mathematics

Article

Rolling Geodesics, Mechanical Systems and Elastic Curves

Velimir Jurdjevic

Department of Mathematics, University of Toronto, Toronto, ON M5S 3G3, Canada; jurdj@math.toronto.edu

Abstract: This paper defines a large class of differentiable manifolds that house two distinct optimal problems called affine-quadratic and rolling problem. We show remarkable connections between these two problems manifested by the associated Hamiltonians obtained by the Maximum Principle of optimal control. We also show that each of these Hamiltonians is completely intergrable, in the sense of Liouville. Finally we demonstrate the significance of these results for the theory of mechanical systems.

Keywords: Lie groups; Lie algebras; homogeneous manifolds; Hamiltonians; Poisson bracket; mechanical tops

MSC: 49J15; 53A17; 53A35; 58A05; 58A30; 70B15

Citation: Jurdjevic, V. Rolling Geodesics, Mechanical Systems and Elastic Curves. *Mathematics* **2022**, *10*, 4827. https://doi.org/10.3390/math10244827

Academic Editor: Margarida Camarinha

Received: 20 September 2022
Accepted: 12 December 2022
Published: 19 December 2022

1. Introduction

This paper is a continuation of my long-standing interest in the role of Lie groups and Lie algebras in the theory of integrable systems and the equations of mathematical physics. The interest in this topic originated in two seemingly unrelated phenomena, the presence of elastica in the theory of rolling spheres ([1,2]), and the presence of the heavy top in the equations describing the equilibrium configurations of an elastic rod ([3,4]). My interest in these phenomena was further renewed by the subsequent studies ([5–7]) that showed intriguing connections between rolling problems, elastic curves and problems in mechanics. These studies also identified a class of variational problems on Lie groups, called affine-quadratic that not only played a pivotal role in this theory, but also made a significant impact on the theory of integrable systems ([8], Chapters 9, 10 and 11).

In this paper, we will shift emphasis to a new class of rolling problems associated with homogeneous Riemannian spaces rolling isometrically on their tangent planes (based on our recent study [9,10]). We will show that each such isometric rolling has a well defined length which then leads to natural definition for a rolling geodesic. The rolling problem then consists of finding some necessary differential conditions that the rolling geodesics must satisfy.

We will show that each rolling problem can be recast as a left-invariant optimal control problem on a Lie group, and consequently, we will be able to regard the rolling geodesics as the projections of the extremal curves generated by a suitable Hamiltonian obtained through Pontryagin's Maximum Principle. We will show several remarkable properties of the aforementioned Hamiltonian. First we will show that any such Hamiltonian is completely integrable, and secondly, we will show that the Hamiltonian system associated with an affine-quadratic system may be regarded as an invariant subsystem of the Hamiltonian differential system associated with the rolling problem. This discovery sheds new light on the geometric origins of the affine-quadratic systems and their connections to mechanical systems ([11,12]). These findings seem particularly remarkable considering the fact that the control functions that define these optimal problems lie in mutually orthogonal spaces of each other.

The general setting of the paper in which the above-mentioned problems will be analyzed is defined by a semi-simple Lie group G and a compact subgroup K with a finite centre. Any such pair (G, K) is reductive in the sense that the Lie algebra $\mathfrak{g}$ of G admits a splitting $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ where $\mathfrak{p}$ is a vector space complementary to the Lie algebra $\mathfrak{k}$ of K. In this paper, $\mathfrak{p}$ will be the orthogonal complement of $\mathfrak{k}$ relative to the Killing form $Kl(X, Y) = Tr(adX \circ adY)$. We recall that the Killing form is non-degenerate on G and also satisfies

$$Kl(X, [Y, Z]) = Kl([X, Y], Z)$$

for any elements X, Y, Z in $\mathfrak{g}$. This implies that $[\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}$. We shall make another assumption that $\mathfrak{k}$ and $\mathfrak{p}$ satisfy strong Cartan's Lie algebraic conditions

$$[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}, [\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}, [\mathfrak{k}, \mathfrak{k}] \subseteq \mathfrak{k}. \tag{1}$$

Finally, we will assume that the Killing form is of definite sign on $\mathfrak{p}$. This last condition is automatically satisfied when G is compact and is also satisfied by irreducible symmetric Riemannian pairs (G, K) in the theory of symmetric spaces.

Let us now recall the definition of the affine-quadratic problem in this general setting ([8]).

1.1. Affine-Quadratic Problem

Any element A in $\mathfrak{p}$ generates an affine set $\Gamma = \{A + U : U \in \mathfrak{k}\}$ in $\mathfrak{g}$, and this set defines a left invariant differential system

$$\frac{dg}{dt} = g(t)(A + U(t)), g(t) \in G, \tag{2}$$

where $U(t)$ is a bounded and measurable curve in $\mathfrak{k}$. We will think of (2) as a control system with $U(t)$ playing the role of control functions. We will assume that A is regular $\mathfrak{p}$, that is, that the set of elements in $\mathfrak{p}$ that commute with A forms an abelian subalgebra in $\mathfrak{g}$. Our assumption $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}$ implies that $\mathfrak{g} = \mathfrak{g}_1 \oplus \mathfrak{g}_2 \cdots \oplus \mathfrak{g}_m$ where each factor $\mathfrak{g}_i$ is a simple ideal of the form $\mathfrak{g}_i = \mathfrak{p}_i + [\mathfrak{p}_i, \mathfrak{p}_i]$. It then follows that the projection of a regular element A on each factor $\mathfrak{g}_i$ in (4) is non-zero which, in turn, implies that (2) is controllable, in the sense that for any two points g_0, g_1 in G there is a solution $g(t)$ on an interval $[0, T]$ that satisfies $g(0) = g_0$ and $g(T) = g_1$ (see [8], page 162 for a proof). Since any two Cartan subalgebras in $\mathfrak{p}$ are Ad_K conjugate, so are the systems defined by any two regular elements A_1 and A_2.

We will now let $\frac{1}{2}\int_0^T \langle U(t), U(t) \rangle \, dt$ be the energy functional associated with any solution $g(t)$ of (2) generated by a control $U(t)$, where $\langle A, B \rangle = -Kl(A, B)$. Note that the Killing form is negative semi-definite on the Lie algebra $\mathfrak{k}$ of K when K is compact, and is strictly negative when K has a finite centre (2). Therefore, our energy functional is positive for any non-zero control $U(t)$. This energy functional is called canonical relative to a more general one $\frac{1}{2}\int_0^T \langle \mathcal{P}(U(t)), U(t) \rangle \, dt$ defined by any positive linear operator $\mathcal{P}$ on $\mathfrak{k}$.

The above data induce a natural optimal control problem: find the solutions $g(t)$ of (2) that satisfy the given boundary conditions $g(0) = g_0$, $g(T) = g_1$ for which the energy of transfer $\frac{1}{2}\int_o^T \langle \mathcal{P}(U(t)), U(t) \rangle \, dt$ is minimal. The above optimal control problem will be referred to as the affine-quadratic problem (reminiscent of linear-quadratic problems in the control theory literature). In this paper we shall be interested only in the canonical case $\mathcal{P} = I$.

As we mentioned earlier, the pair (G, K) is reductive. Any reductive semi-simple Lie algebra $\mathfrak{g}$ also carries along a "hidden" semi-direct product $\mathfrak{g}_0 = \mathfrak{p} \rtimes \mathfrak{k}$ for the following reasons. Since $[\mathfrak{p}, \mathfrak{k}] \subseteq \mathfrak{p}$, K acts linearly on $\mathfrak{p}$ by the adjoint action $h \to \mathrm{Ad}_h|_{\mathfrak{p}}, h \in K$, and induces the semi-direct product $G_0 = \mathfrak{p} \rtimes K$ with the group operation $(A_1, h_1)(A_2, h_2) = (A_1 + \mathrm{Ad}_{h_1}(A_2), h_1 h_2)$. Then the Lie algebra $\mathfrak{g}_0$ of G_0 is equal to $\mathfrak{p} \rtimes \mathfrak{k}$ with the Lie bracket given by

$$[(A_1, B_1), (A_2, B_2)] = ([B_1, A_2] - [B_2, A_1], [B_1, B_2]), (A_i, B_i) \in \mathfrak{p} \times \mathfrak{k}.$$

We will identify elements $(A, B) \in \mathfrak{p} \times \mathfrak{k}$ with the sums $A + B$ under the identification $(A, B) = (A, 0) + (0, B) = A + B$, in which case the Lie brackets in $\mathfrak{g}_0$ are identified with

$$[A_1 + B_1, A_2 + B_2] = [B_1, A_2] - [B_2, A_1] + [B_1, B_2].$$

Thus, $\mathfrak{g}$ as a vector space carries two Lie brackets:

$$[A_1 + B_1, A_2 + B_2]_s = [B_1, A_2] - [B_2, A_1] + s[A_1, A_2] + [B_1, B_2],$$

defined by a single parameter s: $s = 0$ in the semi-direct case, and $s = 1$ in the semi-simple case.

It follows that every affine space $\Gamma = \{A + U : U \in \mathfrak{k}\}$ that defines an affine left-invariant system on G also defines a corresponding left-invariant affine system on the semi-direct product G_0. Thus, behind every affine quadratic optimal problem on G there is a corresponding affine-quadratic "shadow" problem on the semi-direct product G_0.

When K is a compact group with finite centre, then the above optimal problems are well defined in the sense that for any set of boundary points g_0 and g_1 there exists an optimal trajectory that satisfies $g(0) = g_0$ and $g(T) = g_1$ for some $T > 0$.

Remarkably, the Hamiltonian associated with the shadow problem is particularly relevant in the theory of mechanical systems (see [8], Ch. 10 for the mechanical problem of Neumann on the sphere [13], Ch. 11 for Jacobi's problem on the ellipsoid, and Ch. 13 for the elastic problem and the pendulum). This phenomenon raises a natural question: what is the geometric origin behind the affine-quadratic problem that properly accounts for its relevance for the above mentioned problems? This question was partly addressed in the literature on integrable systems where the drift vector was associated with a linear potential V associated to an abstract "rigid body" with a Hamiltonian $H(g, L) = \frac{1}{2}\langle \mathcal{P}^{-1}(L), L\rangle + V(g)$ on the tangent bundle of a Lie group G ([14]) but that association raised its own questions, and at the end proved to be more enigmatic than useful.

In this paper, we will show that the Poisson systems generated by the canonical affine-quadratic problem and the rolling problem provide new and original answers to the above query: we will show that the Poisson system associated with the affine-quadratic problem is an invariant subsystem of the Poisson system generated by the rolling problem on a coadjoint orbit where the drift element A appears as a constant of motion for the rolling problem (Propositions 5 and 6).

With this goal in mind, we will now turn our attention to the quotient space G/K and the rolling problem.

1.2. Homogeneous Riemannian Manifolds

We will first need to introduce the Riemannian structure on the homogeneous manifold $M = G/K$ defined by G and K. To begin with we will regard G as a semi-Riemannian manifold (in the sense of O-Neill [15]) with the left-invariant metric $\langle\langle gX, gY\rangle\rangle_g = \langle X, Y\rangle, X, Y \in \mathfrak{g}$ induced by a scalar multiple of the Killing form $\langle\, ,\, \rangle$ that is positive definite on $\mathfrak{p}$. Such a choice is possible by our assumption. On compact Lie groups G, this multiple will be a negative multiple of the Killing form and then the above metric on G coincides with the canonical bi-invariant metric. However, on non-compact Lir groups, the Killing form is indefinite and the above metric is semi-Riemannian. Here gX is a shorthand notation for the left-invariant vector field $X(g) = d_e L_g(X)$, where L_g is the left translation $L_g(h) = gh$. The same shorthand notation applies to the right-invariant vector fields with $X(g) = Xg = d_e R_g(X), R_g h = hg$. We also recall that the Killing form is invariant under any linear automorphism of $\mathfrak{g}$ and hence the quadratic form $\langle\, ,\, \rangle$ is Ad_G invariant.

In order to make an easy passage to the techniques of optimal control, we will assume that all curves are absolutely continuous, and all differential equations involving such curves will be understood to be true only up to sets of measure zero without explicitly saying so. With that convention in mind any curve $g(t)$ in G is a solution of $\frac{dg}{dt} = g(t)U(t)$ for some bounded and measurable curve $U(t) \in \mathfrak{g}$. When $U(t)$ takes values in $\mathfrak{p}$, $g(t)$ is called horizontal, and when $U(t)$ takes values in $\mathfrak{k}\, g(t)$ is called vertical. Correspondingly, the left-invariant distributions $\mathcal{H}(g) = \{gX : X \in \mathfrak{p}\}$ and $\mathcal{V}(g) = \{gX : X \in \mathfrak{k}\}$ will be called horizontal and vertical, respectively. Thus, horizontal curves are tangent to $\mathcal{H}$ in the sense that $\frac{dg}{dt} \in \mathcal{H}(g(t))$. Likewise vertical curves are tangent to $\mathcal{V}$. It follows that

$$\mathcal{H}(g) \oplus \mathcal{V}(g) = T_g G, g \in G. \tag{3}$$

We shall assume that G/K is endowed with a manifold structure so that the natural projection $\pi(g) = gK$ is a smooth surjection (such a structure exists ([15])). Then G/K with this manifold structure will be denoted by M and o will denote the point in M such that $\pi(e) = o$, where e is the group identity in G.

A curve $g(t)$ in G is called a lift of a curve $p(t) \in M$ if $\pi(g(t)) = p(t)$. Such a lift is said to be horizontal when $g(t)$ is a horizontal curve. The projection $p(t)$ of a vertical curve $g(t)$ is a single point $\pi(g(0))$ in M because any solution of $\frac{dg}{dt} = g(t)U(t), U(t) \in \mathfrak{k}$ is of the form $g(t) = g(0)h(t), h(t) \in K$.

