EUROPEAN JOURNAL OF PURE AND APPLIED MATHEMATICS
Vol. 5, No. 4, 2012, 567-583
ISSN 1307-5543 – www.ejpam.com

Casimirs and Lax Operators from the Structure of Lie Algebras

Carol Linton[1,*], William Holderbaum[1], James Biggs[2]

[1] *School of Systems Engineering, University of Reading, Reading. UK*
[2] *Department of Mechanical Engineering, University of Strathclyde, Glasgow, UK*

Abstract. This paper uses the structure of the Lie algebras to identify the Casimir invariant functions and Lax operators for matrix Lie groups. A novel mapping is found from the cotangent space to the dual Lie algebra which enables Lax operators to be found. The coordinate equations of motion are given in terms of the structure constants and the Hamiltonian.

2010 Mathematics Subject Classifications: 17B45, 53D17, 17B63

Key Words and Phrases: Casimir invariants, Lax operators, structure constants, matrix Lie algebras, Poisson manifolds

1. Introduction

Lie groups are used to plan trajectories (in the widest sense). For example, the attitude of a satellite is controlled by 3 rotations [3], quantum computing needs to control the electron states [7], and underwater vehicles use the rotations and translations of the Euclidean group [14]. The conservation laws and geometric constraints determine which Lie group is appropriate for the system under consideration. These invariants can include momentum and energy. Conversely, the structure of a Lie group determine the invariant functions. This paper uses the base matrices of a Lie algebra to

- identify the structure of the Lie algebra by calculating the structure constants arising from the curvature of the space and non-associative action of the tangent fields
- identify the prerequisite of any invariant function arising from the structure of the Lie algebra (known as Casimir invariants), and then finding these Casimirs for a range of algebras
- produce the differential equations of motion from a Hamiltonian, which incorporates the geometric structure

*Corresponding author.

Email addresses: c.l.linton@pgr.reading.ac.uk (C.Linton), w.holderbaum@reading.ac.uk (W.Holderbaum), james.biggs@strath.ac.uk (J.Biggs)

For a range of low dimensional Lie algebras, the opportunity is taken to list possible base matrices, the structure constants and Casimir invariant functions.

The action of the vector field arising from an invariant function C has no effect and is known as a Lax operator L. The action of L and any vector field X is associative (the order is irrelevant) and the equation $[L,X]=0$ is often used to incorporate the geometric constraints into the mathematical system. A mapping is identified which enables a Lax operator L to be found from any Casimir functions C for matrix Lie groups, through the action of the gradient

$$L=\sum \frac{\partial C}{\partial p_i} e_i = \nabla C$$

The novelty of the paper is in providing this simple mapping from Casimir function C to Lax operator L.

The Casimir functions, equations of motion and Lax operators are the foundations of the method used to determine trajectories by Jurdjevic [11], Biggs [3] and Abazari [1].

2. Lie Theory and Vector Spaces

In this section, Lie theory and vector spaces are described very briefly to introduce some elementary ideas and the notation. A fuller explanation of Lie theory is given in many text books such as [8], while [5] covers the Lie algebra and Lie bracket. The rotation group in 3 dimensions $SO(3)$ will be used to provide examples throughout the paper. Other Lie groups are included in the Appendix.

A matrix Lie group G is a set of matrices that can represent a configuration (a position within the group). For example, a configuration in $SO(3)$ is given by 3 rotations $\{\theta_i\}$ for $i \in \{1,2,3\}$ about 3 orthogonal axes as

$$g=\exp \begin{bmatrix} 0 & -\theta_3 & \theta_2 \\ \theta_3 & 0 & -\theta_1 \\ -\theta_2 & \theta_1 & 0 \end{bmatrix}$$

The matrix exponential function is similar to the scalar function but $\exp(X+Y) \neq \exp(X)\exp(Y)$ in most cases because matrix multiplication is not associative. A one-parameter subgroup of the group represents a trajectory. A trajectory is given by a function $g:\mathbb{R} \to G$ where g is continuous, $g(0)=I$ the identity matrix and $g(s+t)=g(s)g(t)$. For each such $g(t)$, there is a unique matrix X such that $g(t)=\exp(Xt)$. Differentiating this gives the same result as differentiating the scalar exponential function

$$g^{-1}(t)\frac{dg}{dt}(t)=X \tag{1}$$

X is the tangent matrix at the origin, having been pulled back to the origin by the action of g^{-1}.

A Lie algebra $\mathfrak{g}$ is the set of all matrices X such that $\exp(Xt) \in G$ for all t. This is a vector space with the Lie bracket to define the action of one element on another as in

$$[X,Y]=XY-YX \tag{2}$$

The Lie bracket is antisymmetric, bi-linear and satisfies the Jacobi identity. Elements of a vector space can always be written in component form as

$$v = \sum_i v_i e_i$$

where $\{e_i\}$ is the set of base matrices and $\{v_i\}$ are the components. Since the Lie bracket is bi-linear, in component form it becomes

$$\left[v_i e_i, x_j e_j\right] = v_i x_j \left[e_i, e_j\right]$$

The structure of the Lie algebra determines how vectors interact and can be expressed through the Lie bracket as

$$\left[e_i, e_j\right] = c_{ij}^k e_k \tag{3}$$

The Einstein convention on summation and range is used, so that the expression on the right is summed over all k, and the equation applies to all combinations of i and j.

