CASIMIR ELEMENTS AND OPTIMAL CONTROL

V. JURDJEVIC

Department of Mathematics, University of Toronto
100, St George Street, Toronto, Ontario, M5S 1A1 Canada

Introduction. Optimal problems on Lie groups are an important class of problems linking modern control theory with its classical predecessors mechanics and differential geometry.

My previous papers [5], [6] and [7] showed that even the most classical problems of mechanics, such as the motion of the rigid body, or its geometric companion, the equations describing the equilibrium configurations of an elastic rod can be effectively analyzed on Lie groups through the Maximum Principle and its associated Hamiltonian formalism. This paper will further illustrate the importance of modern geometric techniques by concentrating on time optimal control problems of mechanical systems recently studied in [3], [10] and [12]. Each of these studies are inspired by an important paper of Dubins [4] who considered and solved, the following geometric problem:

Among all C^1 curves $\gamma(t)$ in the plane, which are parametrized by arc length, and which further satisfy the condition that $\frac{d^2\gamma}{dt^2}(t)$ is a measurable function with $||\frac{d^2\gamma}{dt^2}(t)|| \leq 1$ almost everywhere, find the curve of minimal length which connects two arbitrary points in the plane and has prescribed tangent vectors at these points.

Following H. J. Sussmann [12] we shall refer to this problem as Dubins' problem. Each admissible curve $\gamma(t)$ in the problem of Dubins can be parametrized by the angle θ, which the tangent vector makes with the horizontal direction to yield:

$$(1) \qquad \frac{dx}{dt} = \cos\theta(t), \quad \frac{dy}{dt} = \sin\theta(t) \quad \text{and} \quad \frac{d\theta}{dt}(t) = v(t)$$

with

$$|v(t)| = \left\| \frac{d^2\gamma}{dt^2}(t) \right\| \quad \text{and} \quad |v(t)| \le 1.$$

As shown in the paper of Sussmann, Dubins problem can then be regarded either as a time optimal problem in R^3, or more naturally, as an optimal problem on $R^2 \times S^1$ with $v(t)$ playing the role of a control.

As a time optimal control problem, Dubins' problem can be restated as follows: Consider a car in the plane whose speed is always one, and which can either turn left or right with its turning radius of curvature greater than or equal to one. Starting from an arbitrary initial position of the car having on arbitrary initial orientation, find the shortest path for the car which transfers the car to a fixed terminal location and a fixed terminal car orientation at this terminal location.

This point of view was taken by Reeds and Shepp in their remarkable paper [11], in which they allowed the car to also reverse its direction. The reversal of orientation introduces another control u which can only take values ± 1 in equation (1) to yield:

$$(2) \qquad \frac{dx}{dt} = u(t)\cos\theta(t), \quad \frac{dy}{dt} = u(t)\sin\theta(t), \quad \frac{d\theta}{dt} = v(t).$$

Inspired by the papers of Sussmann [12], and Laumonde and Souriat [10], J. D. Boissonat, A. Cereso and J. LeBlond considered in [3] the following interesting variation of Dubins' problem:

$$(3) \qquad \frac{dx}{dt} = u(t)\cos\theta(t), \quad \frac{dy}{dt} = u(t)\sin\theta(t), \quad \frac{d\theta}{dt}(t) = k(t), \quad \frac{dk}{dt} = v(t)$$

under the constraint that $|v(t)| \le B$. B is an arbitrary bound which we will for simplicity take to be equal to 1. They regard (3) as a control system in R^4 with coordinates (x, y, θ, k), and ask for time optimal solutions with $v(t)$ an arbitrary integrable control function bounded by $|v(t)| \le 1$.

The starting point for this paper is to rephrase these problems as time optimal problems on Lie groups, and show that much of the geometry of their solutions can be analyzed in terms of the Lie group symmetries and the appropriate conservation laws described by the Casimir elements of these groups. We shall also further illustrate the importance of this formalism by considering the analogous problems on the sphere and the hyperbolic plane. We begin by describing the basic Lie group formalism.

1. Lie groups. A Lie group G is an analytic manifold such that the group multiplication and the operation of taking the inverse are analytic operations relative to the manifold structure of G. The most well known example of a Lie group is the group of all non singular $n \times n$ matrices with real entries. Such a group is called the general linear group of R^n and is usually denoted by $GL_n(R)$. We shall now consider $GL_n^+(R)$, the connected component of $GL_n(R)$ which contains

the identity. Elements of $GL_n^+(R)$ are all non-singular matrices whose determinant is positive.

The Lie algebra of $GL_n^+(R)$ is equal to its tangent space at the identity. It is well known that the Lie algebra of $GL_n(R)$ is equal to the vector space of all $n \times n$ matrices.

We shall use $M_n(R)$ to denote this space. It is also well known that any closed subgroup of a Lie group G is also a Lie group. In particular closed subgroups of $GL_n^+(R)$ are Lie groups.

In the sequel we will mainly be interested in the following subgroups of $GL_n(R)$.

(i) E_2, the group of motions of the plane. This group consists of all 3×3 matrices of the form

$$\begin{pmatrix} 1 & 0 & 0 \\ x_1 & \alpha & \beta \\ x_2 & -\beta & \alpha \end{pmatrix} \quad \text{with } (x_1, x_2) \in R^2, \text{ and } \alpha^2 + \beta^2 = 1.$$

(ii) The group of rotations of R^n. This group will be denoted by $O_n(R)$, and its connected component through the identity will be denoted by $SO_n(R)$. The elements of $SO_n(R)$ are all the elements of $GL_n^+(R)$ which leave the Euclidean product $(x, y) = \sum_{i=1}^n x_i y_i$ in R^n invariant. That is, $R \in SO_n(R)$ if and only if $\langle Rx, Ry \rangle = \langle x, y \rangle$ for all x and y in R^n, and $\det R > 0$.