If $g(t)$ is any lift of a curve $p(t)$, then $\frac{dg}{dt} = g(t)U(t) = g(t)(U_{\mathfrak{p}}(t) + U_{\mathfrak{k}}(t))$ where $U_{\mathfrak{p}}(t)$ and $U_{\mathfrak{k}}(t)$ are the orthogonal projections of $U(t)$ on $\mathfrak{p}$ and $\mathfrak{k}$. Then, $d_{g(t)}\pi(g(t)U(t)) = d_{g(t)}\pi(U_{\mathfrak{p}}(t)) = \frac{dp}{dt}$. The above shows that $\tilde{g}(t)$, the solution of $\frac{d\tilde{g}(t)}{dt} = \tilde{g}(t)U_{\mathfrak{p}}(t), \tilde{g}(0) = g(0)$, is a horizontal curve that projects on $p(t)$, and secondly, it shows that $d_g\pi(gU(t)) = \frac{dp}{dt}$ for any horizontal lift $g(t)$ of $p(t)$. The isomorphism $\mathcal{H}(g) \to T_{\pi(g)}M$ can then be used to induce a metric on M

$$(d_g\pi(gV), d_g\pi(gW))_{\pi(g)} = \langle\langle gV, gW\rangle\rangle_g, V, W \in \mathfrak{p}. \tag{4}$$

Let now $\{\tau_g : g \in G\}$ denote the group of diffeomorphisms on M defined by the group action

$$\pi(L_g(h)) = \tau_g(\pi(h)), h \in G, L_g(h) = gh. \tag{5}$$

Since G acts transitively on M, M can be represented by the orbit $\{\tau_g(o) : g \in G\}$. It follows that $\pi(\exp(tU)g) = \tau_{\exp(tU)}\pi(g)$ for any $U \in \mathfrak{g}$. Note that $g \to \exp(tU)g$ is the flow generated by the right-invariant vector field $U_r(g) = Ug$. The above equality shows that the flow of U_r is π-related to the flow $\{\tau_{\exp(tU)}, t \in R\}$ in M.

In what follows, $\vec{U}$ will denote the infinitesimal generator of the flow $\{\tau_{\exp(tU)}, t \in R\}$, and $\mathcal{F}$ will denote the family of vector fields $\{\vec{U} : U \in \mathfrak{g}\}$. The correspondence $U_r(g) \to \vec{U}(\pi(g))$ is one to one and onto $T_{\pi(g)}M$. Since the Lie brackets of vector fields related by a mapping F are also F-related ([16]) , the Lie brackets $[U_r, V_r]$ are $d\pi$-related to $[\vec{U}, \vec{V}]$. Therefore the correspondence $U_r(g) \to \vec{U}(\pi(g))$ is a Lie algebra homomorphism, and hence $\mathcal{F} = \{\vec{U} : U \in \mathfrak{g}\}$ is a finite dimensional Lie algebra of vector fields that satisfies $\mathcal{F}(p) = T_pM$ for each $p \in M$. Elements of $\mathcal{F}$ are generally known as the vector fields generated by the group action.

Note that $\pi(\exp(tU)) = \tau_{e(tU)}(o) = \exp(t\vec{U})(o)$ and therefore $d_e\pi(U) = \vec{U}(o)$. Then $\pi(g) = \tau_g\pi(e)$ implies that

$$d_g(\pi(gU)) = d_o\tau_g d_e\pi(U) = d_o\tau_g\vec{U}(o). \tag{6}$$

Furthermore,

$$(\vec{U}(o), \vec{V}(o))_o = \langle\langle U_r(e), V_r(e)\rangle\rangle_e = \langle\langle U_l(e), V_l(e)\rangle\rangle_e = \langle U, V\rangle. \tag{7}$$

Hence,

$$(\vec{U}(o), \vec{V}(o))_o = \langle U, V\rangle = \langle\langle gU, gV\rangle\rangle_g =$$
$$(d_g\pi(gU), d_g\pi(gV))_{\pi(g)} = (d_o\tau_g\vec{U}(o), d_o\tau_g\vec{V}(o))_{\pi(g)}.$$

It follows that

$$(d_g\tau_g(V(p), d_g\tau_g(W(p))_{\tau_g(p)} = (V(p), W(p))_p, \tag{8}$$

for any $g \in G$ and any tangent vectors $V(p)$ and $W(p)$ in T_pM.

Therefore, $\{\tau_g : g \in G\}$ acts on M by isometries, and consequently each vector field $\vec{U}$ in $\mathcal{F}$ is a Killing vector field. Recall that the isometry group of M is a subgroup of Diff(M) that leaves the metric invariant, also recall that a vector field is a Killing vector field if its flow acts on M by isometries (the flow of $\vec{U}$ is given by $\tau_{\exp(tU)}, t \in R$). See [15] for additional details.

A homogeneous manifold $M = G/K$ defined by the above data will be referred to as semi-simple (it is defined by a semi-simple Lie group G, a compact subgroup K, and the metric induced by the Killing form). It can be shown that any symmetric Riemannian space with no Euclidean factors can be reduced to a semi-simple manifold (so that $[\mathfrak{p}, \mathfrak{p}] = \mathfrak{k}$ holds). Conversely, if G is simply connected then every semi-simple manifold is symmetric (see [17], Proposition 6.27). In any event, the present exposition makes no use of geodesic symmetry so there is no need to get distracted with the theory of symmetric spaces.

On semi-simple manifolds, parallel transport and covariant derivative are given by nice formulas inherited from G. To elaborate, note that any semi-simple Lie group G with its left-invariant metric a scalar multiple of the Killing form is a semi-Riemannian group in the terminology of O'Neill ([15], p. 305) because the Killing form is Ad_G invariant (it is only in the compact case that this semi-metric is Riemannian, i.e., equal to the canonical bi-invariant metric on G).

Relative to this left-invariant semi-metric, $\nabla_X Y = \frac{1}{2}[X, Y]$, X and Y left-invariant, is the (unique) bi-invariant affine connection that preserves the inner product and is torsion free ([15]). The associated covariant derivative $\frac{D_{g(t)}}{dt}V(t)$ of a vector field $g(t)V(t)$ defined along a curve $g(t)$ in G is given by

$$\frac{D_{g(t)}}{dt}V(t) = g(t)(\frac{dV}{dt}(t) + \frac{1}{2}[V(t), U(t)]), g^{-1}(t)\frac{dg}{dt}(t) = U(t). \tag{9}$$

Since the metric on M is the pull-back of the metric on G, the covariant derivative and parallel transport in M can be described in terms of the lifted objects in $\mathfrak{g}$ via the following formulas ([9]): any curve of tangent vectors $v(t)$ along a curve $p(t)$ in M can be represented by $v(t) = d_{g(t)}\pi(g(t)V(t))$ in terms of a unique curve $V(t) \in \mathfrak{p}$, where $g(t)$ denotes a horizontal curve in G that projects onto $p(t)$. It follows that $\frac{dg}{dt} = g(t)U(t), U(t) \in \mathfrak{p}$ and $d_{g(t)}\pi(g(t)U(t)) = \frac{dp}{dt}$. Then the covariant derivative $\frac{D_{p(t)}}{dt}v(t)$ of $v(t)$ along $p(t)$ is given by

$$\frac{D_{p(t)}}{dt}v(t) = d_{g(t)}\pi(g(t)(\frac{dV}{dt} + \frac{1}{2}[U(t), V(t)]_{\mathfrak{p}})) = d_{g(t)}\pi(g(t)\frac{dV}{dt}). \tag{10}$$

where $[U(t), V(t)]_{\mathfrak{p}}$ denotes the orthogonal projection of $[U, V]$ on $\mathfrak{p}$ (because of our assumption $[\mathfrak{p}, \mathfrak{p}] \subseteq \mathfrak{k}$, the orthogonal projection of $[U, V]$ on $\mathfrak{p}$ is zero). Hence, $v(t)$ is parallel along $p(t)$ whenever $v(t)$ is the projection of a curve $g(t)V(t)$ with $V(t)$ a constant in $\mathfrak{p}$. With this background at our disposal we will now come to the rolling problem.

1.3. The Rolling Problem

The most direct route to the rolling problem is via the intrinsic definition of rolling, introduced by R. Bryant and L. Hsu in ([18]), and later used by A. Agrachev in ([19]), Y. Chitour in ([20,21]) and Godoy Molina in ([22]). According to this definition a curve $\alpha(t)$ on a Riemannian manifold M rolls on a curve $\hat{\alpha}(t)$ on another Riemannian manifold $\widehat{M}$ if there exists an isometry $A(t) : T_{\alpha(t)}M \to T_{\hat{\alpha}(t)}\widehat{M}$ that satisfies:

$$\frac{d\hat{\alpha}}{dt} = A(t)\frac{d\alpha}{dt}, \tag{11}$$

and also satisfies the condition that $A(t)v(t)$ is a parallel vector field in $\widehat{M}$ along $\hat{\alpha}(t)$ for each parallel vector field $v(t)$ along $\alpha(t)$ in M. The triple $(\alpha(t), \hat{\alpha}(t), A(t))$ is called a rolling curve. It is clear that rolling is reflexive in the sense that if $\alpha(t)$ is rolled on $\hat{\alpha}(t)$ by an isometry $A(t)$ then $\hat{\alpha}(t)$ is rolled on $\alpha(t)$ by the isometry $A^{-1}(t)$, and therefore $(\hat{\alpha}(t), \alpha(t), A^{-1}(t))$ is also a rolling curve. We will take $\widehat{M} = T_oM$ which we regard as a Euclidean space with its metric $(u, v)_o$ defined by (4) and we address the rolling of curves in M on curves in $\widehat{M}$. Recall that in any semi-Euclidean vector space parallel transport $v(t)$ along a curve $\hat{\alpha}(t)$ in $\widehat{M}$ is done only by constant vector fields (translations).

Any curve $\alpha(t)$ in M is the projection of a horizontal curve $g(t)$, that is, $\frac{dg}{dt} = g(t)U(t), U(t) \in \mathfrak{p}$ and $\alpha(t) = \pi(g(t)) = \tau_{g(t)}(o)$. Then,

$$\frac{d\alpha(t)}{dt} = d_{g(t)}\pi(g(t)U(t)) = d_o\tau_{g(t)}\vec{U}(t)(o), \tag{12}$$

where $\vec{U}(t)$ denotes the curve of Killing vector fields in T_oM defined by $U(t)$. If we now let $\hat{\alpha}(t)$ be any solution in $\widehat{M}$ of $\frac{d\hat{\alpha}(t)}{dt} = \vec{U}(t)(o)$ and let $A(t) = d_o\tau_{g(t)}$ then $A(t)$ is an isometry that rolls $\hat{\alpha}(t)$ on $\alpha(t)$ since the parallel transport condition is fulfilled (by Equation (10)). Of course, then $A^{-1}(t)$ rolls $\alpha(t)$ on $\hat{\alpha}(t)$.

It follows that each horizontal curve $g(t)$ in G defines a family of curves $\hat{\alpha}(t)$ in $\widehat{M}$, each a solution of $\frac{d\hat{\alpha}}{dt} = \vec{U}(t)(o)$ associated with $U(t) = g^{-1}(t)\frac{dg}{dt}$, that roll on $\alpha(t) = \pi(g(t))$. Conversely, every solution $(g(t), \hat{\alpha}(t))$ of the differential system

$$\frac{dg}{dt} = g(t)U(t), \frac{d\hat{\alpha}(t)}{dt} = \vec{U}(t)(o), U(t) \in \mathfrak{p} \tag{13}$$

defines a curve $\alpha(t) = \pi(g(t))$ in M on which $\hat{\alpha}(t)$ in $\widehat{M}$ is rolled by the isometry $d_o\tau_{g(t)}$.

The rolling problem will be defined on the configuration space $\mathbf{G} = G \times \widehat{M}, \widehat{M} = T_oM$, which will be regarded as a Lie group with the group operation $\mathbf{gh} = (g, p)(h, q) = (gh, p + q)$, for all $\mathbf{g} = (g, p)$ and $\mathbf{h} = (h, q)$ in $\mathbf{G}$. Then the Lie algebra $\mathcal{G}$ of $\mathbf{G}$ will be naturally identified with $\mathfrak{g} \times T_oM$ with the Lie bracket $[(X, \vec{U}(o)), (Y, \vec{V}(o))] = ([X, Y], 0)$.

Let now $\mathcal{H}(g, p)$ denote the left invariant distribution defined by $\Gamma = \{(U, \vec{U}(o)) : U \in \mathfrak{p}\}$ that is,

$$\mathcal{H}(g, p) = \{(gU, \vec{U}(o)) : U \in \mathfrak{p}\}, (g, p) \in \mathbf{G}. \tag{14}$$

The distribution $\mathcal{H}$ will be referred to as the *rolling distribution* and its integral curves will be called *rolling motions*. Any rolling motion $\mathbf{g}(t) = (g(t), p(t))$ is a solution of

$$\frac{dg}{dt} = g(t)U(t), \frac{dp}{dt} = \vec{U}(t)(o), \tag{15}$$

and can be associated with the rolling curve $(\hat{\alpha}(t), \alpha(t)), d_o\tau_{g(t)})$, where $\alpha(t) = \tau_{g(t)}(o)$. The reader may want to show that this intrinsic definition of rolling agrees with the extrinsic descriptions [23] based on the formalism in [24].

Since Γ is a vector subspace in $\mathcal{G}$ that satisfies

$$\Gamma + [\Gamma, \Gamma] + [\Gamma, [\Gamma, \Gamma]] = \mathcal{G}, \tag{16}$$

the Lie algebra generated by the left-invariant vector fields tangent to $\mathcal{H}$ is equal to $\mathcal{G}$, and therefore, any two points in $\mathbf{G}$ can be connected by a rolling motion, and each rolling motion inherits a natural length $\int_0^T \sqrt{\langle U(t), U(t)\rangle}\, dt$ from G. To put the matter in a control theoretic context, let $A_1, \ldots, A_m$ be an orthonormal basis in $\mathfrak{p}$ so that $(A_i, \vec{A}_i(o))$ is an orthonormal basis in Γ. Then an absolutely continuous curve $\mathbf{g}(t) = (g(t), p(t))$ is a rolling motion if and only if

$$\frac{dg}{dt} = g(t)(\sum_{i=1}^{m} u_i(t)A_i), \frac{dp}{dt} = \sum_{i=1}^{m} u_i(t)\vec{A}_i(o), \tag{17}$$

for some bounded and measurable control functions $u_1(t),\ldots,u_m(t)$, in which case the length of $\mathbf{g}(t)$ is given by $\int_0^T \sqrt{u_1^2(t)+\cdots+u_m^2(t)}dt$. It then follows from (16) that the Lie algebra generated by the left-invariant vector fields $X_i(g,p)=(gA_i,\vec{A}_i(o)), i=1,\ldots,m$ is of full rank in $\mathbf{G}$. Since each left-invariant vector field X_i is complete, any pair of points in $\mathbf{G}$ can be connected by an integral curve of $\mathcal{H}$ of minimal length ([19]). An integral curve $\mathbf{g}(t)$ of $\mathcal{H}$ is called a *rolling geodesic* if for any t_0 and t_1, sufficiently close to each other, the length of $\mathbf{g}(t)$ in the interval $[t_0,t_1]$ is minimal among all other integral curves of $\mathcal{H}$ that connect $\mathbf{g}(t_0)$ to $\mathbf{g}(t_1)$.

