] = 0 \quad \forall Y \in \mathcal{L}\}$$

since $d\sigma(X)(f) = \frac{d}{dt} \left. Ad^*_{\exp -tX} f \right|_{t=0}$

$$= -\, ad^*_X f$$

thus $d\sigma(X)(f)(Y) = -ad^*_X f(Y) = \langle f, [X, Y]\rangle$

and $\mathcal{L}_f = \{X \in \mathcal{L} : d\sigma(X)(f) = 0\}$.

If $X \in Z(\mathcal{L})$ obviously $X \in \mathcal{L}_f$ as $d\sigma(X)(f)(Y) = \langle f, [X, Y]\rangle = 0$, $\forall\, Y \in \mathcal{L}$.

Thus $\{0\} = V_x^{\perp} = \mathcal{L}_f(x) \supset Z(\mathcal{L})(x)$, $\forall\, x \in M$

So $Z(\mathcal{L}) = \{0\}$ which contradicts proposition 3.1.1.

Q.E.D.

Now

$$V_x^{\perp} = \{X_{\lambda(X)}(x) \in T_xM : \omega(x)\left(X_{\lambda(X)}(x), \xi(x)\right) = 0 \ \forall \xi \in V_x\}$$

From (3.3.2.) if $X_{\lambda(X)} \in V_x^{\perp}$ then $d\sigma(Y)(f)(X) = 0, \ \forall X \in \mathcal{L}$.

By definition of $\mathcal{L}_f = \{X \in \mathcal{L} : d\sigma(X)(f) = 0\}$ thus $V_x^{\perp} = \mathcal{L}_f(x)$ where $\mathcal{L}_f(x)$ is the corresponding subspace of the tangent space of M at x.

Theorem 3.3.9 If $\phi : G \times M \longmapsto M$ is a nilpotent transitive action on a symplectic G-space (M, ω) then this action is not Poisson.

Proof. Assume the action is Poisson, so there exists a moment map $\tau : M \longmapsto \mathcal{L}^*$. Since the action is transitive $V_x = T_xM \ \forall x \in M$ which implies that $V_x^{\perp} = \{0\}$.

Let $Z(\mathcal{L})$ denote the centre of $\mathcal{L}$ i.e.

$$Z(\mathcal{L}) = \{X \in \mathcal{L} : [X, Y] = 0 \quad \forall Y \in \mathcal{L}\}$$

since $d\sigma(X)(f) = \frac{d}{dt} \left. Ad^*_{\exp -tX} f \right|_{t=0}$

$$= - ad^*_X f$$

thus $d\sigma(X)(f)(Y) = -ad^*_X f(Y) = \langle f, [X, Y]\rangle$

and $\mathcal{L}_f = \{X \in \mathcal{L} : d\sigma(X)(f) = 0\}$.

If $X \in Z(\mathcal{L})$ obviously $X \in \mathcal{L}_f$ as $d\sigma(X)(f)(Y) = \langle f, [X, Y]\rangle = 0, \forall Y \in \mathcal{L}$.

Thus $\{0\} = V_x^{\perp} = \mathcal{L}_f(x) \supset Z(\mathcal{L})(x), \forall x \in M$

So $Z(\mathcal{L}) = \{0\}$ which contradicts proposition 3.1.1.

Q.E.D.

It is possible to avoid this problem by extending the Lie algebra as follows

If $X_1, \ldots, X_m$ is a basis for $\mathcal{L}$ and $d\phi(X_i) = X_{H_i}$. Let $\mu_o(X_i) = H_i$ and define $\mu_o(\sum_{i=1}^{m} a_i H_i) = \sum_{i=1}^{m} a_i \mu_o(X_i)$, $\forall a_i \in \underset{\sim}{R}$.

Then $\mu_o : \mathcal{L} \longmapsto C(M)$ and $X_{\mu_o(X)} = d\phi(X)$.

If μ_o were a Lie algebra homomorphism it would be a lift. Assume it is not a lift and let

$$\tilde{\mathcal{L}} = \mathcal{L} \times R$$

and define

$$[(X, t), (Y, s)] = \big([X, Y], \beta(X, Y)\big)$$

where $\beta(X, Y) = \{\mu_o(X), \mu_o(Y)\} - \mu_o([X, Y])$ (3.3.3)

Then $\mathcal{L}$ is a Lie algebra. Define

$$\lambda(X, t)(x) = \mu_o(X)(x) + t$$

then $\{\lambda(X, t), \lambda(Y, s)\} = \{\mu_o(X), \mu_o(Y)\}$

$$= \mu_o([X, Y]) + \beta(X, Y)$$

$$= \lambda\big([(X, t), (Y, s)]\big)$$

thus λ is a Lie algebra homomorphism.

Let $\tilde{G}$ be the connected simply connected Lie group with Lie algebra $\tilde{\mathcal{L}}$. Let H be the connected subgroup with Lie algebra $\{(0, t) : t \in \underset{\sim}{R}\}$. then $\tilde{G}/H$ is the simply connected Lie group with Lie algebra $\mathcal{L}$. Let

$$\nu : \tilde{G}/H \longmapsto G$$

be the covering map.

Let $\mu : \tilde{G} \longmapsto G$ be defined by $\mu(\tilde{g}) = \nu(\tilde{g}\,H)$.
Then $\tilde{G}$ acts on M by $\tilde{g}.x = \mu(\tilde{g}).x$, $x \in M$ and $\lambda : \tilde{\mathcal{L}} \longmapsto C(M)$ is a lift.

In this case, see Wallach [25] for example, there exists the following.

Theorem 3.3.10 If (M, ω) is a simply connected homogeneous G-space, then there exists a β of the form (3.3.3) and $f \in \tilde{\mathcal{L}}^*$ so that the moment map $\tau : M \longmapsto \tilde{\mathcal{L}}^*$ is a morphism from (M, ω) to $(0, \omega_0)$ and a covering map, where $0 = Ad^*_{\tilde{G}}f$.

Essentially theorems 3.3.7 and 3.3.10 allow all work to be done on the coadjoint orbit and then by use of the moment map restated on the symplectic manifold (M, ω).

The nilpotent case is now considered in depth using the work of Kirillov [21].

Definition 3.3.11 Let $\mathcal{L}$ be a Lie algebra and let $f \in \mathcal{L}^*$. Then a polarizing subalgebra of $\mathcal{L}$ for f, sometimes called a polarization of f, is a subalgebra $\mathfrak{h}$ of $\mathcal{L}$ satisfying,

(i) $f([\mathfrak{h}, \mathfrak{h}]) = 0$

(ii) If $V \subset \mathfrak{L}$ is a subspace so that $f([X, Y]) = 0$ for $X, Y \in V$ and $V \supset \mathfrak{h}$ then $V = \mathfrak{h}$.

In general there does not exist a polarization for f for all Lie algebras. For nilpotent Lie algebras, however, there exists polarizations for arbitrary $f \in \mathfrak{L}^*$.

Let V be a finite dimensional real vector space. Let $B : V \times V \longmapsto \underset{\sim}{R}$ be bilinear and skew-symmetric, although not necessarily non-degenerate. If W is a subspace of V, set

$$W^{\perp} = \{v \in V : B(v, w) = 0 \quad , \quad \forall\, w \in W\}$$

Then

$\dim W + \dim W^{\perp} = \dim V + \dim (V^{\perp} \cap W)$ since for $v \in V$, set $\tilde{v}(u) = B(v, u)$, let $\zeta(v) = \tilde{v}|_W$ then $\zeta : V \longmapsto \left(W/_{W \cap V^{\perp}}\right)^*$, $\ker \zeta = W^{\perp}$. Thus, $\dim V = \dim W - \dim (W \cap V^{\perp}) + \dim W^{\perp}$

A subspace W of V is said to be isotropic if

$$B|_{W \times W} = 0 \qquad \text{i.e. } W \subset W^{\perp}$$

If W is maximal as an isotropic subspace then $V^{\perp} \subset W$, hence $2 \dim W \leq \dim V + \dim V^{\perp}$. More precisely, see Wallach [25],

<u>Lemma 3.3.12</u> If $W \subset V$ is an isotropic subspace, then W is maximal isotropic if and only if

$$\dim W = \frac{1}{2} (\dim V + \dim V^{\perp})$$

Lemma 3.3.13 Let $V' \subset V$ be a subspace of codimension one. Let $N' = \{v \in V' : B(v, v') = 0 \;\forall v' \in V'\}$. If $V^{\perp} \subset V'$ then $V^{\perp}$ is of codimension one in N'. If $W \subset V'$ is maximal isotropic in V', then W is maximal isotropic in V.

Proof. If $V^{\perp} \subset V'$ then since

$$V^{\perp} = \{v \in V : B(v, u) = 0, \forall u \in V\}$$

it can rewritten as $V^{\perp} = \{v \in V' : B(v, u) = 0, \forall u \in V\} \subset N'$ (3.3.4)

Let $x \in V$ be so that $x \notin V'$. Then it is claimed $V^{\perp} = \mathfrak{R}$ where

$$\mathfrak{R} = \{v \in N' : B(x, v) = 0\}.$$

Let $v_1 \in V^{\perp}$ then $v_1 \in N'$ so $B(v_1, V') = 0$ from (3.3.4) and $B(v_1, x) = 0$ so $V^{\perp} \subset \mathfrak{R}$. If $v \in \mathfrak{R}$, $B(x, v) = 0$, $V = V' + \mathbb{R} x$. Also $v \in N'$ so $B(v, V') = 0 \Rightarrow B(v, \mathbb{R} x + V') = 0 \Rightarrow B(v, V) = 0 \Rightarrow v \in V^{\perp}$. So claim is proved.

So $\dim (N'/V^{\perp}) \leq 1$. But since $N' = V'^{\perp}$ with respect to the vector space V' there exists a non-degenerate 2-form on $V'/V'^{\perp} = V'/N'$ and so the $\dim (V'/N')$ is even. Also, $\dim (V/V^{\perp})$ is even. Thus

$\dim V/V^{\perp} - \dim V'/N'$

$= \dim V - \dim V^{\perp} - \dim V' + \dim N'$

$= \dim V - \dim V' + \dim N' - \dim V^{\perp}$

$= 1 + \dim N'/V^{\perp}$

is even, and since $\dim (N'/V^{\perp}) \leq 1$ it follows that $\dim (N'/V^{\perp}) = 1$.

By lemma 3.3.12, W is maximal if and only if $\dim W = \frac{1}{2}(\dim V + \dim V^{\perp})$ thus, $\frac{1}{2}(\dim V + \dim V^{\perp}) = \frac{1}{2}((\dim V-1) + (\dim V^{\perp} + 1)) = \frac{1}{2}(\dim V' + \dim N')$

$$= \frac{1}{2}(\dim V' + \dim V'^{\perp})$$

Hence the result.

Q.E.D.

Lemma 3.3.14 Let $\mathcal{L}$ be a nilpotent Lie algebra. Suppose that the centre of $\mathcal{L}$, $Z(\mathcal{L})$, is one dimensional. That is,

$$Z(\mathcal{L}) = \{X \in \mathcal{L} : [X, \mathcal{L}] = 0\}$$

and $\dim (Z(\mathcal{L})) = 1$.

Suppose that $\mathbb{R}Z = Z(\mathcal{L})$. Then there exists $X, Y \in \mathcal{L}$ such that $[X, Y] = Z$ and if

$$S = \{W \in \mathcal{L} : [Y, W] = 0\}$$

then S is a codimension one subalgebra of $\mathcal{L}$.

Proof. By Engel's theorem 3.1.2 and by considering the adjoint representation of $\mathcal{L}$ i.e. $ad_{\mathcal{L}} : \mathcal{L} \longmapsto \mathcal{L}$, $\mathcal{L}$ can be decomposed as,

$$\mathcal{L} = \mathcal{L}_n \supset \mathcal{L}_{n-1} \supset \ldots . \supset \mathcal{L}_1 \supset \{0\}$$

a sequence of ideals with $[\mathcal{L}, \mathcal{L}_i] \subset \mathcal{L}_{i-1}$ and $\dim \mathcal{L}_i = i$.

Since $\dim Z(\mathcal{L}) = 1 \Rightarrow Z(\mathcal{L}) = \mathcal{L}_1$.

Let $Y \in \mathcal{L}_2$ be such that

$$\mathcal{L}_2 = \mathbb{R}\, Y + \mathbb{R}\, Z$$

If $W \in \mathcal{L}$, then

$\mathrm{ad}_W Y = [W, Y] = f(W)Z$, $f : \mathcal{L} \mapsto \mathbb{R}$ linear. Claim $f \neq 0$, since if $f = 0$ $[W, \mathbb{R}\, Y + \mathbb{R}\, Z] = 0$ $\forall\, W \in \mathcal{L} \Rightarrow \mathbb{R}\, Y + \mathbb{R}\, Z \subset Z(\mathcal{L})$ which is a contradiction, since $\dim Z(\mathcal{L}) = 1$.

Let $X \in \mathcal{L}$ be such that $f(X) = 1$. Then

$$S = \{W \in \mathcal{L} : [W, Y] = 0\}$$

Then $S = \ker f$

Hence $\dim \mathcal{L} = \dim S + 1$

Q.E.D.

Theorem 3.3.15 [Kirillov [21]] If $\mathcal{L}$ is a nilpotent Lie algebra and $f \in \mathcal{L}^*$, then there exists a polarization of f.

Proof. The result is proved by induction on the dimension of $\mathcal{L}$. If $\dim \mathcal{L} = 1$, then $\mathcal{L}$ is a polarization of f is immediate from the definition.

Suppose the theorem has been proved for all nilpotent Lie algebras of dimension $\leq n - 1$ and $\dim \mathcal{L} = n$. Let $Z(\mathcal{L})$ be the centre of $\mathcal{L}$, choose f, a real linear functional on $\mathcal{L}$ such that $Z(\mathcal{L}) \not\subset \ker f$, and denote $Z_o(\mathcal{L})$ the subspace of $Z(\mathcal{L})$ consisting of all Z_o such that $f(Z_o) = 0$. Clearly $Z_o(\mathcal{L}) \subset \ker f$ so,

$$\dim Z_o(\mathcal{L}) \leq \dim (\ker f).$$

Consider $f|_{Z(\mathcal{L})}$ then

$$\dim Z(\mathcal{L}) = 1 + \dim (\ker f|_{Z(\mathcal{L})})$$

$$\geq 1 + \dim Z_o(\mathcal{L})$$

$\Rightarrow \dim Z_o(\mathcal{L}) \leq \dim Z(\mathcal{L}) - 1.$

Case 1 : $\dim Z_o(\mathcal{L}) \geq 1 \Rightarrow \dim Z(\mathcal{L}) \geq 2$. So there exists $Y \in Z(\mathcal{L})$ such that $f(Y) = 0$, $Y \neq 0$. Then f defines an element of $\left(\mathcal{L}/_{\mathbb{R}Y}\right)^*$, $\bar{f}$. Let $\bar{\mathfrak{h}}$ be a polarization of $\bar{f}$. Obviously it is possible to choose $\mathbb{R}Y \in \mathfrak{h}$, $\mathfrak{h} \subset \mathcal{L}$ and if $\mathfrak{h} \to \bar{\mathfrak{h}}$ under the canonical map, then it can be shown that $\mathfrak{h}$ is a polarization of f as follows.

Set $B_f(X_1, X_2) = -f([X_1, X_2])$, $\forall X_1, X_2 \in \mathcal{L}$.

Clearly, $\mathfrak{h}$ is isotropic for B_f, since $Y \in Z(\mathcal{L})$. If $\mathfrak{h} \subset V$ with V an isotropic subspace of $\mathcal{L}$ relative to B_f, then $V \longmapsto \bar{V} \supset \bar{\mathfrak{h}}$ under the canonical map and $\bar{V}$ is isotropic for $B_{\bar{f}}$. Hence by definition $\bar{V} = \bar{\mathfrak{h}}$. But $\mathbb{R}Y \subset \mathfrak{h}$ hence $V = \mathfrak{h}$.

Case 2 : $\dim Z_o(\mathcal{L}) = 0 \Rightarrow \dim Z(\mathcal{L}) \geq 1$. The only case not yet covered is $\dim Z(\mathcal{L}) = 1$.

Let $Z(\mathcal{L}) = \mathbb{R}Z$, $f(Z) = 1$. Let $X, Y \in \mathcal{L}$ be such that $[X, Y] = Z$ and

$$S = \{W \in \mathcal{L} : [Y, W] = 0\}$$

as in lemma 3.3.14.

It has already been seen that dim $S = n - 1$. Let $\mathfrak{h} \subset S$ be a polarization for $f|_S$. Again, set $B_f(X_1, X_2) = -f([X_1, X_2])$. than it is claimed, $\mathcal{L}^{\perp} = \{W \in \mathcal{L} : B_f(W, \mathcal{L}) = 0\} \subset S$.
Since $\mathcal{L} = \underset{\sim}{R} X \oplus S$ and if $W = c\,X + s$, $c \in R$, $s \in S$, and $W \in \mathcal{L}^{\perp}$ then

$$\begin{aligned} B_f(c\,X + s, Y) &= c\,B_f(X, Y) + B_f(s, Y) \\ &= -c\,f([X, Y]) + f([Y, s]) \\ &= -c\,f(Z) + f(0) \\ &= -c \qquad\qquad , Y \in \mathcal{L} \end{aligned}$$

Hence $c = 0$. Thus $W \in S$ which implies $\mathcal{L}^{\perp} \subset S$. Now lemma 3.3.13 implies that $\mathfrak{h}$ is a polarization for f on $\mathcal{L}$.

Q.E.D.

<u>Corollary 3.3.15</u> If $\mathfrak{h}$ is a polarization with respect to f then $Ad_g \mathfrak{h}$ is a polarization with respect to $Ad_g^* f$.

<u>Proof</u>. Clearly $\langle Ad_g^* f, [Ad_g \mathfrak{h}, Ad_g \mathfrak{h}]\rangle = 0$ since Ad_g is a Lie algebra homomorphism and $\mathfrak{h}$ is a polarization of f. So it is sufficient to show that $Ad_g \mathfrak{h}$ is maximal isotropic.

Suppose W is maximal isotropic at $Ad_g^* f$ then $Ad_{g^{-1}} W$ is isotropic at f which implies that

$$Ad_{g^{-1}} W \subset \mathfrak{h} \text{ or } W \subset Ad_g \mathfrak{h}$$

$\Rightarrow$ $Ad_g \mathfrak{h}$ is maximal isotropic.

Q.E.D.

Obviously $\mathcal{L}_f = \{X \in \mathcal{L} : \langle f, [X, Y]\rangle = 0 \;\; \forall\, Y \in \mathcal{L}\}$ is a subalgebra contained in $\mathcal{L}$ since $\langle f, [X, Y]\rangle = 0 \;\; \forall\, X, Y \in \mathcal{L}_f$.

Let

$\sigma : G \times O \longmapsto O$

as before with $O = Ad^*_G f$, for some fixed $f \in \mathcal{L}^*$.

The following lemma can be found in Kirillov [21].

Lemma 3.3.16 If $\mathfrak{h}$ is a polarization at $f \in \mathcal{L}^*$ in the nilpotent Lie algebra $\mathcal{L}$ and O is the orbit in $\mathcal{L}^*$ containing f, then

$$\dim \mathfrak{h} = \dim \mathcal{L} - \frac{1}{2} \dim O \qquad (3.3.5)$$

This lemma along with theorem 3.3.14 says there exists a Lagrangian subspace of $T_f O$, $f \in O$ since (3.3.5) implies

$$\dim \mathfrak{h}/\mathcal{L}_f = \dim \mathcal{L} - \frac{1}{2}(\dim \mathcal{L} - \dim \mathcal{L}_f) - \dim \mathcal{L}_f$$

$$= \frac{1}{2}\dim \mathcal{L} - \frac{1}{2}\dim \mathcal{L}_f = \frac{1}{2}\dim \mathcal{L}/\mathcal{L}_f$$

So this subspace is given by

$$d\sigma(\mathfrak{h})(f) \subset T_f O.$$

Note $\mathfrak{h}$ is a nilpotent Lie subalgebra of $\mathcal{L}$. Also it is known that

$V/\mathcal{L}_f \cap V = (\mathcal{L}_f + V)/\mathcal{L}_f \qquad V \subset \mathcal{L}$, therefore it is possible to construct the following sequences of quotient algebras,

$$\{0\} = \mathcal{L}_f/\mathcal{L}_f \subset \mathcal{L}_f + \mathfrak{h}^m/\mathcal{L}_f \subset \ldots \subset \mathcal{L}_f + \mathfrak{h}^1/\mathcal{L}_f \qquad (3.3.6)$$

where $\mathfrak{h}^1 = \mathfrak{h}$, $\mathfrak{h}^i = [\mathfrak{h}, \mathfrak{h}^{i-1}]$, $\mathfrak{h}^{n+1} = 0$.

$$\{0\} = \mathcal{L}_f/\mathcal{L}_f \subset \mathfrak{h} \cap \mathcal{L}^n + \mathcal{L}_f/\mathcal{L}_f \subset \ldots \subset \mathfrak{h} \cap \mathcal{L}^1 + \mathcal{L}_f/\mathcal{L}_f \qquad (3.3.7)$$

where $\mathcal{L}^1 = \mathcal{L}$, $\mathcal{L}^i = [\mathcal{L}, \mathcal{L}^{i-1}]$, $\mathcal{L}^{n+1} = 0$.