We shall also be interested in the group $SO(2, 1)$ which is the group which leaves the Lorentz form $\langle x, y \rangle = -x_1 y_1 + x_2 y_2 + x_3 y_3$ invariant.

If H is a closed subgroup of a Lie group G then the Lie algebra of H will be denoted by $\mathcal{L}(H)$ and is defined as $\mathcal{L}(H) = \{A \in \mathcal{L}(G) : e^{tA} \in H \text{ for all } t \in R\}$.

It easily follows that $\mathcal{L}(E_2)$ consists of all matrices of the form

$$A = \begin{vmatrix} 0 & 0 & 0 \\ \alpha_1 & 0 & -\alpha_3 \\ \alpha_2 & \alpha_3 & 0 \end{vmatrix}, \quad (\alpha_1, \alpha_2, \alpha_3) \in R^3.$$

The Lie algebra of $SO_n(R)$ consist of all $n \times n$ antisymmetric matrices, while the Lie algebra of $SO(2, 1)$ consist of all matrices of the form

$$\begin{vmatrix} 0 & \alpha_1 & \alpha_2 \\ \alpha_1 & 0 & -\alpha_3 \\ \alpha_2 & \alpha_3 & 0 \end{vmatrix}, \quad (\alpha_1, \alpha_2, \alpha_3) \in R^3.$$

For simplicity of exposition we will only consider Lie groups G which are subgroups of $GL_n^+(R)$. Then the tangent space of G at an arbitrary point g of G will be denoted by $T_g G$, and can be defined in two ways: either as $\{Ag : A \in \mathcal{L}(G)\}$ or as $\{gA : A \in \mathcal{L}(G)\}$. We shall use the second definition. For a fixed element $A \in \mathcal{L}(G)$, the vector field $g \to gA$ will be denoted by $\vec{A}(g)$. Such a vector field is called *left invariant* because its value at g is obtained by left-group multiplication of its value at the group identity which is equal to A. An integral curve of $\vec{A}$ is a

curve $g(t)$ which is a solution of $\frac{dg}{dt} = g(t)A$. It follows that $g(t) = g_0 e^{At}$ for some $g_0 \in G$. Since $g(0) = g_0$, it follows that $g(t)$ is the integral curve which passes through g_0 at $t = 0$. It also follows that the integral curve of $\vec{A}$ through any point g_0 is obtained as the left translate of the integral curve of $\vec{A}$ through the group identity, the property which finally justifies the terminology for $\vec{A}$. (The other choice in the representation of tangent spaces naturally leads to right invariant vector fields.)

We shall now return to the problem of Dubins, and show that it is as a time optimal problem on the group of motions E_2.

Let $\gamma(t)$ be any C^1 curve in R^2 parametrized by arc length. Denote by $\vec{a}_1(t)$ the tangent vector $\frac{d\gamma}{dt}$, and denote by $\vec{a}_2(t)$ the unit vector normal to $a_1(t)$ oriented counterclock wise from $a_1(t)$. Let $R(t)$ denote the rotation matrix defined by $R(t)\binom{x}{y} = x\vec{a}_1(t) + x_2\vec{a}_2(t)$. The curve $R(t)$ belongs to $SO_2(R)$, and hence its derivative is a tangent vector to $SO_2(R)$ at $R(t)$. Using the left-invariant description of the tangent spaces, we get that

$$\frac{dR}{dt}(t) = R(t)A(k(t)) \quad \text{where} \quad A(k(t)) = \begin{vmatrix} 0 & -k(t) \\ k(t) & 0 \end{vmatrix}.$$

Since $a_i(t) = R(t)e_i$, $i = 1, 2$ where e_1, e_2 denotes the standard basis in R^2, we get that $\frac{d\gamma}{dt} = a_1(t) = R(t)e_1$, and that $\frac{da_1}{dt} = \frac{dR}{dt}e_1 = R(t)A(k)e_1 = k(t)a_2(t)$. Thus $|k(t)| = \|\frac{d^2\gamma}{dt^2}(t)\|$. $k(t)$ is called the geodesic curvature associated with the curve $\gamma(t)$.

Curve $\gamma(t)$ along with its frame $R(t)$ may be viewed as a curve $g(t)$ in E_2 defined by

$$g(t) = \begin{vmatrix} 1 & 0 \\ \gamma(t) & R(t) \end{vmatrix}.$$

Then

$$(4) \quad \frac{dg}{dt}(t) = g(t) \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & -k(t) \\ 0 & k(t) & 0 \end{vmatrix} = g(t) \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} + k(t)g(t) \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix}.$$

Set

$$L_1 = \begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad L_2 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{vmatrix}, \quad L_3 = \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{vmatrix}.$$

Using these notations equation (4) can be written as:

$$(5) \quad \frac{dg}{dt} = \vec{L}_1(g) + k(t)\vec{L}_3(g).$$

Equation (5) is the Serret-Frenet equation associated with the curve γ. This equation is a coordinate free analogue equation of (1), as can be easily seen from the following parametrization of E_2:

Since $R(t)$ is a rotation in $SO_2(R)$ it can be parametrized by an angle θ defined by

$$R(t) = \begin{vmatrix} \cos\theta(t) & -\sin\theta(t) \\ \sin\theta(t) & \cos\theta(t) \end{vmatrix}.$$

Then

$$\frac{dR}{dt} = \begin{vmatrix} -\sin\theta(t) & -\cos\theta(t) \\ \cos\theta(t) & -\sin\theta(t) \end{vmatrix} \frac{d\theta}{dt} = R(t) \begin{vmatrix} 0 & -k(t) \\ k(t) & 0 \end{vmatrix}.$$

Hence, $\frac{d\theta}{dt} = k(t)$. If we now denote $\gamma(t)$ by $\binom{x(t)}{y(t)}$, it follows that

$$\frac{d\gamma}{dt} = R(t)e_1 = \begin{pmatrix} \cos\theta(t) \\ \sin\theta(t) \end{pmatrix}$$

and hence our equations reduce to equation (1), providing that $k(t) = v(t)$.