The rolling problem consists of characterizing the rolling geodesics in $\mathbf{G}$ induced by $\mathcal{H}$. Since each rolling geodesic is also a sub-Riemannian geodesic on the configuration space $\mathbf{G}$ relative to the above length, the rolling problem can be equivalently phrased as a sub-Riemannian problem in $\mathbf{G}$ where one looks for the solutions $\mathbf{g}(t)=(g(t),p(t))$ on a fixed time interval $[0,T]$ that satisfy the given boundary conditions $\mathbf{g}(0)=\mathbf{g}_0$ and $\mathbf{g}(T)=\mathbf{g}_1$ along which the energy of transfer $\frac{1}{2}\int_0^T\sum_{i=1}^m u_i^2(t)\,dt$ is minimal.

Return now briefly to the affine-quadratic problem introduced earlier with its dynamics

$$\frac{dg}{dt}=g(t)(A+U(t)),U(t)\in\mathfrak{k} \tag{18}$$

and the energy (sometimes called the cost in the literature on optimal control) $E=\frac{1}{2}\int_0^T\langle\mathcal{P}(U(t)),U(t)\rangle\,dt$, induced by a positive definite operator $\mathcal{P}$ relative to the scalar product $\langle\,,\,\rangle$. Since $\mathcal{P}$ can be diagonalized by an orthonormal basis $B_1,\ldots,B_k$ in $\mathfrak{k}$, the affine-quadratic problem can be restated as an optimal problem over the system

$$\frac{dg}{dt}=g(t)(A+\sum_{i=1}^k u_i(t)B_i)=X_0(g)+\sum_{i=1}^k u_i(t)X_i(g), \tag{19}$$

with $X_0(g)=gA, X_i(g)=gB_i, i=1,\ldots k$, and $E=\frac{1}{2}\int_0^T\sum_{i=1}^k\lambda_i u_i^2(t)\,dt$ the energy of transfer ($\lambda_1,\ldots,\lambda_n$ are the eigenvalues of $\mathcal{P}$). In the canonical case $E=\frac{1}{2}\int_0^T\sum_{i=1}^k u_i^2(t)\,dt$.

Let us now single out some examples that are relevant for the results that follow.

1.4. Some Notable Examples

1. $G=SL(n), K=SO(n)$. In this situation we will assume that the Lie algebra $\mathfrak{g}=sl(n)$, that consists of $n\times n$ matrices having zero trace, is endowed with the scalar product $\langle X,Y\rangle=\frac{1}{2}Tr(XY)$. Then $\mathfrak{k}=so(n)$ is the Lie algebra of K, and $\mathfrak{p}=sym_0(n)$ is the space of symmetric matrices in $\mathfrak{g}$. It is easy to verify that $\langle\,,\,\rangle$ is positive on $\mathfrak{p}$ and negative on $\mathfrak{k}=so(n)$. Therefore G with its left-invariant metric induced by $\langle\,,\,\rangle$ is a semi-Riemannian manifold.

Then the quotient space $M=G/K$ will be identified with $\mathcal{P}^n$, the space of positive-definite matrices of determinant one, through the action $\tau_g(P)=gPg^T, g\in SL(n), P\in\mathcal{P}^n$, where g^T is the matrix transpose of g. Since any positive definite matrix P with $Det(P)=1$ can be written as $P=SS^T$ for some $S\in SL(n)$ the action is transitive, and $\mathcal{P}^n$ can be identified with the orbit through the identity I. Since the identity matrix I is both an element of $\mathcal{P}^n$ and the group identity in G, it is equal to the point o ($\pi(e)=o$). Horizontal curves are the solutions of $\frac{dg}{dt}=g(t)U(t), U(t)\in sym_0(n)$. Any curve $\alpha(t)$ in $\mathcal{P}^n$ is the projection of a horizontal curve $g(t)$ and the length of $\alpha(t)$ is given by $\int_0^T\sqrt{\langle U(t),U(t)\rangle}\,dt$. Killing vector fields are given by $\vec{U}(P)=UP+PU^T$, $U\in sl(n)$ and $P\in M$. The rolling distribution is given by

$$\frac{dg}{dt}=g(t)U(t),U(t)\in sym_0(n),\frac{dp}{dt}=\vec{U}(t)(o)=2U(t) \tag{20}$$

The case $n=2$ is somewhat special, for then $\mathcal{P}^2$ is isometrically diffeomorphic to the Poincaré upper half plane $\mathcal{P}=\{z+iy:y>0\}$ with its metric $\frac{1}{y}\sqrt{\dot{x}^2+\dot{y}^2}$. To elaborate, note that every $g \in SL(2)$ can be written as $g = PR$ where P is upper triangular and R is a rotation matrix. In fact, if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is an element of $SL(2)$ then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \frac{1}{\sqrt{c^2+d^2}} \begin{pmatrix} 1 & ac+bd \\ 0 & c^2+d^2 \end{pmatrix} \frac{1}{\sqrt{c^2+d^2}} \begin{pmatrix} d & -c \\ c & d \end{pmatrix}. \tag{21}$$

Let now

$$F(x+iy) = gg^T = PP^T = \frac{1}{y} \begin{pmatrix} x^2+y^2 & x \\ x & 1 \end{pmatrix}, \tag{22}$$

where $P = \begin{pmatrix} \frac{y}{\sqrt{y}} & \frac{x}{\sqrt{y}} \\ 0 & \frac{1}{\sqrt{y}} \end{pmatrix}$. We will now show that F is an isometry from $\mathcal{P}$ with its Poincaré hyperbolic metric onto $\mathcal{P}^2$ with its G-invariant metric. If $\tilde{\alpha}(t) = F(\alpha(t)$ then

$$\dot{\tilde{\alpha}}(t) = \dot{P}P^T + P\dot{P}^T = P(P^{-1}\dot{P} + \dot{P}^T(P^{-1})^T)P^T,$$

and therefore, $||\dot{\tilde{\alpha}}(t)|| = ||P^{-1}\dot{P} + \dot{P}^T(P^{-1})^T||$. If $Y = \frac{y}{\sqrt{y}}$ and $X = \frac{x}{\sqrt{y}}$, then an easy calculation shows that

$$P^{-1}\dot{P} + \dot{P}^T(P^T)^{-1} = \begin{pmatrix} 2\frac{\dot{Y}}{Y} & \frac{X\dot{Y}+\dot{X}Y}{Y^2} \\ \frac{X\dot{Y}+\dot{X}Y}{Y^2} & -2\frac{\dot{Y}}{Y} \end{pmatrix} = \frac{1}{y} \begin{pmatrix} \dot{y} & \dot{x} \\ \dot{x} & -\dot{y} \end{pmatrix},$$

and hence $||\dot{\tilde{\alpha}}(t)|| = \frac{1}{y}\sqrt{\dot{x}^2+\dot{y}^2}$. It follows that $||\dot{\alpha}(t)|| = ||\dot{\tilde{\alpha}}(t)||$ and therefore F is an isometry.

It then follows that the rolling distribution has its isometric analogue on $\mathcal{P}$ rolling on the tangent space at i. In this scenario $SL(2)$ acts on $\mathcal{P}$ via the Moebius transformations $\tau_g(z) = \frac{az+b}{cz+d}, g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and $\mathcal{P}$ is represented by the orbit $\{\tau_g(i) : g \in SL(2)\}$. Horizontal curves are the solutions of $\frac{dg}{dt} = g(t)\begin{pmatrix} u_1(t) & u_2(t) \\ u_2(t) & -u_1(t) \end{pmatrix}$ and their projections on $\mathcal{P}$ are given by $z(t) = g(t)(i)$. Then

$$\frac{dz(t)}{dt}|_{t=0} = \frac{d}{dt}g(t)(i)|_{t=0} = \frac{d}{dt}\frac{1}{c^2+d^2}(bd+ac+i)|_{t=0} = 2i(u_1 - iu_2).$$

Therefore, rolling motions are the solutions of

$$\frac{dg}{dt} = g(t)\begin{pmatrix} u_1(t) & u_2(t) \\ u_2(t) & -u_1(t) \end{pmatrix}, \frac{dw}{dt} = 2i(u_1(t) - iu_2(t)). \tag{23}$$

2. $G = SO_\epsilon(n+1), K = \{1\} \times SO(n), \epsilon = \pm 1$, where $SO_\epsilon(1,n)$ denotes the connected component of $SO(1,n)$ that contains the group identity when $\epsilon = -1$, and $SO_\epsilon(n+1) = SO(n+1)$, when $\epsilon = 1$. Both cases can be treated in a uniform manner as follows.

Let V_ϵ denote R^{n+1} with the scalar product $(x,y)_\epsilon = x_0y_0 + \epsilon\sum_{i=1}^n x_iy_i$. Each $SO_\epsilon(n+1)$ acts on V_ϵ by matrix multiplications, and each group is defined as the matrix group whose elements have a positive determinant and preserve the bilinear form $(\,,\,)_\epsilon$. It follows that each $g \in SO_\epsilon(n+1)$ satisfies $g^TDg = D$ where D is a diagonal matrix with its diagonal entries equal to $(1,\epsilon,\dots,\epsilon)$. Therefore, $Det(g^T)Det(g)Det(D) = Det(D)$ which implies that $Det(g) = 1$. This shows that each of $SO_\epsilon(n+1)$ is a subgroup of $SL(n+1)$.

We will let S^n_ϵ denote the Euclidean unit sphere when $\epsilon = 1$ and the hyperboloid $\{x \in R^{n+1} : x_0^2 = 1 + \sum_{i=1}^n x_i^2, x_0 > 0\}$ when $\epsilon = -1$. In each case, $SO_\epsilon(n+1)$ acts on S^n_ϵ by the left matrix multiplications on the points of S^n_ϵ written as column vectors. It can be shown that this action is transitive. When S^n_ϵ is represented by the orbit through e_0 then the isotropy group $K = \{g \in SO_\epsilon(n+1) : ge_0 = e_0\}$ is equal to $\{1\} \times SO(n)$. Therefore,

$$S^n_\epsilon = SO_\epsilon(n+1)/K.$$

with the natural projection π given by $\pi(g) = ge_0 = \tau_g(e_0)$.

We will regard $G = SO_\epsilon(n+1)$ as a semi-Riemannian subgroup of $SL(n+1)$ with its left-invariant metric introduced through the bilinear form $\langle X, Y\rangle_\epsilon = -\frac{\epsilon}{2}Tr(XY)$ (this metric is indefinite on $\mathfrak{g}_\epsilon$ when $\epsilon = -1$ and is positive when $\epsilon = 1$).

The following notations will be useful in describing the Cartan factors $\mathfrak{k}_\epsilon$ and $\mathfrak{p}_\epsilon$. If a and b are any points in R^{n+1} then $a \otimes_\epsilon b$ will denote the matrix defined by $(a \otimes_\epsilon b)x = (a, x)_\epsilon b, x \in \mathbb{R}^{n+1}$, and then $a \wedge_\epsilon b$ will denote the matrix $a \otimes_\epsilon b - b \otimes_\epsilon a$. Since $((a \wedge_\epsilon b)x, y)_\epsilon + (x,(a \wedge_\epsilon b)y)_\epsilon = 0$, $a \wedge_\epsilon b$ belongs to $\mathfrak{so}_\epsilon(n+1)$ for any a, b in $\mathbb{R}^{n+1}$.

It is easy to show that the Lie algebra $\mathfrak{k}$ of K and its orthogonal complement $\mathfrak{p}_\epsilon$ are given by the following expressions:

$$\mathfrak{p}_\epsilon = \{U = u \wedge_\epsilon e_0 : (u, e_0)_\epsilon = 0\}, \tag{24}$$

$$\mathfrak{k} = \{V = v \wedge_\epsilon w : (v, e_0)_\epsilon = (w, e_0)_\epsilon = 0\}, \tag{25}$$

The preceding matrices can be also written as

$$U = \begin{pmatrix} 0 & -\epsilon u^* \\ u & 0 \end{pmatrix}, V = \begin{pmatrix} 0 & 0 \\ 0 & v \wedge w \end{pmatrix}, u, v, w \text{ in } \mathbb{R}^n.$$

Horizontal curves are the solutions of

$$\frac{dg}{dt} = g(t)U(t), U(t) = u(t) \wedge_\epsilon e_o, u(t) \perp e_0,$$

that satisfy

$$\langle\langle g(t)U(t), g(t)U(t)\rangle\rangle_\epsilon = \langle U(t), U(t)\rangle_\epsilon = \sum_{i=1}^n u_i^2(t).$$

Then $||\dot\alpha(t)||^2_\epsilon = \epsilon(\dot\alpha, \dot\alpha)_\epsilon = \epsilon\dot\alpha_0^2 + \sum_{i=1}^n \dot\alpha_i^2$ is the natural metric on S^n_ϵ. We then have

$$||\dot\alpha(t)||^2_\epsilon = \epsilon(\dot\alpha(t), \dot\alpha(t))_\epsilon = \epsilon(g(t)u(t), g(t)u(t))_\epsilon = \sum_{i=1}^n u_i^2(t) = \langle U(t), U(t)\rangle_\epsilon,$$

hence the metric is $SO_\epsilon(n+1)$ invariant, and S^n_ϵ with this metric is a semi-simple homogeneous manifold. It follows that the rolling distribution is given by

$$\frac{dg}{dt} = g(t)u(t) \wedge_\epsilon e_0), \frac{dp}{dt} = (u(t) \wedge_\epsilon e_o)e_0 = u(t), \tag{26}$$

which agrees with 2.4 in ([6]).

2. Symplectic Background, Hamiltonian Systems

Let us now turn our attention to the extremal curves associated with our main problems. Because of the constraints present in these problems, the Maximum Principle of optimal control, rooted in the Hamiltonian formalism, is the only tool available for arriving to the appropriate extremal equations. However, in order to make an effective use of the Maximum Principle we will need to work with the symplectic form in a special system of coordinates that is well adapted for left-invariant optimal control problems (described in [3,8]) which calls for a brief review of symplectic geometry. Below is a brief summary of the symplectic material required for the main results.