It is worthwhile observing that the proceeding theory is valid whether sequence (3.3.6) or (3.3.7) is used. However, the calculations are performed for (3.3.7) as they will appear in chapter 4.2.

It is now possible to construct a sequence of isotropic distributions as in theorem 3.2.8. But first some notation.

Let Δ be a distribution on (O, ω_o) then it is said to be G-invariant if

$$Ad^*_{g\,*}\,\Delta \subset \Delta, \quad \forall\, g \in G$$

i.e. $Ad^*_{g_1\,*}\,\Delta(Ad^*_{g_2} f) \subset \Delta(Ad^*_{g_1 g_2} f) \qquad \forall\, g_1, g_2 \in G.$

<u>Lemma 3.3.17</u> On the Poisson G-space (O, ω_o) there exists a sequence of G-invariant isotropic involutive distributions, if G is nilpotent, of the form

$$\{0\} = \Delta_o \subset \Delta_1 \subset \ldots \subset \Delta_{n-1} \subset \Delta_n \text{ with } \Delta_n \text{ Lagrangian.}$$

<u>Proof</u>. Define Δ_i, $0 \leq i \leq n$ by

$$\Delta_i\,(Ad^*_g f) = d\sigma\,(Ad_g(\mathcal{L}_f \cap \mathcal{L}^{n-i+1} + \mathcal{L}_f))\,(Ad^*_g f)$$

note ker $d\sigma\,(Ad^*_g f) = Ad_g \mathcal{L}_f$, $\forall\, g \in G$, $f \in O$, thus $\Delta_i\,(Ad^*_g f)$ is isomorphic to $Ad_g\,(\mathcal{L}_f \cap \mathcal{L}^{n-i+1} + \mathcal{L}_f)/_{Ad_g \mathcal{L}_f}$ then from (3.3.7)

$$\Delta_i\,(Ad^*_g f) \subset \Delta_{i+1}\,(Ad^*_g f), \ \forall\, g \in G,\ f \in O.$$

Which defines the sequence of distributions

$$\{0\} = \Delta_o \subset \Delta_1 \subset \ldots \subset \Delta_{n-1} \subset \Delta_n$$

Since $\mathcal{L}$ is nilpotent $[\mathcal{L}^i, \mathcal{L}^j] \subset \mathcal{L}^{i+j}$ by induction i.e. $[\mathcal{L}, \mathcal{L}^j] = \mathcal{L}^{j+1}$ by definition for all j, suppose $[\mathcal{L}^{i-1}, \mathcal{L}^j] \subset \mathcal{L}^{i+j-1}$ then $[\mathcal{L}^i, \mathcal{L}^j]$ can be rewritten using the Jacobi identity as

$$[[\mathcal{L}, \mathcal{L}^{i-1}], \mathcal{L}^j] = [[\mathcal{L}, \mathcal{L}^j], \mathcal{L}^{i-1}] + [\mathcal{L}, [\mathcal{L}^{i-1}, \mathcal{L}^j]]$$

$$\subset \mathcal{L}^{i+j} + [\mathcal{L}, \mathcal{L}^{i+j-1}]$$

$$= \mathcal{L}^{i+j}$$

Also $d\sigma$ and Ad_g are Lie algebra homomorphisms so Δ_i is involutive, for let $X, Y \in \mathcal{L}_f \cap \mathcal{L}^{n-i+1} + \mathcal{L}_f$ then

$$[d\sigma\,(Ad_g X)\,(Ad_g^* f),\ d\sigma\,(Ad_g Y)\,(Ad_g^* f)] = d\sigma\big(Ad_g\,([X, Y])\big)(Ad_g^* f)$$

$$\subset \Delta_i\,(Ad_g^* f).$$

For G-invariance it is necessary to prove that

$$Ad^*_{g\,*}\,\Delta_i \subset \Delta_i,\ g \in G,\quad 1 \leq i \leq n$$

i.e. $Ad^*_{g_1 *}\,\Delta_i\,(Ad^*_{g_2} f) \subset \Delta_i\,(Ad^*_{g_1 g_2} f),\quad 1 \leq i \leq n$

Let $Y \in \mathcal{L}$, $X \in \mathcal{L}_f \cap \mathcal{L}^{n-i+1} + \mathcal{L}_f$, $g_1, g_2 \in G$

$$Ad^*_{g_1 *}\, d\sigma(Ad_{g_2} X)\,(Ad^*_{g_2} f)\,(Y)$$

$= Ad^*_{g_1} \; d\sigma(Ad_{g_2} X)\,(Ad^*_{g_2} f)\,(Y) \qquad (Ad^*_{g_1} \text{ is linear})$

$= \langle Ad^*_{g_1} \; ad^*_{Ad_{g_2} X} \; Ad^*_{g_2} f,\; Y\rangle$

$= \langle Ad^*_{Ad_{g_2} X} \; Ad^*_{g_2} f,\; Ad_{g_1^{-1}} Y\rangle$

$= - \langle Ad^*_{g_2} f,\; [Ad_{g_2} X,\; Ad_{g_1^{-1}} Y]\rangle$

$= - \langle f,\; Ad_{g_2^{-1}} [Ad_{g_2} X,\; Ad_{g_1^{-1}} Y]\rangle$

$= - \langle f,\; Ad_{g_2^{-1}} Ad_{g_1^{-1}} [Ad_{g_1} Ad_{g_2} X,\; Y]\rangle$

$= - \langle Ad^*_{g_1 g_2} f,\; [Ad_{g_1 g_2} X,\; Y]\rangle$

$= d\sigma\,(Ad_{g_1 g_2} X)\,(Ad^*_{g_1 g_2} f)\,(Y)$

therefore since Y was arbitrary

$$Ad^*_{g_1}{}_* \; \Delta_i \,(Ad^*_{g_2} f) \subset \Delta_i \,(Ad^*_{g_1 g_2} f)$$

as required.

The fact that each Δ_i, $0 \leq i \leq n - 1$ is isotropic and Δ_n is Lagrangian follows from theorem 3.3.14, corollary 3.3.15 and lemma 3.3.16.

Q.E.D.

Observe that Δ_n is a real polarization of $(0, \omega_o)$ in the sense of definition 3.2.6.

For the case in question, an orbit 0 in $\mathcal{L}^*$ given by $0 = Ad_G^* f$, for some $f \in \mathcal{L}^*$, the invariant two form, df, is given by

$$df(X, Y) = - \langle f, [X, Y] \rangle,$$

see chapter 5.3 of Guillemin and Sternberg [22], and the invariant polarization on the orbit through f corresponds to the subalgebra constructed above.

As the two-form ω_o on 0 is invariant under σ_g, $\forall g \in G$, i.e. $\sigma_g^* \omega_o = \omega_o$, df is equal to ω_o.

The next proposition due to Kostant, see Guillemin and Sternberg [22], is of particular importance.

<u>Proposition 3.3.18</u> Let f be a one-form on M, a symplectic manifold such that its symplectic form is given by $df = \omega$ and such that the set $\{x : f|_x = 0\}$ is a n-dimensional submanifold, L. Then L is Lagrangian and there is a unique polarization defined on M in a neighbourhood of L which is transversal to L and whose associated one - form is f.

<u>Lemma 3.3.19</u> The sequence of distributions defined on a homogeneous strongly symplectic G-space $(0, \omega_o)$ in lemma 3.3.17 satisfy the conditions (1) - (3) of theorem 3.2.8.

<u>Proof.</u> Take the Lagrangian submanifold L to be that constructed in proposition 3.3.18 and the sequence of distributions are then that of lemma 3.3.17. By assumption $d\sigma(X) \subset \text{Ham}\ (0, \omega_o)$ - Hamiltonian vector

fields on O - and thus by (3.3.7) and transitivity there exists Hamiltonian vector fields.

$$d\sigma(Ad_g\, Y_1^1)\ (Ad_g^* f), \ldots, d\sigma(Ad_g\, Y_{r_1}^1)\ (Ad_g^* f), \ldots,$$

$$d\sigma(Ad_g\, Y_1^i)\ (Ad_g^* f), \ldots, d\sigma(Ad_g\, Y_{r_i}^i)\ (Ad_g^* f)$$

which spans $\Delta_i\ (Ad_g^* f)$, $\forall\ g \in G$, $1 \leq i \leq n$

and so condition (2) is satisfied. Condition (3) follows immediately from the facts that $d\sigma$ and Ad_g are Lie algebra homomorphisms and $[\mathcal{L}^i, \mathcal{L}^j] \subset \mathcal{L}^{i+j} \subset \mathcal{L}^j$, $j \geq i$

Q.E.D.

Before investigating condition (4) some results from Kostant [19] are required.

Definition 3.3.20 If M and N are manifolds and $\psi : M \longmapsto N$ is an analytic map, $X \in V(M)$, $Y \in V(N)$ X is said to be ψ - related to Y if

$$\psi_* X(x) = Y(\psi(x)),\ \forall\ x \in M.$$

Lemma 3.3.21 If $\psi : M \longmapsto N$ and X_i is ψ - related to Y_i, i =1, 2 then $[X_1, X_2]$ is ψ - related to $[Y_1, Y_2]$.

Proposition 3.3.22 (Kostant [19]). Let (M, ω) be a simply connected homogeneous symplectic G - space with a Poisson action

$$\phi : G \times M \longmapsto M$$

Fix $x_o \in M$, let $\tau : M \longmapsto 0 \subset \mathcal{L}^*$ be the moment map with $0 = Ad^*_G\ \tau\ (x_o)$ and Poisson action on 0 given by

$$\sigma : G \times 0 \longmapsto 0$$

$$\sigma\ (g,\ f) = Ad^*_g f.$$

Then $d\ \phi\ (X)$ is τ - related to $d\sigma\ (X)$.

<u>Proof.</u> Need to show that $\tau_* d\phi\ (X)\ (x) = d\sigma\ (X)\ \tau\ (x)$.

By definition $d\sigma\ (X)\ (f)\ (\psi) = \frac{d}{dt}\Big|_{t=o} \psi\ (\exp - t\ X.f)\ \psi \in\ C\ (\mathcal{L}^*)$

Let $Y \in \mathcal{L}$ and define $\psi^Y : \mathcal{L}^* \longmapsto \underset{\sim}{R}$ by $\psi^Y\ (f) = \langle f,\ Y\rangle$

$$\text{then } d\sigma\ (X)\ (f)\ (\psi^Y) = \psi^{[X,\ Y]} \tag{3.3.8}$$

since

$$\begin{aligned} d\sigma(X)\ (f)\ (\psi^Y) &= \frac{d}{dt}\Big|_{t=o} \psi^Y\ (\exp - t\ X.f) \\ &= \frac{d}{dt}\Big|_{t=o} \langle \exp - t\ X.f,\ Y\rangle \\ &= -\langle X.f,\ Y\rangle \\ &= -\langle f,\ [X,\ Y]\rangle \\ &= \psi^{[X,\ Y]} \end{aligned}$$

By the definition of the moment map i.e.

$$\tau(x)\ (X) = \lambda(X)(x)$$

then $\psi^X\ (\tau(x)) = \lambda(X)(x)$

Let $\quad v = \tau_* \, d\phi(Y)(x)$

then $v\,\psi^X = \tau_* \, d\phi(Y)(x)(\psi^X)$

$= \psi^X{}_* \, \tau_* \, d\phi(Y)(x)$

$= \lambda(X)_* \, d\phi(Y)(x)$

$= \lambda(X)_* \, X_{\lambda(Y)} \, (x)$

$= \{\lambda(Y), \lambda(X)\}(x)$

$= \lambda([Y, X])(x)$

So $\; d\phi(Y)\,\lambda\,(X) = \{\lambda(Y), \lambda(X)\} = \lambda([Y, X])$.

But $\lambda([Y, X])(x) = \tau(x)([Y, X])$

$= \psi^{[Y, X]} \, \tau(x)$

Thus $v\,\psi^X = \psi^{[Y, X]} \, \tau(x) \qquad \forall \, X \in \mathcal{L}$.

$\Rightarrow \tau_* \, d\phi(Y)(x)\psi^X = d\sigma(Y)(\tau(x))\psi^X \quad$ (by $\;$ (3.3.8))

$\Rightarrow \tau_* \, d\phi(Y) = d\sigma(Y)$

Hence result

Q.E.D.

So by lemma 3.3.21 it is immediate that

$[\tau_* \, d\phi(X), \tau_* \, d\phi(Y)] = \tau_*[d\phi(X), d\phi(Y)]$.

From theorem 3.3.9 it is necessary for nilpotent group actions to go to the extension in order for a Poisson action to exist. Notice that the extended Lie algebra $\tilde{\mathcal{L}}$ described after theorem 3.3.9 is nilpotent if $\mathcal{L}$ is nilpotent.

<u>Theorem 3.3.23</u> If (M, ω) is a simply connected symplectic manifold with a transitive Poisson action by the connected nilpotent Lie group $\tilde{G}$, with nilpotent Lie algebra $\tilde{\mathcal{L}}$, given by

$$\tilde{\phi} : \tilde{G} \times M \longmapsto M$$

$$\tilde{\phi}\,(\tilde{g}, x) = \phi(\mu\,(\tilde{g}), x) \qquad (3.3.9)$$

where $\mu : \tilde{G} \longmapsto G$ and $\phi : G \times M \longmapsto M$ the action given by

$$\phi(\exp t_1X_1 {}_{o\ldots o} \exp t_kX_k, x) = \exp t_1X_1 {}_{o\ldots o} \exp t_kX_k . x$$

where $X_1, \ldots, X_k \in \mathcal{L}$. Then there exists a sequence of distributions on M satisfying the conditions (1) - (4) of theorem 3.2.8 and theorem 3.2.9.

<u>Proof.</u> Take the sequence of distributions constructed in lemma 3.3.17 and then take $\tau_*^{-1}\,\Delta_i$ where τ is the moment map, $\tau : M \longmapsto O \subset \tilde{\mathcal{L}}^*$. By proposition 3.3.22 τ_*^{-1} can be taken through the brackets and the validity of conditions (1) - (3) of theorem 3.2.8 follow trivially from lemma 3.3.19.

$\tilde{G}$ - invariance i.e. $\tilde{\phi}_{\tilde{g}\,*}\,(\tau_*^{-1}\,\Delta_i) \subset \tau_*^{-1}\,\Delta_i$ follows from lemma 3.3.5. Since,

$$\tau_o\,\tilde{\phi}_{\tilde{g}} = Ad^*_{\tilde{g}}\,{}_o\,\tau, \quad \forall\, \tilde{g} \in \tilde{G}$$

$$\Rightarrow \tau_*\,\tilde{\phi}_{\tilde{g}\,*}\,\tau_*^{-1}\,\Delta_i = Ad^*_{\tilde{g}\,*}\,\Delta_i \subset \Delta_i \quad \forall\, \tilde{g} \in \tilde{G} \text{ by lemma 3.3.17}$$

$$\Rightarrow \tilde{\phi}_{\tilde{g}\,*}\,(\tau_*^{-1}\,\Delta_i) \subset \tau_*^{-1}\,\Delta_i, \quad \forall\, \tilde{g} \in \tilde{G}.$$

Then by application of the Campbell-Baker-Hausdorff formula i.e.

$$\exp - t X_* Y (\exp t X.x) = \sum_{i=o}^{\infty} \frac{t^i}{i!} \operatorname{ad}_X^i Y (x) \text{ and (3.3.9) it}$$

follows that $\tilde{G}$ - invariance is equivalent to

$$[X_H, \tau_*^{-1} \Delta_i] \subset \tau_*^{-1} \Delta_i \quad o \leq i \leq n \quad \forall X_H \in \mathcal{L} .$$

Condition (4) of theorem 3.2.8 now follows trivially by $\tilde{G}$ - invariance and transitivity since each X_k $1 \leq k \leq n$ of theorem 3.2.8 is a Hamiltonian vector field and $X_\ell^i \in \tau_*^{-1} \Delta_i$, $1 \leq i \leq n$.

Q.E.D.

Similarly the existence of polarization for solvable Lie algebras have been studied by Auslander and Kostant [26]. For this it is necessary to introduce complex polarizations. The questions arising from solvable Lie algebras will not be gone into in full detail but an interesting lemma found in the above paper will be stated for the purpose of application in the next chapter.

Let $\mathcal{L}$ be an arbitrary Lie algebra over $\underset{\sim}{R}$ and let G be the corresponding connected Lie group. As before define the alternating bilinear from B_f on $\mathcal{L}$ by

$$B_f (X, Y) = - \langle f, [X, Y] \rangle, \; f \in \mathcal{L}^*, \; X, Y \in \mathcal{L} \quad \text{and}$$

$$\mathcal{L}_f = \{X \in \mathcal{L} \; : B_f (X, Y) = 0 \;\; \forall Y \in \mathcal{L} \; \}$$

<u>Definition 3.3.24</u> Let $\mathcal{L}_{\underset{\sim}{C}} = \mathcal{L} + i\mathcal{L}$ and consider $f \in \mathcal{L}^*$ as a complex-valued linear functional on $\mathcal{L}_{\underset{\sim}{C}}$. A polarization at f is a complex sub-algebra $\mathfrak{h} \subset \mathcal{L}_{\underset{\sim}{C}}$ such that

(i) $\mathcal{L}_f \subset \mathfrak{h}$ and $\mathcal{L}_f$ is invariant under Ad_{G_f}

(ii) $\dim_{\underset{\sim}{C}} \mathcal{L}_{\underset{\sim}{C}}/\mathfrak{h} = \frac{1}{2} \dim_{\underset{\sim}{R}} \mathcal{L}/\mathcal{L}_f$

(iii) $\langle f, [\mathfrak{h},\mathfrak{h}]\rangle = 0$

(iv) $\mathfrak{h} + \bar{\mathfrak{h}}$ is a Lie algebra of $\mathcal{L}_{\underset{\sim}{C}}$

$\mathfrak{h}$ can be reinterpreted as a maximal isotropic subspace of $\mathcal{L}_{\underset{\sim}{C}}$ relative to B_f.

Let G be a connected, simply connected solvable Lie group with Lie algebra $\mathcal{L}$ and let S denote the nil-radical of $\mathcal{L}$ i.e. maximal nilpotent ideal of $\mathcal{L}$, with Lie group N then the following lemma is proved in Auslander and Kostant [26].

Lemma 3.3.25 Let N be a simply connected nilpotent Lie group and let S be it's Lie algebra. Let Aut S be the group of all Lie algebra automorphisms of S so that Ad_N is a subgroup of Aut S. Regard Aut S as operating by contragradience on the dual S^*. Let $f \in S^*$. Assume F is a group and a homomorphism $F \mapsto$ Aut S is given (so that F operates on S and S^*) such that (i) the commutator subgroup F^1 maps into Ad_N and (ii) F. f = f. Then there exists a polarization $\mathfrak{h}$ at f which is invariant under F.

Essentially this lemma can be used along with all the preceding theory on the construction of sequences of involutive distributions to extend to a special case of a solvable group action on a symplectic manifold which will be of importance in the next chapter.

Chapter 4

HAMILTONIAN REALIZATIONS OF FINITE VOLTERRA SERIES

Chapter 4 uses the theory developed in chapter 3 to find necessary and sufficient conditions for the existence of Hamiltonian realizations of stationary finite Volterra series. Then using the structure of chapter 3.3, section two contains a canonical coordinate realization for minimal linear analytic Hamiltonian realizations as well as an algorithm for computing the Volterra kernels if such a system is given. In the last section it is shown that such realizations are closely related to the concept of interconnection as found in electrical network theory.