It is now easy to see that the dynamics for the Reeds-Shepp car can be expressed by

$$(6) \qquad \frac{dg}{dt} = u(t)\vec{L}_1(g) + v\vec{L}_3(g)$$

with $u(t) = \pm 1$ and $|v(t)| \leq 1$.

The natural state space for the problem of Boissonat-Cereso-LeBlond is $G = E_2 \times R$ in which the dynamic equations are given by

$$(7) \qquad \frac{dg}{dt} = \vec{L}_1(g) + k(t)\vec{L}_3(g), \quad \frac{dk}{dt} = v(t), \quad |v(t)| \leq 1$$

where $\widehat{g}(t) = (g(t), k(t))$ denotes the solution curve in G.

2. Controllability. Controllability of control systems on Lie groups, which are affine in controls, has been studied in considerable detail some time ago ([2], [8], [9]).

That both systems (5) and (6) are controllable follows from easy geometrical arguments. In this context, it might be interesting to recall the controllability results of [2] which state that any left invariant family of vector fields on any semi-direct product of a compact group with a vector space is controllable provided that the Lie algebra generated by the family is equal to the Lie algebra of the semi-direct product. Controllability of (5) is a particular case of this theorem since E_2 is a semi-direct product of $SO_2(R)$ with R^2.

It will be convenient for further analysis to note that both (5) and (6) are left-invariant control systems, and that therefore their reachable sets from an arbitrary point g in E_2 are the left translates of the reachable set through the identity. That means that

$$\mathcal{A}(g) = g\mathcal{A}(e) = \{gh : h \in \mathcal{A}(e)\},$$

where $\mathcal{A}(g)$ denotes the reachable set from g.

The preceding observation shows that in both Dubins' and Reeds-Shepp's problem it suffices only to consider the shortest path which originates at the

group identity e. For, if $g(t)$ is the shortest path which connects g_0 to g_1, then $h(t) = g_0^{-1}g(t)$ is the shortest path which connects e to $g_0^{-1}g_1$. Evidently, the converse must also be true.

We now consider controllability of the Boissonat-Cereso-LeBlond problem. To begin with, note that it is not a left-invariant problem on G because the system is not invariant in k-direction. That is, $(\vec{L}_1(g) + k\vec{L}_3(g), v) \neq (g, k)(\vec{L}_1(0) + 0 \cdot \vec{L}_3(0), v)$ for any v. Since however

$$(\vec{L}_1(g) + k\vec{L}_3(g), v) = (g, 0)(\vec{L}_1(0) + k\vec{L}_3(0), v).$$

the system is still invariant relative to E_2. It then follows that its reachable set $\mathcal{A}(g, k)$ from $(g, k) \in G$ satisfies the following invariance property $\mathcal{A}(g, k) = (g, 0)\mathcal{A}e, k)$.

The drift vector field $X(g, k)$ is equal to $(\vec{L}_1(g) + k\vec{L}_3(g), 0)$. Its integral curves are given by

$$\frac{dg}{dt} = \vec{L}_1(g) + k(t)\vec{L}_3(g) \quad \text{and} \quad \frac{dk}{dt} = 0.$$

Hence $k(t) = \text{constant}$, and $R(t) = \begin{vmatrix} \cos kt & -\sin kt \\ \sin kt & \cos kt \end{vmatrix}$, and $x(t) = \frac{1}{k}\sin kt$, $y(t) = \frac{1}{k}(1 - \cos kt)$. It then follows that $(x(t), y(t))$ moves along a circle.

Thus each integral curve of the drift vector field is closed, and hence, control system (7) has a periodic drift. But then it is well known ([8] or [9]) that such a system is controllable whenever its Lie algebra is of full rank. So in order to show controllability we need only to verify the rank condition.

Following the general convention that the Lie bracket of any vector fields X and Y is given by the formula $[X, Y]f = Y(Xf) - X(Yf)$, it can be proved that the Lie bracket of two left invariant vector fields $\vec{A}(g)$ and $\vec{B}(g)$ is given by $[\vec{A}, \vec{B}](g) = \overrightarrow{[A, B]}(g)$ where $[A, B] = BA - AB$. Notice that our convention implies that $[A, B]$ is the negative of the usual commutator rule.

We then get the following Lie bracket table for the Lie algebra of E_2:

$$[L_1, L_2] = 0, \quad [L_1, L_3] = L_2, \quad \text{and} \quad [L_3, L_2] = L_1.$$

Let $X = (\vec{A}(k), a\frac{\partial}{\partial k})$ and $Y = (\vec{B}(k), b\frac{\partial}{\partial k})$ be two vector fields in G such that for each fixed k, $\vec{A}(k)$ and $\vec{B}(k)$ are left invariant on E_2. Then

$$[X, Y]f = Y(Xf) - X(Yf)$$
$$= \left(\vec{B}(k), b\frac{\partial}{\partial k}\right)\left(\vec{A}(k)f + a\frac{\partial f}{\partial k}\right) - \left(\vec{A}(k), a\frac{\partial}{\partial k}\right)\left(\vec{B}(k)f + b\frac{\partial f}{\partial k}\right)$$
$$= (\vec{B}(k)\vec{A}(k) - A(k)\vec{B}(k))f + b\frac{\partial}{\partial k}\vec{A}(k)f - a\frac{\partial}{\partial k}\vec{B}(k)f.$$

In particular, now take $X = (\vec{L}_1 + k\vec{L}_3, 0)$ and $Y = (0, \frac{\partial}{\partial k})$. Then $[X, Y](f) = (\vec{L}_3 f, 0) = X_1$, $[X, [X, Y]] = [\vec{L}_1, \vec{L}_3] = \vec{L}_2 = X_2$. Finally $[X_1, X_2] =$

$([\vec{L}_3, \vec{L}_2], 0) = (\vec{L}_1, 0)$. Therefore the Lie algebra is of full rank at each point of G, and therefore (7) is controllable.