The dual of the Lie algebra $\mathfrak{g}^*$ is the space of co-vectors v^* so that $v^* : v \to \mathbb{F}$ where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. It is created using the identity form which is a non-degenerate bi linear symmetric form so that

$$\mathbb{I} e_i = e^i \tag{4}$$

where $\left\{e^i\right\}$ is the set of base matrices for $\mathfrak{g}^*$ and $\mathbb{I}$ is the unit matrix. The base matrices of the dual algebra look the same as the Lie algebra basis, and hence the structure is also the same.

For $SO(3)$, any tangent matrix can be written as

$$v = \sum_{i=1}^{3} v_i e_i = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}$$

This expression enables the base matrices of $\mathfrak{so}(3)$ to be identified as the set $\{e_i\}$. Matrix multiplication of the base matrices is used to identify the structure constants using equation (3) as

$$\begin{array}{cccc} c_{23}^1 & = c_{31}^2 & = c_{12}^3 & = 1 \\ c_{32}^1 & = c_{13}^2 & = c_{21}^3 & = -1 \end{array}$$

After this very concise introduction to Lie groups, its Lie algebra and dual algebra, the next section concentrates on functions on the dual of the Lie algebra.

3. Dual Lie Algebra and the Poisson Bracket

This paper is concerned with two type of functions on the dual space; constant functions which arise from the structure of the space and Hamiltonains which induces a trajectory on the base manifold. Having defined a function on the dual space, the Poisson bracket is used to find vector fields arising from a function. There is a close relationship between the Poisson bracket and the Lie bracket. They both describes the structure of the Lie algebra in a similar manner. The relationships are proved in this section in preparation to finding the invariant functions and the differential equations of motion.

Definition 1. *A function G on the dual space is an operation that assigns a scalar value $\in \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ to every point $\left(p, q\right)$ on the dual space; $G : T^*M \to \mathbb{F}$ where p is a cotangent or one-form corresponding to a vector in the tangent space for the position q.*

The structure of the dual space is reflected in the action of the Poisson bracket, defined here.

Definition 2. *The Poisson bracket $\{\cdot,\cdot\}$ is defined as satisfying the following conditions [see p20 of 9]*

1. *bi-linear $\{\lambda F, G+E\} = \lambda \{F, G\} + \lambda \{F, E\}$*
2. *skew symmetric $\{E, F\} = -\{F, E\}$,*
3. *satisfies the Leibniz rule $\{FG, E\} = \{F, G\} E + F \{G, E\}$*
4. *satisfies the Jacobi identity $\{F, \{G, E\}\} + \{G, \{E, F\}\} + \{E, \{F, G\}\} = 0$*

where E, F, G are functions on the dual space.

The canonical form of the Poisson bracket

$$\{F, E\} = \frac{\partial F}{\partial q}\frac{\partial E}{\partial p} - \frac{\partial F}{\partial p}\frac{\partial E}{\partial q}$$

satisfies this definition with $F(p, q)$ and $E\left(p, q\right)$ being functions on the cotangent space when the canonically conjugate coordinates satisfy

$$\left\{q_i, p_j\right\} = \delta_{ij} \tag{5}$$

Definition 3. *A vector field arising from a function F and the Poisson bracket [see p21 of 9] is defined as*

$$X_F = \{\cdot, F\} = \frac{\partial F}{\partial p}\frac{\partial}{\partial q} - \frac{\partial F}{\partial q}\frac{\partial}{\partial p} \tag{6}$$

Theorem 1. *Poisson brackets can be expressed in terms of alternative coordinates $\{z_i\} = \left\{p_i, q_j\right\}$ [compare p 49 of 9] as*

$$\{F, E\} = \frac{\partial F}{\partial z_i}\left\{z_i, z_j\right\}\frac{\partial E}{\partial z_j}$$

with implicit summation over all i, j.

Proof. From the definition

$$\begin{aligned}\{F,E\} =&\frac{\partial F}{\partial q}\frac{\partial E}{\partial p}-\frac{\partial F}{\partial p}\frac{\partial E}{\partial q}\\ =&\frac{\partial F}{\partial z_i}\frac{\partial z_i}{\partial q}\frac{\partial E}{\partial z_j}\frac{\partial z_j}{\partial p}-\frac{\partial F}{\partial z_i}\frac{\partial z_i}{\partial p}\frac{\partial E}{\partial z_j}\frac{\partial z_j}{\partial q}\\ =&\frac{\partial F}{\partial z_i}\left(\frac{\partial z_i}{\partial q}\frac{\partial z_j}{\partial p}-\frac{\partial z_i}{\partial p}\frac{\partial z_j}{\partial q}\right)\frac{\partial E}{\partial z_j}\\ =&\frac{\partial F}{\partial z_i}\left\{z_i,z_j\right\}\frac{\partial E}{\partial z_j}\end{aligned}$$

The dual Lie algebra is the cotangent space pulled back to the origin so there is no positional dependencies and the coordinates are written as $\{p_i\}$. If F and E are functions on the dual of the Lie algebra, then they are dependent only on $\{p_i\}$. The Poisson bracket on dual Lie algebras can thus be written as

$$\{F(p),E(p)\}=\frac{\partial F(p)}{\partial p_i}\left\{p_i,p_j\right\}\frac{\partial E(p)}{\partial p_j} \tag{7}$$

The structure of the vector space influences the Poisson bracket and Lie bracket so one might expect a close relationship. This is shown in the two theorems which follow.