4.1 On finite Volterra Series which admit Hamiltonian realizations

This section begins with a brief review of the work of Brockett [7] and Crouch [5] on nonlinear systems whose input-output behaviour is represented by a Volterra series - in particular linear analytic systems and finite Volterra Series.

So the type of system investigated here is of the form (2.1.1) with the relevant conditions as described in chapter 2.1. The formal Volterra series will be written as

$$y(t) = W_o(t) + \int_o^t W_1(t, \sigma_1)\, u(\sigma_1)\, d\sigma_1 + \int_o^t \int_o^{\sigma_1} W_2(t, \sigma_1, \sigma_2)\, u(\sigma_1)\, u(\sigma_2)\, d\sigma_2\, d\sigma_1 + \ldots$$

Theorem 4.1.1. (Brockett [7]). Given $T > 0$, $\exists\ \varepsilon > 0$ such that for all locally integrable u satisfying $\int_0^T |u(s)|\, ds < \varepsilon$, the following Volterra series converges uniformly and absolutely on $[0, T]$ to the input-output map of system (2.1.1)

$$y(t) = W_0(t) + \sum_{i=1}^{\infty} \int_0^t \cdots \int_0^{\sigma_{i-1}} W_i(t, \sigma_1, \ldots, \sigma_i)\, u(\sigma_i) \ldots u(\sigma_1)\, d\sigma_i \ldots d\sigma_1$$

for unique analytic kernel functions

$$t, \sigma_1, \ldots, \sigma_i \longmapsto W_i(t, \sigma_1, \ldots, \sigma_i)$$

A series terminating with the term involving the r^{th} kernel is called a finite Volterra series of length r.

In Krener and Lesiak [27], the kernels are shown to be given inductively by the equations

$$W_0^{\,i}(t, x) = h_i(\gamma_f(t)\, x)$$

$$W_n^{j_0 j_1 \ldots j_n}(t, \sigma_1, \ldots, \sigma_n, x) = \gamma_f(-\sigma_n)_* \, g_{j_n}(\gamma_f(\sigma_n)\, x)$$

$$(W_{n-1}^{j_0 \ldots j_{n-1}}(t, \sigma_1, \ldots, \sigma_{n-1}, \cdot)) \qquad (4.1.2)$$

where $\gamma_f : \mathbb{R} \times M \longmapsto M$ is the flow of $f \in V(M)$.

The kernels include $x \in M$ since each initial state defines an input-output map and so the kernels depend on these initial states.

In the case of a finite Volterra series Brockett [7] gives necessary and sufficient conditions in order that the finite Volterra series has a linear analytic realization.

The kernel $W_r(t, \sigma_1,\dots,\sigma_r)$ is said to be separable if it can be expressed as a finite sum

$$W_r(t, \sigma_1, \sigma_2,\dots,\sigma_r) = \sum_{i=1}^{N} \gamma_o^i(t)\, \gamma_1^i(\sigma_1)\dots\gamma_r^i(\sigma_r)$$

If the component functions γ_j^i are differentiable then it is called differentiably separable.

Further, it is said to be stationary if

$$W_r(t, \sigma_1,\dots,\sigma_r) = W_r(0, \sigma_1 - t,\dots,\sigma_r - t)$$

that is the kernels satisfy

$$\partial/_{\partial t} + \sum_{i=1}^{r} \partial/_{\partial\sigma_i} \quad W_r(t, \sigma_1,\dots,\sigma_r) \equiv 0. \tag{4.1.3}$$

A linear analytic realization is said to be stationary if f, g and h do not depend explicitly on time and $f(x(0)) = 0$.

<u>Theorem 4.1.2</u> (Brockett [7]). A finite Volterra series has a (stationary) linear analytic realization if and only if the kernels are (stationary and) differentiably separable if and only if it has a (stationary) bilinear realization.

Further in Brockett [7] for the bilinear realization

$$\dot{x} = Ax + u\,B\,x \qquad x(0) = x_o$$
$$y = C'\,x \qquad x \in \underset{\sim}{R}^n$$

it may be assumed that the subalgebra generated by the matrices $ad_A^k(B)$ is nilpotent. In fact the bilinear realizations constructed there are such that the matrices $ad_A^k(B)$ are strictly lower triangular.

As has been already mentioned, many physical systems evolve on a symplectic manifold, Abraham and Marsden [1] and Arnold [2] being two of the more recent texts which treat this subject from a Symplectic geometric view point. In both these texts only free systems are considered. This then motivates the investigation of linear analytic systems of the form

$$\dot{x} = X_f\,(x) + u\,X_g\,(x) \qquad x(0) = x_o \qquad x \in (M, \omega)$$

$$y = g(x) \tag{4.1.4}$$

where X_f and X_g are complete analytic Hamiltonian vector fields on a 2n - dimensional analytic connected symplectic manifold (M, ω).

The question to be answered here is under what conditions on the kernels of a stationary finite Volterra series does there exist a realization of the form (4.1.4) and conversely? See Crouch and Irving [28].

Recall that the Poisson bracket of functions on a symplectic manifold (M, ω) is given by

$$\{a, b\} = \omega(X_a, X_b) = X_b\,(a)\ ,\ a, b \in C(M) \tag{4.1.5}$$

Theorem 4.1.3 (Crouch [29]) For the system (4.1.4)

$$W_n\ (t,\ \sigma_1,\ldots,\sigma_n,x) = \{\ldots\{h(t),\ h(\sigma_1)\},\ h(\sigma_2)\ldots\}h(\sigma_n)\}(x) \qquad (4.1.6)$$

$$W_1\ (t,\ \sigma_1,\ x) = \{h(t),\ h(\sigma_1)\}(x)$$

where $h(\sigma)(x) = g_o\ \gamma_f(\sigma)(x)$, $\gamma_f(\sigma)$ the flow of X_f

Proof. From (4.1.2) and Jacobi's theorem 3.2.3

$$W_n\ (t,\ \sigma_1,\ldots,\sigma_n,x) = X_{g_o\gamma_f(\sigma_n)}\ (x)\ (W_{n-1}\ (t,\ \sigma_1,\ldots,\sigma_{n-1},\ x))$$

which by (4.1.5) becomes

$$\omega(x)\ (X_{W_{n-1}}(x),\ X_{h(\sigma_n)}(x)) = \{W_{n-1}(t,\ldots,\sigma_{n-1},x),\ h(\sigma_n)(x)\}$$

So for $n = 1$ applying Jacobi's theorem to (4.1.2) it becomes $\{h\ (t),\ h(\sigma_1)\}\ (x)$. The result now follows easily from induction using the above.

Q.E.D.

In Crouch [5] the following bracket operation applied to Volterra kernels was given, defined inductively by

$$W_n(\sigma_o\ldots\sigma_r[\ldots[\sigma_{r+1},\ \sigma_{r+2}]\ldots\sigma_q]\sigma_{q+1}\ldots\sigma_n)$$

$$= W_n\ (\sigma_o\ldots\sigma_r[\ldots[\sigma_{r+1},\sigma_{r+2}]\ldots\sigma_{q-1}]\sigma_q\ \sigma_{q+1}\ldots\sigma_n)$$

$$- W_n(\sigma_o\ldots\sigma_r\ \sigma_q[\ldots[\sigma_{r+1},\ \sigma_{r+2}]\ldots\sigma_{q-1}]\sigma_{q+1}\ldots\sigma_n)$$

These relations are used to give the following necessary condition.

Lemma 4.1.4 The Volterra kernels of a Hamiltonian system (4.1.3) satisfy the following for $n \geqslant 0$

$$W_n\,([\ldots[t,\ \sigma_1]\ldots\sigma_q]\ \sigma_{q+1}\ldots\sigma_n,x) \tag{4.1.7}$$

$$= (q+1)\ W_n\ (t,\ \sigma_1\ldots\sigma_q,\ \sigma_{q+1}\ldots\sigma_n,x) \text{ for } 1 \leqslant q \leqslant n$$

Proof. The proof is by induction on q. By (4.1.6) for q = 1

$$W_n\ ([t,\ \sigma_1],\ \sigma_2\ldots\sigma_n,x) = W_n\ (t,\ \sigma_1,\ \sigma_2\ldots\sigma_n,x)$$

$$- W_n\ (\sigma_1,\ t,\ \sigma_2\ldots\sigma_n,x)$$

$$= \{\ldots\{h\ (t),\ h(\sigma_1)\}\ h\ (\sigma_2)\ldots\}\ h(\sigma_n)\}(x)$$

$$- \{\ldots\{h\ (\sigma_1),\ h\ (t)\}\ h\ (\sigma_2)\ldots\}\ h(\sigma_n)\}(x)$$

$$= 2\ \{\ldots\{h\ (t),\ h(\sigma_1)\}\ h\ (\sigma_2)\ldots\}\ h\ (\sigma_n)\}(x)$$

by the anti-symmetry of Poisson brackets, $= 2\ W_n(t,\ \sigma_1,\ \sigma_2\ldots\sigma_n x)$

Assuming (4.1.7) is true for q then by definition and induction

$$W_n([\ldots[t, \sigma_1]\ldots\sigma_{q+1}]\sigma_{q+2}\ldots\sigma_n, x)$$

$$= (q+1)\, W_n(t, \sigma_1 \ldots \sigma_q\, \sigma_{q+1}\ldots\sigma_n, x)$$

$$- W_n(\sigma_{q+1}, [\ldots[t, \sigma_1]\ldots\sigma_q]\sigma_{q+2}\ldots\sigma_n, x)$$

By (4.1.6) and the Jacobi identity, applied to Poisson brackets, it is clear that

$$W_n(t, \sigma_1\ldots\sigma_r[\sigma_{r+1}, \sigma_{r+2}]\sigma_{r+3}\ldots\sigma_n, x)$$

$$= \{\ldots\{h(t), h(\sigma_1)\},\ldots\}, (\{h(\sigma_{r+1}, h(\sigma_{r+2})\})\}\ldots h(\sigma_n)\}(x)$$

Then by repeated application of the Jacobi identity

$$W_n(\sigma_{q+1}, [\ldots[t, \sigma_1]\ldots\sigma_q]\sigma_{q+2}\ldots\sigma_n, x)$$

$$= -\{\ldots(\{h(t), h(\sigma_1)\}\ldots h(\sigma_q)\}), h(\sigma_{q+1})\}\ldots h(\sigma_n)\}(x)$$

$$= - W_n(t, \sigma_1,\ldots,\sigma_q, \sigma_{q+1},\ldots,\sigma_n, x)$$

This verifies (4.1.7) for q+1, and hence result.

Q.E.D.

To prove the converse some results are required from the theory of permutation groups and their associated group algebras. These results were proved in Crouch and Irving [28] by Crouch, where they can be found in detail. For the purpose of this thesis these results will only be stated in order to give a proof of the main theorem, the consequences of which will be used in a later section.

Let S^{n+1} denote the symmetric group on (n+1) letters $\{0,1,\ldots,n\}$ and $\underset{\sim}{R}^{n+1}$ the space of (n+1) - tuples $(\sigma_o, \sigma_1,\ldots,\sigma_n)$, then the action of S^{n+1} on $\underset{\sim}{R}^{n+1}$ is given by

$$\rho(\sigma_o, \sigma_1,\ldots,\sigma_n) = (\sigma_{\rho(o)}, \sigma_{\rho(1)},\ldots,\sigma_{\rho(n)})$$

Let A^{n+1} be the group algebra generated by S^{n+1}, and F^{n+1} the space of real valued functions on $\underset{\sim}{R}^{n+1}$

$$(\sigma_o, \sigma_1,\ldots,\sigma_n) \longmapsto f\ (\sigma_o, \sigma_1,\ldots,\sigma_n).$$

A^{n+1} is represented as an algebra of linear operators on F^{n+1} by setting $(\rho f)\ (\sigma^n) = f(\rho(\sigma^n))$ for $\rho \in S^{n+1}$ and $\sigma^n = (\sigma_o, \sigma_1,\ldots,\sigma_n)$

Further, A^{k+1} may be considered as a subalgebra of A^{n+1} for $k < n$, by defining the action of A^{k+1} on $\underset{\sim}{R}^{n+1}$ to be that obtained by fixing $\sigma_{k+1},\ldots,\sigma_n$.

Let $\rho_k \in S^{k+1}$ be the permutation defined by the cycle $(k, k-1,\ldots,1,0)$, for $k \geq 1$ and set ρ_o to be the identity in S^{k+1}.

If $\rho_k \in A^{k+1}$ considered as subalgebra of A^{n+1} then

$$\rho_k\ (\sigma_o\ \sigma_1\ldots\sigma_k, \sigma_{k+1},\ldots,\sigma_n) = (\sigma_k\ \sigma_{k-1}\ldots\sigma_o\ \sigma_{k+1}\ldots\sigma_n)$$

<u>Lemma 4.1.5</u> If $\beta_n = (\rho_o-\rho_1)(\rho_o-\rho_2)\ldots(\rho_o-\rho_n) \subset A^{n+1}$ then $\beta_n^2 = (n+1)\ \beta_n$.

Now let $B_k^{a,\ b\ldots c}$ - for a, b,...,c belonging to some index set - be the function

$$\sigma^k \longmapsto \gamma_a(\sigma_o)\ \gamma_b\ (\sigma_1)\ldots\gamma_c(\sigma_k)$$

and set $W_k^{a,\ b\ldots c} = \beta_k\ B_k^{a,\ b\ldots c}$, then

<u>Theorem 4.1.6</u> If $W_n\ (\sigma_o,\ \sigma_1,\ldots,\sigma_n,x)$ is a differentiably separable kernel satisfying the conditions of lemma 4.1.4 then

$$W_n\ (\sigma_o,\sigma_1\ldots\sigma_n,x) = \sum_{i=1}^{N} W_n^{a_i b_i\ldots c_i}(\sigma_o,\ldots,\sigma_n,x)$$

for differentiable functions $\gamma_{a_i}\ldots\gamma_{c_i}$ $1 \leq i \leq N$. Moreover, this is uniquely defined.

Now with the aid of lemma 4.1.5 and theorem 4.1.6 the main result may be proven.

<u>Theorem 4.1.7</u> A finite Volterra series in which the kernels W_n $1 \leq n \leq r$, satisfying the stationarity condition (4.1.3) has a symplectic realization in the form of system (4.1.4) if and only if

(i) The kernels are differentiably separable

(ii) The kernels satisfy condition (4.1.7) of lemma 4.1.4, that is

$$W_n([\ldots[t,\sigma_1]\ldots\sigma_q]\sigma_{q+1}\ldots\sigma_n x) = (q+1)\ W_n(t,\ \sigma_1,\ldots,\sigma_n,\ x);\quad 1 \leq q \leq n$$

Proof. Implication follows from theorem 4.1.2 and lemma 4.1.4

Conversely, if the kernels are differentiably separable then by Brockett [7] the Volterra series has a bilinear realization. It therefore follows that the kernels may be written in the form

$$W_n(t, \sigma_1,\dots,\sigma_n) = \sum_{i=1}^{N} \gamma_n^i (t - \sigma_1)\, \gamma_{n-1}^i(\sigma_1 - \sigma_2)\dots\gamma_1^i(\sigma_{n-1} - \sigma_n) \tag{4.1.8}$$

where each function γ_k^i has the form

$$\gamma_k^i (\sigma) = c_{ik}' \; e^{A_{ik}\sigma} \; b_{ik}$$

for suitable c_{ik}, $b_{ik} \in \underset{\sim}{R}^s$ and $A_{ik} \in \underset{\sim}{R}^{s\times s}$ for some $s > 0$. This sum may be split into minimal groups still having the form given by the right hand side of equation (4.1.8), and also satisfying condition (4.1.7). Clearly each group is also differentiably separable. Select one such group and apply theorem 4.1.6 to see that it must be expressible in the form,

$$W_n^*(t, \sigma_1,\dots,\sigma_n) = \beta_n\gamma_n(t - \sigma_1)\, \gamma_{n-1}(\sigma_1 - \sigma_2)\dots\gamma_1(\sigma_{n-1} - \sigma_n)$$

for suitable functions $\gamma_1\dots\gamma_n$. This is the sum of 2^n terms, with each term constructed from the product of the functions $\gamma_1\dots\gamma_n$.

Let $\gamma_i(t) = c_i' \; e^{A_i t} \; b_i$, $1 \leq i \leq n$. Now a realization of the input-output map

$$g^* (t) = \int_0^t \dots \int_0^{\sigma_{n-1}} W_n^* (t, \sigma_1\dots,\sigma_n)\, u (\sigma_n)\dots u(\sigma_1) d\sigma_n \dots d\sigma_1 \tag{4.1.9}$$

is constructed.

With W_n^* replaced by

$$\gamma_n(t - \sigma_1)\, \gamma_{n-1}(\sigma_1 - \sigma_2)\ldots\gamma_1(\sigma_{n-1} - \sigma_n) \tag{4.1.10}$$

the above input-output map realizes the following system, see Brockett [7].

$$\begin{aligned}
\dot{q}_1 &= A_1 q_1 + b_1 u & q_1(0) &= 0\\
\dot{q}_2 &= A_2 q_2 + b_2 c_1' q_1 u & q_2(0) &= 0\\
&\vdots\\
\dot{q}_n &= A_n q_n + b_n c_{n-1}' q_{n-1} u & q_n(0) &= 0\\
y_1 &= c_n' q_n
\end{aligned}$$

Now, consider the composite system obtained by adjoining these equations to the equations

$$\begin{aligned}
-\dot{p}_1 &= A_1' p_1 + c_1 b_2' p_2 u & p_1(0) &= 0\\
-\dot{p}_2 &= A_2' p_2 + c_2 b_3' p_3 u & p_2(0) &= 0\\
&\vdots\\
-\dot{p}_n &= A_n' p_n + c_n u & p_n(0) &= 0
\end{aligned}$$

This composite system may be expressed as a Hamiltonian system by defining the Hamiltonian functions.

$$H(q, p) = p_1' A_1 q_1 + \ldots + p_n' A_n q_n$$

$$H_u(q, p) = p_1' b_1 + c_n' q_n + c_1' q_1 b_2' p_2 + \ldots + c_{n-1}' q_{n-1} b_n' p_n$$

If $q_i \in \underset{\sim}{R}^{n_i}$ and $\sum_{i=1}^{n} n_i = N$, the following system on $\underset{\sim}{R}^{2N}$ coincides with the composite system above.

$$\begin{aligned} \dot{q} &= Aq + u\,Nq + bu \qquad && q(0) = 0 \\ -\dot{p} &= A'p + uN'q + cu && p(0) = 0 \\ y_2 &= b'p + c'q + p'Nq && \end{aligned} \tag{4.1.11}$$

where

$$A = \begin{pmatrix} A_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & A_n \end{pmatrix}, \quad N = \begin{pmatrix} 0 & & & & 0 \\ b_2c_1' & 0 & & & \\ & b_3c_2' & 0 & & \\ & & \ddots & \ddots & \\ 0 & & & b_nc_{n-1}' & 0 \end{pmatrix}$$

$$b = \begin{pmatrix} b_1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \quad c = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ c_n \end{pmatrix}, \quad q = \begin{pmatrix} q_1 \\ \cdot \\ \cdot \\ \cdot \\ q_n \end{pmatrix}, \quad p = \begin{pmatrix} p_1 \\ \cdot \\ \cdot \\ \cdot \\ p_n \end{pmatrix}$$

With respect to the canonical symplectic structure J on $\underset{\sim}{R}^{2N}$ this is a Hamiltonian system, since

$$\dot{q} = \frac{\partial H}{\partial p}(q, p)' + \frac{\partial H_u}{\partial p}(q, p)'$$

$$-\dot{p} = \frac{\partial H}{\partial q}(q, p)' + \frac{\partial H_u}{\partial q}(q, p)'$$

$$y_2 = H_u(q, p).$$

It now remains to show that the input-output map of system (4.1.11) coincides with that in equation (4.1.9). Note that the output function in system (4.1.11) contains the term $c'q = c_n' q_n = y_1$, and so the input-output map of system (4.1.11) contains a term with Volterra kernel given by equation (4.1.10). It is an easy calculation to check that if $b_{r+1}' p_{r+1}$ and $c_r' q_r$ are viewed as outputs of system (4.1.11) then

$$c_r' q_r = \int_0^t \cdots \int_0^{\sigma_{r-1}} \gamma_r(t - \sigma_1)\, \gamma_{r-1}(\sigma_1 - \sigma_2) \cdots \gamma_1(\sigma_{r-1} - \sigma_r)$$

$$u(\sigma_r) \cdots u(\sigma_1)\, d\sigma_r \cdots d\sigma_1$$

$$b_{r+1}' p_{r+1} = \int_0^t \cdots \int_0^{\sigma_{n-r-1}} \gamma_{r+1}^*(t - \sigma_1) \cdots \gamma_n^*(\sigma_{n-r-1} - \sigma_{n-r})$$

$$u(\sigma_{n-r}) \cdots u(\sigma_1)\, d\sigma_{n-r} \cdots d\sigma_1$$

where $\gamma_k^*(\sigma) = -\gamma_k(-\sigma)$. It follows that $y_2(t)$ is a Volterra series with one term involving n iterated integrals of u, and associated Volterra kernel constructed from a sum of terms, each of which is in turn constructed from a product of all the functions $\gamma_1 \ldots \gamma_n$. It follows from Crouch [30], that the contribution of the term

$c'_r\, q_r\, b'_{r+1}\, p_{r+1}$, $1 \leq r \leq n-1$ is exactly $n!/r!\,(n-r)!$ such terms. Thus the Volterra defining $y_2(t)$ as a Volterra series consists of

$$1 + \sum_{r=1}^{n-1} n!/r!\,(n-r)! + 1 = 2^n$$

such terms. As system (4.1.11) is Hamiltonian, lemma 4.1.4 and the uniqueness of theorem 4.1.6 imply that the input-output map for system (4.1.11) coincides with the input-output map in equation (4.1.9) as claimed.