R e m a r k. Let (x, y, θ, k) be the system of coordinates described earlier. Then the vector field $\vec{L}_1(g) + k\vec{L}_3(g) = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y} + k\frac{\partial}{\partial \theta} + 0 \cdot \frac{\partial}{\partial k}$. In terms of these coordinates $\vec{L}_3(g) = \frac{\partial}{\partial \theta}$ and

$$[X, Y] = [\cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y} + k\frac{\partial}{\partial \theta}, \frac{\partial}{\partial k}] = \frac{\partial}{\partial \theta},$$

$$[X, [X, Y]] = -\sin\theta \frac{\partial}{\partial x} + \cos\theta \frac{\partial}{\partial y} = \vec{L}_2 \quad \text{and}$$

$$\left[\frac{\partial}{\partial \theta}, -\sin\theta \frac{\partial}{\partial x} + \cos\theta \frac{\partial}{\partial y} \right] = \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y} = \vec{L}_1.$$

Thus our calculations agree with those in [3].

3. The cotangent bundle of a Lie group. The cotangent space of G at g, which will be denoted by $T_g^*(G)$, is the space of all linear functions on $T_g(G)$. The cotangent bundle of G, denoted by T^*G, is the union of all $T_g^*(G)$ as g varies over G. We will regard T^*G as the product $\mathcal{L}^* \times G$ where $\mathcal{L}^*$ denotes the dual of the Lie algebra $\mathcal{L}$ of G. This realization of T^*G as $\mathcal{L}^* \times G$ is accomplished through the left translations $L_g(x) = gx$ for all $x \in G$, and $g \in G$, and goes as follows:

Let dL_g denote the tangent map of L_g at x. Then $dL_g : T_x g \to T_{L_g(x)} G$. The dual map dL_g^* of dL_g maps $T_{gx}^*(G)$ onto T_x^*G. In particular $dL_{g^{-1}}^*$ maps T_e^*G onto T_g^*G. The correspondence $(\ell, g) \to dL_g^{-1}(\ell)$ realizes $\mathcal{L}^* \times G$ as T^*G. It then follows that the tangent bundle $T(T^*G)$ of the cotangent bundle T^*G can be identified with

$$T(\mathcal{L}^*(G) \times G) = T(\mathcal{L}^*(G)) \times TG = (\mathcal{L}^* \times \mathcal{L}^*) \times (\mathcal{L} \times G).$$

For convenience of notation we will write $T(T^*G) = (\mathcal{L}^* \times G) \times (\mathcal{L}^* \times \mathcal{L})$ with the understanding that each element $((p, g), (W^*, V))$ of $(\mathcal{L}^* \times G) \times (\mathcal{L}^* \times \mathcal{L})$ denotes a tangent vector (W^*, V) at a basepoint (p, g).

The cotangent bundle T^*M of each manifold is equipped with a natural differential form θ. In terms of the above representation of $T(T^*M)$, $\theta_{(p,g)}(W^*, V) = p(V)$. The symplectic form ω is equal to the exterior derivative $d\theta$ of θ.

It follows that

(9) $\qquad \omega_{(g,p)}((W_1^*, V_1), (W_2^*, V_2)) = W_1^*(V_2) - W_2^*(V_1) - p([V_1, V_2])$

as can be easily deduced from the general formula

$$d\theta(X_1, X_2) = X_1(\theta(X_2)) - X_2(\theta(X_1)) - \theta([X_1, X_2]).$$

Each function H on T^*G corresponds to a vector field $\vec{H}$, defined by the formula $dH_x(v) = \omega_x(v, \vec{H}(x))$ for each point x on T^*G. This correspondence is unique up to constant functions. $\vec{H}$ is called the Hamiltonian vector field of H,

and H is called a Hamiltonian of $\vec{H}$. More explicitly this formalism leads to the following equations:

Denote $x = (p, g)$, and $\vec{H}(p, g) = (Y^*, X)$. Using $dH_{(p,g)}(W^*, V) = W^*(X) - Y^*(V) + p[X, V]$, we get that an integral curve $p(t), g(t)$ of $\vec{H}$ must satisfy

$$(10) \qquad \begin{cases} W^*(g^{-1}\frac{dg}{dt}) = W^*(X) = dH_{(p(t),g(t))}(0, W^*) \text{and} \\ \frac{dp}{dt}(V) = Y^*(V) = p(t)([X, V]) - dH_{(p(t),g(t))}(V, 0) \end{cases}$$

for all (W^*, V) in $\mathcal{L}^* \times \mathcal{L}$.

4. Hamiltonians, Poisson brackets and Casimir elements. Each vector field X on G, or more generally, on any manifold M has a canonical Hamiltonian lift $\vec{H}_X$. $\vec{H}_X$ is the Hamiltonian vector field which corresponds to $H_X(x) = x(X(g))$ for all $x \in T_g^*G$. In particular when $X(g) = \vec{A}(g)$ is a left invariant vector field on G, then $H_X(p, g) = p(A)$. Thus H_X is a linear function on $\mathcal{L}^*$.

The basic symplectic geometry is naturally expressed in terms of the Poisson bracket. The essential properties of the Poisson bracket $\{,\}$ are summarized by the following proposition.