Recall that a manifold M endowed with a non-degenerate and closed 2-form ω is called *symplectic*. The symplectic form induces a correspondence between functions and vector fields: every function f corresponds to a vector field $\vec{f}$ defined by $\omega(\vec{f}, X) = df(X)$. In this context, $\vec{f}$ is called the Hamiltonian vector field generated by f. Every symplectic manifold is even dimensional, and at each point of M there exists a neighbourhood with coordinates $(x_1, \ldots, x_n, p_1, \ldots, p_n)$ such that the Hamiltonian vector fields are given by

$$\vec{f} = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i} \frac{\partial}{\partial x_i} - \frac{\partial f}{\partial x_i} \frac{\partial}{\partial p_i}. \tag{27}$$

This choice of coordinates in which $\vec{f}$ is represented by (27) is called *symplectic*.

Any cotangent bundle T^*M is a symplectic manifold endowed with its canonical symplectic form, usually written as $\omega = dp \wedge dx$ relative to a choice of symplectic coordinates $\sum_{i=1}^{n} p_i dx_i$. As a symplectic manifold the cotangent bundle is somewhat special, it is a vector bundle at the same time. For that reason every vector field X on M can be lifted to a unique Hamiltonian vector field $\vec{f}_X$ in T^*M via the function $f_X(\xi) = \xi(X(x))$, $\xi \in T_x^*M$. Vector field $\vec{f}_X$ is called the *Hamiltonian lift of* X. The same procedure is applicable to any time varying vector field, and by extension to any differential system on M. Thus, any differential system in M can be lifted to a Hamiltonian system in T^*M. Then the Maximum Principle singles out the appropriate Hamiltonian lifts that govern the optimal solutions ([8]).

When the base manifold is a Lie group G, and when the underlying differential system is either left or right invariant, then there is privileged system of coordinates based on the realization of T^*G as $G \times \mathfrak{g}^*$, with $\mathfrak{g}^*$ the dual of $\mathfrak{g}$, that preserves the left (or right) invariant symmetries and elucidates the conservation laws of the associated Hamiltonian system. The passage to these coordinates is explained below.

2.1. Left-Invariant Trivializations and the Symplectic Form

Having in mind applications involving left-invariant variational systems, the cotangent bundle T^*G and the tangent bundle TG will be represented as $G \times \mathfrak{g}^*$ and $G \times \mathfrak{g}$ via the left-translations. That is, tangent vectors $v \in T_gG$ will be identified with the pairs $(g, V) \in G \times \mathfrak{g}$ via the relation $v = dL_g V$. Similarly, linear functions $\xi \in T_g^*G$ will be identified with pairs $(g, \ell) \in G \times \mathfrak{g}^*$ via $\xi = dL_g^{-1^*}\ell$, i.e., $\xi(v) = \xi(dL_g V) = \ell(V)$. Then $T(T^*G)$ is naturally identified with the product $(G \times \mathfrak{g}) \times (\mathfrak{g}^* \times \mathfrak{g}^*) \cong (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*)$, with the understanding that an element $((g, \ell), (A, a)) \in (G \times \mathfrak{g}^*) \times (\mathfrak{g} \times \mathfrak{g}^*)$ denotes the tangent vector (A, a) at the base point (g, ℓ).

Note that $G \times \mathfrak{g}^*$ is a Lie group in its own right since $\mathfrak{g}^*$ is an abelian Lie group with the group multiplication given by the vector addition. Then left-invariant vector fields in $G \times \mathfrak{g}^*$ are the left-translates of the pairs (A, a) in the Lie algebra $\mathfrak{g} \times \mathfrak{g}^*$ of $G \times \mathfrak{g}^*$. In this formalism the flow associated with the left-invariant vector field (gA, a) in $G \times \mathfrak{g}^*$ is given by $(g \exp(tA), \ell + ta)$. In terms of left-invariant vector fields $V_1 = (A_1, a_1)$ and $V_2 = (A_2, a_2)$, the canonical symplectic form on T^*G is given by the following formula:

$$\omega_{(g,\ell)}(V_1, V_2) = a_2(A_1) - a_1(A_2) - \ell([A_1, A_2]) \tag{28}$$

The above differential form is invariant under left-translations in $G \times \mathfrak{g}^*$, and is particularly revealing for the Hamiltonian vector fields generated by left-invariant functions on $G \times \mathfrak{g}^*$, that is, functions that satisfy $H(hg, \ell) = H(g, \ell) = H(e, \ell)$ for all $g, h \in G$ and all $\ell \in \mathfrak{g}^*$. Evidently, left-invariant functions on $G \times \mathfrak{g}^*$ are in exact correspondence with functions in $C^\infty(\mathfrak{g}^*)$.

Each left-invariant vector field $X(g) = dL_g X$, $X \in \mathfrak{g}$, lifts to a linear function h_X on $\mathfrak{g}^*$ because

$$h_X(\xi) = \xi(X(g)) = \xi \circ (dL_g)(X) = \ell(X), \xi \in T_g^*G$$

and each function H on $\mathfrak{g}^*$ generates a Hamiltonian vector field $\vec{H}$ on $G \times \mathfrak{g}^*$ whose integral curves are the solutions of

$$\frac{dg}{dt}(t) = g(t)dH_{\ell(t)}, \quad \frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_{\ell(t)}(\ell(t)). \tag{29}$$

Equation (29) can be easily verified by the following argument: when H is a function on $\mathfrak{g}^*$, then its differential at a point ℓ is a linear function on $\mathfrak{g}^*$, hence an element of $\mathfrak{g}$, because $\mathfrak{g}^*$ is a finite dimensional vector space. If $\vec{H}_{(g,\ell)} = (A(g,\ell), a(g,\ell))$ for some vectors $A(g,\ell) \in \mathfrak{g}$ and $a(g,\ell) \in \mathfrak{g}^*$, then

$$b(dH_\ell) = b(A) - a(B) - \ell[A,B],$$

must hold for any tangent vector (B,b) at (g,ℓ). This implies that $A(g,\ell) = dH_\ell$, and $a = -\mathrm{ad}^* dH_\ell(\ell)$, where $(\mathrm{ad}^* A)(B)(\ell) = \ell[A,B]$ for all $B \in \mathfrak{g}$. Hence, (29) holds.

In a more general case where H is a function of both g and ℓ, the equations for $\vec{H}$ are given by

$$\frac{dg}{dt}(t) = g(t)dH_{\ell(t)}, \quad \frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_{\ell(t)}(\ell(t)) - dH_g \circ dL_g, \tag{30}$$

as can be easily verified through the relations

$$b(dH_\ell) + dH_g \circ dL_g B = b(A) - a(B) - \ell[A,B].$$

This situation typically occurs in problems of mechanics in the presence of potential functions. For instance, the motion of a three-dimensional rigid body with a potential function $V : SO(3) \to R$ is described by the Hamiltonian

$$H(R,\ell) = H_0(\ell) + V(\alpha_1, \alpha_2, \alpha_3)$$

where $\alpha_1, \alpha_2, \alpha_3$ denote the columns of the matrix transpose R^T of the rotation R in $SO(3)$. If $R(t) = Re^{tX}$ is a curve in $SO(3)$ defined by an element $X \in \mathfrak{so}(3)$, then $\alpha_i(t) = R(t)^T e_i = e^{-tX} R^T e_i = e^{-tX}\alpha_i$. Therefore,

$$dV(RX) = \sum_{i=1}^{3} (\frac{\partial V}{\partial \alpha_i}, \frac{d\alpha_i}{dt})|_{t=0} = \sum_{i=1}^{3} (\frac{\partial V}{\partial \alpha_i}, -X\alpha_i) = \sum_{i=1}^{3} \langle \frac{\partial V}{\partial \alpha_i} \wedge \alpha_i, X \rangle$$

where $\langle\, , \,\rangle$ is the standard inner product $-\frac{1}{2}Tr(XY)$ in $\mathfrak{so}(3)$. Thus, $dH_g \circ dL_g = \sum_{i=1}^{3} \frac{\partial V}{\partial \alpha_i} \wedge \alpha_i$ is the external torque exerted by V. The corresponding equations of motion are given by

$$\frac{dg}{dt}(t) = g(t)dH_0(\ell(t)), \quad \frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_0(\ell(t))(\ell(t)) + \sum_{i=1}^{3} \alpha_i \wedge \frac{\partial V}{\partial \alpha_i}. \tag{31}$$

These equations extend to an "n-dimensional rigid body" $H(R,\ell) = H_0(\ell) + V(\alpha_1, \dots, \alpha_n))$ with the external torque $\sum_{i=1}^{n} \alpha_i \wedge \frac{\partial V}{\partial \alpha_i}$. This system of equations is usually written on the tangent bundle of $SO(n)$, represented as the product $SO(n) \times so(n)$, as

$$\begin{gathered} \tfrac{dR}{dt} = R(t)\Omega(t), \tfrac{dM}{dt} = [\Omega(t), M(t)] + \textstyle\sum_{i=1}^{n} \alpha_i \wedge \tfrac{\partial V}{\partial \alpha_i} \\ \mathcal{P}(\Omega(t)) = M(t), \alpha_i(t) = R^T(t)e_i, i = 1, \dots, n. \end{gathered} \tag{32}$$

In this context, $M(t)$ is the generalization of the angular momentum, $\Omega(t)$ is the generalization of the angular velocity, and $\mathcal{P}$ is the generalization of the inertia tensor.

2.2. Poisson Manifolds, Coadjoint Orbits

We will now address the Poisson structure on $\mathfrak{g}^*$ inherited from the symplectic form ω given by (28). Recall that a manifold M together with a bilinear, skew-symmetric form

$$\{\,,\,\} : C^\infty(M) \times C^\infty(M) \to C^\infty(M)$$

that satisfies

$$\{fg,h\} = f\{g,h\} + g\{f,h\}, \text{(Leibniz's rule), and}$$
$$\{f,\{g,h\}\} + \{h,\{f,g\}\} + \{g,\{h,f\}\} = 0, \text{ (Jacobi's identity),}$$

for all functions f, g, h on M, is called a *Poisson manifold*.

Every symplectic manifold is a Poisson manifold with the Poisson bracket defined by $\{f,g\}(p) = \omega_p(\vec{f}(p), \vec{g}(p)), p \in M$. However, a Poisson manifold need not be symplectic, because it may happen that the Poisson bracket is degenerate at some points of M. Nevertheless, each function f on M induces a Poisson vector field $\vec{f}$ through the formula $\vec{f}(g) = \{f,g\}$. It is known that every Poisson manifold is foliated by the orbits of its family of Poisson vector fields, and that each orbit is a symplectic submanifold of M with its symplectic form $\omega_p(\vec{f},\vec{h}) = \{f,h\}(p)$. (This foliation is known as a *the symplectic foliation of* M).

Proposition 1. *The dual $\mathfrak{g}^*$ of a Lie algebra $\mathfrak{g}$ is a Poisson manifold with the Poisson bracket*

$$\{f,h\}(\ell) = \ell([dh,df], f, h \textit{ in } C^\infty(\mathfrak{g}^*).$$

Proof. Functions on $\mathfrak{g}^*$ coincide with the left-invariant functions on $G \times \mathfrak{g}^*$. Hence,

$$\omega_{(g,\ell)}(\vec{f},\vec{h}) = \omega_{(g,\ell)}((df,0),(dh,0)) = -\mathrm{ad}^*([df,dh])(\ell) = \ell([dh,df]).$$

It follows that the Poisson bracket on $\mathfrak{g}^*$ is the restriction of the canonical Poisson bracket on $G \times \mathfrak{g}^*$ to the left-invariant functions. As such it automatically satisfies the properties of a Poisson manifold. □

In the literature on integrable systems, Poisson bracket $\{f,h\}(\ell) = \ell([df,dh])$ is often referred as *the Lie-Poisson bracket* ([14]). We have taken its negative so that Poisson vector fields agree with the projections of the Hamiltonian vector fields generated by left-invariant functions (and also agree with the sign convention in [7,8]).

It follows that each function H on $\mathfrak{g}^*$ defines a Poisson vector field $\vec{H}$ on $\mathfrak{g}^*$ through the formula $\vec{H}(f)(\ell) = \{H,f\}(\ell) = -\ell([dH_\ell, df]) = -ad^* dH_l(df)$. The integral curves of $\vec{H}$ are the solutions of

$$\frac{d\ell}{dt}(t) = -\mathrm{ad}^* dH_{\ell(t)}(\ell(t)) \tag{33}$$

That is, each function H on $\mathfrak{g}^*$ may be considered both as a Hamiltonian on T^*G, as well as a function on the Poisson space $\mathfrak{g}^*$. It follows that the Poisson equations of the associated Poisson field are the projections of the Hamiltonian Equation (29) on $\mathfrak{g}^*$.

Solutions of Equation (33) are intimately linked with the coadjoint orbits of G. We recall that the coadjoint orbit of G through a point $\ell \in \mathfrak{g}^*$ is given by $\mathrm{Ad}^*_g(\ell) = \{\ell \circ \mathrm{Ad}_{g^{-1}},\ g \in G\}$.

The following proposition is a paraphrase of A.A. Kirillov' fundamental contributions to the Poisson structure of $\mathfrak{g}^*$ ([25]).

Proposition 2. *Let $\mathcal{F}$ denote the family of Poisson vector fields on $\mathfrak{g}^*$ and let $M = \mathcal{O}_\mathcal{F}(\ell_0)$ denote the orbit of $\mathcal{F}$ through a point $\ell_0 \in \mathfrak{g}^*$. Then M is equal to the connected component of the coadjoint orbit of G that contains ℓ_0. Consequently each coadjoint orbit is a symplectic submanifold of $\mathfrak{g}^*$.*

The fact that the Poisson equations evolve on coadjoint orbits implies useful reductions in the theory of Hamiltonian systems with symmetries. Our main results will make use of this fact.

2.3. Representation of Coadjoint Orbits on Lie Algebras- Semi-Simple vs. Semi-Direct

On semi-simple Lie groups, the Killing form, or any scalar multiple of it $\langle\,,\,\rangle$, is non-degenerate, and can be used to identify linear functions ℓ on $\mathfrak{g}$ with points $L \in \mathfrak{g}$ via the formula $\langle L, X\rangle = \ell(X)$, $X \in \mathfrak{g}$. Then Poisson Equation (33) can be expressed dually on $\mathfrak{g}$ as

$$\frac{dL}{dt} = [dH, L]. \tag{34}$$

The argument is simple:

$$\langle \frac{dL}{dt}, X\rangle = \frac{d\ell}{dt}(X) = -\ell([dH, X]) = \langle L, [X, dH]\rangle = \langle [dH, L], X\rangle.$$

Since X is arbitrary, Equation (34) follows.