The proof is completed by noting that given a finite Volterra series satisfying the hypothesis of the theorem, it may be split into a finite number of terms, each of which is realizable by a Hamiltonian system of the form given by system (4.1.11). Obviously, the composite system is still Hamiltonian, and realizes the given finite Volterra series.

Q.E.D.

System (4.1.11) may be rewritten with the introduction of two new variables q_o and p_o as

$$\begin{bmatrix} \dot q_o \\ \dot q \\ \dot p_o \\ \dot p \end{bmatrix} = \begin{bmatrix} 0 & 0 & & \\ 0 & A & & 0 \\ & & 0 & 0 \\ & 0 & 0 & -A' \end{bmatrix} \begin{bmatrix} q_o \\ q \\ p_o \\ p \end{bmatrix} + u \begin{bmatrix} 0 & 0 & & \\ b & N & 0 & \\ 0 & -c' & 0 & -b' \\ -c & 0 & 0 & -N' \end{bmatrix} \begin{bmatrix} q_o \\ q \\ p_o \\ p \end{bmatrix}$$

$$y = p^{*'} N^* q^* + \frac{1}{2} q^{*'} Q^* q$$

$$q_o(0) = 1 \qquad\qquad p_o(0) = p^o \tag{4.1.12}$$

where $N^* = \begin{pmatrix} 0 & 0 \\ b & N \end{pmatrix}$, $Q^* = \begin{pmatrix} 0 & c' \\ c & 0 \end{pmatrix}$, $q^* = \begin{pmatrix} q_o \\ q \end{pmatrix}$, $p^* = \begin{pmatrix} p_o \\ p \end{pmatrix}$

System (4.1.12) has the same input-output map as system 4.1.11, since

$$p^{*'} N^* q^* + \frac{1}{2} q^{*'} Q^* q = (p_o' \ p') \begin{pmatrix} 0 & 0 \\ b & N \end{pmatrix} \begin{pmatrix} q_o \\ q \end{pmatrix}$$

$$+ \frac{1}{2} (q_o' \ \ q') \begin{pmatrix} 0 & 0 \dots c_n' \\ 0 & \\ \vdots & 0 \\ c_n & \end{pmatrix} \begin{pmatrix} q_o \\ q \end{pmatrix}$$

$$= (p_1'b_1 \ p'N) \begin{pmatrix} q_o \\ q \end{pmatrix} + \frac{1}{2} (q_n'c_n \ 0 \dots 0 \ q_o' \ c_n') \begin{pmatrix} q_o \\ q \end{pmatrix}$$

$$= p_1'b_1q_o + p'Nq + \frac{1}{2} q_n'c_nq_o + \frac{1}{2} q_o'c_n'q_n$$

$$= p_1'b_1 + p'Nq + c_n'q_n \qquad \text{Since } q_o = 1$$

$$= y_2$$

So system (4.1.12) is exactly the form found in chapter 3.1 if the Lie algebra $\{ad^k_{\overline{A}} \ \overline{B}\}_{L.A.}$ is assumed to be nilpotent. Here, of course, the matrices are block diagonal or block lower/upper triangular whereas before each was an individual element. Where

$$\overline{A} = \begin{pmatrix} 0 & 0 & 0 & \\ 0 & A & & \\ & 0 & 0 & 0 \\ & & 0 & -A' \end{pmatrix} , \quad \overline{B} = \begin{pmatrix} N^* & 0 \\ -Q^* & -N^{*'} \end{pmatrix}$$

Example 4.1.8 Consider the system

$$\dot{q}_1 = q_1 + u \qquad q_1(0) = 0$$

$$\dot{q}_2 = q_2 + u\, q_1 \qquad q_2(0) = 0$$

$$-\dot{p}_1 = p_1 + u\, p_2 \qquad p_2(0) = 0$$

$$-\dot{p}_2 = p_2 + u \qquad p_2(0) = 0$$

$$y = p_2 q_1 + p_1 + q_2$$

on the symplectic manifold $(T^*\underset{\sim}{R}^2, \sum_{i=1}^{2} dq_i \wedge dp_i)$

This is clearly a Hamiltonian system i.e. it is of the form of equation (4.1.4), with Hamiltonian functions given by

$$f = p_1 q_1 + p_2 q_2$$

$$g = p_2 q_1 + p_1 + q_2$$

and $\quad X_f = q_1 \,{}^{\partial}/_{\partial q_1} + q_2 \,{}^{\partial}/_{\partial q_2} - p_1 \,{}^{\partial}/_{\partial p_1} - p_2 \,{}^{\partial}/_{\partial p_2}$

$$X_g = {}^{\partial}/_{\partial q_1} + q_1 \,{}^{\partial}/_{\partial q_2} - p_2 \,{}^{\partial}/_{\partial p_1} - {}^{\partial}/_{\partial p_2}$$

$S = \{ad_{X_f}, X_g\}_{L.A.}$ is the ideal of $\mathcal{L}$ generated by

$$X_g = \partial/\partial q_1 + q_1\, \partial/\partial q_2 - p_2\, \partial/\partial p_1 - \partial/\partial p_2,\ \partial/\partial q_1 + \partial/\partial p_2,$$

$$\partial/\partial q_2 + \partial/\partial p_1,\ \partial/\partial q_1 - \partial/\partial p_2,\ \partial/\partial q_2 - \partial/\partial p_1$$

which is obviously strongly accessible since $\dim S(x) = T_x T^*\underset{\sim}{R}^2 \ \forall\ x \in T^*R^2$ and thus by Van der Schaft [31] weakly observable. Note also that S is a nilpotent Lie algebra.

Since,

$$q_1(t) = \int_o^t e^{t - \sigma_1} u(\sigma_1)\, d\sigma_1$$

$$q_2(t) = \int_o^t \int_o^{\sigma_1} e^{t - \sigma_2} u(\sigma_2)\, u(\sigma_1)\, d\sigma_2\, d\sigma_1$$

$$p_1(t) = \int_o^t \int_o^{\sigma_1} e^{-(t - \sigma_2)} u(\sigma_2)\, u(\sigma_1)\, d\sigma_2\, d\sigma_1$$

$$p_2(t) = -\int_o^t e^{-(t - \sigma_1)} u(\sigma_1)\, d\sigma_1$$

the Volterra series has the form

$$y(t) = \int_o^t \int_o^{\sigma_1} (e^{-(t - \sigma_2)} + e^{t - \sigma_2} - e^{\sigma_2 - \sigma_1} - e^{\sigma_1 - \sigma_2})\, u(\sigma_2)\, u(\sigma_1)\, d\sigma_2\, d\sigma_1$$

So $W_2(t, \sigma_1, \sigma_2, x) = e^{\sigma_2 - t} + e^{t - \sigma_2} - e^{\sigma_2 - \sigma_1} - e^{\sigma_1 - \sigma_2}$

which is seen to satisfy condition (4.1.7) of lemma 4.1.4 as follows,

$$W_2([t, \sigma_1], \sigma_2) = W_2(t, \sigma_1, \sigma_2) - W_2(\sigma_1, t, \sigma_2)$$

$$= e^{\sigma_2 - t} + e^{t-\sigma_2} - e^{\sigma_2-\sigma_1} - e^{\sigma_1-\sigma_2}$$

$$- e^{\sigma_2-\sigma_1} - e^{\sigma_1-\sigma_2} + e^{\sigma_2-t} + e^{t-\sigma_2}$$

$$= 2\,(e^{\sigma_2-t} + e^{t-\sigma_2} - e^{\sigma_2-\sigma_1} - e^{\sigma_1-\sigma_2})$$

$$= 2\,W_2(t, \sigma_1, \sigma_2)$$

and $W_2([[t, \sigma_1]\,\sigma_2]]) = W_2([t, \sigma_1], \sigma_2) - W_2(\sigma_2[t, \sigma_1])$

$$= 2\,W_2(t, \sigma_1, \sigma_2) - W_2(\sigma_2, t, \sigma_1)$$

$$+ W_2(\sigma_2, \sigma_1, t)$$

$$= 2\,W_2(t, \sigma_1, \sigma_2) - e^{\sigma_1-\sigma_2} - e^{\sigma_2-\sigma_1} + e^{\sigma_1-t}$$

$$+ e^{t-\sigma_1} + e^{t-\sigma_2} + e^{\sigma_2-t} - e^{t-\sigma_1} - e^{\sigma_1-t}$$

$= 3\,W_2(t, \sigma_1, \sigma_2)$ as required.

4.2 A Canonical Realization

In Crouch [5] dynamical realizations of finite Volterra series are investigated. In particular canonical coordinate charts are chosen to express strongly accessible, observable linear analytic realizations of finite Volterra series. Here the work of chapter 3 is applied to find realizations of the above type which include a symplectic structure.

Sussmann [9] guarantees the existence and uniqueness of minimal realizations i.e. for analytic systems accessible and observable realizations. However, for Hamiltonian systems the following observations from Goncalves [16] must be made.

Let G be a connected Lie group and suppose there exists a Poisson action of G on a strongly symplectic G-space, (M, ω). Note that (M, ω) is strongly symplectic if for example M is simply connected. The fact that the action is Poisson means that there exists a moment map τ.

Let $(\sum, x_o)$ represent an initialized analytic system of complete vector fields on the symplectic manifold (M, ω), and let O be the orbit in $\mathcal{L}^*$, the Lie algebra of G, of $\tau(x_o)$ by the coadjoint action of G. Let $(\sum_O, \tau(x_o))$ be the corresponding system on the co-adjoint orbit. Then with strong equivalence as defined in chapter 2.1 it is proved in Goncalves [16],

<u>Theorem 4.2.1</u> $(\sum_O, \tau(x_o))$ is strongly equivalent to $(\sum, x_o)$, and moreover it is an accessible Hamiltonian system.

It is noted in [16] that $\sum_O$ is accessible but may not be observable or weakly observable. However, the following theorem may be found in either Goncalves [16] or Van der Schaft [31].

Theorem 4.2.2 If Σ_o is strongly accessible it is weakly observable.

The following results are to be found in Crouch [5], and will prove useful throughout this section.

Lemma 4.2.3 Given a strongly accessible realization of a finite Volterra series of length n then

(i) The Volterra series has length n when evaluated at any point in the state space M.

(ii) The kernel W_n depends only on the time parameters t, $\sigma_1,\ldots,\sigma_n$, not on the state $x \in M$.

Theorem 4.2.4 Given a strongly accessible weakly observable realization of a finite Volterra series of length n, S is a nilpotent Lie algebra, with a descending central series of length less than or equal to n, and the Lie algebra $\mathcal{L}$ is solvable and finite dimensional.

Recall that S is the ideal in $\hat{\mathcal{L}}$ generated by the vector fields $g_1,\ldots,g_r$ of (2.1.1).

Since the systems under consideration are strongly accessible, the connected Lie subgroup N of G, corresponding to S, acts transitively on M and therefore M is analytically diffeomorphic to the homogeneous space N/N_{x_o} where N_{x_o} is the isotropy subgroup

$$N_{x_o} = \{g \in N : g.\, x_o = x_o\}$$

As N is nilpotent this is a nilmanifold. For further details of this point see Helgason [15].

One of the most important theorems in Crouch [5] states,

Theorem 4.2.5 A strongly accessible observable realization of a finite Volterra series has a state space which is diffeomorphic to a Euclidean space.

This shows that the underlying manifold is simply connected.

Finally from [5],

Corollary 4.2.6 A strongly accessible analytic realization of a finite Volterra series is observable if and only if it is weakly observable.

Now the results of chapter 3 are utilized to find a canonical realization of a linear analytic Hamiltonian strongly accessible, observable realization.

As N acts as a nilpotent transitive action on M by strong accessibility and theorem 4.2.4, then theorem 3.3.9 says it must be extended in order to make the action Poisson i.e. let $\tilde{N}$ be the nilpotent connected Lie group of the Lie subalgebra $\tilde{S} = S \times \mathbb{R}$. Further note that (M, ω) is always strongly symplectic as M in this case is simply connected by theorem 4.2.5.

Recall that the extended action was defined as follows, let

$$\mu \; : \; \tilde{N} \longmapsto N$$

and if $\tilde{g} \in \tilde{N}$, $x \in M$ then the action is

$$\tilde{g}.\, x = \mu(\tilde{g}).x$$

So obviously $\tilde{N}$ acts transitively on M.

A particularly interesting case of when accessibility and strong accessibility are equivalent is found when considering the concept of stationarity of linear analytic systems as described in Brockett [7].

Consider a linear analytic Hamiltonian system of the form

$$\dot{x} = X_H(x) + \sum_{i=1}^{r} u_i X_{H_{u_i}}(x) \qquad x \in (M, \omega)$$

$$x(0) = x_o$$

$$y^i = H_{u_i}(x) \qquad (4.2.1)$$

where (M, ω) is a 2m - dimensional symplectic manifold.

Then it is stationary if $X_H(x_o) = 0$, which is equivalent to $\text{expt}\, X_H . x_o = x_o$ where $\text{expt}\, X_H$ is the flow of X_H. By Sussmann [32] it then follows that $\dim S(x) = \dim \mathcal{L}(x)$, $\forall\, x \in M$ and thus accessibility and strong accessibility are equivalent.

Furthermore, if this system is lifted to the corresponding system on $(0, \omega_o)$ it too has this property, since if $X_H \in S$, let $\tilde{X}_H \in \tilde{S}$ then by definition of the action it follows that

$$\text{expt}\, \tilde{X}_H . x_o = \mu(\text{expt}\, \tilde{X}_H).x_o = \text{expt}\, X_H . x_o = x_o$$

so by Ad^* - equivalence (see lemma 3.3.5)

$$Ad^*_{\exp t\,\tilde{X}_H}{}_{o}\ \tau(x_o) = \tau_o \exp t\,\tilde{X}_H \cdot x_o$$

$$\Rightarrow \quad Ad^*_{\exp t\,\tilde{X}_H}\ f = \tau(x_o) = f$$

$$\Rightarrow \quad \sigma_{\exp t\,\tilde{X}_H}\ (f) = f$$

where $\sigma : \tilde{N} \times 0 \longmapsto 0$ as in chapter 3.3. Differentiating gives $d\sigma\ (\tilde{X}_H)\ (f) = 0$ as required.

Thus from theorem 4.2.1, theorem 4.2.2 and the fact that a stationary finite Volterra series which admits a linear analytic Hamiltonian realization is such that the linear analytic system is stationary, see Brockett [7], then the following corollary holds.

<u>Corollary 4.2.7</u> A stationary linear analytic Hamiltonian realization of a stationary finite Volterra series on a symplectic manifold (M, ω) is strongly equivalent to a minimal i.e. accessible and observable stationary linear analytic Hamiltonian realization on $(0, \omega_o)$.

The remainder of this section is mostly concerned with systems of the form (4.2.1) which have nilpotent Lie groups associated to them, and to which the results of chapter 3.3 apply. There are basically two cases to be considered. The first corresponds to the case where the system is strongly accessible and $\mathcal{L} = S$. The second case corresponds to the system being strongly accessible but $\mathcal{L} \neq S$.

First an in depth consideration of the case $S = \mathcal{L}$ on a strongly accessible, observable linear analytic Hamiltonian realization of the form 4.2.1 of a finite Volterra series. Theorem 4.2.5 gives S nilpotent and theorems 3.3.23 and 3.2.9 give rise to the following coordinate system.

Let $(q_1,\dots,q_n,\ p_1,\dots,p_n)$ be canonical coordinates on the symplectic manifold $(M^{2m},\ \omega)$ with

$$q_i = q^i_1,\dots,q^i_{r_i} \qquad 1 \leq i \leq n$$

$$p_i = p^i_1,\dots,p^i_{r_i} \qquad 1 \leq i \leq n$$

then the realization can be given by

$$\begin{array}{lll}
\dot{q}_1 = f_1\ (q_1) & & g^i_1\ (q_1) \\
\dot{q}_2 = f_2\ (q_1,\ q_2) & & g^i_2(q_1,\ q_2) \\
\vdots & + \sum\limits_{i=1}^{r} & u_i\ g^i_j\ (q_1,\dots,q_j) \\
\dot{q}_n = f_n\ (q_1,\dots,q_n) & & g^i_n\ (q_1,\dots,q_n)
\end{array}$$

$$-\dot{p}_1 = \frac{\partial f'_1}{\partial q_1}\, p_1 + \dots + \frac{\partial f'_n}{\partial q_1}\, p_n + \frac{\partial f'_o}{\partial q_1}$$

$$+ \sum_{i=1}^{r} u_i \quad \frac{\partial g^{i\,\prime}_1}{\partial q_1}\, p_1 + \dots + \frac{\partial g^{i\,\prime}_n}{\partial q_1}\, p_n + \frac{\partial g^{i\,\prime}_o}{\partial q_1}$$

$$\vdots$$

$$-\dot{p}_n = \frac{\partial f'_n}{\partial q_n}\, p_n + \frac{\partial f'_o}{\partial q_n} + \sum_{i=1}^{r} u_i \quad \frac{\partial g^{i\,\prime}_n}{\partial q_n}\, p_n + \frac{\partial g^{i\,\prime}_o}{\partial q_n}$$

where $x = (q_1,\dots,q_n, p_1,\dots,p_n)'$ $x(0) = x_o$

and $H = p_1' f_1(q_1)+\dots+p_n' f_n (q_1,\dots,q_n) + f_o (q_1,\dots,q_n)$

$y^i = H_{u_i} = p_1' g_1^i (q_1)+\dots+ p_n' g_n^i (q_1,\dots,q_n) + g_o^i (q_1,\dots,q_n)$.

where f_j and g_j^i are analytic functions in $q_1,\dots,q_j$, $1 \leq j \leq n$ and f_o and g_o^i are analytic functions in $q_1,\dots,q_n$.

In fact under further scrutiny the following theorem is obtained.