PROPOSITION 1. *Suppose that H and F are smooth functions on T^*M. Denote by $\{\exp t\vec{H}\}$ the one-parameter group of diffeomorphisms generated by the Hamiltonian field $\vec{H}$ of H. Then*

(i) $\{F, H\}(x) = \frac{d}{dt}F \circ \exp t\vec{H}(x)|_{t=0}$ *for each $x \in T^*M$.*

(ii) *The Lie bracket $[\vec{F}, \vec{H}]$ of two Hamiltonian vector fields $\vec{F}$ and $\vec{H}$ is a Hamiltonian vector field, and $[\vec{F}, \vec{H}] = \overrightarrow{\{F, H\}}$.*

(iii) *If H_X and H_Y are the Hamiltonian functions on T^*M which correspond to smooth vector fields X and Y on M then $\{H_X, H_Y\} = H_{[X,Y]}$.*

R e m a r k. All of these statements along with their proofs can be found in Arnold's book ([1], Ch. 8) with a slight transposition of the terminology: the Poisson bracket of vector fields in Arnold is equal to the Lie bracket of this paper is the same as the Poisson bracket of functions of Arnold.

It follows immediately from (iii) of the above proposition that the Hamiltonians of left invariant vector fields on G form a Lie algebra under the Poisson bracket. This algebra is isomorphic to $\mathcal{L}(G)$. It also follows from (i) that $\{F, H\}(x) = \omega_x(\vec{F}(x), \vec{H}(x))$ for all $x \in T^*G$. This relation easily yields the Leibniz rule: $\{F_1F_2, H\} = F_1\{F_2, H\} + F_2\{F_1, H\}$ for any functions F_1, F_2 and H.

We shall now restrict our attention to functions on $\mathcal{L}^*$. Each such function may be regarded as a Hamiltonian which is constant over G. It then follows from equations (10) that $\vec{f}(p, g) = ((ad\, df(p))^*p, df(p))$, where $(ad\, df(p))^*p$ dentes the linear function $V \to p[df(p), V]$. The Poisson bracket of two functions f and h on $\mathcal{L}^*$ is given by $\{f, h\}(p) = p[df(p), dh(p)]$ as can be easily verified from the relation

$\{f, h\}(p) = w_{(p,g)}, (\vec{f}(p, g), \vec{h}(p, g))$. In the literature on Hamiltonian systems this expression is often called the Poisson-Lie bracket.

Suppose now that $A_1, \ldots, A_n$ is any basis in the Lie algebra of G. Denote by $h_1, \ldots, h_n$ the Hamiltonians of $\vec{A}_1, \ldots, A_n$. As we have already remarked $h_1, \ldots, h_n$ form a Lie algebra under their Poisson bracket, since $\{h_i, h_j\}(p) = p([A_i, A_j])$. Let $\mathcal{P}(h_1, \ldots, h_n)$ denote the algebra of all polynomials in the variables $h_1, \ldots, h_n$. Each element H of $\mathcal{P}(h_1, \ldots, h_n)$ is a function on $\mathcal{L}^*$, and the preceding remarks apply. $\mathcal{P}(h_1, \ldots, h_n)$ a Poisson-Lie algebra, since $\{H, F\}$ is an element of $\mathcal{P}(h_1, \ldots, h_n)$ for any H and F in $\mathcal{P}(h_1, \ldots, h_n)$.

Each element of the center of $\mathcal{P}(h_1, \ldots, h_n)$ is called a Casimir element. Thus M is a Casimir element if and only if $\{M, h_i\} = 0$ for each $i = 1, 2, \ldots, n$. The essential property of Casimir elements is expressed by the following

PROPOSITION 2. *Suppose that $H_1, \ldots, H_k$ are any elements of $\mathcal{P}(h_1, \ldots, h_n)$, and suppose that $\xi(t)$ is any absolutely continuous curve defined on an interval $[0, T]$ which further satisfies $\frac{d\xi}{dt}(t) = \sum_{i=1}^{k} u_i(t)\vec{H}_i(\xi(t))$ a.e. on $[0, T]$ for some integrable functions $u_1, \ldots, u_m$. Then, $M \circ \xi(t) = $ constant for any Casimir element M.*

Proof.

$$\frac{d}{dt} M \circ \xi(t) = \Big\{ M, \sum_{i=1}^{k} u_i(t) H_i \Big\} \xi(t)$$

$$= \sum_{i=1}^{k} u_i(t)\{M, H_i\}(\xi(t)) = 0 \quad \text{a.e.in } [0, T],$$

because $\{M, H_i\} = 0$ for each $i = 1, 2, \ldots, k$. Hence $M \circ \xi(t) = $ constant. ∎

5. Applications to control problems. Let us now return to the problems of Dubins, and Reeds-Shepp mentioned earlier, and interpret some of their optimality results in the context of the above formalism. In order to further emphasize the importance of the Lie algebraic formalism we will consider analogous problems simultaneously on $E_2, SO_3(R)$ and $SO(2, 1)$. Their Lie algebras are expressed by the following Lie bracket table:

Table 1

[,]	L_1	L_2	L_3
L_1	0	$-\epsilon L_3$	L_2
L_2	ϵL_3	0	$-L_1$
L_3	$-L_2$	L_1	0

$\epsilon = 0, \quad$ for E_2

$\epsilon = 1, \quad$ for $SO_3(R)$

$\epsilon = -1, \quad$ for $SO(2, 1)$.

$SO_3(R)$ is the isometry group for the Riemannian sphere S^2, while $SO(2, 1)$ is the isometry group for the hyperboloid $H^2 = \{x^2 - (y^2 + z^2) = 1, x > 0\}$ with the hyperbolic metric (Minkowski model for the hyperbolic geometry). ϵ is the curvature of these spaces. Each of the control problems mentioned earlier can

be naturally stated on S^2 and H^2, with the ordinary derivative replaced by the covariant derivative.

Following the notation of the previous section let h_1, h_2 and h_3 denote respectively the Hamiltonian of $\vec{L}_1, \vec{L}_2$ and $\vec{L}_3$. Their Poisson algebra is isomorphic to the Lie algebra, given in Table 1 and hence:

Table 2

[,]	h_1	h_2	h_3
h_1	0	$-\epsilon h_3$	h_2
h_2	ϵh_3	0	$-h_1$
h_3	$-h_2$	h_1	0

It is easy to check that $M = h_1^2 + h_2^2 + \epsilon h_3^2$ is a Casimir element for $\mathcal{P}(h_1, h_2, h_3)$. It is also easy to verify that there are no other functionally independent Casimir elements for these groups.

