Symmetries and Integrals of Motion in Optimal Control [1]

H. J. SUSSMANN[2], Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08903.

1 INTRODUCTION

A very important result in classical mechanics is *Noether's Theorem*, according to which, whenever a symplectic action of a Lie group G consists of symmetries of a Hamiltonian H, then every element of the Lie algebra of G gives rise to an *integral of motion*, i.e. a function that is conserved along trajectories of the Hamiltonian vector field $\vec{H}$ associated to H. This is often a very useful tool for solving the equations of motion, because finding integrals of motion makes it possible to reduce the dimension of the system and, in extreme cases for which the group of symmetries is very large, solve it completely.

In this note we state and prove a control-theory version of Noether's theorem, which says that every one-parameter group of symmetries of a control system gives rise to a conservation law that is valid along biextremals. Equivalently, if M denotes the state space of the system and G is a Lie group that acts as a group of symmetries, then the "momentum map" μ from the cotangent bundle T^*M to the dual $L(G)^*$ of the Lie algebra of G is a vector-valued integral of motion along biextremals. (The precise definition of "biextremals" is given in Section 4 below. They are pairs consisting of a trajectory of the system and an extremal adjoint vector, i.e. an "adjoint vector that minimizes the Hamiltonian." In control theory, these curves are sometimes known as "extremals," or "Pontryagin extremals," although there is some confusion in the literature as to whether an "extremal" is a trajectory of the system for which there

[1] Part of this work was done while the author was a visitor at Institute for Mathematics and its Applications (I.M.A.), University of Minnesota, Minneapolis, Minnesota 55455, U.S.A.

[2] Partially supported by NSF grant DMS92-02554.

exists an extremal adjoint vector, or a pair consisting of a trajectory and an extremal adjoint vector. We choose to make precise the distinction between these two concepts, by reserving the word "extremal" for the former, and using "biextremal" for the latter.)

As in the case of classical mechanics, integrals of motion can be used to reduce the dimension of the system of equations that characterize the biextremals, and in some cases yield very detailed information on their structure. This phenomenon is illustrated in Section 5 by analyzing the solution of the "Markov-Dubins problem" in dimension 3, which exhibits some interesting features, such as the simultaneous occurrence of two families of optimal arcs, one of which consists of smooth arcs, while the other one consists of nonsmooth arcs whose nonsmoothness has a special structure, corresponding to the existence of a conserved quantity that is the product of two functions that are not separately conserved.

The paper is organized as follows: in Section 2 we spell out our notational conventions regarding manifolds and Hamiltonians, and review various versions of the classical Noether Theorem, including a detailed discussion of the case when the Hamiltonian H is only locally Lipschitz. In Section 3 we extend Noether's Theorem to a result about minimizing trajectories of a family of Hamiltonians of class C^1, and explain why this result does not extend to families of Lipschitz Hamiltonians. In Section 4 we introduce the relevant definitions from control theory and state and prove the control theory version of Noether's Theorem. Finally, in Section 5 we apply our result to the study of the Markov-Dubins problem.

2 REVIEW OF NOETHER'S THEOREM

Throughout this paper, the word "smooth" will always mean "of class C^∞." A *manifold* is always, by definition, smooth, finite-dimensional, Hausdorff, second countable, and without boundary. If M is a manifold, we use TM, T^*M to denote, respectively, the tangent and cotangent bundles of M. If $x \in M$, then T_xM, T_x^*M denote, respectively, the tangent and cotangent spaces of M at x. Vector fields and 1-forms are continuous by definition, and all additional smoothness requirements will be explicitly indicated. We use $C^k(M)$ to denote the space of all real-valued functions of class C^k on M, and $\Gamma^k(E)$ to denote the space of sections of class C^k of a bundle E, so for example $\Gamma^k(TM)$ and $\Gamma^k(T^*M)$ are the spaces of vector fields and 1-forms of class C^k on M. In particular, $\Gamma^\infty(TM)$, equipped with the usual Lie bracket operation for vector fields, is a Lie algebra.