Theorem 2. *The relationships between the Lie and Poisson bracket are [see p69 of 9]:*

$$\left[X_F,X_E\right]=-X_{\{F,E\}} \tag{8}$$

$$X_G=\left[X_F,X_E\right]\ \text{whenever}\ G=\{F,E\} \tag{9}$$

Proof. To prove equation (8), $\left[X_F,X_E\right]=X_FX_E-X_EX_F$ definition of Lie bracket
$=\{F,\cdot\}\{E,\cdot\}-\{E,\cdot\}\{F,\cdot\}$ definition of vector fields
$=\{F,\{E,\cdot\}\}-\{E,\{F,\cdot\}\}$ substitution
$=-\{\{F,E\},\cdot\}$ using Jacobi identity
$=-X_{\{F,E\}}$
To prove equation (9), $\left[X_F,X_E\right]f=X_FX_Ef-X_EX_Ff$ from definition of the Lie bracket
$=\left\{\{f,F\},E\right\}-\left\{\{f,E\},F\right\}$ definition of vector fields
$=-\left\{\{E,f\},F\right\}-\left\{\{F,E\},f\right\}-\left\{\{f,E\}F\right\}$
$=\left\{f,\{F,E\}\right\}$ using Jacobi identity
$=X_{\{F,E\}}f$ and hence the result.

Theorem 3. *If $\{p_i\}$ are the coordinates using the basis $\left\{e^i\right\}$ and the structure constants are defined as $\left[e^i,e^j\right]=c^k_{ij}e^k$, then the relationships between Poisson bracket and structure constants are*

$$\left\{p_i,p_j\right\}=-c^k_{ij}p_k \tag{10}$$

$$\{F,E\} = -c_{ij}^k p_k \frac{\partial F}{\partial p_i}\frac{\partial E}{\partial p_j} \tag{11}$$

(This is the Lie-Poisson bracket - [see p131 of 4])

Proof. To prove equation (10), define a linear map $\hat{p}\left(e^i\right) = \hat{p}_{e^i} = p_i e^i$. Using Theorem 2 and substituting $X_i = e^i$ gives

$\hat{p}_{\{e^i,e^j\}} = \hat{p}\left(\left\{e^i,e^j\right\}\right)$ notation

$= -\left[\hat{p}\left(e^i\right),\hat{p}\left(e^j\right)\right]$ see above

$= -\hat{p}\left(\left[e^i,e^j\right]\right)$ $\hat{p}$ is a linear operator

$= -\hat{p}\left(c_{ij}^k e^k\right)$ definition of structure constants

$= -c_{ij}^k \hat{p}\left(e^k\right)$

$= -c_{ij}^k \hat{p}_k$

Hence $\left\{p_i,p_j\right\} = -c_{ij}^k p_k$

To prove equation (11), take $\{F,E\} = \frac{\partial F}{\partial p_i}\left\{p_i,p_j\right\}\frac{\partial E}{\partial p_j}$ using equation (7)

$= -\frac{\partial F}{\partial p_i} c_{ij}^k p_k \frac{\partial E}{\partial p_j}$ using equation (10).

This section has introduced functions on the dual Lie algebra and the vector fields which they induce. The close relationship of the Poisson bracket with the structure constants shows how the functions interact. In the next section, functions are found which don't interact with other functions. These are the invariant Casimirs. The action of a Hamiltonian function is covered in Section 5.

4. Casimir Invariants

The Casimir invariant functions are functions on the dual of the Lie algebra that depend only on the algebra involved. Examples of their use can be found in [p 314 of 9], [12], [16] and [2]. They invoke the symmetry of the group which reflects the inherent conservation laws, and determine the format of the solutions to any problems posed in that symmetry.

The defining property of a Casimir functions is provided. A necessary requirement of them is identified as well as a theorem on finding further examples. This information is used to find the Casimir functions for some 3 dimensional Lie algebras. The basic Casimirs functions for other algebras are shown in the Appendix. The maximal number of Casimirs is determined by the rank of the Lie algebra [p53 of 11]. The section ends by identifying the associated invariant vectors.

Definition 4. *Casimir invariant functions C (also known as distinguished functions or Casimir functions) are defined as functions which Poisson commute with all other functions on that space. That is $\{C,F\} = 0$ for all functions F on $\mathfrak{g}^*$ [p132 of 4].*

Since on $\mathfrak{g}^*$, a Casimir has the property that

$$\{C,F\} = -c_{ij}^k p_k \frac{\partial C}{\partial p_i}\frac{\partial F}{\partial p_j} = 0$$

for all F, the requirement is for

$$\sum_{i,k} c_{ij}^k p_k \frac{\partial C}{\partial p_i} = 0 \tag{12}$$

for all j, where all the indices range over the dimension of the algebra.

This can be used to find the basic Casimir functions, some of which as given in the Appendix. As an example, the next theorem determines the Casimir for $\mathfrak{so}(3)$ which is the conservation of angular momentum.