Under the above identification coadjoint orbits are identified with the adjoint orbits $\mathcal{O}(L_0) = \{gL_0g^{-1} : g \in G\}$, and the Poisson vector fields $\vec{f}_X(\ell) = -\mathrm{ad}^*X(\ell)$ are identified with vector fields $\vec{X}(L) = [X, L]$. Each vector field $[X, L]$ is tangent to $\mathcal{O}(L_0)$ at L, and $\omega_L([X, L], [Y, L]) = \langle L, [Y, X]\rangle$, X, Y in $\mathfrak{g}$ is the symplectic form on each orbit $\mathcal{O}(L_0)$.

In a reductive semi-simple Lie group G with a subgroup K there is also the semi-direct product $G_0 = \mathfrak{p} \rtimes K$, described earlier in the introduction. Then Poisson equations on $\mathfrak{g}_0^* = (\mathfrak{p} \rtimes \mathfrak{k})^*$ can be also represented on $\mathfrak{g}_0$ via the quadratic form $\langle\,,\,\rangle$ as in the semi-simple case, but the resulting expression takes on a slightly different form. To see the difference, let $dH = dH_\mathfrak{p} + dH_\mathfrak{k}$ and $L = L_\mathfrak{p} + L_\mathfrak{k}$ denote the decompositions of dH and L onto the factors $\mathfrak{p}$ and $\mathfrak{k}$. On the semi-direct product,

$$\begin{gathered}\langle \tfrac{dL_\mathfrak{p}}{dt}, X_\mathfrak{p}\rangle + \langle \tfrac{dL_\mathfrak{k}}{dt}, X_\mathfrak{k}\rangle = \langle \tfrac{dL}{dt}, X\rangle = \tfrac{d\ell}{dt}(X) = -\ell([dH, X]) = \\ -\langle L, [dH, X]\rangle = -\langle L, [dH_\mathfrak{p}, X_\mathfrak{k}] + [dH_\mathfrak{k}, X_\mathfrak{p}] + [dH_\mathfrak{k}, X_\mathfrak{k}]\rangle \\ = -\langle L_\mathfrak{p}, [dH_\mathfrak{p}, X_\mathfrak{k}] + [dH_\mathfrak{p}, X_\mathfrak{k}]\rangle - \langle L_\mathfrak{k}, [dH_\mathfrak{k}, X_\mathfrak{k}] = \\ \langle X_\mathfrak{k}, [dH_\mathfrak{k}, L_\mathfrak{k}] + [dH_\mathfrak{p}, L_\mathfrak{p}]\rangle + \langle X_\mathfrak{p}, [dH_\mathfrak{k}, L_\mathfrak{p}]\rangle.\end{gathered}$$

Hence, the Poisson equations are given by

$$\frac{dL_\mathfrak{k}}{dt} = [dH_\mathfrak{k}, L_\mathfrak{k}] + [dH_\mathfrak{p}, L_\mathfrak{p}], \frac{dL_\mathfrak{p}}{dt} = [dH_\mathfrak{k}, L_\mathfrak{p}]. \tag{35}$$

This equation can be combined with the equations for the semi-simple case in terms of the parameter s with

$$\frac{dL_\mathfrak{k}}{dt} = [dH_\mathfrak{k}, L_\mathfrak{k}] + [dH_\mathfrak{p}, L_\mathfrak{p}], \frac{dL_\mathfrak{p}}{dt} = [dH_\mathfrak{k}, L_\mathfrak{p}] + s[dH_\mathfrak{k}, L_\mathfrak{p}], s = 0, 1. \tag{36}$$

One can show that the coadjoint orbit through $P_0 \in \mathfrak{p}, Q_0 \in \mathfrak{k}$ under the action of $G_0 = \mathfrak{p} \rtimes K$ consists of pairs (P, Q)

$$P = Ad_h(P_0), Q = [Ad_h(P_0), X] + Ad_h(Q_0), (X, h) \in G_0, \tag{37}$$

when $\ell_0 \in \mathfrak{g}_s^*$ is identified with $L_0 = P_0 + Q_0$ in $\mathfrak{g}_0$, and when $\ell = \mathrm{Ad}^*_{(X,h)}(\ell_0)$ is identified with $L = P + Q$ ([8]).

The adjoint orbits of a non-compact semi-simple Lie group G are often symplectomorphic with the cotangent bundles of manifolds ([26]). It appears that the same is true for coadjoint orbits under the action of semi-direct products. We will now single out two such situations which are relevant for the connections to mechanical tops.

Return now to $G = SO_\epsilon(n+1)$ and $K = \{1\} \times SO(n)$ introduced in Example 2.

Proposition 3. *The coadjoint orbit $\mathcal{O}(P_0)$ through $P_0 = p_0 \wedge_\epsilon e_0, (p_0, e_0)_\epsilon = 0, Q_0 = 0$ under the action of the semi-direct product $\mathfrak{p}_\epsilon \rtimes K$ is diffeomorphic to the tangent bundle of the connected component of the "sphere" $S^n_\epsilon = \{p \in \mathbb{R}^{n+1} : (p,p)_\epsilon = (p_0,p_0)_\epsilon\}$ that contains p_0.*

Proof. Let $h \in K$, and $X = x \wedge e_o, (x, e_0)_\epsilon = 0$. Then

$$P = Ad_h(P_0) = h(p_0) \wedge_\epsilon h(e_0) = p \wedge_\epsilon e_0, p = h(p_0)$$
$$Q = [Ad_h(P_0), X] = [p \wedge_\epsilon e_0, x \wedge_\epsilon e_0] = p \wedge_\epsilon x = p \wedge_\epsilon x_p^\perp,$$

where $x_p^\perp$ is the projection of x on the orthogonal complement of p in $\mathbb{R}^{n+1}$. Therefore,

$$(p, x_p^\perp) \Rightarrow p \wedge_\epsilon e_0 + p \wedge x_p^\perp$$

is the desired diffeomorphism from the tangent bundle of the connected sphere S^n_ϵ onto the coadjoint orbit $\{Ad_h(P_0) + [Ad_h(P_0), X], (X, h) \in \mathfrak{p}_\epsilon \rtimes K\}$. □

The above diffeomorphism is actually a symplectomorphism from the cotangent bundle of either the Euclidean sphere S^n when $\epsilon = 1$, or the hyperboloid of one sheet when $\epsilon = -1$, to the appropriate coadjoint orbit, but we will not go into these details. ([8]).

We will now turn our attention to the reductive pair $G = SL(n), K = SO(n)$ (Example 1) and the coadjoint orbit through a symmetric matrix P_0 with distinctive non-zero eigenvalues $\alpha_1, \ldots, \alpha_k$ under the action of $G_o = \mathfrak{p} \rtimes SO(n)$. We recall that $sl(n) = so(n) \oplus \mathfrak{p}$ where $\mathfrak{p}$ is the space of symmetric $n \times n$ matrices of trace zero. Every symmetric $n \times n$ matrix S can be written as $S = S_0 + \frac{Tr(S)}{n} I$, $S_0 \in \mathfrak{p}$. An easy inspection of (37) shows that the orbit through S differs by a constant factor $\frac{Tr(S)}{n} I$ from the orbit through S_0. So the zero-trace requirement is inessential for the structure of coadjoint orbits.

Proposition 4. *The coadjoint orbit through P_0 given by*

$$P = Ad_h(P_0), Q = [Ad_h(P_0), X], (X, h) \in \mathfrak{p} \rtimes SO(n)$$

is diffeomorphic to the tangent bundle of the flag manifold $\mathbb{F}(1, 2, \ldots, k)$ consisting of subspaces $V_1 \subset V_2 \cdots \subset V_k$ with $dim V_i = i$.

Sketch of the proof: Let P_0 denote a symmetric matrix with distinct non-zero eigenvalues $\alpha_1 < \alpha_2 \cdots < \alpha_k$. Then P_0 can be identified with a point $(V_1 \subset V_2 \cdots \subset V_k)$ in $\mathbb{F}(1, \ldots, k)$ where each subspace V_i is equal to the linear span of unit eigenvectors $a_1, \ldots, a_i$ of P_0. If P_0 is represented by the matrix $\sum_{i=1}^k \alpha_i (a_i \otimes a_i)$, then $Ad_h(P_0)$ is represented by the matrix $\sum_{i=1}^k \alpha_i (h(a_i) \otimes h(a_i))$ that corresponds to the point $F_h = (hV_1 \subset hV_2 \cdots \subset hV_k)$ in $\mathbb{F}(1, \ldots, k)$. The correspondence $Ad_h(P_0) \to F_h$ is a diffeomorphism from the orbit $\{Ad_h(P_0), h \in SO(n)\}$ onto $\mathbb{F}(1, \ldots, k)$.

Let now St^n_k denote the Stiefel manifold of k-orthonormal frames $[a_1, \ldots, a_k]$ in $\mathbb{R}^n$. Points of St^n_k can be represented by $n \times k$ matrices M with columns $a_1, \ldots, a_k$ that satisfy $M^T M = I_k$, where M^T denotes the matrix transpose of M, and where I_k is the k-dimensional identity matrix. Let $\phi : St^n_k \to \mathbb{F}(1, \ldots, k)$ be the embedding

$$M = [a_1, \ldots, a_k] \to F_M = (V_1 \subset V_2 \cdots \subset V_k), V_i = < a_1, \ldots, a_i > .$$

Then $\phi^{-1}(F_M) = MD$, where D is a diagonal $k \times k$ matrix with its diagonal entries equal to ± 1. Therefore, $\mathbb{F}(1, \ldots, k)$ is a covering space for St^n_k, and hence $\mathbb{F}(1, \ldots, k)$ and St^n_k are locally diffeomorphic, that is, every point $M \in St^n_k$ admits an open neighbourhood U such that the restriction of ϕ to U is a diffeomorphism onto $\phi(U)$. It follows that tangent vectors at a point M can be identified with $n \times k$ matrices $\dot{M}$ that satisfy $\dot{M}^T M + M^T \dot{M} = 0$.

Let now U be an open set in St_k^n such that ϕ restricted to U is a diffeomorphism onto $\phi(U)$. For every $F_h \in \phi(U)$, $Ad_h(P_0)$ is identified with $M = [m_1, \ldots, m_k], m_i = h(a_i), i = 1, \ldots, k$. Then

$$Q = [\mathrm{Ad}_h(P), X] = \sum_{i=1}^{k} [\alpha_i(m_i \otimes m_i), X] = \sum_{i=1}^{k} y_i \wedge m_i,$$

with $y_i = \alpha_i X(m_i)$. Since X is symmetric, $\alpha_j(y_i, m_j) = \alpha_i(m_i, y_j)$. Moreover, y_i could be replaced by its orthogonal projection on $m_i^{\perp}$ without altering the value of Q. So we may assume that $(y_i, m_i) = 0, i = 1, \ldots, k$.

It follows that $M^T Q M + M^T Q^T M = 0$, hence $\dot{M} = Q^T M$ is a tangent vector at M. The pairs $(M, Q^T M)$ are parametrized by the entries of M and the entries of the matrix Y. The columns $y_i = \alpha_i X m_i$ of Y satisfy $k(k+1)$ constraints $\alpha_j(y_i, m_j) = \alpha_i(y_j, m_i), i \neq j$, and $(y_i, m_i) = 0$. This implies that the manifold of pairs of $n \times k$ matrices (M, Y) subject to the constraints

$$(m_i, m_j) = \delta_{ij}, \alpha_j(y_i, m_j) = \alpha_i(y_j, m_i), i \neq j, (y_i, m_i) = 0, \tag{38}$$

is of the same dimension as the tangent bundle of St_k^n. Therefore, the correspondence $\sum_{i=1}^{k} \alpha_i(m_i \otimes m_i), \sum_{i=1}^{k} y_i \wedge m_i \to (M, Q^T M)$ is one to one and onto the sub-bundle TU over U.

Corollary 1. *If P_0 is the orthogonal projection on a k-dimensional vector space, i.e., if $P_0 = \sum_{i=1}^{k} a_i \otimes a_i$, for some orthonormal vectors $a_1, \ldots, a_k$, then the coadjoint orbit through P_0 under the action of the semi-direct product $\mathfrak{p} \rtimes SO_n$ is diffeomorphic to the tangent bundle of the oriented Grassmannian Gr_k^n.*

Here P_0 is identified with the flag consisting of a single k-dimensional vector space V_k spanned by $a_1, \ldots, a_k$. Then $\{(hV_k), h \in SO(n)\}$ is diffeomorphic to the oriented Grassmannians Gr_k^n.

Note 1. *Proposition 4 is a correction to Proposition 10.2 on page 170 in [8] which incorrectly states that the coadjoint orbit through P_0 is the Steifel St_k^n rather than the flag manifold $\mathbb{F}(1, 2, \ldots, k)$.*

3. Hamiltonian and Poisson Systems: Extremal Curves

We now come to the central part of the paper, the Hamiltonian systems associated with our optimal control problems,

3.1. Rolling Hamiltonians

Recall the rolling problem Equation (17),

$$\frac{dg}{dt} = g(t)(\sum_{i=1}^{m} u_i(t) A_i), \frac{dp}{dt} = \sum_{i=1}^{m} u_i(t) \vec{A}_i(o),$$

and the associated optimal control problem of minimizing the energy function $\frac{1}{2}\int_0^T \sum_{i=1}^{m} u_i^2(t)dt$. Our immediate aim is to use the Maximum Principle to obtain the equations for the extremal curves in the cotangent bundle $T^*\mathbf{G}$ of the configuration space $\mathbf{G}$. To emphasize the structure of the problem, we will rewrite (17) as

$$\frac{d\mathbf{g}}{dt} = \sum_{i=1}^{m} u_i(t) X_i(\mathbf{g}), \tag{39}$$

where each X_i a left-invariant vector field $X_i(\mathbf{g}) = (gA_i, \vec{A}_i(o))$, $\mathbf{g} = (g, p)$. If $\mathbf{g}(t)$ is an optimal trajectory then, according to the Maximum Principle, $\mathbf{g}(t)$ is the projection of an extremal curve $\xi(t)$ in $T^*\mathbf{G}$ along which the cost extended Hamiltonian

$$-\frac{\lambda}{2}\sum_{i=1}^{m} u_i^2(t) + \sum_{i=1}^{m} u_i(t)H_i(\xi(t)), \lambda = 0, 1$$

is maximal relative to all competing controls. In this notation, each H_i is the Hamiltonian lift of X_i, i.e., $H_i(\xi(t)) = \xi(t)(X_i(\mathbf{g}(t))$. In the abnormal case, $\lambda = 0$, the Maximum principle results in the constraints

$$H_i(\xi(t)) = 0, i = 1, \ldots, m, \tag{40}$$

while in the normal case, $\lambda = 1$, the maximality condition implies that the optimal controls are of the form $u_i(t) = H_i(\xi(t))$, in which case the corresponding optimal solutions are the projections of the solution curves of a single Hamiltonian vector field $\vec{H}$ generated by the Hamiltonian

$$H(\xi) = \frac{1}{2}\sum_{i=1}^{m} H_i^2(\xi). \tag{41}$$

This Hamiltonian is left-invariant in the representation $T^*\mathbf{G} = \mathbf{G} \times \mathcal{G}^*$ and hence its Hamiltonian equations are given by the Equation (29), that is,

$$\frac{d\mathbf{g}}{dt} = \sum_{i=1}^{n} H_i(\ell(t))X_i(\mathbf{g}(t)), \frac{d\ell}{dt} = -ad^* dH(\ell(t))(\ell(t))$$