The words "Lie algebra" mean "real, not necessarily finite-dimensional, Lie algebra" If L is a Lie algebra, then L^* denotes the algebraic dual of L. A *smooth action* of a Lie algebra L on a manifold M is a Lie algebra homomorphism θ from L to $\Gamma^\infty(TM)$.

A *symplectic manifold* is a manifold N equipped with a smooth closed nonsingular 2-form Ω.

From now on, N will always denote a symplectic manifold, and Ω will denote the symplectic form of N.

If $k \geq 1$, then to every function $H \in C^k(N)$ there is associated a vector field $\vec{H} \in \Gamma^{k-1}(TN)$ —the *Hamilton vector field* of H— characterized by the property that $\Omega(\vec{H}, X) = \langle dH, X \rangle$ for all $X \in \Gamma^0(TN)$. A vector field $X \in \Gamma^0(TN)$ is *Hamiltonian* if $X = \vec{H}$ for some function $H \in C^1(N)$, and *locally Hamiltonian* if every $x \in N$ has a neighborhood U on which X is Hamiltonian. We use $\Gamma^k_{ham,loc}(TN)$ to denote the set of all locally Hamiltonian vector fields on N, and $\Gamma^k_{ham}(TN)$ for the set of those $Y \in \Gamma^k_{ham,loc}(TN)$ that are globally Hamiltonian. If $H, K \in C^1(N)$, then the *Poisson bracket* $\{H, K\}$ is defined to be the function $\vec{H}K$, i.e. the derivative of K along the integral curves of $\vec{H}$. The Poisson bracket is skew-symmetric (i.e. such that $\{H, K\} = -\{K, H\}$ for all $H, K \in C^1(N)$), and satisfies the *Jacobi identity* $\{H_1, \{H_2, H_3\}\} = \{\{H_1, H_2\}, H_3\} + \{H_2, \{H_1, H_3\}\}$ whenever the H_i belong to $C^2(N)$. This implies that $C^\infty(N)$, equipped with the Poisson bracket, is a Lie algebra. Moreover, if $H_1, H_2 \in C^\infty(N)$ and $K = \{H_1, H_2\}$, then $\vec{K} = [\vec{H_1}, \vec{H_2}]$, so the map $H \to \vec{H}$ from $C^\infty(N)$ to $\Gamma^\infty(TN)$ is a Lie-algebra homomorphism, whose image is obviously $\Gamma^\infty_{ham}(TN)$. It follows that both $\Gamma^\infty_{ham}(TN)$ and $\Gamma^\infty_{ham,loc}(TN)$ are Lie subalgebras of $\Gamma^\infty(TN)$.

More generally, Ω gives rise to a bijective correspondence $J_\Omega : \Gamma^0(T^*N) \to \Gamma^0(TN)$ between 1-forms and vector fields on N. The vector field $J_\Omega(\xi)$ that corresponds to a 1-form $\xi \in \Gamma^0(T^*N)$ is characterized by the fact that $\Omega(J_\Omega(\xi), X) = \langle \xi, X \rangle$ for all $X \in \Gamma^0(TN)$. The Hamiltonian vector fields are precisely those that correspond to exact 1-forms, and the locally Hamiltonian ones are those associated to closed 1-forms. (A continuous 1-form ξ is *closed* if it is locally exact.) The locally Hamiltonian vector fields can also be characterized as infinitesimal generators of one-parameter groups of symplectic diffeomorphisms, provided that they are smooth enough for these groups to exist. More precisely, for a C^1 vector field X on N, the following conditions are equivalent: (i) $J_\Omega^{-1}(X)$ is closed, (ii) X is locally Hamiltonian, (iii) $\mathcal{L}_X \Omega \equiv 0$, (iv) $(e^{tX})^*(\Omega) = \Omega$ for all $t \in \mathbb{R}$. (Here $\mathcal{L}_X$ denotes Lie differentiation in the direction of X, $\{e^{tX}\}_{t \in \mathbb{R}}$ is the flow of X, $(e^{tX})^*(\Omega)$ denotes the pullback of Ω by the map e^{tX}, and the equality $(e^{tX})^*(\Omega) = \Omega$ is understood to hold on the domain of e^{tX}, which need not be the whole manifold N, since X need not be complete.) If X is only of class C^0, then (i), (ii) and (iii) are still equivalent. (Recall that the Lie derivative $\mathcal{L}_X T$ is always well defined as a distributional tensor field if T is a smooth tensor field and X is a distributional vector field.) However, (iv) need not make sense, since X may fail to have unique trajectories.