Theorem 4. *For $\mathfrak{so}(3)$, the Casimir invariant is*

$$C_2 = \sum_{i=1}^{3} p_i^2$$

Proof. With $C = \sum_{i=1}^{3} p_i^2$, for j=1,

$$\sum_{i,k} c_{ij}^k p_k \frac{\partial C}{\partial p_i} = 0$$

So

$$p_2 \frac{\partial C}{\partial p_3} - p_3 \frac{\partial C}{\partial p_2} = p_2 p_3 - p_3 p_2 = 0$$

For $j = 2$,

$$p_3 \frac{\partial C}{\partial p_1} - p_1 \frac{\partial C}{\partial p_3} = p_3 p_1 - p_1 p_3 = 0$$

For $j = 3$,

$$p_1 \frac{\partial C}{\partial p_2} - p_2 \frac{\partial C}{\partial p_1} = p_1 p_2 - p_2 p_1 = 0$$

so the theorem is proved.

It is possible to find further Casimir functions from the basic ones found using structure constants.

Theorem 5. *Any function C^* of C_n where $\{C_n\}$ are basic Casimir functions is also a Casimir function.*

Proof. Consider a function $C^*(C_n)$, then, for any function F on the dual of Lie algebra,

$$\begin{aligned}
\{C^*, F\} &= \frac{\partial C^*}{\partial p} \frac{\partial F}{\partial q} - \frac{\partial C^*}{\partial q} \frac{\partial F}{\partial p} \\
&= \sum_i \frac{\partial C^*}{\partial C_n} \left(\frac{\partial C_n}{\partial p} \frac{\partial F}{\partial q} - \frac{\partial C_n}{\partial q} \frac{\partial F}{\partial p} \right) \\
&= \sum_i \frac{\partial C^*}{\partial C_n} \{C_n, F\} = 0
\end{aligned}$$

since $\{C_n, F\} = 0$ for all n.

A constant function is also a Casimir, as is any polynomial combination of basic Casimirs, for example, $a + bC_1 + C_1 C_2^2$.

Definition 5. *A Casimir invariant vector X_C is defined as a vector that does not change the action of any other vector X. That is; $[X_C, X] = 0$ for all X.*

If C is a Casimir invariant function then $\{C, F\} = 0$ for all F. Theorem 2 gives the relationship

$$X_G = [X_C, X_E] \text{ whenever } G = \{C, E\}$$

so

$$[X_C, X_F] = 0 \text{ for all } X_F \tag{13}$$

where X_F and X_C are related to the functions F and C by equation (6)

$$X_F = \frac{\partial F}{\partial p_i} \{p_j, p_i\} \frac{\partial}{\partial p_j}$$

Using equation (10), the Casimir invariant vector X_C for the dual Lie algebra becomes

$$X_C = \frac{\partial C}{\partial p_i} c_{ij}^k p_k \frac{\partial}{\partial p_j} \tag{14}$$

In this section, the requirement of a Casimir invariant function has been stated in terms of structure constants in equation (12). It was used to find the Casimir function for $SO(3)$. The same method can be used for other algebras. A list of basic functions is provided in the Appendix, enabling further functions to be identified using Theorem 5. The format of a Casimir invariant vector has been found in equation (14) and will be used to find Lax operators in Section 6.

5. Trajectory Induced by a Hamiltonian

In the previous section, the concentration was on functions which are dependent only on the group structure. In this section, a Hamiltonian function induces a trajectory though the action of the Hamiltonian vector field, if the Lie Poisson structure is imposed - see [6]. The differential equations of motion, which depend on the structure and the Hamiltonian, are found.

Bloch [4] on p121 defines X_H, the Hamiltonian vector field of H, as the unique vector field such that

$$\dot{k} = X_H(k) \equiv \langle dk, X_H \rangle = \{k, H\} \text{ for all } k \in \mathscr{F}(P) \tag{15}$$

where P is a Poisson manifold and $\mathscr{F}(P)$ is the set of all functions on P.

If the Lie-Poisson structure as identified in Section 3 is imposed [6], then this equation (15) can be used on any Lie algebra. For a Lie algebra, there is no position dependence and the Hamiltonian is dependent only on p. The Poisson bracket part of equation (15) is written using the alternative coordinates Theorem 1.

$$\begin{aligned}\dot{k} &= \{k,H\}\\ &= \frac{\partial H}{\partial z_i}\left\{z_j,z_i\right\}\frac{\partial k}{\partial z_j} \text{ for all coordinates } \left\{z_i\right\}=\left\{q_j,p_k\right\}\\ &= \frac{\partial H}{\partial q_j}\left\{q_k,q_j\right\}\frac{\partial k}{\partial q_k}+\frac{\partial H}{\partial q_j}\left\{p_k,q_j\right\}\frac{\partial k}{\partial p_k}\\ &\quad+\frac{\partial H}{\partial p_j}\left\{p_k,p_j\right\}\frac{\partial k}{\partial p_k}+\frac{\partial H}{\partial p_j}\left\{q_k,p_j\right\}\frac{\partial k}{\partial q_k}\end{aligned}$$