We will now concentrate on the solutions of the associated Poisson equation

$$\frac{d\ell}{dt} = -ad^* dH(\ell(t))(\ell(t)) \tag{42}$$

Let us first expand on the structure of the coadjoint orbits in this situation. Since $\widehat{M}$ is a Euclidean vector space, its tangent space at the origin can be identified with $\widehat{M}$. Then the Lie algebra $\mathcal{G}$ can be identified with $\mathfrak{g} \times \widehat{M}$, and its dual can be identified with $\mathcal{G}^* = \mathfrak{g}^* \oplus \widehat{M}^*$, where

$$\mathfrak{g}^* = \{\ell \in \mathcal{G}^* : \ell(\dot{p}) = 0, \dot{p} \in \widehat{M}\}, \widehat{M}^* = \{\ell \in \mathcal{G}^* : \ell(\mathfrak{g}) = 0\}.$$

It then follows that every $\ell \in \mathcal{G}^*$ can be written as $\ell = \ell_1 + \ell_2$ with $\ell_1 \in \mathfrak{g}^*$ and $\ell_2 \in \widehat{M}^*$. Since $\widehat{M}$ is a vector space, and therefore an abelian algebra, the projection ℓ_2 on $\widehat{M}^*$ is constant on each coadjoint orbit of $\mathbf{G}$. The argument is straightforward: if $\mathbf{g} = (g, p)$, then

$$Ad^*_{\mathbf{g}}(\ell)(X + \dot{p}) = \ell(Ad_{\mathbf{g}^{-1}}(X + \dot{p})) = \ell(Ad_{g^{-1}}(X) + \dot{p}) = \ell_1(Ad_{g^{-1}}(X)) + \ell_2(\dot{p}),$$

It follows that the coadjoint orbits in $\mathcal{G}$ are of the form

$$\{Ad^*_g(\ell_1) : g \in G\} + \ell_2, \text{ for any } \ell = \ell_1 + \ell_2.$$

This fact can be also verified directly from Equation (42): we have

$$\frac{d\ell}{dt}V = -\ell[dH, V], \text{for any } V = X + \dot{p} \text{ in } \mathcal{G},$$

where $dH = \sum_{i=1}^{m} H_i(\ell)(A_i + \vec{A}_i(o))$ and $H_i(\ell) = \ell_1(A_i) + \ell_2(\vec{A}_i(o))$. Therefore,

$$\frac{d\ell_1}{dt}(X) + \frac{d\ell_2}{dt}(\dot{p}) = -(\ell_1 + \ell_2)([dH, X + \dot{x}]) = -\sum_{i=1}^{m} H_i(\ell_i)[A_i, X].$$

from which follows that

$$\frac{d\ell_1}{dt}(X) = -\sum_{i=1}^{n} H_i(\ell_i)[A_i, X], X \in \mathfrak{g}, \frac{d\ell_2}{dt}(\dot{p}) = 0.$$

Since $\dot{p}$ is arbitrary $\frac{d\ell_2}{dt} = 0$.

To uncover other constants of motion, identify $\mathcal{G}^*$ with $\mathcal{G}$ via the natural quadratic forms on each of the factors, and then recast the preceding equations on $\mathcal{G}$. More precisely, identify each ℓ_2 in $\widehat{M}^*$ with a tangent vector $l = \sum_{i=1}^{m} l_i \vec{A}_i(o)$ via the formula $\ell_2(\dot{p}) = (l, \dot{p}), \dot{p} \in \widehat{M}$. Similarly, identify $\ell_1 \in \mathfrak{g}^*$ with $L \in \mathfrak{g}$ via the formula $\ell_1(X) = \langle L, X \rangle, X \in \mathfrak{g}$. Then decompose $L \in \mathfrak{g}$ into the sum $L = L_{\mathfrak{p}} + L_{\mathfrak{k}}$, $L_{\mathfrak{p}} \in \mathfrak{p}$ and $L_{\mathfrak{k}} \in \mathfrak{k}$. Relative to the basis $A_1, \ldots, A_m$ in $\mathfrak{p}$, $L_{\mathfrak{p}} = \sum_{i=1}^{m} P_i A_i$ where $P_i = \ell_1(A_i) = \langle L, A_i \rangle$. It follows that

$$H_i(\xi) = \ell(A_i + \vec{A}_i(o)) = \ell_1(A_i) + \ell_2(\vec{A}_i(o)) = P_i + l_i,$$

and

$$\frac{d\ell_1}{dt}(X) = \langle \frac{dL}{dt}, X \rangle = -\langle L, [\textstyle\sum_{i=1}^{m}(l_i + P_i)A_i, X] \rangle = -\langle [L, \textstyle\sum_{i=1}^{m}(l_i + P_i)A_i], X \rangle,$$
$$(\frac{dl}{dt}, \dot{p}) = \frac{d\ell_2}{dt}(t)(\dot{p}) = 0$$

Since X and $\dot{p}$ are arbitrary,

$$\frac{dL}{dt} = [\sum_{i=1}^{m}(l_i + P_i)A_i, L] = [A + L_{\mathfrak{p}}, L], A = \sum_{i=1}^{m} l_i A_i, \frac{dl}{dt} = 0. \tag{43}$$

Equation (43) constitutes the Poisson equations on $\mathcal{G}$ generated by the Hamiltonian $H = \frac{1}{2}\sum_{i=1}^{m} H_i^2 = \frac{1}{2}\sum_{i=1}^{m}(l_i + P_i)^2$. Note that in this identification of the Lie algebras with their duals, coadjoint orbits $\{Ad_g^*(\ell_1) + \ell_2 : g \in G\}$ are identified with the affine sets $\{Ad_g(L) + l : g \in G\}$. Coupled with

$$\frac{dg}{dt} = g(t)(A + L_{\mathfrak{p}}), \frac{dp}{dt} = \sum_{i=1}^{n}(l_i + P_i)\vec{A}_i(o), \tag{44}$$

Equation (43) constitutes the extremal equations for the rolling geodesics. Each extremal curve projects onto a geodesic $\mathbf{g}(t) = (g(t), p(t))$, and each geodesic further projects onto the pair of curves $\alpha(t) = \tau_{g(t)}(o)$ in M and $\beta(t) = p(t)$ in $\widehat{M}$ that are rolled upon each other by $g(t)$.

3.2. Affine-Quadratic Hamiltonian

Similar to the rolling problem, the Maximum Principle reveals that the normal extremals of the affine-quadratic system (18) are the integral curves of the Hamiltonian vector field $\vec{H}$ associated with the Hamiltonian function

$$H(L) = \frac{1}{2}\langle \mathcal{P}^{-1}L_{\mathfrak{k}}, L_{\mathfrak{k}} \rangle + \langle A, L_{\mathfrak{p}} \rangle,$$

where as before $L = L_{\mathfrak{p}} + L_{\mathfrak{k}}$ is the decomposition of $L \in \mathfrak{g}$ onto the factors $\mathfrak{p}$ and $\mathfrak{k}$. In the canonical case $\mathcal{P} = I$, and in the representation $T^*G = G \times \mathfrak{g}^*$, the Hamiltonian equations generated by H are then given by

$$\frac{dg}{dt} = g(t)(A + U(t)), U(t) = L_{\mathfrak{k}}(t), \frac{dL}{dt} = [dH, L] = [A + L_{\mathfrak{k}}, L], \tag{45}$$

The Poisson equation $\frac{dL}{dt} = [dH, L]$ can be written in expanded form as

$$\frac{dL_{\mathfrak{k}}}{dt} = [A, L_{\mathfrak{p}}], \frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{k}}, L_{\mathfrak{p}}] + [A, L_{\mathfrak{k}}] = [A - L_{\mathfrak{p}}, Ł_{\mathfrak{k}}] \tag{46}$$

The "shadow" problem generates an analogous Hamiltonian on the tangent bundle of the semi-direct product $G_0 = \mathfrak{p} \rtimes K$ with its extremal equations given by:

$$\frac{dx}{dt} = Ad_{R(t)}A, \frac{dR}{dt} = R(t)L_{\mathfrak{k}}(t), \frac{dL_{\mathfrak{k}}}{dt} = [A, L_{\mathfrak{p}}], \frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{k}}, L_{\mathfrak{p}}]. \tag{47}$$

Here $g(t) = (x(t), R(t))$, and $\frac{dg}{dt} = g(t)(A + L_{\mathfrak{k}}(t))$ is the same as $\frac{dx}{dt} = Ad_{R(t)}A$, $\frac{dR}{dt} = R(t)L_{\mathfrak{k}}(t)$.

The propositions below reveal a remarkable fact that the Poisson equations of a canonical affine-quadratic Hamiltonian can always be regarded as an invariant subsystem of the Poisson equations associated with a rolling Hamiltonian. We will use bold letters when referring to the variables in the rolling Hamiltonian in contrast to the variables in the affine-quadratic Hamiltonian.

Proposition 5. *Let* $\mathbf{g}(t) = (g(t), p(t)), \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)$ *be any integral curve of the rolling Hamiltonian* $\mathbf{H} = \frac{1}{2}||\mathbf{A} + \mathbf{L}_{\mathfrak{p}}||^2$*, that is,*

$$\frac{dg}{dt} = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)), \frac{dp}{dt} = \sum_{i=1}^m (l_i + P_i)\vec{A}_i(o),$$
$$\frac{d\mathbf{L}_{\mathfrak{k}}}{dt} = [\mathbf{A}, \mathbf{L}_{\mathfrak{p}}], \frac{d\mathbf{L}_{\mathfrak{p}}}{dt} = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}, \mathbf{L}_{\mathfrak{k}}], \mathbf{A} = \sum_{i=1}^m l_i\mathbf{A}_i$$

Then

$$\tilde{g}(t) = g(t)h(t), L_{\mathfrak{p}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)), L_{\mathfrak{k}} = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{k}}(t)) \tag{48}$$

is an integral curve of the affine Hamiltonian $H = \frac{1}{2}\langle L_{\mathfrak{k}}, L_{\mathfrak{k}}\rangle + \langle A, L_{\mathfrak{p}}\rangle$*, where* $A = Ad_{h^{-1}(t)}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t))$*, and* $h(t)$ *is the solution of* $\frac{dh}{dt} = \mathbf{L}_{\mathfrak{k}}(t)h(t)$ *with* $h(0) = I$*.*

Moreover, $\tilde{g}(t) = (x(t), R(t))$ *in* $\mathfrak{p} \rtimes K$ *with* $R(t) = h(t)$ *and* $x(t)$ *a solution of* $\frac{dx}{dt} = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$ *is the projection of an extremal curve*

$$L_{\mathfrak{k}}(t) = Ad_{h^{-1}(t)}\mathbf{L}_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)) - A, A = Ad_{h^{-1}(t)}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t))$$

associated with the shadow Hamiltonian $H = \frac{1}{2}\langle L_{\mathfrak{k}}, L_{\mathfrak{k}}\rangle + \langle A, L_{\mathfrak{p}}\rangle$*.*

Proof. If A is any element in $\mathfrak{p}$ then $\frac{d}{dt}Ad_{h(t)}(A) = [Ad_{h(t)}(A), \mathbf{L}_{\mathfrak{k}}]$. Since $\frac{d}{dt}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)) = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)]$, $Ad_{h(t)}(A)$ and $\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$ are the solutions of the same differential equation they will be equal to each other whenever $Ad_{h(0)}(A) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(0)$, that is, when $A = \mathbf{A} + L_{\mathfrak{p}}(0)$.

Assume that $Ad_{h(t)}(A) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$. Then,

$$\frac{d\tilde{g}}{dt} = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t))h(t) + g(t)\mathbf{L}_{\mathfrak{k}}(t)h(t) =$$
$$\tilde{g}(t)(Ad_{h^{-1}(t)}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)) + Ad_{h^{-1}(t)}\mathbf{L}_{\mathfrak{k}}(t)) = \tilde{g}(t)(A + L_{\mathfrak{k}}(t)).$$

Additionally,

$$\frac{dL_{\mathfrak{p}}}{dt} = \frac{d}{dt}Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)) = Ad_{h^{-1}(t)}([\mathbf{L}_{\mathfrak{k}}, \mathbf{L}_{\mathfrak{p}}]) + Ad_{h^{-1}(t)}([\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)])$$
$$= Ad_{h^{-1}(t)}[\mathbf{A}, \mathbf{L}_{\mathfrak{k}}(t)] = [Ad_{h^{-1}(t)}\mathbf{A}, Ad_{h^{-1}(t)}\mathbf{L}_{\mathfrak{k}}(t)] = [A - Ad_{h^{-1}(t)}\mathbf{L}_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t)] =$$
$$[A - L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t)],$$

and

$$\frac{dL_{\mathfrak{k}}}{dt} = \frac{d}{dt}Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{k}}(t)) = Ad_{h^{-1}(t)}[\mathbf{A}, \mathbf{L}_{\mathfrak{p}}(t)] = [Ad_{h^{-1}(t)}\mathbf{A}, Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)] =$$
$$[A - Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)), Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t))] = [A, L_{\mathfrak{p}}(t)].$$

As to the proof of the second statement, note that $\dot{x}(t) = Ad_{R(t)}A = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$ and $\frac{dR}{dt} = \mathbf{L}_{\mathfrak{k}}(t)R(t) = R(t)L_{\mathfrak{k}}(t)$ is a solution of $\frac{d\tilde{g}}{dt} = \tilde{g}(t)(A + L_{\mathfrak{k}}(t))$ as remarked in Equation (47). An argument identical to the one above shows that

$$\frac{dL_{\mathfrak{k}}}{dt} = [A, L_{\mathfrak{p}}(t)], \frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t)].$$

□

The converse also holds as this proposition demonstrates.