Theorem 4.2.8 Any strongly accessible observable linear analytic Hamiltonian realization of a finite Volterra series which satisfies condition 4.1.7 of lemma 4.1.4 and such that $S = \mathcal{L}$ has a canonical realization as follows, let $(q_1,\dots,q_n, p_1,\dots,p_n)$ be canonical coordinates on the state space (M^{2m}, ω) where

$$q_i = q_1^i,\dots,q_{r_i}^i , \qquad 1 \leq i \leq n$$

$$p_i = p_1^i,\dots,p_{r_i}^i , \qquad 1 \leq i \leq n$$

then the realization is expressed by

$$\begin{array}{lcl}
\dot{q}_1 = f_1 & & g_1^i \\
\dot{q}_2 = f_2 (q_1) & & g_2^i (q_1) \\
\vdots & +\sum\limits_{i=1}^{r} & u_i\, g_j^i (q_1,\dots,q_{j-1}) \\
\dot{q}_n = f_n (q_1,\dots,q_{n-1}) & & g_n^i (q_1,\dots,q_{n-1})
\end{array}$$

$$- \dot{p}_1 = \frac{\partial f'_2}{\partial q_1} p_2 + \dots + \frac{\partial f'_n}{\partial q_1} p_n + \frac{\partial f'_o}{\partial q_1}$$

$$+ \sum_{i=1}^{r} u_i \frac{\partial g^{i\,'}_2}{\partial q_1} p_2 + \dots + \frac{\partial g^{i\,'}_n}{\partial q_1} p_n + \frac{\partial g^{i\,'}_o}{\partial q_1}$$

$$\vdots$$

$$- \dot{p}_n = \frac{\partial f^1_o}{\partial q_n} + \sum_{i=1}^{r} u_i \frac{\partial g^{i\,'}_o}{\partial q_n} \qquad (4.2.2)$$

set $x = (q_1,\dots,q_n, p_1,\dots,p_n)'$ then $x(0) = x_o$

$$H = p'_1 f_1 + p'_2 f_2 (q_1) + \dots + p'_n f_n (q_1,\dots,q_{n-1}) + f_o (q_1,\dots,q_n)$$

$$y^i = H_{u_i} = p'_1 g^i_1 + p'_2 g^i_2 (q_1) + \dots + p'_n g^i_n (q_1,\dots,q_{n-1}) + g^i_o(q_1,\dots,q_n)$$

where f_o and g^i_j are analytic functions in $q_1,\dots,q_{j-1}$, $1 \le j \le n$ and f_o and g^i_o are analytic functions in $q_1,\dots,q_n$. Also $2\sum_{i=1}^{n} r_i = 2m$ the dimension of M.

<u>Proof.</u> Construct the coadjoint orbit on $\tilde{S}^* = \tilde{\mathcal{L}}^*$ as described previously.

From lemma 3.3.17 the involutive distributions on (O, ω_o) were constructed from a sequence of quotient algebras of the form

$$\{0\} = \tilde{S}_f / \tilde{S}_f \subset \dots \subset \tilde{\mathfrak{h}} \cap \tilde{S}^{n-i+1} + \tilde{S}_f / \tilde{S}_f \subset \dots \subset \tilde{\mathfrak{h}} \cap \tilde{S} + \tilde{S}_f / \tilde{S}_f \qquad (4.2.3)$$

The above sequence (4.2.3) can now be used to rewrite the sequence of distributions as follows

$$\Delta_i\,(Ad^*_{\tilde g}f) = \Delta_n\,(Ad^*_{\tilde g}f) \cap \bar{\Delta}_i\,(Ad^*_{\tilde g}f) \quad \forall\, \tilde g \in \tilde N$$

where $\bar{\Delta}_i\,(Ad^*_{\tilde g}f)$ is defined by $d\sigma(Ad_{\tilde g}(\tilde S^{n-i+1} + \tilde S_f))\,(Ad^*_{\tilde g}f)$

Let $\tilde X \in \tilde S$, $\tilde Y \in \tilde S^{n-i+1} + \tilde S_f$

so $\tilde Y = \tilde Y_1 + \tilde Y_2$ where $\tilde Y_1 \in \tilde S^{n-i+1}$, $\tilde Y_2 \in \tilde S_f$

Now $d\sigma(Ad_{\tilde g}\,\tilde Y_2)\,(Ad^*_{\tilde g}f) = 0$ by definition of

$\tilde S_f = \{\tilde X \in \tilde S : d\sigma(\tilde X)(f)(\tilde Z) = 0,\ \forall\, \tilde Z \in \tilde S\}$ alternatively $\tilde Y_2 \in \tilde S_f$ if

$$\langle f, [\tilde Y_2, \tilde Z]\rangle = 0,\ \forall\, \tilde Z \in \tilde S \tag{4.2.4}$$

Thus $\langle Ad^*_{\tilde g}f, [Ad_{\tilde g}\,\tilde Y_2, \tilde Z]\rangle$

$$= \langle f, [\tilde Y_2, Ad_{\tilde g}\,\tilde Z]\rangle = 0 \text{ by (4.2.4) so } Ad_{\tilde g}\,\tilde Y_2 \in \tilde S_{Ad^*_{\tilde g}f}$$

Hence $[d\sigma(Ad_{\tilde g}\,\tilde X),\ d\sigma(Ad_{\tilde g}\,\tilde Y)]\ (Ad^*_{\tilde g}f)$

$$= [d\sigma(Ad_{\tilde g}\,\tilde X),\ d\sigma(Ad_{\tilde g}\,\tilde Y_1)]\ (Ad^*_{\tilde g}f)$$

$$= d\sigma(Ad_{\tilde g}\,[\tilde X, \tilde Y_1])\,(Ad^*_{\tilde g}f) \in d\sigma(Ad_{\tilde g}\,\tilde S^{n-i+2})\,(Ad^*_{\tilde g}f)$$

$$= \bar{\Delta}_{i-1}\,(Ad^*_{\tilde g}f) \quad \forall\, \tilde g \in \tilde N$$

Thus if $\tilde Y \in \hat{\mathfrak{h}} \cap \tilde S^{n-i+1} + \tilde S_f$ then by $\tilde N$ - invariance on Δ_n and the above calculations

$$[\tau_*^{-1} d\sigma\ (Ad_{\tilde{g}}\ \tilde{X}),\ \tau_*^{-1} d\sigma\ (Ad_{\tilde{g}}\ \tilde{Y})]\ (Ad_{\tilde{g}}^{*} f)$$

$$= \tau_*^{-1}\ [d\sigma(Ad_{\tilde{g}}\ \tilde{X}),\ d\sigma(Ad_{\tilde{g}}\ \tilde{Y})]\ (Ad_{\tilde{g}}^{*} f)$$

$$\in\ \tau_*^{-1}(\Delta_n \cap \overline{\Delta}_{i-1})(Ad_{\tilde{g}}^{*} f)) \qquad \forall\ \tilde{g} \in \tilde{N} \tag{4.2.5}$$

So by construction there exists a basis of Hamiltonian vector fields

$$d\sigma(Ad_{\tilde{g}}\ \tilde{Y}_1^1)\ \ (Ad_{\tilde{g}}^{*} f), \ldots, d\sigma(Ad_{\tilde{g}}\ \tilde{Y}_r^1)\ (Ad_{\tilde{g}}^{*} f), \ldots, d\sigma(Ad_{\tilde{g}}\ \tilde{Y}_1^i)(Ad_{\tilde{g}}^{*} f), \ldots$$

$$\ldots,\ d\sigma(Ad_{\tilde{g}}\ \tilde{Y}_{r_i}^i)\ (Ad_{\tilde{g}}^{*} f), \qquad 1 \leq i \leq n$$

which span $\Delta_i\ (Ad_{\tilde{g}}^{*} f)$, $\forall\ \tilde{g} \in \tilde{N}$ on $(O,\ \omega_o)$. Let $d\sigma(\tilde{Y}_1), \ldots, d\sigma(\tilde{Y}_m)$ be the Hamiltonian vector fields, which exist by transitivity, that span the involutive distribution transversal to Δ_n.

Notice that it is possible to construct canonical coordinates on $(O,\ \omega_o)$ similar to that on $(M,\ \omega)$ in theorem 3.3.23 by simple application of the moment map τ, i.e. the condition

$$[X_k,\ X_j^i] \in\ \tau_*^{-1}\ \Delta_i$$

was required, let $\tau_*^{-1}\ d\sigma\ (\tilde{Y}_k) = X_k$ and $\tau_*^{-1}\ d\sigma\ (\tilde{Y}_j^i) = X_j^i$ then

$$[d\sigma\ (\tilde{Y}_k),\ d\sigma\ (\tilde{Y}_j^i)] \in\ \Delta_i$$

and there exists canonical coordinates

$(\tilde{q}_1, \ldots, \tilde{q}_m,\ \tilde{p}_1, \ldots, \tilde{p}_m)$ on $(O,\ \omega_o)$ such that

$\partial/\partial\tilde{p}_1, \ldots,\ \partial/\partial\tilde{p}_i$ spans $\Delta_i \quad 1 \leq i \leq n$

where

$$q_i = (q_1^i, \ldots, q_{r_i}^i) \qquad 1 \leq i \leq n$$

$$p_i = (p_1^i, \ldots, p_{r_i}^i) \qquad 1 \leq i \leq n$$

Observe that $d\sigma\,(Ad_{\tilde{g}}\tilde{X})\,(Ad_{\tilde{g}}^{*}f) = Ad_{\tilde{g}\,*}^{*}\; d\sigma(\tilde{X})\,(f)$

since

$$d\sigma\,(Ad_{\tilde{g}}\tilde{X})\,(Ad_{\tilde{g}}^{*}f)\,(\tilde{Z})$$

$$= \langle ad^{*}_{Ad_{\tilde{g}}\tilde{X}}\;\; Ad_{\tilde{g}}^{*}f,\ \tilde{Z}\rangle \qquad \tilde{g} \in \tilde{N}, \quad \tilde{Z} \in \tilde{\mathcal{L}}$$

$$= -\langle Ad_{\tilde{g}}^{*}f,\ [Ad_{\tilde{g}}\tilde{X},\ \tilde{Z}]\rangle$$

$$= -\langle f,\ Ad_{\tilde{g}}^{-1}\ [Ad_{\tilde{g}}\tilde{X},\ \tilde{Z}]\rangle$$

$$= -\langle f,\ [\tilde{X},\ Ad_{\tilde{g}}^{-1}\ \tilde{Z}]\rangle$$

$$= \langle ad_{\tilde{X}}^{*}f,\ Ad_{\tilde{g}^{-1}}\ \tilde{Z}\rangle$$

$$= Ad_{\tilde{g}\,*}^{*}\ d\sigma(\tilde{X})\ (f)\ (\tilde{Z})$$

thus

$$[d\sigma(Ad_{\tilde{g}}\tilde{X})\ (Ad_{\tilde{g}}^{*}f),\ d\sigma\ (Ad_{\tilde{g}}\ \tilde{Y}_j^i)\ (Ad_{\tilde{g}}^{*}f)]$$

$$= Ad_{\tilde{g}\,*}^{*}\ [d\sigma\ (\tilde{X}),\ d\sigma\ (\tilde{Y}_j^i)](f) \qquad (4.2.6)$$

Also from the construction of the proof of theorem 3.2.8

$$d\sigma\ (\tilde{Y}_j^i)(f) = \partial/_{\partial\, \tilde{p}_j^i}$$

and by invariance

$$Ad^*_{\tilde{g}} * d\sigma\,(\tilde{Y}^i_j)\,(f) = \partial/\partial\tilde{p}^i_j$$

i.e. $$Ad^*_{\tilde{g}} * \partial/\partial\tilde{p}^i_j = Ad^*_{\tilde{g}} * X_{\tilde{q}^i_j}\,(f) = X_{\tilde{q}^i_j} \circ Ad^*_{\tilde{g}^{-1}}\,(Ad^*_{\tilde{g}} f)$$

$$= \partial/\partial\tilde{p}^i_j$$

where $X_{\tilde{q}^i_j}$ is the Hamiltonian vector field of $\partial/\partial\tilde{p}^i_j$.

So as $d\sigma\,(\tilde{X}) \in$ Ham (O, ω_o), locally

$$\Big[\sum_{\ell=1}^{r_k} \sum_{k=1}^{n} \frac{\partial\tilde{H}}{\partial\tilde{p}^k_\ell}\, \partial/\partial\tilde{q}^k_\ell - \frac{\partial\tilde{H}}{\partial\tilde{q}^k_\ell}\, \partial/\partial\tilde{p}^k_\ell\,,\ \partial/\partial\tilde{p}^i_j \Big]$$

$$\in \text{span}\,\{\partial/\partial\tilde{p}_i, \ldots, \partial/\partial\tilde{p}_{i-1}\} \qquad \text{(by 4.2.5)}$$

After straightforward computation exactly of the form of theorem 3.2.9 with

$X_{\tilde{H}} = d\sigma\,(\hat{X})$ it is seen that

$$\tilde{H} = \sum_{i=1}^{n} \tilde{p}'_i\, \tilde{f}_i\,(\tilde{q}_1, \ldots, \tilde{q}_{i-1}) + f_o(\tilde{q}_1, \ldots, \tilde{q}_n)$$

similarly for $\tilde{H}_{u_k}$. Now using the moment map τ, canonical coordinates $(q^1_1, \ldots, q^n_{r_n}, p^1_1, \ldots, p^n_{r_n})$ can be found on (M, ω) with $H = \tau^*\tilde{H}$ and

$H_{u_k} = \tau^*\tilde{H}_{u_k}$

then

$$H = \sum_{i=1}^{n} p_i' \; f_i \; (q_1,\dots,q_{i-1}) + f_o \; (q_1,\dots,q_n)$$

and also

$$H_{u_k} = \sum_{i=1}^{n} p_i' \; g_i^k \; (q_1,\dots,q_{i-1}) + g_o^k \; (q_1,\dots,q_n) \quad \text{for } 1 \leq k \leq r$$

as required.

Q.E.D.

Theorem 4.1.3 gave the following formula for calculating the Volterra series of a Hamiltonian system of the form (4.1.4) i.e.

$$W_n \; (t, \sigma_1,\dots,\sigma_n, x) = \{W_{n-1}(t, \sigma_1,\dots,\sigma_{n-1}, x), h(\sigma_n)(x)\}$$

$$W_1 \; (t, \sigma_1, x) = \{h \; (t), h(\sigma_1)\}(x)$$

where $h(\sigma)(x)) = g_o \; \gamma_f(\sigma)(x)$

By use of the Campbell-Baker-Hausdorff formula and known identities of Hamiltonian vector fields on symplectic manifolds this can be rewritten in a form more amenable for calculation. The Campbell-Baker-Hausdorff formula is given by

$$\gamma_f(-\sigma)_* \; X_g \; (\gamma_f(\sigma)x) = \sum_{i=o}^{\infty} \frac{\sigma^i}{i!} \; ad_{X_f}^i \; X_g \; (x)$$

where $ad_{X_f}^o \; X_g = X_g$, $ad_{X_f}^i \; X_g = ad_{X_f}^{i-1} \; [X_f, X_g]$ and $\gamma_f(t)$ is the flow of X_f. Also note,

$$\mathrm{ad}_{X_f} X_g = [X_f, X_g] = X_{\{f,g\}}$$

Now $\{h(t)(x), h(\sigma_1)(x)\}$

$$= \{g_o \gamma_f(t), g_o \gamma_f(\sigma_1)\}(x)$$

$$= \omega(x)\,(X_{g_o\gamma_f(t)}(x), X_{g_o\gamma_f(\sigma_1)}(x))$$

$$= \omega(x)\,(\gamma_f(-t)_* X_g(\gamma_f(t)\,x), \gamma_f(-\sigma_1)_* X_g(\gamma_f(\sigma_1)\,x)$$

$$= \sum_{i=o}^{\infty} \sum_{j=o}^{\infty} \frac{t^j \sigma_1^i}{j!\, i!}\, \omega(x)\,(\mathrm{ad}^j_{X_f} X_g(x), \mathrm{ad}^i_{X_f} X_g(x)) \qquad (4.2.7)$$

The following detailed example of a strongly accessible, weakly observable linear analytic symplectic realization of a finite Volterra series shows the relevant computation.

<u>Example 4.2.9</u> Consider the following system on $(\underset{\sim}{R}^4, \sum_{i=1}^{2} dq_i \wedge dp_i)$

$$\dot{q}_1 = u \qquad q_1(0) = 0$$

$$\dot{q}_2 = q_1^2 + q_1 + u \qquad q_2(0) = 0$$

$$-\dot{p}_1 = (2q_1+1)p_2 + q_2 \qquad p_1(0) = 0$$

$$-\dot{p}_2 = q_1 \qquad p_2(0) = 0$$

$$y(q,p) = p_1 + p_2$$

Here $f = (q_1^2 + q_1)\, p_2 + q_1 q_2$

$g = p_1 + p_2$

and

$$X_f = (q_1^2 + q_1)\,{}^{\partial}/_{\partial q_2} - (2q_1 + 1)\,p_2\,{}^{\partial}/_{\partial p_1} - q_2\,{}^{\partial}/_{\partial p_1} - q_1\,{}^{\partial}/_{\partial p_2}$$

$$X_g = {}^{\partial}/_{\partial q_1} + {}^{\partial}/_{\partial q_2}$$

This system is strongly accessible since

$$[X_g\,,\ X_f] = (2q_1 + 1)\,{}^{\partial}/_{\partial q_2} - 2p_2\,{}^{\partial}/_{\partial p_1} - {}^{\partial}/_{\partial p_1} - {}^{\partial}/_{\partial p_2}$$

$$[X_g,\ [X_g,\ X_f]] = 2\,{}^{\partial}/_{\partial q_2}$$

$$[X_f\,,\ [X_g,\ X_f]] = 2q_1\,{}^{\partial}/_{\partial p_1}$$

$$[X_g,\ [X_f\ [X_g,\ X_f]]] = 2\,{}^{\partial}/_{\partial p_1}$$

and thus $S = \{ad_{X_f}\,,\ X_g\}_{L.A.}$ is such that $\dim S(x) = \dim T_x\underset{\sim}{R}^4\ \ \forall\, x \in \underset{\sim}{R}^4$ and S is nilpotent.

It is weakly observable either by simple computation to check the algebraic conditions of Krener and Hermann [13] i.e.

$$dg = dp_1 + dp_2$$

$$d\,L_{X_f} g = -d((2q_1 + 1)\,p_2 + q_2 + q_1)$$

$$= -\,(2q_1\,dp_2 + 2p_2\,dq_1 + dq_2 + dq_1)$$

$$d\, L_{[X_g, X_f]}\, g = -\, d\,(2p_2 + 2)$$

$$= -\, 2\, dp_2$$

$$d\, L_{[X_f[X_g, X_f]]}\, g = d\,(2q_1)$$

$$= 2\, dq_1$$

and from this it is clear that $\dim d\mathcal{X}(x) = \dim T_x^* \underset{\sim}{R}^4 \quad \forall\, x \in \underset{\sim}{R}^4$ and thus the system is weakly observable. Alternatively, it is easy to show that a Hamiltonian system of the form (4.1.4) is strongly accessible if and only if it is weakly observable from the observation that if $h \in \mathcal{X}$

$$X_f h \; = \; L_{X_f} h \; = \; \{f,\, h\}$$

and $d\, L_{X_f} h = d\,\{f,\, h\} = \overline{\omega}\;\; ([X_f,\, X_h]\,)$

where $\overline{\omega}\;\; : X \longmapsto \omega(X,\, \cdot\,)$. Thus if $\{ad_{X_f},\; X_g\}_{L.A.}$ spans so does $d\,\mathcal{X}$. For further details see Van der Schaft [31]. Moreover, if such a system is weakly observable by Crouch [5] it is observable.