So, for the dual of the Lie algebra where $H=H\left(p_i\right)$, the rate of change of any function k is given by

$$\dot{k}=\frac{\partial H}{\partial p_j}\left\{p_k,p_j\right\}\frac{\partial k}{\partial p_k}+\frac{\partial H}{\partial p_j}\left\{q_k,p_j\right\}\frac{\partial k}{\partial q_k}$$

Setting $k=p_i$ gives

$$\dot{p}_i=-\frac{\partial H}{\partial p_j}\left\{p_j,p_i\right\} \tag{16}$$

The structure constant relationship for the Poisson bracket (equation (11)) enables the coordinate differential equation to be written as

$$\dot{p}_i=-\frac{\partial H}{\partial p_j}c_{ij}^k p_k \tag{17}$$

For example $SO(3)$, if the total energy is defined as the Hamiltonian $H=\frac{1}{2}\sum_{i=1}^3\frac{p_i^2}{m_i}$ where $\left\{m_i\right\}$ are the inertia terms, then the equations of motion become

$$\dot{p}_i=p_j\frac{p_k}{m_k}-p_k\frac{p_j}{m_j}$$

for $\{i,j,k\}$ permuted over $\{1,2,3\}$ (see page 9 of [15]). The first part of equation (15) is $\dot{p}_i=X_H p_i$ so the Hamiltonian vector field can be expressed as

$$X_H=-\frac{\partial H}{\partial p_j}c_{ij}^k p_k\frac{\partial}{\partial p_i} \tag{18}$$

Setting $k=q_i$ and since $\left\{q_j,p_k\right\}=\delta_{jk}$, this gives

$$\dot{q}_i=X_H q_i=\frac{\partial H}{\partial p_i} \tag{19}$$

This is pulled back to the origin by the action of q^{-1} so that

$$q^{-1}\dot{q}=X_H=\sum_i\frac{\partial H}{\partial p_i}e^i$$

The inverse of the bi-linear form $\mathbb{I}$ used in equation (4) lowers this equation to the Lie algebra as

$$\mathbb{I}^{-1}X_H = \sum_i \frac{\partial H}{\partial p_i} e_i = \nabla H$$

The inverse of this is

$$X_H = \mathbb{I}\nabla H$$

which is proved in various ways by, for example, [15] and [10].

In this section, the action of a Hamiltonian has been used to identify the differential equations of motion on the Lie algebra. The Hamiltonian vector field has been expressed in terms of the structure constants and as the gradient of Hamiltonian.

$$X_H = -\frac{\partial H}{\partial p_j} c_{ij}^k p_k \frac{\partial}{\partial p_i} = \mathbb{I}\nabla H \tag{20}$$

This will be used in the next section in deriving the Lax operator.

6. Lax Operator and Casimir Invariants

Invariant functions generate invariant vectors, which offers a method of incorporating the conservation laws into the mathematical equations and enable systems in Geometric Control and Quantum control to be resolved. This is done through Lax pairs, which are pairs of time dependent operators (L,X). If $\dot{g}(t) = g(t)X$ and $\dot{L} = [L(t),X] = 0$ then $g(t)L(t)g^{-1}(t)$ is independent of time.

In the previous section, an important relationship was found which is now used to find the invariant Lax operator. First the Lax pair theorem for left invariant systems is presented based on the work by Peter Lax in 1968 [13]. After that the invariant Lax operators are found from the Casimir invariants found earlier.

Theorem 6 (Lax pair theorem for left invariant systems)**.** *Given that the inverse of g exists and that g and L are differentiable, then*

$$\begin{aligned} \dot{g} &= gX \\ \dot{L} &= [L,X] \end{aligned} \tag{21}$$

if and only if $g(t)L(t)g^{-1}(t)$ is independent of time.

Proof. $\Longrightarrow$If $\dot{L} = [L,X] = LX - XL$ then

$$g\dot{L}g^{-1} = gLXg^{-1} - gXLg^{-1}$$

Since $\dot{g} = gX$

$$g\dot{L}g^{-1} = -gL\dot{g}^{-1} - \dot{g}Lg^{-1}$$

$$g\dot{L}g^{-1} + gL\dot{g}^{-1} + \dot{g}Lg^{-1} = \frac{d}{dt}\left(gLg^{-1}\right) = 0$$

Hence $g(t)L(t)g^{-1}(t)$ is independent of time.

$\Longleftarrow$ If gLg^{-1} is independent of time, then

$$\frac{d}{dt}\left(gLg^{-1}\right) = 0$$
$$\dot{g}Lg^{-1} + g\dot{L}g^{-1} + gL\dot{g^{-1}} = 0$$

Since $\dot{g} = gX$ and $\frac{d}{dt}g^{-1} = -g^{-1}\dot{g}g^{-1} = -g^{-1}gXg^{-1} = -Xg^{-1}$

$$gXLg^{-1} + g\dot{L}g^{-1} - gLXg^{-1} = 0$$
$$XL + \dot{L} - LX = 0$$
$$\dot{L} = [L,X]$$

and the converse is proved.

A general relationship between a Lax operator and Casimir invariant functions has been proved in previous papers [see 17, 19, 18] and relies on a nondegenerate bi-linear form and the Cartan decomposition. The situation is easier when considering matrix Lie algebras with an imposed Lie-Poisson structure although a nondegenerate bi-linear form $\mathbb{I}$ is assumed in dually $\mathfrak{g}$ - see equation (4).