The above characterization implies that to every local smooth one-parameter group $\{g_t\}$ of symplectic diffeomorphisms of N there correspond, locally, smooth functions h such that $g_t = \mathrm{e}^{t\vec{h}}$. To make this completely precise, we should establish a bijective correspondence between $\Gamma^k_{ham,loc}(TN)$ and the set of sections of an appropriate sheaf

$\mathcal{S}^k$. (The correct choice of $\mathcal{S}^k$ is the sheaf of germs of functions of class C^{k+1} modulo the constant germs.) In the specific example of interest to us, we can avoid general sheaf language, and define the correspondence directly, as follows. We define a *local function* on N to be a continuous real-valued function K whose domain $\mathrm{Dom}(K)$ is an open subset of N. We then associate to each vector field $Y \in \Gamma^0(N)$ the collection $\nu(Y)$ of all local functions K such that $K \in C^1(\mathrm{Dom}(K))$ and $\vec{K} = Y$ on $\mathrm{Dom}(K)$. Then $Y \in \Gamma^0_{ham,loc}(N)$ if and only if the domains of the local functions in $\nu(Y)$ cover N.

The identity $\{H, K\} = -\{K, H\}$, says that $\{H, K\}$ is also equal to minus the derivative of H along the integral curves of $\vec{K}$. This implies, in particular, that

(I) *if H, K are functions of class C^1 on N such that H is constant along all trajectories of $\vec{K}$, then K is constant along all trajectories of $\vec{H}$.*

Properly reinterpeted, the above statement remains true if H is only locally Lipschitz, as we now show. A locally Lipschitz function H has a well defined *generalized gradient* ∂H in the sense of F. Clarke (cf. [3]). By definition, ∂H is a set-valued map that assigns to each $x \in N$ the subset $\partial H(x)$ of $T^*_x N$ defined as follows: we let $\mathcal{D}(H)$ denote the set of all points $y \in N$ such that H is differentiable at y, and let $\partial H(x)$ be the convex hull of the set of all $z \in T^*_x N$ that can be expressed as a limit $z = \lim_{j\to\infty} dH(x_j)$, for some sequence $\{x_j\}$ in $\mathcal{D}(H)$ such that $x_j \to x$ as $j \to \infty$. (The fact that many such sequences exist follows from Rademacher's Theorem, according to which H is differentiable almost everywhere.) Then it is well known that $\partial H(x)$ is a nonempty compact convex subset of $T^*_x N$. We then define $\vec{H}(x) = J_\Omega^{-1}(\partial H(x))$, so $\vec{H}(x)$ is a nonempty, compact convex subset of $T_x N$. A *trajectory* of $\vec{H}$ is now a solution of the differential inclusion $\dot{x} \in \vec{H}(x)$, i.e. a locally absolutely continuous curve $\gamma : I \to N$, defined on an interval $I \subseteq \mathbb{R}$, such that $\dot{\gamma}(t) \in \vec{H}(\gamma(t))$ for almost all $t \in I$.