On $SO_3(R)$, $M = $ constant is a sphere in the h_1, h_2, h_3 space, usually called the angular momentum sphere. The reader might recall the equations for the Euler's top whose solutions are curves in the intersection of the energy ellipsoid with the momentum sphere. Of course, M is an integral of motion, not just for the Euler top, but for any Hamiltonian which is invariant on G, i.e., a polynomial (or more generally an analytic function) on $\mathcal{L}^*$. This observation applies in particular to the Hamiltonians of our time optimal control problems. Let us begin with the problem of Dubins:

The Hamiltonian lift of control system (5) is given by $H = -\lambda + h_1 + k(t)h_3$. According to the Maximum Principle each time optimal trajectory $g(t)$ defined on an interval I, corresponding to the control $k(t)$ is the projection of an integral curve $\xi(t)$ of $\vec{H}$ defined for t in I such that:

(1°) λ is either 0 or 1. If $\lambda = 0$, then $\xi(t) \neq 0$ for any t in I.
(2°) $H(\xi(t)) = 0$ for almost all t in I.
(3°) $H(\xi(t)) = \sup_{|u| \leq 1}(-\lambda + h_1(\xi(t)) + uh_3(\xi(t)))$ for almost all t in I.

We shall say that an absolutely continuous curve $\xi(t)$ defined on an interval I is an extremal curve (for the problem of Dubins) if there exists a measurable function $k(t)$ with $|k(t)| \leq 1$ a.e. on I such that $\xi(t)$ is an integral curve of $\vec{H}_0 = \vec{h}_1 + k(t)\vec{h}_3$, along which (1°), (2°) and (3°) hold. The extremals which correspond to $\lambda = 0$ are called *abnormal*, while the extremals which correspond to $\lambda = 1$ are called *regular*. The only difference between regular and abnormal extremals for the problems of Dubins type is that they are constrained to different "energy" levels of H. The Maximum Principle says that every optimal trajectory is the projection of an extremal curve.

Suppose now that $(\xi(t), k(t))$ is an extremal pair with $\xi(t)$ being an extremal defined by the control function $k(t)$. Then, the maximality condition (3°) implies that $k(t) = \operatorname{sgn} h_3(\xi(t))$. i.e., $k(t) = 1$ for all t for which $h_3(\xi(t)) > 0$ and that $k(t) = -1$ for all t for which $h_3(\xi(t)) > 0$. $h_3 = 0$ is called the *switching sur-*

face. Extremals can be further classified according to its type of crossing of the switching surface.

We shall first consider the following situation:

(i) An extremal $\xi(t)$ does not cross the switching surface. Then the corresponding control is always constant and equal to ± 1. The induced trajectory $g(t)$ in G is an integral curve of

$$\frac{dg}{dt} = \vec{L}_1(g) \pm \vec{L}_3(g), \quad \text{or} \quad \frac{dg}{dt} = g(t) \begin{pmatrix} 0 & -\epsilon & 0 \\ 1 & 0 & \mp 1 \\ 0 & \pm 1 & 0 \end{pmatrix}.$$

Let

$$A = \begin{pmatrix} 0 & -\epsilon & 0 \\ 1 & 0 & \mp 1 \\ 0 & \pm 1 & 0 \end{pmatrix}.$$

Then $g(t) = g_0 e^{At}$ with $g(0) = g_0$. For $\epsilon = 0$

$$e^{At} = \begin{pmatrix} 1 & 0 \\ x(t) & R(t) \end{pmatrix} \quad \text{with} \quad R(t) = \begin{pmatrix} \cos t & \mp \sin t \\ \pm \sin t & \cos t \end{pmatrix},$$

and $x_1(t) = \sin t$, $x_2(t) = \pm(1 - \cos t)$. Thus $x(t)$ traverses circles passing through the origin with centers at $(0, \pm 1)$. Each circle is traversed with speed 1; the circle which corresponds to $k = 1$ is traversed in the counter clockwise direction, while the other is traversed with the opposite orientation. Any other solution is a translate and a rotation of this solution.

For $\epsilon = 1$, e^{At} is the rotation around $\omega = \pm e_1 + e_3$ with $\theta(t) = \sqrt{2}t$. The projection of e^{At} on S^2 is given by $e^{At}e_1$ and is a circle

$$(x_1 \mp \tfrac{1}{2})^2 + x_2^2 + (x_3 - \tfrac{1}{2})^2 = \tfrac{1}{2}.$$

Any other solution is a rotation (from the left) of this circle.

For $\epsilon = -1$,

$$A = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & \mp 1 \\ 0 & \pm & 0 \end{vmatrix}.$$

It follows that

$$A^2 = \begin{vmatrix} 1 & 0 & \mp 1 \\ 0 & 0 & 0 \\ \pm 1 & 0 & -1 \end{vmatrix}$$

and $A^3 = 0$. Hence,

$$e^{At} = I + tA + \frac{t^2}{A^2} = \begin{vmatrix} 1 + \frac{t^2}{2} & t & \mp \frac{t^2}{2} \\ t & 1 & \mp t \\ \pm \frac{t^2}{2} & \pm t & 1 - \frac{t^2}{2} \end{vmatrix}.$$

The projection of e^{At} on the hyperboloid $x^2 - (y^2 + z^2) = 1$, $x > 0$ is given by $e^{At}e_1$. It follows that $x(t) = 1 + \frac{t^2}{2}$, $y(t) = t$, and $z(t) = \pm \frac{t^2}{2}$. Such a curve is

called a horo-cycle (it is the projection of a one-parameter group generated by a nilpotent element of $\mathcal{L}(G)$). It is noncompact, since $y^2(t) = \pm 2z(t)$.