Earlier it was shown that $\left[X_C, X\right] = 0$ for all vector fields $X \in \mathfrak{g}^*$ where

$$X_C = \frac{\partial C}{\partial p_i} c_{ij}^k p_k \frac{\partial}{\partial p_j}$$

From the derivation of the Hamiltonian vector field, the following expression was found

$$X_H = -\frac{\partial H}{\partial p_j} c_{ij}^k p_k \frac{\partial}{\partial p_i} = \mathbb{I}\nabla H$$

This can be applied to any function such as a Casimir so that

$$X_C = -\frac{\partial C}{\partial p_j} c_{ij}^k p_k \frac{\partial}{\partial p_i} = \mathbb{I}\nabla C$$

(The first equality was proved earlier. The second equality uses the mapping $-c_{ij}^k p_k \frac{\partial}{\partial p_i} = e^j$ which is independent of the function). The expression $\left[X_C, X\right] = 0$ is lowered to the Lie algebra using the inverse form $\mathbb{I}^{-1}$ to give

$$\mathbb{I}^{-1}\left[X_C, X\right] = 0 \text{ for all } X \in \mathfrak{g}^*$$
$$\left[\mathbb{I}^{-1}X_C, \mathbb{I}^{-1}X\right] = 0$$
$$\left[\nabla C, X^\sharp\right] = 0 \text{ for all } X^\sharp \in \mathfrak{g}$$

In particular, if $L = \nabla C$ then $\frac{dL}{dt} = [L,X] = 0$ for X such that $\dot{g(t)} = g(t)X$. The Lax operators are given by $L = \nabla C$ for all the Casimir functions of the Lie group.

For example, the Lax operator for $\mathfrak{so}(3)$ is

$$\nabla C = \nabla \sum_{i=1}^{3} p_i^2 = 2 \sum_{i=1}^{3} p_i e_i$$

By limiting the scope of the paper to matrix Lie algebras of finite dimensions with the Lie-Poisson structure imposed, Lax operators are easily found from the Casimir functions.

Alternatively, the Lax operators can be found using the structure constants and the explicit matrices, using the method outlined below.

1. Assume that the Lax pair take the form

   $$\frac{dL}{dt} = [L, X]] \text{ with } L = \frac{1}{2} \nabla C$$

   where C is any Casimir invariant as found above, and $X = \sum u_i e_i$.

2. Expand the expression for L and differentiate

3. Substitute for the values of $\dot{p}_i$ using equation (17)

4. Then expand the expression $2\frac{dL}{dt} = 2[L, X]$ using the structure constants, and show that the two are equivalent.

This final section has introduced two novel findings:

1. the relationship between the Lax operators L and any Casimir invariant function C, $L = \nabla C$

2. the mapping $-c_{ij}^k p_k \frac{\partial}{\partial p_i} = e^j$ which applies to matrix Lie algebras with the Lie-Poisson structure imposed

7. Conclusion

By imposing the Lie-Poisson structure (if it does not automatically apply) on the matrix Lie algebras, it is possible to derive formula for several useful functions and operators, and a mapping between alternative bases for the Lie algebra. This structure is automatic for semi-simple groups, but needs to be imposed for other groups.

The Casimir functions are invariant and incorporate conservation laws (such as conservation of momentum) into the system being considered. They are well known for the common algebras, but are listed in the Appendix for reference. The differential equations of motion are also well known but their origin from the structure constants is worth repeating, especially for $SE(3)$ with the imposed structure.

The simple relationship between any Casimir operator and a Lax operator is novel. The methodology applies to finite dimensional matrix Lie algebras with an imposed Lie-Poisson structure. It avoids the complex differential geometry necessary for the tangent and cotangent space (by restriction to the origin), and includes algebras which are not semi-simple. The Lax pair is an alternative method to the Casimir functions for including the conservation laws in the mathematical formulation of the system, and for evaluating trajectories on Lie groups.

Finally, a mapping between the alternative bases for the Lie algebra was found.

References

[1] N. Abazari and I. Sager. Planning rigid body motions and optimal control problem on Lie group SO(2,1). *Engineering and Technology*, 64:448–452, 2010.

[2] J.D. Biggs and W. Holderbaum. The Geometry of Optimal Control Solutions on some Six Dimensional Lie Groups. *Proceedings of the 44th IEEE Conference on Decision and Control*, (2):1427–1432, 2005.

[3] J.D. Biggs and N. Horri. Optimal geometric motion planning for spin-stabilized spacecraft. *System and Control Letters*, 2012.

[4] A.M. Bloch. *Nonholonomic Mechanics and Control: With the Collaboration of J.Baillieul, P.Crouch and J.Marsden (Interdisciplinary Applied Mathematics)*. Springer, 2003.

[5] F. Bullo and A.D. Lewis. *Geometric Control of Mechanical Systems: Modeling, Analysis, and Design for Simple Mechanical Control Systems (Texts in Applied Mathematics)*. Springer, 2005.

[6] M. Craioveanu, C. Pop, A. Aron, and C. Petri. An optimal control problem on the special Euclidean group SE (3 , R). In *International Conference "Differential Geometry-Dynamical Systems 2009"*, pages 68–78, 2009.