Now, if H is constant along trajectories of K, then $\langle dH(x), \vec{K}(x)\rangle = 0$ whenever $x \in \mathcal{D}(H)$, because $\langle dH(x), \vec{K}(x)\rangle = \frac{d}{dt}\Big|_{t=0} \Big(H(\gamma(t))\Big)$, where γ is any integral curve of $\vec{K}$ such that $\gamma(0) = x$. If $x \in N$, and $\{x_j\}$ is a sequence in $\mathcal{D}(H)$ such that $x_j \to x$ and $\{dH(x_j)\}$ converges to a limit $z \in T^*_x N$, then $\langle z, \vec{K}(x)\rangle = 0$, since $\vec{K}$ is continuous. Therefore $\langle z, \vec{K}(x)\rangle = 0$ for all $z \in \partial H(x)$. So $\Omega(v, \vec{K}(x)) = 0$ for all $v \in \vec{H}(x)$, so that $vK = 0$ whenever $v \in \vec{H}(x)$. If γ is a trajectory of $\vec{H}$, it follows that the derivative of the function $t \to K(\gamma(t))$ vanishes for almost all t. So K is constant along γ, since $K \circ \gamma$ is locally absolutely continous.

So we have shown that

(II) *if H is a locally Lipschitz function on N, and K is a function of class C^1 on N such that H is constant along all trajectories of $\vec{K}$, then K is constant along all trajectories of $\vec{H}$.*

From now on, we will be primarily interested in functions K of class C^∞. For an arbitrary function $H : N \to \mathbb{R}$, define Λ_H (or $\Lambda(N, \Omega, H)$ if we wish to spell out the explicit dependence on N and Ω) to be the set of all $K \in C^\infty(N)$ such that H is constant along all integral curves of $\vec{K}$. We then have

Proposition 2.1. *Λ_H is a Lie subalgebra of $C^\infty(N)$.*

PROOF. If $H \in C^2(N)$, then the conclusion follows easily from the basic properties of the Poisson bracket, since $K \in \Lambda_H$ if and only if $\{K, H\} = 0$. If H is an arbitrary function, then we cannot use Poisson brackets, so we use instead the orbit teorem of [10]. According to this result, if S is any set of vector fields on a smooth manifold M, then M is partitioned into connected, immersed (but not necessarily embedded) submanifolds —the *S-orbits*— such that (a) two points x_1, x_2 of M belong to the same S-orbit if and only if x_2 can be reached from x_1 by means of a finite concatenation of integral curves of members of S, and (b) every integral curve γ of an $X \in S$ is smooth as a map into the S-orbit $\mathcal{O}$ that contains γ. From these properties it follows that every $X \in S$ is tangent to all the S-orbits. If we let $L(S)$ denote the Lie algebra of vector fields generated by S, then every $X \in L(S)$ is tangent to all the S-orbits. Since the S-orbits constitute a partition of M, it follows that every integral curve of an $X \in L(S)$ is entirely contained in an S-orbit. Now, if $\varphi : M \to \mathbb{R}$ is an arbitrary function which is constant along all integral curves of members of S, then Property (a) implies that φ is constant on every S-orbit, so φ is constant along all integral curves of members of $L(S)$. If we apply this with $M = N$, $S = \Lambda_H$, $\varphi = H$, we find that $L(\Lambda_H) \subseteq \Lambda_H$, and our conclusion follows. □

We define a map $\mu_H : N \to \Lambda_H^*$ by letting

$$\mu_H(x)(K) = K(x) \quad \text{for} \quad K \in \Lambda_H \,, x \in N \,. \tag{2.1}$$

Then (II) implies

(III) *if H is a locally Lipschitz function on N, then $\mu_H(x)$ is constant along all trajectories of $\vec{H}$.*