The following computations of Poisson brackets are required

$$\{f,\, g\} = (2q_1 + 1)p_2 + q_2 + q_1$$

$$\{g\;\; \{f,\, g\}\} = -\, 2p_2 - 1$$

$$\{f\;\{f,\, g\}\} = q_1\;(2q_1 + 1)$$

$\{\{f, g\}, \{f\{f, g\}\}\} = 0$

These are needed to calculate the Volterra series for this system.

$$W_o(t, x) = g(\gamma_f(t)x)$$

$$W_1(t, \sigma_1, x) = \{h(t)(x), h(\sigma_1)(x)\} \qquad (4.2.8)$$

where $h(\sigma) = g_o \gamma_f(\sigma)$

Using (4.2.7), (4.2.8) becomes

$$\sum_{i=o}^{\infty} \sum_{j=o}^{\infty} \frac{t^j \sigma_1^i}{j!\, i!} \omega(ad^j_{X_f} X_g, ad^i_{X_f} X_g)(x)$$

Now $\omega(X_f, X_g) = \{f, g\}$ and $\omega(X_f, [X_g, X_h]) = \{f\{g, h\}\}$

so $(4.2.8) = \sigma_1 \omega(X_g, ad_{X_f} X_g)(x)$

$$+ t\, \omega(ad_{X_f} X_g, X_g)(x) + \frac{t\sigma_1^2}{2} \omega(ad_{X_f} X_g, ad^2_{X_f} X_g)(x)$$

$$+ \frac{t^2\sigma_1}{2} \omega(ad^2_{X_f} X_g, ad_{X_f} X_g)(x) + \ldots$$

$$= \sigma_1 \omega(X_g, ad_{X_f} X_g)(x) + t\, \omega(ad_{X_f} X_g, X_g)(x)$$

(all other terms vanish)

$$= -\sigma_1 (2p_2 + 1) + t(2p_2 + 1)$$

$$= (t - \sigma_1)(2p_2 + 1)$$

By (4.2.8)

$$W_2\ (t,\ \sigma_1,\ \sigma_2,\ x) = \{W_1\ (t,\ \sigma_1,\ x),\ h\ (\sigma_2)(x)\}$$

$$= \omega(x)\ (X_{W_1(t,\ \sigma_1)}(x),\ \ X_{h(\sigma_2)}\ (x))$$

Again using the Campbell-Baker-Hausdorff formula this can be rewritten as

$$= \sum_{i=1}^{\infty} \frac{\sigma_2^i}{i!}\ \omega\ (X_{W_1(t,\ \sigma_1)}\ ,\ ad^i_{X_f}\ X_g)\ (x) \qquad (4.2.9)$$

$$X_{W_1} = 2\ (t - \sigma_1)\ {}^{\partial}/_{\partial q_2}$$

$$\{W_1,\ g\} = 0$$

$$\{W_1,\ \{f,\ g\}\} = 2\ (t - \sigma_1)$$

$$\{W_1,\ \{f,\ \{f,\ g\}\}\} = 0$$

Computing (4.2.9)

$$W_2\ (t,\ \sigma_1,\ \sigma_2,\ x) = 2\ \sigma_2\ (t - \sigma_1)$$

Hence the Volterra series for this system corresponds to

$$y(t) = \int_o^t\ (2p_2 + 1)\ (t - \sigma_1)\ u\ (\sigma_1)\ d\sigma_1$$

$$+ \int_o^t \int_o^{\sigma_1} 2\ \sigma_2\ (t - \sigma_1)\ u\ (\sigma_2)\ u\ (\sigma_1)\ d\sigma_2\ d\sigma_1.$$

Notice that $W_1(t, \sigma_1, x)$ does depend on the state variable but $W_2(t, \sigma_1, \sigma_2, x)$ does not. This is as expected from lemma 4.2.3 as S is given by

$$S = \{ \partial/_{\partial q_1} + \partial/_{\partial q_2},\ (2q_1 + 1)\ \partial/_{\partial q_2} - 2\,p\ \partial/_{\partial p_1} - \partial/_{\partial p_1} - \partial/_{\partial p_2},$$
$$2q_1\ \partial\ {}_{\partial p_1}\}_{L.A.}$$

which has step length 2 i.e. $ad_X^3 = 0 \ \forall\ X \in S$.

Further observe that Δ_2 is spanned by the Hamiltonian vector fields $[X_g, X_f]$ and $[X_g[X_f[X_g\ X_f]]]$ and Δ_1 is spanned by $[X_g[X_f[X_g\ X_f\ ]]]$.

<u>Corollary 4.2.10</u> With the same conditions as theorem 4.2.7 but with $S \neq \mathcal{L}$ and stationarity assumed then the canonical realization is of the form.

$$\begin{aligned}
\dot{q}_1 &= A_1 q_1 & & g_1^i \\
\dot{q}_2 &= A_2 q_2 + f_2(q_1) & & g_2^i(q_1) \\
&\vdots & + \sum_{i=1}^{r} u_i \quad & g_j^i\ (q_1, \ldots, q_{j-1}) \\
\dot{q}_n &= A_n q_n + f_n(q_1, \ldots, q_{n-1}) & & g_n^i\ (q_1, \ldots, q_{n-1})
\end{aligned}$$

$$-\dot{p}_1 = A_1' p_1 + \frac{\partial f_2'}{\partial q_1} p_2 + \ldots + \frac{\partial f_n'}{\partial q_1} p_n + \frac{\partial f_o'}{\partial q_1}$$

$$+ \sum_{i=1}^{r} u_i \quad \frac{\partial g_2^{i'}}{\partial q_1} p_2 + \ldots + \frac{\partial g_n^{i'}}{\partial q_1} p_n + \frac{\partial g_o^{i'}}{\partial q_1}$$

$$\vdots$$

$$-\dot{p}_n = A_n' p_n + \frac{\partial f_o'}{\partial q_n} + \sum_{i=1}^{r} u_1 \frac{\partial g_o^{i'}}{\partial q_n}$$

$$x = (q_1, \ldots, q_n, p_1, \ldots, p_n)', \qquad x(0) = x_o$$

$$H = p_1' A_1 q_1 + p_2' (A_2 q_2 + f_2 (q_1)) + \ldots + p_n' (A_n q_n + f_n (q_1, \ldots, q_{n-1}))$$

$$+ f_o (q_1, \ldots, q_n)$$

$$g^i = H_{u_i} = p_1' g_1^i + p_2' g_2^i (q_1) + \ldots + p_n' g_n^i (q_1, \ldots, q_{n-1})$$

$$+ g_o^i (q_1, \ldots, q_n)$$

$$1 \leq i \leq r$$

Proof. Follows on application of lemma 3.3.25 by letting F = expt $\tilde{X}_H$ where X_H is the free vector field on (M, ω), $\tau(z_o) = f$ so $Ad^*_{\text{expt}\ \tilde{X}_H} f = f$

follows from Ad^* - equivariance as previously shown. Also as the commutator in this case for F = expt $\tilde{X}_H$ obviously vanishes, expt $\tilde{X}_H$ is a valid choice for an F which satisfies the conditions of lemma 3.3.25. Furthermore,

$$ad_{\tilde{X}_H} : \tilde{S} \longmapsto \tilde{S}$$

and so clearly

$$\mathrm{ad}_{\tilde{X}_H} : \tilde{S}^i \longmapsto \tilde{S}^i .$$

Now $\tilde{S}_f$ can be interpreted as those vector fields which vanish at f, thus

$$\mathrm{ad}_{\tilde{X}_H} : \tilde{S}_f \longmapsto \tilde{S}_f$$

so $$\mathrm{ad}_{\tilde{X}_H} : \tilde{S}^i + \tilde{S}_f \longmapsto \tilde{S}^i + \tilde{S}_f \qquad (4.2.10)$$

A sequence of distributions satisfying the conditions of theorem 3.2.8 can now be found exactly as in the previous case using the polarization of lemma 3.3.25.

As in theorem 4.3 of Crouch [5] there is a decomposition of $G = V\,N$ where V is the one parameter subgroup with generator X_{H_1} such that $X_{H_1}(x_o) = 0$ and $X_H = X_{H_1} + X_{H_2}$, $\quad X_{H_2} \in S$.

So as above

$$\mathrm{ad}_{\tilde{X}_{H_1}} : \tilde{S}^i + \tilde{S}_f \longrightarrow \tilde{S}^i + \tilde{S}_f \qquad (4.2.11)$$

which is seen to be a linear endomorphism from Crouch [5].

Thus using the same basis for $\Delta_i\ (\mathrm{Ad}^*_{\tilde{g}} f)$ as in theorem 4.2.8., (4.2.11) gives that

$$[d\sigma(Ad_{\tilde g}\,\tilde X_{H_1}),\ d\sigma(Ad_{\tilde g}\,\tilde Y^i_j)]\ (Ad^*_{\tilde g}f)$$

$$= d\sigma\ Ad_{\tilde g}\ ([\tilde X_{H_1},\ \tilde Y^i_j])\ (Ad^*_{\tilde g}f)$$

$$= d\sigma\ (Ad_{\tilde g}\ (\sum_{k=1}^{i}\ \sum_{\ell=1}^{r_k}\ \tilde a_{k\ell}\ \hat Y^k_\ell))\ (Ad^*_{\tilde g}f)$$

Thus, with the introduction of canonical coordinates as in theorem 4.2.8 on $(0,\ \omega_o)$, this becomes

$$[\ \sum_{t=1}^{n}\ \sum_{s=1}^{r_t}\ \frac{\partial\tilde H_1}{\partial\tilde p^t_s}\ {}^{\partial}/_{\partial\tilde q^t_s}\ -\ \frac{\partial\tilde H_1}{\partial\tilde q^t_s}\ {}^{\partial}/_{\partial\tilde p^t_s}\ ,\ {}^{\partial}/_{\partial\tilde p^i_j}\ ]$$

$$= \sum_{k=1}^{i}\ \sum_{\ell=1}^{r_k}\ \tilde a_{k\ell}\ {}^{\partial}/_{\partial\hat p^k_\ell} \qquad 1 \leq i \leq n$$

$$1 \leq j \leq r_i$$

$\tilde a_{k\ell}$ constants.

$$\Rightarrow \tilde H_1 = \tilde p'\ \tilde A\tilde q$$

where $\tilde q = (\tilde q^1_1,\ldots,\tilde q^n_{r_n})$ $p = (\tilde p^1_1,\ldots,\tilde p^n_{r_n})$

As the other vector fields belong to $\tilde S$ their structure is determined by theorem 4.2.8. The required structure is obtained by observing that

$$H = \tau^*\tilde H = \tau^*\ (\tilde H_1 + \tilde H_2)$$

which means

$$H\ (q_1,\ldots,q_n,\ p_1,\ldots,p_n)$$
$$= \sum_{i=1}^{n} p_i'\ (A_i q_i + f_i\ (q_1,\ldots,q_{i-1}) + f_o(q_1,\ldots,q_n)$$

with $f_1 = 0$ by stationarity.

Q.E.D.

4.3. Interconnections

In this section the systems of § 4.1 are shown to be intimately connected with systems that arise in the theory of electrical networks.

The results here depend on a system theoretical framework described by Willems [32] and Van der Schaft [33]. The basic motivation behind this approach is that it is often difficult to determine which space is actually the input space and which the output. Hence the following definition.

<u>Definition 4.3.1</u> A differential system in state space form is described by:

(i) analytic manifolds X, W

(ii) an analytic bundle $\pi : B \longmapsto X$

(iii) an analytic map $f : B \longmapsto TX \times W$ for which the diagram

$$\begin{tikzcd} B \arrow[rr, mapsto, "f"] \arrow[dr, "\pi"'] & & TX \times W \arrow[dl, "\pi_X"] \\ & X & \end{tikzcd}$$

commutes.

The system itself is defined by $\sum := \{(x, w) : \underset{\sim}{R} \longmapsto X \times W : x$ absolutely continuous and $(\dot{x}(t), w(t)) \in f(\pi^{-1}(x(t)))$ a.e.$\}$ and will be denoted by $\sum(X, W, B, f)$.

Obviously analyticity in the above definition may be replaced by a less stringent assumption but that is not necessary here.

To make such a system Hamiltonian the appropriate extra structure must be added to give, Definition 4.3.2 A Hamiltonian system in state space form is described by:

$$\begin{tikzcd} B \arrow[r, mapsto, "f"] \arrow[d, "\pi"'] & TM \times W \arrow[r, mapsto, "\pi_2"] \arrow[dl, "\pi_M"] \arrow[d, "\pi_1"] & W \\ M & TM & \end{tikzcd}$$

commutes

(i) (M, ω) is a symplectic manifold of dimension 2n.

(ii) (W, ω^e) is a symplectic manifold of dimension 2m.

It can be seen, see for example Abraham and Marsden [1], that TM x W can be made into a symplectic manifold by defining the symplectic form to be

$$\Omega = \pi_1^* \dot{\omega} - \pi_2^* \omega^e \quad \text{on } TM \times W$$

where $\dot{\omega}$ is the induced symplectic form on the symplectic manifold TM constructed from M, that is as ω is nondegenerate it defines a bundle isomorphism $\bar{\omega}$ from TM to T^*M by setting

$$\bar{\omega}(X) = i(X)\omega \text{ for } X \in TM.$$

as T^*M is symplectic and has a naturally defined 1 - form θ then $\bar{\omega}^* \theta$ is a 1-form on TM and let $\dot{\omega} = d\,\bar{\omega}^* \theta$ which is now clearly a symplectic form on TM.

Definition 4.3.3 If $\sum(M, W, B, f)$ is a system with M and W symplectic manifolds then it is said to be (i) full Hamiltonian if f(B) is a

Lagrangian submanifold of $(TM \times W, \Omega)$ and (ii) degenerate Hamiltonian if there exists a full Hamiltonian system $\sum'(M, W, B', f')$ such that $f(B)$ is a submanifold of $f'(B')$.

Definition 4.3.4 Let (W_i, ω_i^e) $i=1,\ldots,k$ be symplectic manifolds. Then $(W_1 x\ldots xW_k, \pi_1^* \omega_1^e + \ldots + \pi_k^* \omega_k^e)$ is a symplectic manifold. An interconnection of $(W_i)_{i=1,\ldots,k}$ is a submanifold $I \subset W_1 x\ldots xW_k$. An interconnection is full (degenerate) Hamiltonian if I is a Lagrangian (co-isotropic) submanifold of $(W_1 x\ldots xW_k, \pi_1^* \omega_1^e + \ldots + \pi_k^* \omega_k^e)$.

After a number of restrictions are placed upon the above definition it is shown in Van der Shaft [33] that Kirchhoff's laws are a special type of interconnection. Also in Van der Schaft [33] it is shown that the following theorem holds.

Theorem 4.3.5 Let $\sum_i (M_i, W_i, B_i, f_i)$ $i=1,\ldots,k$ be Hamiltonian systems interconnected by a Hamiltonian interconnection $I \subset W_1 x\ldots xW_k$. Then the resulting system is a Hamiltonian system

$$\sum_I (M_1 x\ldots xM_k, W_1 x\ldots xW_k, B_I, f_I).$$

Proof. The product system is constructed given by $\sum (M_1 x\ldots xM_k, W_1 x\ldots xW_k, B_x, f_x)$ of the systems $\sum_i (M_i, W_i, B_i, f_i)$, $i=1,\ldots,k$. Let $x = (x_1,\ldots,x_k) \in M_1 x\ldots xM_k$. Since $\pi_i : B_i \longmapsto M_i$ are fibre bundles there exists neighbourhoods $U_i \subset M_i$ of x_i such that $\pi_i^{-1}(U_i) \cong U_i \times F_i$, F_i is the standard fibre, B_x is now defined locally as $\pi_x : (U_1 x\ldots xU_k) \times (F_1 x\ldots xF_k) \longmapsto U_1 x\ldots xU_k$.

Next define $f_x : B_x \mapsto T(M_1 x \ldots x M_k) \times W_1 x \ldots x W_k$ locally as $f_x = (f_1, \ldots, f_k)$.
Obviously if $\sum_i (M_i, W_i, B_i, f_i)$ are Hamiltonian then
$\sum (M_1 x \ldots x M_k, W_1 x \ldots x W_k, B_x, f_x)$ is Hamiltonian either full or degenerate.
Also the interconnected system must satisfy $f_I(B_I) \subset f_x(B_x)$. Thus definition 4.2.3 gives result.

Q.E.D.

The main result of this section may now be given.

Theorem 4.3.6 A stationary linear analytic symplectic realization of a stationary finite Volterra series can be decomposed into a set of linear Hamiltonian systems with a full Hamiltonian interconnection.

Proof. Theorem 4.1.7 gives the conditions on the Volterra series for a stationary linear analytic symplectic realization to exist. Furthermore it is clear from the constructions given in the proof of theorem 4.1.7 that a Volterra series has a stationary linear analytic symplectic realization if and only if it has a stationary bilinear symplectic realization. This bilinear realization is in the form of system (4.1.11) i.e.

$$\dot{q} = Aq + u\,Nq + bu \qquad q(0) = 0$$

$$-\dot{p} = A'p + u\,N'p + cu \qquad p(0) = 0$$

$$y = b'p + c'q + p'Nq$$

where $A = \begin{bmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_n \end{bmatrix}$, $\quad N = \begin{bmatrix} 0 & & & 0 \\ b_2 c_1' & 0 & & \\ & \ddots & \ddots & \\ 0 & & b_n c_{n-1}' & 0 \end{bmatrix}$

$$b = \begin{bmatrix} b_1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ c_n \end{bmatrix}, \quad q = \begin{bmatrix} q_1 \\ \cdot \\ \cdot \\ \cdot \\ q_n \end{bmatrix}, \quad p = \begin{bmatrix} p_1 \\ \cdot \\ \cdot \\ \cdot \\ p_n \end{bmatrix}$$

This can be decomposed into n linear Hamiltonian systems as follows,

$$(1) \quad \begin{aligned} \dot{q}_1 &= A_1 q_1 + b_1 u_1^1, \; q_1(0) = 0, \\ -\dot{p}_1 &= A_1' p_1 + c_1 u_2^1 \; p_1(0) = 0, \end{aligned} \quad \begin{bmatrix} y_1^1 \\ y_2^1 \end{bmatrix} = \begin{bmatrix} 0 & b_1' \\ c_1' & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ p_1 \end{bmatrix}$$

$$(2) \quad \begin{aligned} \dot{q}_2 &= A_2 q_2 + b_2 u_1^2, \; q_2(0) = 0, \\ -\dot{p}_2 &= A_2' p_2 + c_2 u_2^2, \; p_2(0) = 0, \end{aligned} \quad \begin{bmatrix} y_1^2 \\ y_2^2 \end{bmatrix} = \begin{bmatrix} 0 & b_2' \\ c_2' & 0 \end{bmatrix} \begin{bmatrix} q_2 \\ p_2 \end{bmatrix}$$

$$\vdots$$

$$(n) \quad \begin{aligned} \dot{q}_n &= A_n q_n + b_n u_1^n, \; q_n(0) = 0, \\ -\dot{p}_n &= A_n' p_n + c_n u_2^n, \; p_n(0) = 0, \end{aligned} \quad \begin{bmatrix} y_1^n \\ y_2^n \end{bmatrix} = \begin{bmatrix} 0 & b_n' \\ c_n' & 0 \end{bmatrix} \begin{bmatrix} q_n \\ p_n \end{bmatrix}$$

with interconnection I satisfying

$$\begin{aligned} u_1^1 &= u \\ u_2^1 &= u \;\; y_1^2 \\ u_1^2 &= u \;\; y_2^1 \\ u_2^2 &= u \;\; y_1^3 \\ &\vdots \end{aligned}$$

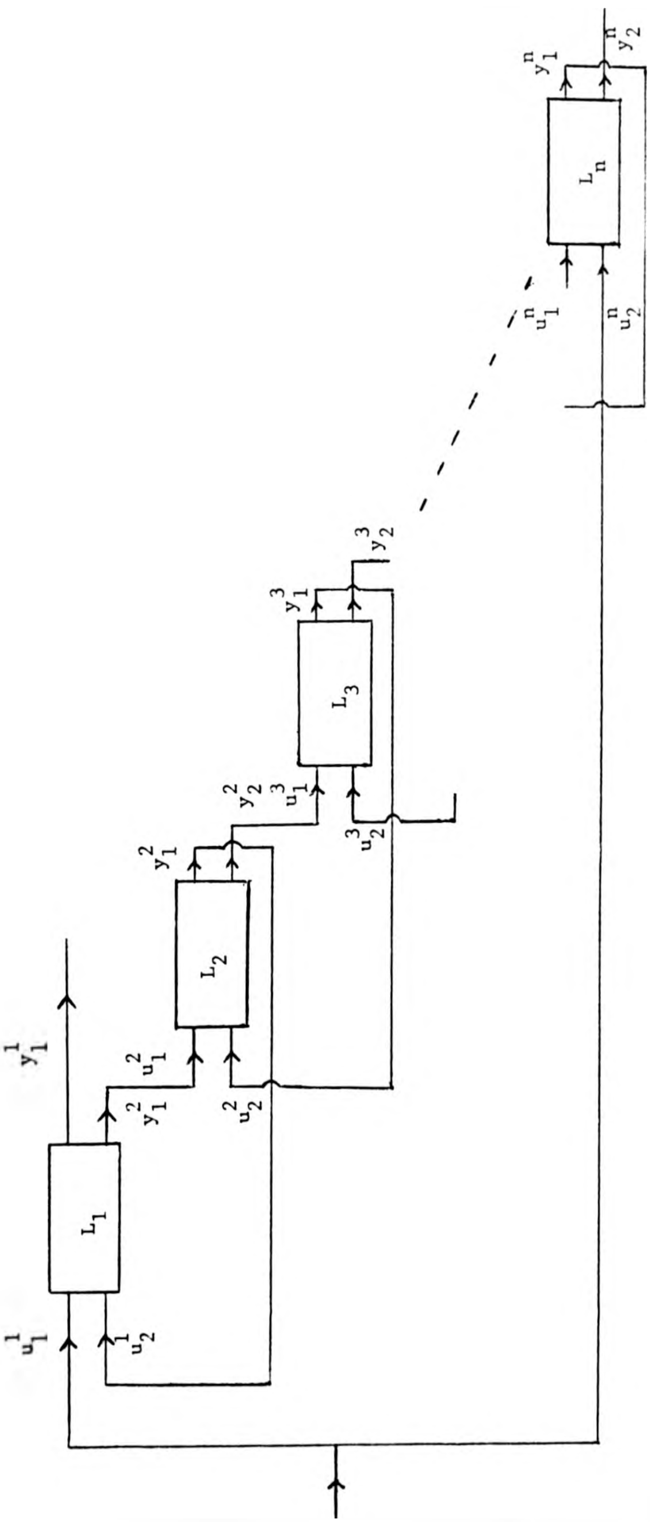
$$\begin{aligned}
&\vdots \quad \vdots\\
&u_1^i = u \quad y_2^{i-1}\\
&u_2^i = u \quad y_1^{i+1}\\
&\vdots \quad \vdots\\
&u_1^n = u \quad y_2^{n-1}\\
&u_2^n = u
\end{aligned}$$

with the usual symplectic form on $W_1 x \ldots x W_n$ given locally by $dy_1^1 \wedge du_1^1 + dy_2^1 \wedge du_2^1 + \ldots + dy_1^n \wedge du_1^n + dy_2^n \wedge du_2^n$. So on I this form is, with the usual assumption that $u \in \mathcal{U}_{p.c.}$,
$u\, dy_1^2 \wedge dy_2^1 + u\, dy_2^1 \wedge dy_1^2 + u\, dy_1^2 \wedge dy_2^2 + u\, dy_2^2 \wedge dy_1^3 + \ldots + u\, dy_1^n \wedge dy_2^{n-1}$
$+ u\, dy_2^{n-1} \wedge dy_1^n = 0.$

Hence I is a Lagrangian submanifold since it is described by 2N constraints on a 4N - dimensional symplectic manifold and hence a full Hamiltonian interconnection.