[7] D. D'Alessandro. Algorithms for quantum control based on decompositions of Lie groups. *Proceedings of the 39th IEEE Conference on Decision and Control (Cat. No.00CH37187)*, pages 967–968, 2000.

[8] B. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction (Graduate Texts in Mathematics)*. Springer, 2004.

[9] D.D. Holm. *Geometric Mechanics, Part I: Dynamics and Symmetry*. Imperial College Press, 2008.

[10] V. Jurdjevic. *Geometric Control Theory (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 1997.

[11] V. Jurdjevic. Integrable Hamiltonian Systems on Complex Lie Groups. *Memoirs of the American Mathematical Society*, 178(838), 2005.

[12] E.W. Justh and P.S. Krishnaprasad. Optimal natural frames. *Communications in Information and Systems*, 11(1):17–34, 2011.

[13] P.D. Lax. Integrals of Nonlinear Equations of Evolution and Solitary. Technical Report January, Courant Institute of Mathematical Sciences, 1968.

[14] N. E. Leonard and P.S. Krishnaprasad. High Order Averaging onLie Groups and Control of an Autononmous Underwater Vehicle. *Proceedings of the American Control Conference*, June(1):2–7, 1994.

[15] J.E. Marsden, S.T. Ratiu, F. Scheck, and M.E. Mayer. *Introduction to Mechanics and Symmetry and Mechanics: From Newton's Laws to Deterministic Chaos*. American Institute of Physics, 1998.

[16] C.C. Remsing. Control and Integrability on SO (3). In *World Congress on Engineering*, volume III, 2010.

[17] A.G. Reyman and M.A. Semenov-Tian-Shansky. Reduction of Hamiltonian systems, affine Lie algebras and Lax equations. *Inventiones Mathematicae*, 54(1):81–100, February 1979.

[18] A.G. Reyman and M.A. Semenov-Tian-Shansky. Integrable Systems II: Group-Theoretical Methods in the Theory of Finite-Dimensional Integrable Systems. In *Dynamical systems. VII, Encyclopaedia of Mathematical Sciences, vol. 16*, volume 1, page 341. Springer, 1994.

[19] O.K. Sheinman. Lax equations and knizhnik-zamolodchikov connection. *aiXiv.1009.4706v2*, pages 1–21, 2011.

Appendix A. Basis, Structure Constants, Casimir Functions

This appendix has been included to provide a reference for several low dimension Lie algebras. In each case, the following information is listed

- a set of base matrices $\{e_i\}$
- corresponding structure constants $c_{i,j}^k$ defined by $c_{i,j}^k e_k = \left[e_i, e_j\right]$
- the basic Casimir functions C_n from which other invariant functions C^* can be found: $C^*\left(C_n\right)$
- the Lax operators L derived from $L = \nabla C$ such that
  - if $g^{-1}\frac{dg}{dt} = \nabla H$ and $[L, \nabla H] = 0$,
  - then $g^{-1}L(t)L(0)gL(t) = L(t)$

The differential equations of motions are $\dot{p}_i = -\frac{\partial H}{\partial p_j} c_{ij}^k p_k$ where H is the Hamiltonian function.

Appendix A.1. $\mathfrak{so}(4)$, $\mathfrak{se}(3)$ and $\mathfrak{so}(3,1)$

A basis for tangent spaces $\mathfrak{so}(4)$, $\mathfrak{se}(3)$ and $\mathfrak{so}(3,1)$, where $\varepsilon = 1$ for $\mathfrak{so}(4)$, $\varepsilon = 0$ for $\mathfrak{se}(3)$ and $\varepsilon = -1$ for $\mathfrak{so}(3,1)$, is given by

$$\sum_{i=1}^{6} \omega_i e_i = \begin{bmatrix} 0 & -\varepsilon\omega_1 & -\varepsilon\omega_2 & -\varepsilon\omega_3 \\ \omega_1 & 0 & -\omega_6 & \omega_5 \\ \omega_2 & \omega_6 & 0 & -\omega_4 \\ \omega_3 & -\omega_5 & \omega_4 & 0 \end{bmatrix}$$

Structure Constants

$$\begin{gathered} c_{15}^3 = c_{26}^1 = c_{34}^2 = c_{42}^3 = c_{45}^6 = c_{53}^1 = c_{56}^4 = c_{61}^2 = c_{64}^5 = 1 \\ c_{16}^2 = c_{24}^3 = c_{35}^1 = c_{43}^2 = c_{46}^5 = c_{51}^3 = c_{54}^6 = c_{62}^1 = c_{65}^4 = -1 \\ c_{12}^6 \quad = c_{23}^4 \quad = c_{31}^5 \quad = \epsilon \\ c_{13}^5 \quad = c_{21}^6 \quad = c_{32}^4 \quad = -\epsilon \end{gathered}$$

Casimir Functions

$$C_2 = \sum_{i=1}^{3} p_i^2 + \varepsilon \sum_{i=4}^{6} p_i^2$$

$$C_3 = \sum_{i=1}^{3} p_i p_{i+3}$$

Lax operators

$$L = \sum_{i=1}^{3} p_i e_i^* + \epsilon \sum_{i=4}^{6} p_i e_i^*$$

$$L = \sum_{i=1}^{3} p_i e_{i+3}^* + \sum_{i=1}^{3} p_{i+3} e_i^*$$

Appendix A.2. $\mathfrak{so}(3)$, $\mathfrak{se}(2)$, $\mathfrak{so}(2,1)$,$\mathfrak{su}(2)$,$\mathfrak{sl}(2)$ and $\mathfrak{sp}(2)$