(ii) An extremal $\xi(t)$ crosses the switching surface. There are two types of crossing, transversal and tangential, depending whether $\frac{d}{dt}h_3(\xi(t)) \neq 0$ for the crossing time t. It follows that,

$$\frac{d}{dt}h_3(\xi(t)) = \{h_3, h_1 + k(t)h_3\}(\xi(t) = \{h_3, h_1\}(\xi(t) = -h_2(\xi(t)).$$

The transversal case occurs when $h_3(\xi(t)) = 0$ and $h_2(\xi(t)) \neq 0$. The tangential case occurs when $h_2(\xi(t)) = h_3(\xi(t)) = 0$.

The remaining case occurs when

(iii) An extremal $\xi(t)$ remains on the switching surface for an open interval I. Then $h_3(\xi(t)) = 0$ for all t in I. After differentiation, $h_2(\xi(t)) = 0$ for all t in I. Upon further differentiation we get that $\{h_2, h_1 + k(t)h_3\} = 0$ for all t in I. Hence, $\{h_2, h_1\}(\xi(t) + k(t)\{h_2, h_3\}(\xi(t)) = 0$ for all t in I, or $\epsilon h_3(\xi(t)) - k(t)h_1(\xi(t)) = 0$ for all t in I. This last relation shows that $\xi(t)$ cannot be an abnormal extremal: for if it were, then $k(t) = 0$ for all t in I, since $h_1(\xi(t))$ cannot vanish anywhere in I by the Maximum Principle. But these two conditions are mutually contradictory since

$$0 = -\lambda + h_1(\xi(t)) + k(t)h_3(\xi(t))$$

and $\lambda = 0$.

Thus $\xi(t)$ is a regular extremal, and therefore $1 = h_1(\xi(t)) + k(t)h_3(\xi(t)) = h_1(\xi(t))$. Hence, $k(t) = 0$ for all t in I. The projection of such an extremal is a geodesic. In the planar case, the geodesics are straight lines, in the elliptic case they are great circles on S^2, and in the hyperbolic case they are hyperbolas on H^2.

All of the types of solutions can be easily described geometrically in terms of the Casimir surface $h_1^2 + h_2^2 + \epsilon h_3^2 = M$ as follows:

Abnormal extremals are the intersections of planes $h_1 \pm h_3 = 0$ with the Casimir surface. These intersections are always transversal to the switching surface $h_3 = 0$. The switching times for the Euclidean case occur in intervals π, for the elliptic case they occur in intervals $\sqrt{2}\pi$, while for the hyperbolic case there is only one switch.

Regular extremals are the intersections of planes $1 = h_1 \pm h_3$ with the Casimir surface. The critical value of M for which switching occurs is $M = 1$. For $M < 1$ there is no switching.

For $M > 1$, the optimal control is of bang-bang type since the value of the control function switches between the extreme values ± 1. $M = 1$ contains the singular arc. The optimal solutions bifurcate at $M = 1$ as they go from no switching, to switching including the singular arc, to bang-bang type.

The Reeds-Shepp problem, although much more singular, can be analyzed in much the same way. It follows that for this problem the solution curves are the

intersections of the planes $\pm h_1 \pm h_3$ with the Casimir surfaces. The switching surfaces are $h_1 = 0$ and $h_3 = 0$.

We shall end this paper by concentrating on some aspects of the Boissonat-Cereso-LeBlond problem (Equation (7)). We shall continue with our earlier notations. In particular, functions h_1, h_2 and h_3 have the same meaning as before. We shall use h_4 to denote the Hamiltonian lift of $\frac{\partial}{\partial k}$. Thus h_1, h_2, h_3 and h_4 are all linear functions on $\mathcal{L}^*$. As in the case of Dubins and Reeds-Shepp, this problem also extends to non-Euclidean situations. We shall continue to use G for any one of $E_2, SO_3(R)$ and $SO(2,1)$. $\widehat{G} = G \times R$ denotes the state space for the equation (7) with solution curves $\widehat{g}(t) = (g(t), k(t))$ evolving in $G \times R$.

Equations (10) can be used to obtain the following differential equations along the extremals $\xi(t)$. Here $h_i(t) = h_i \circ \xi(t)$, $i = 1, 2, 3, 4$.

$$
\begin{aligned}
(11) \qquad
&\frac{dh_1}{dt} = \{h_1, -\lambda + h_1 + kh_3 + vh_4\} = k\{h_1, h_3\} = kh_2, \\
&\frac{dh_2}{dt} = \{h_2, -\lambda + h_1 + kh_3 + vh_4\} = \{h_2, h_1\} + k\{h_2, h_3\} = \epsilon h_3 - kh_1, \\
&\frac{dh_3}{dt} = \{h_3, -\lambda + h_1 + kh_3 + vh_4\} = \{h_3, h_1\} = -h_2, \\
&\frac{dh_4}{dt} = \{h_4, -\lambda + h_1 + kh_3 + vh_4\} = \{h_4, kh_3\} = h_3\{h_4, k\} = -h_3.
\end{aligned}
$$

It is easy to verify that $h_1^2 + h_2^2 + \epsilon h_3^2 = M^2$ is a constant of motion. Let us now calculate the singular arcs: Assume that $h_4(\xi(t)) = 0$ for all t in some open interval I. Then,

$$
\frac{d}{dt}h_4(\xi(t)) = \{h_4, -\lambda + h_1 + kh_3 + v(t)h_4\} = \{h_4, kh_3\}(\xi(t)) = -h_3(\xi(t)) = 0,
$$

$$
\frac{d}{dt}h_3(\xi(t)) = -h_2(\xi(t)) = 0, \quad \text{and} \quad -\frac{dh_2}{dt}(\xi(t)) = kh_1(\xi(t)) = 0.
$$

Thus, $kh_1(\xi(t)) = 0$.