Q.E.D.

Schematically this can be represented by



L_1
L_2
L_3
L_n
u_1^1
u_2^1
y_1^1
y_1^2
u_1^2
u_2^2
y_2^2
u_1^3
u_2^3
y_1^3
y_2^3
u_1^n
u_2^n
y_1^n
y_2^n

A similar decomposition procedure may be obtained for a linear analytic Hamiltonian realization of the form appearing in Corollary 4.2.9. For simplicity assume only one control then this system may be written as n - linear Hamiltonian Systems of the form

$$\text{(1)}\quad \begin{aligned} \dot{q}_1 &= A_1 q_1 + u_1^1,\ q_1(0) = q_1^o, \\ -\dot{p}_1 &= A_1' p_1 + u_2^1,\ p_1(0) = p_1^o, \end{aligned} \qquad \begin{bmatrix} y_1^1 \\ y_2^1 \end{bmatrix} \quad \begin{bmatrix} 0 & I_{r_1} \\ I_{r_1} & 0 \end{bmatrix} \quad \begin{bmatrix} q_1 \\ p_1 \end{bmatrix}$$

$$\text{(2)}\quad \begin{aligned} \dot{q}_2 &= A_2 q_2 + u_1^2,\ q_2(0) = q_2^o, \\ -\dot{p}_2 &= A_2' p_2 + u_2^2,\ p_2(0) = p_2^o, \end{aligned} \qquad \begin{bmatrix} y_1^2 \\ y_2^2 \end{bmatrix} = \begin{bmatrix} 0 & I_{r_2} \\ I_{r_2} & 0 \end{bmatrix} \quad \begin{bmatrix} q_2 \\ p_2 \end{bmatrix}$$

$$\vdots$$

$$\text{(n)}\quad \begin{aligned} \dot{q}_n &= A_n q_n + u_1^n,\ q_n(0) = q_n^o, \\ -\dot{p}_n &= A_n' p_n + u_2^n,\ p_n(0) = p_n^o, \end{aligned} \qquad \begin{bmatrix} y_1^n \\ y_2^n \end{bmatrix} = \begin{bmatrix} 0 & I_{r_n} \\ I_{r_n} & 0 \end{bmatrix} \quad \begin{bmatrix} q_n \\ p_n \end{bmatrix}$$

where I_{r_i} is the $r_i \times r_i$ identity matrix with interconnection I satisfying

$${}^k u_1^1 = \partial V/\partial\,{}^k y_1^1\ , \quad \forall\, k \in \{1,\ldots,r_1,1,\ldots,r_2,\ldots,1,\ldots,r_n\}$$

$${}^k u_2^1 = \partial V/\partial\,{}^k y_2^1$$

$${}^k u_1^2 = \partial V/\partial\,{}^k y_1^2$$

$${}^k u_2^2 = \partial V/\partial\,{}^k y_2^2$$

$$\vdots$$

$${}^k u_1^n = \partial V/\partial\,{}^k y_1^n$$

$${}^k u_2^n = \partial V/\partial\,{}^k y_2^n$$

where $u_j^i = ({}^1u_j^i, \ldots, {}^{r_i}u_j^i)$, $\quad 1 \leq i \leq n$

$1 \leq j \leq 2$

$$y_j^i = ({}^1y_j^i, \ldots, {}^{r_i}y_j^i)$$

and

$$V = y_1^{2\,'} f_2(y_2^1) + \ldots + y_1^{n\,'} f_n(y_2^1, \ldots, y_2^{n-1}) + f_o(y_2^1, \ldots, y_2^n)$$

$$+ u\,(y_1^{1\,'} g_1 + y_1^{2\,'} g_2(y_2^1) + \ldots + y_1^{n\,'} g_n(y_2^1, \ldots, y_2^{n-1}) + g_o(y_2^1, \ldots, y_2^n))$$

The symplectic form on $W_1 x \ldots x W_n$ is given locally by,

$$\sum_{\ell=1}^{n} \sum_{k=1}^{r_\ell} d^k y_1^\ell \wedge d^k u_1^\ell + d^k y_2^\ell \wedge d^k u_2^\ell$$

and on this form I vanishes since

$$\sum_{\ell=1}^{n} \sum_{k=1}^{r_\ell} d^k y_1^\ell \wedge d\left(\frac{\partial V}{\partial^k y_1^\ell}\right) + d^k y_2^\ell \wedge d\left(\frac{\partial V}{\partial^k y_2^\ell}\right)$$

$$= \sum_{\ell=1}^{n} \sum_{k=1}^{r_\ell} \sum_{s=1}^{n} \sum_{j=1}^{r_s} \left(\frac{\partial^2 V}{\partial^k y_1^\ell\, \partial^j y_1^s} d^k y_1^\ell \wedge d^j y_1^s + \frac{\partial^2 V}{\partial^k y_1^\ell\, \partial^j y_2^s} d^k y_1^\ell \wedge d^j y_2^s\right)$$

$$+ \sum_{\ell=1}^{n} \sum_{k=1}^{r_\ell} \sum_{s=1}^{n} \sum_{j=1}^{r_s} \left(\frac{\partial^2 V}{\partial^k y_2^\ell\, \partial^j y_1^s} d^k y_2^\ell \wedge d^j y_1^s \right.$$

$$\left. + \frac{\partial^2 V}{\partial^k y_2^\ell\, \partial^j y_2^s} d^k y_2^\ell \wedge d^j y_2^s\right)$$

is clearly zero as the wedge product ($\wedge$) is anti-symmetric.

Furthermore, as $W_i = \underset{\sim}{R}^{4 r_i}$, $1 \leq i \leq n$ and I is described by $2 \sum_{j=1}^{n} r_j = \frac{1}{2} \dim (W_1 x \ldots x W_n)$ constraints I is a Lagrangian submanifold and thus, by definition, a full Hamiltonian interconnection.

Chapter 5

COMPLETE INTEGRABILITY OF NILPOTENT SYSTEMS

In this chapter a generalization of the concept of a transitive abelian group action on a Lagrangian submanifold of a symplectic manifold as in Liouville's theorem [2] is found by considering a nilpotent group action which is the first stage to a control theoretic version of Liouville's theorem. It is shown that such systems can also be integrated by quadratures.

5.1 Integrability of Hamiltonian Systems

In this section a feasible framework is constructed in order to look at families of Hamiltonian vector fields on a symplectic manifold (M, ω) which have a nilpotent Lie algebra structure. This, in some sense, extends the classical case of complete integrability as given by Liouville's theorem where, essentially, abelian Lie algebras are of interest.

First, however, the symplectic vector space case is investigated.

Let (V, ω) be a $2n$ - dimensional symplectic vector space and $\mathcal{L} \subset gl(V)$ a Lie subalgebra of Hamiltonian endomorphisms. Assume

$U_1 = \{v \in V : H(v) = 0$ for all $X_H \in \mathcal{L}\}$ is an isotropic subspace of V, and $\overline{W}$ is a Lagrangian subspace of V transversal to the level surface U_1 i.e. $\overline{W} \cap U_1 = \{0\}$. U_1 can be extended to a Lagrangian subspace, W, as follows. If U_1 is not already Lagrangian then

$$U_1 \subsetneq U_1^{\perp}$$

It is claimed that $U^{\perp} \subset U_1 + \overline{W}$ for if $U_1^{\perp} \subset U_1 + \overline{W} \Rightarrow U_1^{\perp} \cap \overline{W}^{\perp} \subset U_1$ $\Rightarrow U_1^{\perp} \cap \overline{W} \subset U_1$, so $U_1 \cap \overline{W} = \{0\} \Rightarrow U_1^{\perp} \cap \overline{W} = \{0\} \Rightarrow \dim U_1^{\perp} \leq n$

$$\Rightarrow \dim U_1 \geq n$$

which is a contradiction.

So there exists a $v \in U_1^{\perp}$, $v \notin U_1 + \overline{W}$. Let $U_2 = U_1 + \underset{\sim}{R}v$ then $\dim U_2 = \dim U_1 + 1$. All that is now required is to prove that U_2 is isotropic. So suppose

$$w \in U_2 \cap \overline{W} \Rightarrow w = v_1 + \lambda v \text{ with } v_1 \in U_1 \Rightarrow \lambda v = w - v_1 \in U_1 + \overline{W}$$

but $v \notin U_1 + \overline{W}$ so $\lambda = 0$, this implies that $w \in U_1 \cap \overline{W} = \{0\} \Rightarrow w = 0$. In fact this automatically proves that $\overline{W}$ is transversal to U_2 as well as U_2 being isotropic.

Continue by induction until U_m is reached with $\dim U_m = n = \frac{1}{2} \dim V$. Let $W = U_m$ which is a Lagrangian subspace.

<u>Proposition 5.1.1</u> Let (V, ω) be a $2n$ - dimensional symplectic vector space with W and $\overline{W}$ as above. Suppose $X_H . \overline{W} \subset \overline{W}$ and $X_H . W \subset W$, $\forall X_H \in \mathcal{L}$ and that the restriction X_i of X_{H_i}, $\forall X_{H_i} \in \mathcal{L}$, to W is a nilpotent endomorphism on W. Then there exists a symplectic basis of V such that in canonical coordinates relative to this basis each $X_H \in \mathcal{L}$ can be expressed by

$$X_H = \begin{pmatrix} N & 0 \\ 0 & -N' \end{pmatrix}$$

where $N \in \underset{\sim}{R}^{n \times n}$ is strictly lower triangular.

<u>Proof.</u> By Engel's theorem 3.1.2 there exists an ordered basis for W, say $e_{q_1},\dots,e_{q_n}$ such that in coordinates relative to this basis each X_i can be expressed by a strictly lower triangular matrix. Then, as in theorem 3.1.11 a complementary basis may be found on $\overline{W}$, say $e_{p_1},\dots,e_{p_n}$ such that $(e_{q_1},\dots,e_{q_n}, e_{p_1},\dots,e_{p_n})$ is a symplectic basis on V.

Let $(q_1,\dots,q_n, p_1,\dots,p_n)$ be canonical coordinates relative to this basis then each X_H takes the form

$$X_H = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad \begin{matrix} a, b, c, d \in \underset{\sim}{R}^{n\times n} \\ a = -d', b = b', c = c' \end{matrix}$$

as each $X_H \in \mathcal{L}$ is by assumption infinitesimally symplectic.

Let $(0, \overline{w})' \in \overline{W}$, then since $X_H . \overline{W} \subset \overline{W}$, $\forall X_H \in \mathcal{L}$

$$X_H . (0, \overline{w})' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 \\ \overline{w} \end{pmatrix} = \begin{pmatrix} b\overline{w} \\ d\overline{w} \end{pmatrix} \subset \overline{W}$$

which implies $b \equiv 0$. Thus $X_H = \begin{pmatrix} a & 0 \\ c & -a' \end{pmatrix}$

and so $H = p'a\,q + \frac{1}{2} q'c\,q$.

But $H = 0$ on W, so $c = 0$ on W, as W is described by $p_1 = \dots = p_n = 0$, and thus on V. Hence, $X_H = \begin{pmatrix} a & 0 \\ 0 & -a' \end{pmatrix}$

But X_H restricted to W is strictly lower triangular, thus $a = N$ which gives the result.

Q.E.D.

The following example introduces the concept of transitivity of a nilpotent Lie group on a Lagrangian subspace of a symplectic vector space as well as providing an example of proposition 5.1.1.

<u>Example 5.1.2</u> Consider the following bilinear Hamiltonian system on $(\underset{\sim}{R}^4, J)$ with canonical coordinates (q_1, q_2, p_1, p_2)

$\dot{x} = X_H + uX_{H_u}$, $x(0) = 0$, x $(\underset{\sim}{R}^4, J)$ where $x = (q_1, q_2, p_1, p_2)'$ and $H = p_2q_1$, $H_u = p_2q_1 + p_1$. This can be expanded as,

$$
\begin{aligned}
\dot{q}_1 &= u & q_1(0) &= 0 \\
\dot{q}_2 &= q_1 + u\, q_1 & q_2(0) &= 0 \\
-\dot{p}_1 &= p_2 + u\, p_2 & p_1(0) &= 0 \\
-\dot{p}_2 &= 0 & p_2(0) &= 0
\end{aligned}
\qquad (5.1.1)
$$

Now consider the level surface given by

$$L = \{v \in \underset{\sim}{R}^4 : H_i(v) = 0, \ \forall X_{H_i} \in \mathcal{L} = \{X_H, X_{H_u}\}_{L.A.}\}.$$

Since $[X_{H_u}, X_H] = \partial/\partial q_2 = X_{p_2}$ and all other brackets vanish, L may be described by $p_1 = p_2 = 0$. Thus L is a Lagrangian subspace of $\underset{\sim}{R}^4$ and further the restriction of the vector fields of $\mathcal{L}$ to L span L everywhere and are nilpotent.

Furthermore, by introducing, the new coordinates q_o, p_o, (5.1.1) can be rewritten as in chapter 4.2 as,

$$\begin{bmatrix} \dot{q}_o \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{p}_o \\ \dot{p}_1 \\ \dot{p}_2 \end{bmatrix} = \underbrace{\begin{bmatrix} 0&0&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&0&0 \\ 0&0&0&0&0&-1 \\ 0&0&0&0&0&0 \end{bmatrix}}_{X_{H^e}} \begin{bmatrix} q_o \\ q_1 \\ q_2 \\ p_o \\ p_1 \\ p_2 \end{bmatrix} + u \underbrace{\begin{bmatrix} 0&0&0&0&0&0 \\ 1&0&0&0&0&0 \\ 0&1&0&0&0&0 \\ 0&0&0&0&1&0 \\ 0&0&0&0&0&-1 \\ 0&0&0&0&0&0 \end{bmatrix}}_{X_{H_u^e}} \begin{bmatrix} q_o \\ q_1 \\ q_2 \\ p_o \\ p_1 \\ p_2 \end{bmatrix}$$

on $(\underset{\sim}{R}^6, J^e)$ with canonical coordinates $(q_o, q_1, q_2, p_o, p_1, p_2)$,

$J^e = \begin{pmatrix} 0 & I_3 \\ -I_3 & 0 \end{pmatrix}$ and initial conditions $q_o(0) = 1$, $q_1(0) = q_2(0) = 0$,

$p_o(0) = p^o$, $p_1(0) = p_2(0) = 0$.

Now it is necessary to pass to the extended level surface U_1 on $\underset{\sim}{R}^6$ of L on $\underset{\sim}{R}^4$ in order to make a comparison with proposition 5.1.1 where

$$U_1 = \{v \in \underset{\sim}{R}^6 : H_i^e(v) = 0, \forall\ X_{H_i^e} \in \mathcal{L}^e = \{X_{H^e}, X_{H_u^e}\}_{L.A.}\}$$

Now $H^e = p_2 q_1$

$H_u^e = p_2 q_1 + p_1 q_o$

and $\{H_u^e, H^e\} = -p_2 q_o$ with all other Poisson brackets vanishing.
So U_1 is an isotropic subspace described by $p_1 = p_2 = 0$. In this case the obvious extension is a Lagrangian subspace W is

$$W = \{v \in \underset{\sim}{R}^6 : H_i^e(v) = 0, \ \forall X_{H_i^e} \in \mathcal{L}^e \text{ and } p_o = 0\}$$

which may be described by $p_o = p_1 = p_2 = 0$.

Obviously theorem 5.1.1 can be proved with nilpotent replaced by solvable and assuming the field over the Lie subalgebra $\mathcal{L} \subset g\ell\,(V)$ is algebraically closed, then by using Lie's theorem 3.1.14 each element of $\mathcal{L}$ has the matrix representation

$$X_H = \begin{pmatrix} S & 0 \\ 0 & -S' \end{pmatrix} \quad \text{where } S \in \underset{\sim}{R}^{n\text{x}n} \text{ is lower triangular.}$$

The next part of this section considers a globalization of the above.

Let $\{X_{H_o}, X_{H_1}, \ldots, X_{H_m}\}$ be a family of Hamiltonian vector fields on a 2 n - dimensional symplectic manifold (M, ω). Denote by $\mathcal{L}$ the Lie algebra generated by these vector fields i.e. $\mathcal{L} = \{X_{H_o}, X_{H_1}, \ldots, X_{H_m}\}_{L.A.}$

and suppose its corresponding Lie group is given by G.

Let expt X_H represent the flow of the Hamiltonian vector field X_H on M, $\forall X_H \in \mathcal{L}$. Define a group action on M as follows

$$\Phi : G \times M \longmapsto M$$

$$\Phi\,(\text{expt}_1 X_{H_1} \circ \cdots \circ \text{expt}_k X_{H_k}, x)$$

$$= \text{expt}_1 X_{H_1} \circ \cdots \circ \text{expt}_k X_{H_k} . x$$

Consider the level surface

$$N = \{x \in M : H(x) = 0, \quad \forall X_H \in \mathcal{L}\}$$

For the rest of this section N is assumed to be a Lagrangian submanifold of M.

By Kostant's corollary 3.2.5 in a neighbourhood U of N in M, M can be identified with T^*N. As the results in the remainder of this section are local, without loss of generality, M is assumed to be T^*N.

Let $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ be canonical coordinates on T^*N. N can be thought of as being the zero section and thus defined by $p_1=\ldots=p_n = 0$.

Let Δ be a polarization in T^*N transverse to the Lagrangian submanifold N. By proposition 3.2.7 locally Δ is spanned by $\partial/\partial p_1,\ldots,\partial/\partial p_n$.

Therefore the corresponding condition here to that of $X_H \cdot \overline{W} \subset \overline{W}$ in theorem 5.1.1 is

$$[X_H, \Delta] \subset \Delta, \quad \forall\, X_H \in \mathcal{L} \tag{5.1.2}$$

<u>Proposition 5.1.3</u> Each Hamiltonian function, H, of a Hamiltonian vector field on the symplectic manifold (T^*N, ω_N) satisfying (5.1.2), locally has the form $H = \sum_{i=1}^{n} p_i\, f_i\,(q_1,\ldots,q_n) + f_o\,(q_1,\ldots,q_n)$

<u>Proof.</u> It is sufficient to perform the following calculation with respect to the canonical coordinates $(q_1,\ldots,q_n, p_1,\ldots,p_n)$

$$[X_H, f(q,p)\, {}^{\partial}/_{\partial p_j}] = [\sum_{i=1}^{n} \frac{\partial H}{\partial p_i}\, {}^{\partial}/_{\partial q_i} - \frac{\partial H}{\partial q_i}\, {}^{\partial}/_{\partial p_i},\ f(q,p)\, {}^{\partial}/_{\partial p_j}]$$

$$= \sum_{i=1}^{n} f\, \frac{\partial^2 H}{\partial p_j \partial q_i}\, {}^{\partial}/_{\partial p_i} - f\, \frac{\partial^2 H}{\partial p_j \partial p_i}\, {}^{\partial}/_{\partial q_i} + \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i}\, {}^{\partial}/_{\partial p_j} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i}\, {}^{\partial}/_{\partial p_j}$$

$$\subset \Delta \text{ by (5.1.2)} \qquad\qquad 1 \leq j \leq n$$

This implies $\frac{\partial^2 H}{\partial p_i \partial p_j} = 0,\ \forall\ 1 \leq i,\ j \leq n$

$\Rightarrow$ H is linear in each p_i, $1 \leq i \leq n$.