A basis for $\mathfrak{so}(3)$ is given by

$$\sum_{i=1}^{3} \omega_i e_i = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

A basis for $\mathfrak{se}(2)$ and $\mathfrak{so}(2,1)$, where $\varepsilon = 0$ for $\mathfrak{se}(2)$ and $\varepsilon = -1$ for $\mathfrak{so}(2,1)$, is

$$\sum_{i=1}^{3} \omega_i e_i = \begin{bmatrix} 0 & -\varepsilon\omega_2 & -\varepsilon\omega_3 \\ \omega_2 & 0 & -\omega_1 \\ \omega_3 & \omega_1 & 0 \end{bmatrix}$$

A basis for $\mathfrak{su}(2)$ is

$$\sum_{i=1}^{3}\omega_i e_i = \frac{1}{2}\begin{pmatrix} i\omega_1 & \omega_2 + i\omega_3 \\ -\omega_2 + i\omega_3 & -i\omega_1 \end{pmatrix}$$

A basis for $\mathfrak{sl}(2)$ and $\mathfrak{sp}(2)$ is

$$\sum_{i=1}^{3}\omega_i e_i = \frac{1}{2}\begin{bmatrix} \omega_3 & \omega_1 + \omega_2 \\ -\omega_1 + \omega_2 & -\omega_3 \end{bmatrix}$$

Structure constant values for $\mathfrak{so}(3)$ (with $\varepsilon = 1$), $\mathfrak{se}(2)$ (with $\varepsilon = 0$), $\mathfrak{so}(2,1)$ (with $\varepsilon = -1$), $\mathfrak{su}(2)$ (with $\varepsilon = 1$), $\mathfrak{sl}(2)$ (with $\varepsilon = -1$) and $\mathfrak{sp}(2)$ (with $\varepsilon = -1$) are

$$\begin{aligned} c_{31}^{2} &= c_{12}^{3} &= 1 \\ c_{13}^{2} &= c_{21}^{3} &= -1 \\ c_{23}^{1} &= \epsilon \\ c_{32}^{1} &= -\epsilon \end{aligned}$$

Casimir Functions

$$C_2 = \epsilon p_1^2 + \sum_{i=2}^{3} p_i^2$$

Lax operators

$$2L = \nabla C_2 = \epsilon p_1 e_1^* + \sum_{i=2}^{3} p_i e_i^*$$

Appendix A.3. $\mathfrak{h}_{2n+1}$, the Heisenberg Lie Algebra

A basis for $\mathfrak{h}_{2n+1}$, the Heisenberg Lie algebra, is

$$\sum_{i=1}^{n} x_i e_i + \sum_{i=n+1}^{2n} y_i e_{i+n} + z e_{2n+1} = \begin{bmatrix} 0 & (x_1 \dots x_n) & z \\ 0 & 0_n & (y_1 \dots y_n)^T \\ 0 & 0 & 0 \end{bmatrix}$$

Structure constant values for $\mathfrak{h}_{2n+1}$ are for $i \in \{1, n\}$

$$c_{i,(n+i)}^{2n+1} = 1,\ c_{(n+i),i}^{2n+1} = -1,\ c_{i,(2n+1)}^{n+i} = c_{(n+i),(2n+1)}^{i} = 0$$

Structure constant values for $\mathfrak{h}_3$ are

$$c_{12}^{3} = 1,\ c_{21}^{3} = -1,\ c_{23}^{1} = c_{13}^{2} = 0$$

Casimir Functions - $C_i = x + y$ is invariant if x, y are vectors with equal number of non-zero components, since

$$\begin{aligned} \sum_i c_{(n+i),i}^{2n+1} z \frac{\partial C_i}{\partial x_i} + \sum_i c_{i,(n+i)}^{2n+1} z \frac{\partial C_i}{\partial y_i} =& \sum_i c_{(n+i),i}^{2n+1} z + \sum_i c_{i,(n+i)}^{2n+1} z \\ =& z - z = 0 \end{aligned}$$

For $\mathfrak{h}_3$, $C = x + y$ is invariant.

Appendix A.4. $\mathfrak{h}_3^\diamond$, the Oscillator Lie Algebra

A basis for $\mathfrak{h}_3^\diamond$, the oscillator Lie algebra, is

$$\sum_{i=1}^{3} \omega_i e_i = \begin{bmatrix} 0 & -\omega_1 & \omega_2 & -2\omega_3 \\ 0 & 0 & -\omega_4 & 0 \\ 0 & 0 & 0 & \omega_1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Structure constants for $\mathfrak{h}_3^\diamond$ are

$$c_{12}^3 = c_{14}^2 = c_{42}^1 = 1$$
$$c_{21}^3 = c_{41}^2 = c_{24}^1 = -1$$

Casimir Functions

$$C_2 = p_1^2 + p_2^2 - 2p_3p_4$$
$$C_3 = p_3$$