It follows by the Maximum Principle that $h_1(\xi(t)) \neq 0$. Hence the singular arcs are constrained to the submanifolds $k = 0, v = 0$. The corresponding integral curves of (7) project onto geodesics.

We will end this paper by drawing comparisons with [3]. h_3 and h_4 have the same meaning as β and r in [3]. Since $h_1^2 + h_2^2 = M^2$, let $h_1 = M \sin \theta$, and $h_2 = M \cos \theta$. Then

$$
\frac{dh_1}{dt} = -M \sin \theta \frac{d\theta}{dt} = kh_2
$$

(from the equations above). Therefore $\frac{d\theta}{dt} = k$. Thus the angle θ, and the angle α used in [3] in equations (1), differ by a constant. Suppose that this constant is $\frac{\pi}{2} - \phi$, or that $\theta = \frac{\pi}{2} + \alpha - \phi$. Then $h_1 = M \sin(\frac{\pi}{2} + \alpha - \phi) = M \cos(\alpha - \phi)$.

Moreover,

$$\frac{d\beta}{dt} = \frac{dh_3}{dt} = -h_2 = -M\cos\theta = -M\cos\left(\frac{\pi}{2} + \alpha - \phi\right) = M\sin(\alpha - \phi).$$

This correspondence of notation relates equations above with the extremals in [3] (of course, λ in [3] is equal to M in this paper).

We will end this paper by providing a geometric interpretation for a conservation law $\beta - py + qx =$ constant in Lemma 5 of [3]. We shall show that this relation follows from left invariance of equations (7) with respect to the subgroup G of $\widehat{G} = G \times R$.

Equations (10) show that if $dH_{(p,g)}(V,0) = 0$ then $\frac{dp}{dt}(V) = p(t)[X,V] = \mathrm{ad}^* X(V)(p)$.

This equation integrates to yield $p(t)(g^{-1}(t)Vg(t)) =$ constant. We shall identify $\mathcal{L}^*$ with $\mathcal{L}$ via the pairing

$$\left\langle \begin{pmatrix} 0 & 0 & 0 \\ a_1 & 0 & -a_3 \\ a_2 & a_3 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ b_1 & 0 & -b_3 \\ b_2 & b_3 & 0 \end{pmatrix} \right\rangle = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Then the projection $p(t)$ of each extremal curve $\xi(t)$ on $\mathcal{L}^*$ is identified with a curve $P(t)$ in $\mathcal{L}$ with

$$P(t) = \begin{vmatrix} 0 & 0 & 0 \\ h_1(t) & 0 & -h_3(t) \\ h_2(t) & h_3(t) & 0 \end{vmatrix}$$

(see [5] for further details).

Let $g(t)$ denote the projection of $\xi(t)$ on G. We shall write

$$g(t) = \begin{vmatrix} 1 & 0 \\ x(t) & R(t) \end{vmatrix}.$$

Then $(P(t), g^{-1}(t)Vg(t)) =$ constant yields:

$$R(t)\begin{pmatrix} h_1(t) \\ h_2(t) \end{pmatrix} \quad \text{is a constant vector,}$$

and

$$\left(R(t)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x(t) \right) - h_3(t) = \text{ constant.}$$

If we denote the components of the constant vector by (p,q) then the second equation yields $-px_2 + qx_1 - h_3 =$ constant. Except for the change of notation this relation is the same as in Lemma 5 of [3]. Incidentally, this relation is valid even if $p = q = 0$. For in this case $M^2 = p^2 + q^2$. Thus $M = 0$ implies that $h_1 = h_2 = 0$, and therefore $h_3 =$ constant, as can be easily seen from equations (11).

References

[1] V. I. A r n o l d, *Méthodes Mathématiques de la Mécanique Classique*, Editions Mir, 1976.

[2] B. B o n n a r d, V. J u r d j e v i c, I. A. K. K u p k a and G. S a l l e t, *Transitivity of families of vector fields on the semi-direct products of Lie groups*, Trans. Amer. Math. Soc. 275 (1982), 525–535.

[3] J. D. B o i s s o n a t, A. C e r e s o, and J. L e B l o n d, *On shortest paths in the plane subject to a constraint on the derivative of the curvature*, a preprint.

[4] L. E. D u b i n s, *On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal position and tangents*, Amer. J. Math. 79 (1957), 497–516.

[5] V. J u r d j e v i c, *Non-Euclidean elastica*, Amer. J. Math. 117 (1995), 93–124.

[6] —, *The geometry of the plate-ball problem*, Arch. Rational Mech. Anal. 124 (1993), 305–328.

[7] —, *Optimal Control Problems on Lie Groups: Crossroads between Geometry and Mechanics*, in: The Geometry of Non-linear Feedback and Optimal Control, B. Jakubczyk and W. Respondek (eds.), Marcel Dekker, to appear.

[8] V. J u r d j e v i c and I. K u p k a, *Control systems on semi-simple Lie groups and their homogeneous spaces*, Ann. Inst. Fourier (Grenoble) 31 (4) (1981), 151–179.

[9] C. L o b r y, *Controllability of non-linear systems on compact manifolds*, SIAM J. Control 12 (1974), 1–4.

[10] J. P. L a u m o n d e et P. S o u r i a t, *Synthèse des plus courts chemins pour la voiture de Reeds et Shepp*, Repport LAAS/CNRS 92234, Juin 1992.

[11] J. A. R e e d s and R. A. S h e p p, *Optimal paths for a car that goes both forward and backwards*, Pacific J. Math. 145 (1990), 367–393.

[12] H. J. S u s s m a n n and G. T a n g, *Shortest paths for the Reeds-Shepp car: A worked out example of the use of geometric techniques in non-linear optimal control*, a preprint.