Proposition 6. *Suppose that* $(\tilde{g}(t), L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t))$ *is an extremal curve of the affine Hamiltonian* $H = \frac{1}{2}\langle L_{\mathfrak{k}}, L_{\mathfrak{k}}\rangle + \langle A, L_{\mathfrak{p}}\rangle$. *Then*

$$\begin{gathered}\mathbf{g}(t) = ((\tilde{g}(t)h^{-1}(t), p(t)), \tfrac{dp}{dt} = \vec{\mathbf{A}}(o) + \vec{\mathbf{L}}_{\mathfrak{p}}(o), \tfrac{dh}{dt} = h(t)(L_{\mathfrak{k}}(t)) \\ \mathbf{L}_{\mathfrak{p}}(t) = Ad_{h(t)}(L_{\mathfrak{p}}(t)), \mathbf{L}_{\mathfrak{k}}(t) = Ad_{h(t)}(L_{\mathfrak{k}}(t)), \mathbf{A} = Ad_{h(t)}(A - L_{\mathfrak{p}}(t))\end{gathered}$$

is an extremal curve of the rolling Hamiltonian $\mathbf{H} = \frac{\mathbf{1}}{\mathbf{2}}\langle \mathbf{A} + \mathbf{L}_{\mathfrak{p}}, \mathbf{A} + \mathbf{L}_{\mathfrak{p}}\rangle$.

However, if $\tilde{g}(t) = (x(t), R(t))$, $L_{\mathfrak{p}}(t)$ *and* $L_{\mathfrak{k}}(t)$ *is an extremal curve of the shadow Hamiltonian* H, *then*

$$\begin{gathered}\tfrac{dg}{dt} = g(t)Ad_{R(t)}(A)), \tfrac{dp}{dt} = \tfrac{\vec{dx}}{dt}(o) \\ \mathbf{L}_{\mathfrak{p}}(t) = Ad_{R(t)}(A + L_{\mathfrak{p}}(t)), \mathbf{L}_{\mathfrak{k}}(t) = Ad_{R(t)}(L_{\mathfrak{k}}(t))\end{gathered}$$

is an extremal equation of the Hamiltonian $\mathbf{H} = \frac{\mathbf{1}}{\mathbf{2}}\langle \mathbf{A} + \mathbf{L}_{\mathfrak{p}}, \mathbf{A} + \mathbf{L}_{\mathfrak{p}}\rangle$ *with* $Ad_{R(t)}A = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$.

Proof. The proof of the first part is essentially the same as in the previous proposition.

In the second part, we have

$$\frac{dx}{dt} = Ad_R(t)(A), \frac{dR}{dt} = R(t)L_{\mathfrak{k}}(t), \frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{k}}, L_{\mathfrak{p}}], \frac{dL_{\mathfrak{k}}}{dt} = [A, L_{\mathfrak{p}}].$$

Then $\frac{d}{dt}Ad_{R(t)}(L_{\mathfrak{p}}(t)) = Ad_{R(t)}([L_{\mathfrak{p}}, L_{\mathfrak{k}}]) + Ad_{R(t)}([L_{\mathfrak{k}}, L_{\mathfrak{p}}]) = 0$.

Let $Ad_{R(t)}(L_{\mathfrak{p}}(t)) = -\mathbf{A}$ so that $Ad_{R(t)}(A) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$. It follows that $\frac{dg}{dt} = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t))$ and $\frac{dx}{dt} = Ad_{R(t)}(A) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$. Hence,

$$\frac{dp}{dt} = \vec{\mathbf{A}}(o) + \vec{\mathbf{L}}_{\mathfrak{p}}(o) = \frac{\vec{dx}}{dt}(o).$$

Additionally,

$$\begin{aligned}\tfrac{d\mathbf{L}_{\mathfrak{p}}}{dt} &= \tfrac{d}{dt}Ad_{R(t)}(A + L_{\mathfrak{p}}(t)) = Ad_{R(t)}([A + L_{\mathfrak{p}}, L_{\mathfrak{k}}]) + Ad_{R(t)}([L_{\mathfrak{k}}, L_{\mathfrak{p}}]) = \\ &\quad Ad_{R(t)}([A, L_{\mathfrak{k}}]) = [Ad_{R(t)}(A), Ad_{R(t)}(L_{\mathfrak{k}})] = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)], \\ \tfrac{d\mathbf{L}_{\mathfrak{k}}}{dt} &= \tfrac{d}{dt}Ad_{R(t)}(L_{\mathfrak{k}}(t)) = Ad_{R(t)}([A, L_{\mathfrak{p}}(t)]) = [Ad_{R(t)}(A), Ad_{R(t)}(L_{\mathfrak{p}}(t)]) = \\ &\quad [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), -\mathbf{A}] = [\mathbf{A}, \mathbf{L}_{\mathfrak{p}}(t)].\end{aligned}$$

□

The above shows that the Poisson systems generated by any affine-quadratic Hamiltonian are invariant subsystems of the rolling Hamiltonians. To summarize, let $\mathbf{L}(t) = \mathbf{L}_{\mathfrak{p}}(t) + \mathbf{L}_{\mathfrak{k}}(t)$ denote an integral curve of the rolling Hamiltonian $\mathbf{H} = \frac{\mathbf{1}}{\mathbf{2}}\langle \mathbf{A} + \mathbf{L}_{\mathfrak{p}}, \mathbf{A} + \mathbf{L}_{\mathfrak{p}}\rangle$. If $h(t)$ denotes the solution of $\frac{dh}{dt} = \mathbf{L}_{\mathfrak{k}}(t)h(t), h(0) = I$, then define $A \in \mathfrak{p}$ by $Ad_{h(t)}(A) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$. It follows from above that

$$L_{\mathfrak{p}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)), L_{\mathfrak{k}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{k}}(t)) \tag{49}$$

are integral curves of the affine quadratic Hamiltonian $H = \frac{1}{2}\langle L_{\mathfrak{k}}, L_{\mathfrak{k}}\rangle + \langle A, L_{\mathfrak{p}}\rangle$. However, when

$$L_{\mathfrak{p}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{p}}(t)) - A, L_{\mathfrak{k}}(t) = Ad_{h^{-1}(t)}(\mathbf{L}_{\mathfrak{k}}(t)) \tag{50}$$

then $L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t)$ are integral curves of the shadow Hamiltonian H.

3.3. Isospectral Representations and Integrability

An $n \times n$ matrix equation $\frac{dL}{dt} = [M(t), L(t)]$ is called a Lax equation, and (M, L) is called Lax pair. If (M, L) is a Lax pair, then the spectrum of $L(t)$ is constant. The proof is simple: $g(t)L(t)g^{-1}(t) = \Lambda$, where Λ a constant matrix for any solution $\frac{dg}{dt} = g(t)M(t)$ in the general linear group $Gl(n)$. Since the spectrum of Λ is equal to the spectrum of $L(t)$, the spectrum of $L(t)$ must be constant.

It follows that the Poisson equation of any left-invariant Hamiltonian H is a Lax equation on a semi-simple Lie algebra $\mathfrak{g}$ (Equation (34)) and therefore, the eigenvalues of $L(t)$ are constants of motion for any left-invariant Hamiltonian on $\mathfrak{g}$ and hence may be regarded as the conservation laws on $\mathfrak{g}$.

A function h on a Poisson space is said to be invariant if $\{h, f\} = 0$ for any function f. On semi-simple Lie algebras any spectral function is invariant. In particular functions $\phi_k(L) = Tr(L^k), k = 1, 2, \ldots$ form a family of invariant functions.

In some situations, a Lax equation $\frac{dL}{dt} = [M(t), L(t)]$ extends to a Lax equation $\frac{dL_\lambda}{dt} = [M_\lambda(t), L_\lambda(t)]$ with a spectral parameter λ. Then a discrete spectrum of L is replaced by a continuous spectrum of L_λ which results in additional constants of motion. In the case of rolling spheres J. Zimmerman in his PhD thesis (2002, University of Toronto) discovered an extension of the Lax equation which he called *isospectral* ([6]). Remarkably, Zimmerman's extension exists for the rolling problem on any semi-simple homogeneous manifold, for the same reasons as in the rolling sphere problem. In fact, if $X_0(t) = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), X_1(t) = \mathbf{L}_{\mathfrak{k}}(t), X_2(t) = -\mathbf{A}, X_3 = 0$, then the Poisson equations may be written as

$$\frac{dX_i}{dt} = [X_0(t), X_{i+1}(t)], i = 0, 1, 2. \tag{51}$$

This equation is invariant under a dilational change $X_i \to \lambda^{i-1}X_i$. It then follows that

$$\mathbf{L}_\lambda = \sum_{i=0}^{3} \lambda^i X_i = \mathbf{L}_{\mathfrak{p}}(t) + \lambda \mathbf{L}_{\mathfrak{k}}(t) + (1-\lambda^2)\mathbf{A} \tag{52}$$

satisfies the equation

$$\frac{d\mathbf{L}_\lambda}{dt} = [\mathbf{M}_\lambda(t), \mathbf{L}_\lambda(t)], \mathbf{M}_\lambda(t) = \frac{1}{\lambda}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)). \tag{53}$$

Therefore, the spectrum of $\mathbf{L}_\lambda(t)$ is constant. We will refer to $\mathbf{L}_\lambda$ as the spectral curve for $\mathbf{H}$. Of course, the above implies that the Poisson system associated with the affine-quadratic Hamiltonian also admits an isospectral representation. To be specific note that after the substitutions from Equation (49),

$$\begin{gathered}\mathbf{L}_\lambda = Ad_{h(t)}L_{\mathfrak{p}} + \lambda Ad_{h(t)}L_{\mathfrak{k}} + (1-\lambda^2(Ad_{h(t)}(A - L_{\mathfrak{p}}) = \\ Ad_{h(t)}(\lambda^2 L_{\mathfrak{p}} + \lambda L_{\mathfrak{k}} + (1-\lambda^2)A) = Ad_{h(t)}L_\lambda.\end{gathered}$$

Then

$$\begin{aligned}\tfrac{d\mathbf{L}_\lambda}{dt} &= \tfrac{d}{dt}(Ad_{h(t)}(L_\lambda) = Ad_{h(t)}[L_\lambda, L_{\mathfrak{k}}] + Ad_{h(t)}\tfrac{dL_\lambda}{dt} \\ &= [\tfrac{1}{\lambda}(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}), \mathbf{L}_\lambda] = Ad_{h(t)}[\tfrac{1}{\lambda}A, L_\lambda].\end{aligned}$$

Therefore,

$$\frac{dL_\lambda}{dt} = [L_{\mathfrak{k}}, L_\lambda] + [\frac{1}{\lambda}A, L_\lambda] = [\frac{1}{\lambda}A + L_{\mathfrak{k}}, L_\lambda].$$

To be consistent with my earlier publications, replace λ by $-\frac{1}{\lambda}$ to get

$$\frac{dL_\lambda}{dt} = [M_\lambda, L_\lambda], \tag{54}$$

where $M_\lambda = L_{\mathfrak{k}} - \lambda A$, and $L_\lambda = L_{\mathfrak{p}} - \lambda L_{\mathfrak{k}} + (\lambda^2 - 1)A$. Equation (54) agrees with the isospectral representation in ([8]) (obtained by other means).

To get the spectral curve L_λ for the shadow Hamiltonian, use Equation (50). In such a case, $\mathbf{L}_{\mathfrak{k}} = Ad_h(L_{\mathfrak{k}})$, $\mathbf{L}_{\mathfrak{p}} = Ad_h(L_{\mathfrak{p}} + A)$ and $\mathbf{A} = -Ad_h L_{\mathfrak{p}}$ yields

$$\mathbf{L}_\lambda = Ad_h L_\lambda, L_\lambda = \lambda^2 L_{\mathfrak{p}} + \lambda L_{\mathfrak{k}} + A.$$

Then a calculation analogous to the one above gives $\frac{dL_\lambda}{dt} = [\frac{1}{\lambda}A + L_{\mathfrak{k}}, L_\lambda]$. After the rescaling $\lambda \to -\frac{1}{\lambda}$ we get a modified Lax pair

$$\frac{dL_\lambda}{dt} = [M_\lambda, L_\lambda], M_\lambda = L_{\mathfrak{k}} - \lambda A, L_\lambda = L_{\mathfrak{p}} - \lambda L_{\mathfrak{k}} + \lambda^2 A. \tag{55}$$

Each spectral curve $\mathbf{L}_\lambda$ defines a family of functions

$$\mathcal{I} = \{\phi_\lambda^{(k)}(L) = Tr(\mathbf{L}_\lambda^k), k = 1, 2, \dots\} \cup \{f(L) = \langle \mathbf{L}, X\rangle : X \in \mathfrak{k}, [X, \mathbf{A}] = 0\}.$$

Proposition 7. *The family $\mathcal{I}$ is involutive, that is, $\{h, g\} = 0$ for each g and h in $\mathcal{I}$, and in the case that* **A** *is regular, it is also complete, in the sense that it contains a subfamily $\mathcal{I}_0$ that is Liouville integrable on each coadjoint orbit in $\mathfrak{g}$ ([8], pp. 164–165).*

See also related papers also [27–29]).

Since **H** belongs to $\mathcal{I}$, the rolling problem is completely integrable when **A** is regular.

Corollary 2. *Each affine-quadratic Hamiltonian $H = \frac{1}{2}\langle L_{\mathfrak{k}}, L_{\mathfrak{k}}\rangle + \langle A, L_{\mathfrak{k}}\rangle$ is completely integrable on $\mathfrak{g}$ when A is regular.*

4. Symmetric Mechanical Tops

We will now relate the "top-like" equations

$$\frac{dR}{dt} = R(t)(\mathcal{P}^{-1}(M(t))), \alpha_i(t) = R(t)^T e_i, i = 1, \dots, n \tag{56}$$

$$\frac{dM}{dt} = [\mathcal{P}^{-1}(M(t)), M(t)] + \sum_{i=1}^n \alpha_i(t) \wedge \frac{\partial V}{\partial \alpha_i} \tag{57}$$

on the tangent bundle of $SO(n)$, associated with the energy Hamiltonian $H = \frac{1}{2}\langle \mathcal{P}^{-1}(M), M\rangle + V(\alpha_1, \dots, \alpha_n)$, to the rolling equations. For simplicity of exposition, we will assume that the top is maximally symmetric, that is we will assume that all principal moments of inertia are equal, which is the same as $\mathcal{P} = I$. We will first consider the case of linear potentials.