$\Rightarrow$ $H = \sum_{i=1}^{n} p_i f_i\,(q_1,\dots,q_n) + f_o\,(q_1,\dots,q_n)$

Q.E.D.

This structure does appear naturally as the following proposition found in Abraham and Marsden [1] shows.

<u>Proposition 5.1.4</u> Let expt X_H be the flow of a Hamiltonian vector field X_H on a symplectic manifold (T^*N, ω_N), where $\omega_N = \sum_{i=1}^{n} dq_i \wedge dp_i$ locally. Assume that expt $X_H : T^*N \longmapsto T^*N$ preserves the fibres of T^*N, then locally the Hamiltonian function, H, is given by

$$H = \sum_{i=1}^{n} p_i f_i\,(q_1,\dots,q_n) + f_o\,(q_1,\dots,q_n).$$

However, by assumption each H = 0 on N i.e. on $p_1 = \dots = p_n = 0$ which implies $f_o\,(q_1,\dots,q_n) = 0$ on N, and since it only depends on the q_i coordinates, on T^*N. So each $H = p'f\,(q)$ where $q = (q_1,\dots,q_n)'$, $p = (p_1,\dots,p_n)'$ for all $X_H \in \mathcal{L}$.

Again these structures appear quite naturally as the next proposition from Abraham and Marsden [1] shows.

Proposition 5.1.5 Suppose expt X_H is the flow of a Hamiltonian vector field X_H on the symplectic manifold (T^*N, ω_N), and further assume that it preserves the natural 1 - form θ_N, $\omega_N = -d\theta_N$. Then, locally, H is of the form $H = \sum_{i=1}^{n} p_i f_i\ (q_1,\ldots,q_n)$.

The following structure theorem on nilpotent Lie algebras of Matsushima [34] is of importance.

Theorem 5.1.6 Let G be a connected simply connected nilpotent group with corresponding Lie algebra $\mathcal{L}$. Then a basis $X_1,\ldots,X_n$ of $\mathcal{L}$ can be chosen which has the property $\{X_{i+1},\ldots,X_n\}$ is an ideal of $\{X_i,\ldots,X_n\}$ for i = 1,...,n - 1, and hence every element in G may be written uniquely in the form $\exp t_1 X_1 \ldots \exp t_n X_n$.

The Hamiltonian vector fields on M, such that each Hamiltonian function $H = p' f\ (q)$ can be decomposed as follows, let

$$X_{H_1} = \overline{X}_1 + X_1$$

$$X_{H_2} = \overline{X}_2 + X_2$$

$$X_i = f_1^i\ (q_1,\ldots,q_n)\ \partial/\partial q_1 + \ldots + f_n^i (q_1,\ldots,q_n)\ \partial/\partial q_n \quad , i = 1, 2$$

$$\overline{X}_i = \sum_{k=1}^{n} \sum_{j=1}^{n} p_k \frac{\partial f_k^i}{\partial q_j} \ \partial/\partial p_j \quad , i = 1, 2$$

Δ is the involutive distribution spanned by $\partial/\partial p_1,\ldots,\partial/\partial p_n$ and let Δ_n be the involutive distribution spanned by $\partial/\partial q_1, \ldots, \partial/\partial q_n$, then notice that

$[\bar{X}_1, \bar{X}_2] \in \Delta_n$, $[\bar{X}_1, X_2] \in \Delta$, $[X_1, \bar{X}_2] \in \Delta$

and $[X_1, X_2] \in \Delta$.

Let $\bar{\mathcal{L}} = \{\bar{X}_o, \bar{X}_1, \ldots, \bar{X}_m\}_{L.A.}$ with corresponding Lie group $\bar{G}$ which is assumed to be nilpotent and acts transitively on N. Thus by Matsushima's theorem 5.1.4 there exists a basis $\bar{X}_1, \ldots, \bar{X}_n$ such that each element in $\bar{G}$ can be expressed by

$\exp t_1 \bar{X}_1 \ldots \exp t_n \bar{X}_n$.

Let $x_o \in N$ then by assumption $N := \exp t_1 \bar{X}_1 \circ \ldots \circ \exp t_n \bar{X}_n \cdot x_o$

Also since $\bar{X}_1(x), \ldots, \bar{X}_n(x)$ spans $T_x N$, $\forall\, x \in N$ by transitivity, a dual basis $\bar{\omega}_i$ to $\bar{X}_i$ can be chosen, i.e.

$\bar{\omega}_i\,(X_j) = \delta_{ij}$ on N, $\quad 1 \leq i, j \leq n$ such that $\bar{\omega}_i$ spans $T_x^* N$, $\forall\, x \in N$, $1 \leq i \leq n$.

<u>Lemma 5.1.7</u> Let θ be the natural 1- form on T^*N, given locally by $\sum_{i=1}^{n} p_i dq_i$, and let ω_i be the extension of $\bar{\omega}_i$ defined by

$\omega_i(X_{H_j}) = \bar{\omega}_i\,(\bar{X}_j)$ where each X_{H_j} satisfies the above conditions.

Then $\sum_{i=1}^{n} p_i dq_i = \sum_{i=1}^{n} \omega_i\, H_i$

Proof. It is only necessary to consider Δ_n, since both forms vanish on the complement i.e. on Δ. Also, since $\overline{X}_1(x),\ldots,\overline{X}_n(x)$ spans T_x^*N, $\forall\, x \in N$, it is sufficient to consider the Hamiltonian vector fields $X_{H_1},\ldots,X_{H_n}$. Thus

$$\sum_{i=1}^{n} p_i dq_i \,(X_{H_j}) = \sum_{i=1}^{n} p_i \frac{\partial H_j}{\partial p_i} \tag{5.1.3}$$

and

$$\sum_{i=1}^{n} \omega_i H_i \,(X_{H_j}) = \sum_{i=1}^{n} \omega_i \,(X_{H_j})\, H_i = H_j \tag{5.1.4}$$

But since $H_j = \sum_{i=1}^{n} p_i f_i^j(q)$, $H_j = \sum_{i=1}^{n} p_i \frac{\partial H_j}{\partial p_i}$.

Thus (5.1.3) and (5.1.4) are equivalent.

Q.E.D.

Notice that under poisson bracket considerations that $X_{H_j}(H_k) = \{H_k, H_j\}$ and that this vanishes on N i.e.

$$X_{H_j}(H_k)\Big|_N = 0$$

thus the flow of each $\overline{X} \in \overline{\mathcal{L}}$ does indeed stay on N.

Since $\overline{G}$ is a nilpotent transitive Lie group action on N the coordinate system of Crouch [5]. Theorem 2.1.10 can be used to reparameterize the q - coordinates on N.

That is

$$(q_1,\ldots,q_n) \longmapsto (I_1,\ldots,I_n)$$

where

$$\dot{I}_1 = f_1 = \text{constant}$$

$$\dot{I}_2 = f_2\ (I_1)$$

$$\vdots$$

$$\dot{I}_n = f_n\ (I_1,\ldots,I_{n-1})$$

where $f_1,\ldots,f_n$ are polynomials in the I - coordinates.

In $I_1,\ldots,I_n$ coordinates $\overline{X}$ is given by

$$f_1\ {}^{\partial}/_{\partial I_1} + f_2\ (I_1)\ {}^{\partial}/_{\partial I_2} + \ldots + f_n\ (I_1,\ldots,I_{n-1})\ {}^{\partial}/_{\partial I_n} \qquad (5.1.5)$$

Now introduce coordinates $\phi_1,\ldots,\phi_n$ such that $I_1,\ldots, I_n, \phi_1,\ldots,\phi_n$ form canonical coordinates on a neighbourhood of T^*N. The aim is to show that there exists a local symplectomorphism from $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ to $(I_1,\ldots,I_n, \phi_1,\ldots,\phi_n)$..

If H is a Hamiltonian function on T^*N with $X_H = \overline{X} + X$ as before and $\overline{X}$ as in equation (5.1.5), then H must be of the form

$$H = \phi_1 f_1 + \phi_2 f_2\ (I_1) + \ldots + \phi_n f_n\ (I_1,\ldots,I_{n-1}) + f_o\ (I_1,\ldots,I_n)$$

Locally N is defined by $\phi_1 = \ldots = \phi_n = 0$ thus $f_o\ (I_1,\ldots,I_n) = 0$ on T^*N since it only depends on I_i, $1 \leq i \leq n$.

<u>Theorem 5.1.8</u> Let (U, ψ_u) be a chart on T^*N with canonical coordinates $(q_1,\ldots,q_n, p_1,\ldots,p_n)$ and $x_o \in U$. Let (V, ψ_v), $x_o \in V$ be a chart on T^*N with the above canonical coordinates $(I_1,\ldots,I_n, \phi_1,\ldots,\phi_n)$.

Then, after suitable restrictions of U and V there exists a local symplectomorphism $f : U' \longmapsto V'$, $U' \subset U$, $V' \subset V$.

<u>Proof.</u> It is sufficient to prove that

$$\sum_{i=1}^{n} \omega_i H_i = \sum_{i=1}^{n} \phi_i \, dI_i$$

and then applying lemma 5.7.1.

As

$$H_i = \sum_{j=1}^{n} \phi_j \, f_j^i \, (I_1, \ldots, I_{j-1})$$

$$\sum_{i=1}^{n} \omega_i \, H_i = \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_i \, \phi_j \, f_j^i \, (I_1, \ldots, I_{j-1}) \qquad (5.1.6)$$

Also $dI_j\left(X_{H_k}\right) = \dfrac{\partial H_k}{\partial \phi_j} = f_j^k \, (I_1, \ldots, I_{j-1})$

Thus, as $\omega_i \, (X_{H_k}) = \delta_{ik}$,

$$dI_j = \sum_{i=1}^{n} \omega_i \, f_j^i \, (I_1, \ldots, I_{j-1}) \qquad (5.1.7)$$

Combining (5.1.6) and (5.1.7) gives result.

Q.E.D.

The above structure can be understood more fully with the aid of an example.

Example 5.1.9 Take as the state space $(T^*\underset{\sim}{R}^4, \omega)$ where $\omega = \sum_{i=1}^{n} dI_i \wedge d\phi_i$. $(I_1,\ldots,I_4, \phi_1,\ldots,\phi_4)$ being canonical coordinates on $T^*\underset{\sim}{R}^4$.

Suppose $\dot{x} = X_{H_o} + \sum_{i=1}^{3} u_i X_{H_i}$ $x(0) = X_o$, $x \in T^*\underset{\sim}{R}^4$ where

$$H_o = \phi_1$$

$$H_1 = \phi_2 + I_1\phi_3$$

$$H_2 = \phi_3 + I_2\phi_4$$

$$H_3 = \phi_4$$

so $X_{H_o} = {}^{\partial}/_{\partial I_1}$

$$X_{H_1} = {}^{\partial}/_{\partial I_2} + I_1 \, {}^{\partial}/_{\partial I_3} - \phi_3 \, {}^{\partial}/_{\partial \phi_1}$$

$$X_{H_2} = {}^{\partial}/_{\partial I_3} + I_2 \, {}^{\partial}/_{\partial I_4} - \phi_4 \, {}^{\partial}/_{\partial \phi_2}$$

$$X_{H_3} = {}^{\partial}/_{\partial I_4}$$

Let $\mathcal{L} = \{X_{H_o}, X_{H_1}, X_{H_2}, X_{H_3}\}_{L.A.}$

It is clear that $\mathcal{L}$ is a nilpotent Lie algebra since

$[X_{H_o}, X_{H_1}] = {}^{\partial}/_{\partial I_3}$, $[X_{H_1}, X_{H_2}] = {}^{\partial}/_{\partial I_4}$ and all other brackets vanish.

Let $N = \{x \in T^*\underset{\sim}{R}^4 : H(x) = 0 \ , \ \forall X_H \in \mathcal{L} \ \}$

N is obviously Lagrangian defined by $\phi_1 = \phi_2 = \phi_3 = \phi_4 = 0$ and

$$\overline{X}_1 = {}^{\partial}/_{\partial I_1}$$

$$\overline{X}_2 = {}^{\partial}/_{\partial I_2} + I_1 \, {}^{\partial}/_{\partial I_3}$$

$$\overline{X}_3 = {}^{\partial}/_{\partial I_3} + I_2 \, {}^{\partial}/_{\partial I_4}$$

$$\overline{X}_4 = {}^{\partial}/_{\partial I_4}$$

These vector fields span T_xN, $\forall \ x \in N$, so this system is in accordance with the previous theory. The dual forms ω_i, $1 \leqslant i \leqslant 4$, of lemma 5.1.7 i.e. $\omega_i \ (X_{H_j}) = \delta_{ij}$ are

$$\omega_1 = dI_1$$

$$\omega_2 = dI_2$$

$$\omega_3 = dI_3 - I_1 \, dI_2$$

$$\omega_4 = dI_4 - I_2 dI_3 + I_1 I_2 \, dI_2$$

Thus

$$\begin{aligned}\sum_{i=1}^{4} \omega_i H_i &= \phi_1 dI_1 + (\phi_2 + I_1\phi_3) \, dI_2 \\ &\quad + (\phi_3 + I_2 \, d\phi_4) \, (dI_3 - I_1 \, dI_2) \\ &\quad + \phi_4 \, (dI_4 - I_2 \, dI_3 + I_1 \, I_2 dI_2) \\ &= \sum_{i=1}^{4} \phi_i \, dI_i\end{aligned}$$

Hence,

$$\phi_1 \; dI_1 = \omega_1 \; H_1$$

$$\phi_2 \; dI_2 = \omega_2 \; H_2 - \phi_3 I_1 \; dI_2$$

$$\phi_3 \; dI_3 = \omega_3 H_3 \; + \phi_3 I_1 \; dI_2 - \phi_4 \; I_2 \; dI_3 + \phi_4 \; I_1 \; I_2 d \; I_2$$

$$\phi_4 \; dI_4 = \omega_4 H_4 \; + \phi_4 I_2 \; dI_3 - \phi_4 \; I_1 \; I_2 \; dI_2$$

Let $\mu_1 = \phi_3 \; I_1 \; dI_2$, $\mu_2 = \phi_4 I_2 \; (I_1 \; dI_2 - dI_3)$

then $\phi_1 \; dI_1 = \omega_1 H_1$

$$\phi_2 \; dI_2 = \omega_2 H_2 - \mu_1$$

$$\phi_3 \; dI_3 = \omega_3 H_3 + \mu_1 + \mu_2$$

$$\phi_4 \; dI_4 = \omega_4 H_4 - \mu_2$$

So if $\zeta_1 = 0$, $\zeta_2 = - \mu_1$, $\zeta_2 = \mu_1 + \mu_2$, $\zeta_4 = - \mu_2$

$$\sum_{i=1}^{4} \zeta_i = 0 \text{ and so} \sum_{i=1}^{4} \phi_i \; dI_i$$

But each $\phi_i \; dI_i = \omega_i \; H_i + \zeta_i \; (I, \phi)$

This example illustrates the difficulty of calculating the coordinates $\phi_1, \ldots, \phi_n$ in general, since the terms $\zeta_1, \ldots, \zeta_n$ are not necessarily closed, this makes integration normally impossible.

REFERENCES

1. R. Abraham and J.E. Marsden, Foundations of Mechanics, Benjamin/Cummings (1978).

2. V.I. Arnold, Mathematical Methods of Classical Mechanics, Springer-Verlag New York (1978).

3. R.W. Brockett, "Control Theory and Analytical Mechanics", 1976 Ames Research Centre (NASA) Conference on Geometric Control Theory eds. C. Martin and R. Hermann, Math. Sci. Press, pp. 1-48 (1977).

4. A.J. Van der Schaft, "Observability and Controllability for smooth Nonlinear Systems", SIAM J. of Control, Vol. 20, No.3, pp. 338-354 (1982).

5. P.E. Crouch, "Dynamical Realizations of Finite Volterra Series", SIAM J. of Control Vol.19, No.2, pp. 177-202 (1981).

6. R.W. Brockett and A. Rahimi, "Lie algebras and Linear Differential Equations", Ordinary Differential Equations, ed. L. Weiss, Academic Press, pp. 379-386 (1972).

7. R.W. Brockett "Volterra Series and Geometric Control Theory", Automatica 12 pp. 167-176 (1976).

8. P.E. Crouch, Lecture Notes on Geometric Nonlinear Systems Theory, Control Theory Centre Report No. 99, University of Warwick, Coventry (1980).

9. H.J. Sussmann, "Existence and Uniqueness of Minimal Realizations of Nonlinear Systems", Mathematical Systems Theory Vol. 10, pp. 263-284 (1977).

10. R.S. Palais, "A Global Formulation of the Lie Theory of Transitive Groups", Memoirs of the A.M.S., No.22, (1957).

11. T. Nagano, "Linear Differential Systems with Singularities and an Application to Transitive Lie Algebras", J. Math. Soc. Japan, Vol. 18, pp. 398-404 (1966).

12. M. Spivak, A Comprehensive Introduction to Differential Geometry, Vol. 1 (second edition), Publish or Perish, Berkeley, (1979).

13. R. Hermann and A.J. Krener, "Nonlinear Observability and Controllability", I.E.E.E. Transactions on Automatic Control. Vol. AC-22, No.5, pp. 728-740 (1977).

14. H.J. Sussmann and V. Jurdjevic, "Controllability of Nonlinear Systems", J. of Differential Equations, Vol.12, pp. 95-116, (1972).

15. S. Helgason, Differential Geometry, Lie Groups, and Symmetric Spaces, Academic Press, (1978).

16. J.B. Goncalves, Nonlinear Controllability and Observability with applications to gradient systems, Thesis (Ph.D), University of Warwick, (1981).

17. A.M. Vinogradov and B.A. Kuperschmidt, "The Structures of Hamiltonian Mechanics", Russian Math. Surveys 32:4, pp. 177-243 (1977).

18. N. Nehorosev, "Action-angle coordinates and their generalizations", Trud. Mosk. Math. 26, pp. 180-198 (1972).

19. B. Kostant, "Quantization and Unitary Representations", Lecture Notes in Mathematics 170, ed. C.T. Taam, Springer-Verlag, pp. 87-208 (1970).

20. J.M. Souriau, Structure des Systemes Dynamiques, Dunod, Paris, (1970).

21. A.A. Kirillov, "Unitary Representations of Nilpotent Lie Groups", Russian Math. Surveys 17, pp. 53-104 (1962).

22. V. Guillemin and S. Sternberg, Geometric Asymptotics, Mathematical Surveys No.14, A.M.S. (1977).

23. J.E. Marsden and A. Weinstein, "Reduction of Symplectic Manifolds with Symmetry", Rep. Math. Phys. 5, pp. 121-130 (1974).

24. J.E. Humphreys, Introduction to Lie Algebras and Representation Theory, Springer-Graduate Texts in Mathematics, Vol.9, (1972).

25. N.R. Wallach, Symplectic Geometry and Fourier Analysis, Math. Sci. Press, (1977).

26. L. Auslander and B. Kostant, "Polarization and Unitary Representations of Solvable Lie Groups", Inventiones math. 14, pp. 255-354 (1977).

27. A.J. Krener and C.M. Lesiak, "The Existence and Uniqueness of Volterra Series for Nonlinear Systems", Trans. I.E.E.E. Vol. AC-23, No. 6, pp. 1090-1095 (1978).

28. P.E. Crouch and M. Irving, "On Finite Volterra Series which admit Hamiltonian Realizations", Control Theory Centre Report No. 113, University of Warwick, Coventry, (1983).

29. P.E. Crouch, "Geometric structures in systems theory", I.E.E. Proc, Vol. 128, Pt.D, No. 5, pp. 242-252, (1981).

30. P.E. Crouch, "Realizations of a Single Volterra Kernel", Asterique 75-76, (1980) (Analyse Des Systèmes, Bordeaux, 11-16 September 1978).