Linear potentials: $V = -\sum_{i=1}^n c_i(\alpha_i, a)$, where a is a vector in $\mathbb{R}^n$, and $c_1, \dots, c_n$ are constants. Then Equation (56) can be written as

$$\frac{dR}{dt} = R(t)\Omega(t), \frac{dM}{dt} = a \wedge p(t), \frac{dp}{dt} = -\Omega(t)p(t) \tag{58}$$

where $\Omega(t) = M(t)$ and $p(t) = \sum_{i=1}^{n} c_i \alpha_i(t)$. Our proposition below relates Equation (58) to the rolling equations

$$\frac{dg}{dt} = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)),\ \frac{dx}{dt} = \vec{\mathbf{A}}(o) + \vec{\mathbf{L}}_{\mathfrak{p}}(o), \tag{59}$$

$$\frac{d\mathbf{L}_{\mathfrak{p}}}{dt} = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)],\ \frac{d\mathbf{L}_{\mathfrak{k}}}{dt} = [\mathbf{A}, \mathbf{L}_{\mathfrak{p}}(t)]. \tag{60}$$

on $SO_\epsilon(n+1) \times T_o M, \epsilon = \pm 1$, where $M = SO_\epsilon(n+1)/K, K = \{1\} \times SO(n)$.

To set the stage for this proposition, we will need to embed Equation (58) in R^{n+1} via the following embeddings. To begin with, $\hat{v} \in R^{n+1}$ will denote the embedding $\hat{v} = 0e_0 + \sum_{i=1}^{n} v_i e_i$ for any $v = \sum_{i=1}^{n} v_i e_i$. Then $a \in R^n$ will be identified with $A = \hat{a} \wedge_\epsilon e_0$ and $p \in R^n$ will be identified with $L_{\mathfrak{p}} = p \wedge_\epsilon e_0$ in $\mathfrak{p}_\epsilon$. In addition $R \in SO(n)$ will be identified with $h = \{1\} \times R = \begin{pmatrix} 1 & 0 \\ 0 & R \end{pmatrix}$, and Ω will be identified with $L_{\mathfrak{k}} = \begin{pmatrix} 0 & 0 \\ 0 & M \end{pmatrix}$, so that $\frac{dR}{dt} = R(t)\Omega(t)$ is identified with $\frac{dh}{dt} = h(t)L_{\mathfrak{k}}(t)$. Then

$$\frac{dL_{\mathfrak{p}}}{dt}(t) = \frac{d}{dt}(\hat{p}(t) \wedge_\epsilon e_0) = -Ł_{\mathfrak{k}}(t)\hat{p}(t) \wedge_\epsilon e_0 = [L_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t)]$$

is the same as $\frac{dM}{dt} = a \wedge p(t)$. It follows that Equation (58) can be paraphrased as

$$\frac{dh}{dt} = h(t)L_{\mathfrak{k}}(t), \frac{dL_{\mathfrak{p}}}{dt} = [L_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t)]. \tag{61}$$

However, then

$$\frac{d}{dt} Ad_{h(t)} L_{\mathfrak{p}}(t) = Ad_{h(t)}[L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t)] + Ad_{h(t)}[L_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t)] = 0,$$

and therefore $Ad_{h(t)} L_{\mathfrak{p}}(t)$ is constant (same as $\frac{d}{dt}(R(t)p(t)) = R(t)\Omega(t)p(t) - R(t)\Omega(t)p(t) = 0$).

Proposition 8. *Top-like Equation (58) are isomorphic to the Equations (59) and (60) under the identification*

$$\mathbf{A} = -Ad_{h(t)} L_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{p}}(t) = Ad_{h(t)}(A + L_{\mathfrak{p}}(t)), \mathbf{L}_{\mathfrak{k}}(t) = Ad_{h(t)} L_{\mathfrak{k}}(t)),$$
$$\frac{dg}{dt} = g(t) Ad_{h(t)} A,\ \frac{dx}{dt} = \vec{\mathbf{A}}(o) + \vec{\mathbf{L}}_{\mathfrak{p}}(o).$$

Proof. It follows that $Ad_{h(t)} A = \mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)$ and $\frac{dg}{dt} = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t))$. Thus, (60) is satisfied. We also have

$$\frac{d\mathbf{L}_{\mathfrak{p}}}{dt} = Ad_{h(t)}[A + L_{\mathfrak{p}}(t), L_{\mathfrak{k}}(t)] + Ad_{h(t)}[L_{\mathfrak{k}}(t), L_{\mathfrak{p}}(t)] =$$
$$Ad_{h(t)}[A, L_{\mathfrak{k}}(t)] = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{k}}(t)],$$
$$\frac{\mathbf{L}_{\mathfrak{k}}}{dt} = Ad_{h(t)}[A, L_{\mathfrak{p}}(t)] = [Ad_{h(t)} A, Ad_{h(t)} L_{\mathfrak{p}}(t)] = [\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t), -\mathbf{A}] = [\mathbf{A}, \mathbf{L}_{\mathfrak{p}}(t)],$$

and Equation (59) are also satisfied. □

Corollary 3. *An n-dimensional symmetric top with a linear potential is completely integrable.*

Quadratic potentials. We will now show that the rolling geodesic equations on $M = SL(n)/SO(n)$ can be identified with movements of the symmetric top under a quadratic potential. For our purposes, an n-dimensional top with quadratic potential is synonymous with the Hamiltonian

$$H(R, M) = \frac{1}{2}(\mathcal{P}^{-1}(M), M) + \frac{1}{2}\sum_{i=1}^{n} a_i \langle S\alpha_i, \alpha_i \rangle,$$

with $R \in SO(n), M \in so(n), R^T e_i = \alpha_i$, S a symmetric $n \times n$ matrix, and $a_1, \ldots, a_n$ arbitrary numbers. In accordance with (32) the Hamiltonian equations of $\vec{H}$ are given by

$$\frac{dR}{dt} = R(t)\Omega(t), \frac{dM}{dt} = [\Omega(t), M(t)] + \sum_{i=1}^{n} a_i \alpha_i(t) \wedge S\alpha_i(t), \tag{62}$$

$\Omega(t) = \mathcal{P}^{-1}(M(t))$. In the symmetric case $\mathcal{P} = I$ and $[\Omega(t), M(t)] = 0$ and the equations reduce to

$$\frac{dR}{dt} = R(t)\Omega(t), \frac{dM}{dt} = \sum_{i=1}^{n} a_i \alpha_i(t) \wedge S\alpha_i(t). \tag{63}$$

To relate these equations to the rolling equations, let

$$L_{\mathfrak{p}}(t) = \sum_{i=1}^{n} a_i(\alpha_i(t) \otimes \alpha_i(t)) - \frac{1}{n} \sum_{i=1}^{n} a_i I.$$

Recall that $a \otimes a$ is a rank one matrix defined by $(a \otimes a)x = (a, x)a$ where (a, x) is the standard Euclidean inner product in R^n. Therefore each matrix $\alpha_i \otimes \alpha_i$ is a symmetric matrix with its trace equal to one, and consequently $L_{\mathfrak{p}}$ is a symmetric matrix having zero trace. Along each solution of (63)

$$\frac{d}{dt} L_{\mathfrak{p}}(t) = -\sum_{i=1}^{n} a_i(\Omega(t)\alpha_i(t) \otimes \alpha_i(t) + \alpha_i(t) \otimes \Omega(t)\alpha_i(t)) = [\Omega(t), L_{\mathfrak{p}}(t)].$$

Additionally,

$$\frac{d}{dt} Ad_{R(t)} L_{\mathfrak{p}}(t) = Ad_{R(t)}[L_{\mathfrak{p}}(t), \Omega(t)] + Ad_{R(t)}[\Omega(t), L_{\mathfrak{p}}(t)]) = 0.$$

Now let

$$\mathbf{A} = -Ad_{R(t)} L_{\mathfrak{p}}(t), \mathbf{L}_{\mathfrak{p}}(t) = Ad_{R(t)}(S + L_{\mathfrak{p}}(t)), \mathbf{L}_{\mathfrak{k}}(t) = Ad_{R(t)} \Omega(t). \tag{64}$$

We then have

Proposition 9. *Equation (63) are isomorphic to the Poisson equations of the rolling problem on $\mathfrak{g} = sl(n)$ (Equation (43)) associated with the extremal*

$$\frac{dg}{dt} = Ad_{R(t)} S = g(t)(\mathbf{A} + \mathbf{L}_{\mathfrak{p}}(t)), \frac{dp}{dt} = \vec{\mathbf{A}}(o) + \vec{\mathbf{L}}_{\mathfrak{p}}(o).$$

Proof. By a straightforward calculation. □

Corollary 4. *Equations of a symmetric n-dimensional top with quadratic potential are completely integrable.*

See also related results in [30–32]).

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: No conflict of interest.

References

1. Jurdjevic, V. The geometry of the Ball-Plate Problem. *Arch. Ration. Mech. Anal.* **1993**, *124*, 305–328. [CrossRef]
2. Jurdjevic, V. Non-Euclidean Elasticae. *Am. J. Math.* **1995**, *117*, 93–125. [CrossRef]
3. Jurdjevic, V. *Geometric Control Theory*; Cambridge Studies in Advanced Mathematics; Cambridge University Press: New York, NY, USA, 1997; Volume 52.

4. Jurdjevic, V. Integrable Hamiltonian Systems on Lie groups: Kowalewski type. *Ann. Math.* **1999**, *150*, 605–644. [CrossRef]
5. Zimmerman, J.A. Optimal control of the sphere S^n rolling on E^n. *Math. Control Signals Syst.* **2005**, *17*, 14–37. [CrossRef]
6. Jurdjevic, V.; Zimmermann, J. Rolling sphere problems on spaces of constant curvature. *Math. Proc. Camb. Phil. Soc.* **2008**, *144*, 729–747. [CrossRef]
7. Jurdjevic, V. *Integrable Hamiltonian Systems on Complex Lie Groups*; American Mathematical Society: Providence, RI, USA, 2005; Volume 178.
8. Jurdjevic, V. *Optimal Control and Geometry: Integrable Systems*; Cambridge Studies in Advanced Mathematics; Cambridge University Press: Cambridge, UK, 2016.
9. Jurdjevic, V.; Markina, I.; Silva Leite, F. Symmetric spaces rolling on flat spaces. *arXiv* **2022**, arXiv:2205.14636.
10. Jurdjevic, V. Rolling on Affine Tangent Planes: Parallel Transport and the Associated Sub-Riemannian Problems. In Proceedings of the 14th APCA International Conference on Automatic Control and Soft Computing, Bragança, Portugal, 1–3 July 2020; pp. 136–147.
11. Fomenko, A.T.; Mischenko, A.S. Euler equation on finite-dimensional Lie groups. *Izv. Ross. Akad. Nauk. Seriya Mat.* **1978**, *12*, 371–389.
12. Fomenko, A.T.; Trofimov, V.V. Integrability in the sense of Liouville of Hamiltonian systems on Lie algebras. *Uspekhi Mat. Nauk* **1984**, *2*, 3–56. (In Russian)
13. Neumann, C. De problemate quodam mechanico, quod ad primam integralium ultraellipticorum classem revocatum. *J. Reine Angew. Math.* **1859**, *56*, 46–63.
14. Reyman, A.G.; Semenov Tian-Shansky, M.A. Group theoretic Methods in the Theory of finite dimensional Integrable Systems. In *Dynamical Systems VII*; Chapter 2, Encyclopaedia of Mathematical Sciences; Springer: Berlin/Heidelberg, Germany, 1994; Volume 16.
15. O'Neill, B. *Semi-Riemannian Geometry*; Academic Press: Cambridge, MA, USA; Elsevier: Amsterdam, The Netherlands, 1983.
16. Sternberg, S. *Lectures on Differential Geometry*; Prentice- Hall, Inc.: Englewood Cliffs, NJ, USA, 1964.
17. Ziller, W. *Lie Groups. Representation Theory and Symmetric Spaces*; University of Pennsylvania: Philadelphia, PA, USA, 2010.
18. Bryant, R.; Hsu, L. Rigidity of integral curves of rank 2 distributions. *Invent. Math.* **1993**, *114*, 435–461. [CrossRef]
19. Agrachev, A.; Sachkov, Y. *Control Theory from the Geometric Viewpoint*; Springer: Berlin/Heidelberg, Germany, 2004.
20. Chitour, Y.; Kokkonen, P. Rolling Manifolds: Intrinsic Formulation and Controllability. *arXiv* **2011**, arXiv:1011.2925v2.
21. Chitour, Y.; Godoy Molina, M.; Kokkonen, P. The rolling problem: Overview and challenges. In *Geometric Control theory and Sub-Riemannian Geometry*; Springer INdAM Series 5; Springer: Berlin/Heidelberg, Germany, 2014; pp. 103–123.
22. Godoy Molina, M.; Grong, E.; Markina, I.; Silva Leite, F. An intrinsic formulation of the problem on rolling manifolds. *J. Dyn. Control Syst.* **2012**, *18*, 181–214. [CrossRef]
23. Krakowski, K.A.; Machado, L.; Silva Leite, F. Rolling Symmetric Spaces. In Proceedings of the Second International Conference on Geometric Science of Information, Palaiseau, France, 28–30 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 550–557.
24. Sharpe, R.W. *Differential Geometry*; GTM, 166; Springer: New York, NY, USA, 1997.
25. Kirillov, A.A. *Lectures on the Orbit Method*; Graduate Studies in Mathematics, 64; American Mathematical Society: Providence, RI, USA, 2004.
26. Gasparim, E.; Grama, L.; San Martin, L.A. Adjoint orbits of semi-simple Lie groups and Lagrangian submanifolds. *Proc. Edinb. Math. Soc.* **2017**, *60*, 361–385. [CrossRef]
27. Bolsinov, A. A completeness criterion for a family of functions in involution obtained by the shift method. *Soviet Math. Dokl.* **1989**, *38*, 161–165.
28. Reyman, A.G. Integrable Hamiltonian Systems connected with graded Lie algebras. *J. Sov. Math.* **1982**, *19*, 1507–1545. [CrossRef]
29. Perelomov, A.M. *Integrable Systems of Classical Mechanics and Lie Algebras*; Birkhauser: Basel, Switzerland, 1990; Volume 1.
30. Bogoyavlenski, O. New Integrable Problem of Classical Mechanics. *Comm. Math. Phys.* **1984**, *94*, 255–269. [CrossRef]
31. Jurdjevic, V. Affine quadratic problem on Lie groups. *J. Lie Theory* **2020**, *20*, 425–444. [CrossRef]
32. Manakov, S.V. Note on the integration of Euler's equations of the dynamics of an n dimensional rigid body. *Funct. Anal. Appl.* **1976**, *10*, 328–329. [CrossRef]