The above results can be stated in terms of infinitesimal symmetries, by using the relationship between Hamiltonian vector fields and one-parameter groups of sympletic diffemorphisms. We define an *infinitesimal symmetry* of H to be a smooth vector field Y on N such that (i) the flow maps e^{tY} arising from Y are symplectic (i.e. $\left(\mathrm{e}^{tY}\right)^*(\Omega) = \Omega$ on the domain of e^{tY} for all $t \in \mathbb{R}$), and (ii) H is constant along all integral curves of Y. We use L_H to denote the set of all infinitesimal symmetries of H. It is clear that a vector field $Y \in \Gamma^\infty(TN)$ is in L_H if and only if it is locally Hamiltonian and satisfies (ii). Therefore, if $Y \in L_H$ and $K \in \nu(Y)$, then H is constant along all integral curves of $\vec{K}$. So (II) applies, and we get

(IV) *if H is a locally Lipschitz function on N, and Y is an infinitesimal symmetry of H, then (i) every $x \in N$ belongs to the domain of a local function K such that $K \in \nu(Y)$, and (ii) every local function $K \in \nu(Y)$ is constant along all trajectories of $\vec{H}$ that are contained in* $\mathrm{Dom}(K)$.

Now suppose L is a real Lie algebra. A *symplectic action* of L on N is a Lie algebra homomorphism θ from L to $\Gamma^\infty_{ham,loc}(TN)$, i.e. a smooth action of L on N such that every vector field $\theta(X)$, $X \in L$, is locally Hamiltonian. A *local momentum map* for a symplectic action θ is a L^*-valued map μ, defined on an open subset V of N, such that for every $X \in L$ the function $\mu_X : \mathrm{Dom}(\mu) \to \mathbb{R}$ given by $\mu_X(x) = \mu(x)(X)$ is smooth on V and satisfies the identity $\theta(X) = \vec{\mu}_X$ throughout V. (Equivalently, a local momentum map is a way of assigning to each $X \in L$ a smooth function μ_X on some fixed open set V, such that $\vec{\mu}_X = \theta(X)$ on V, and μ_X depends linearly on X.) It is clear that every point $x \in N$ belongs to the domain of a local momentum map, since we can always choose V to be a neighborhood of x which is diffeomorphic to a ball, and define μ_X to be the unique smooth function φ on V such that $\varphi(x) = 0$ and $\vec{\varphi} = \theta(X)$ on V.

An *infinitesimal group of symmetries* of a function $H : N \to \mathbb{R}$ is a symplectic action θ on N of a real Lie algebra L, such that H is constant along all integral curves of the vector fields $\theta(X)$, $X \in L$. Then

(V) *if H is a locally Lipschitz function on N, and $\theta : L \to L_H$ is an infinitesimal group of symmetries of H, then (i) every $x \in N$ belongs to the domain of a local momentum map μ, and (ii) every local momentum map μ is constant along all trajectories of $\vec{H}$ that are contained in* $\mathrm{Dom}(\mu)$.

There is one important situation when there exists a canonical, globally defined momentum map. This is the case when N is the cotangent bundle T^*M of a smooth manifold M, and the symplectic action θ of L on N is the action $T^*(\tau) : L \to \Gamma^\infty_{ham}(T\,T^*M)$ on T^*M obtained by lifting to T^*M a smooth action τ of L on M. (Recall that every vector field $X \in \Gamma^\infty(TM)$ gives rise to a smooth function h_X on T^*M, defined by letting $h_X(x,z) = \langle z, X(x)\rangle$ for $x \in M$, $z \in T^*_xM$. It is easy to verify that the identity $\{h_X, h_Y\} = h_{[X,Y]}$ holds, so the map $X \to h_X$ is a Lie algebra homomorphism from $\Big(\Gamma^\infty(TM), [\cdot,\cdot]\Big)$ to $\Big(C^\infty(T^*M), \{\cdot,\cdot\}\Big)$. Also, the map $h \to \vec{h}$ is a Lie algebra homomorphism from $\Big(C^\infty(T^*M), \{\cdot,\cdot\}\Big)$ to $\Big(\Gamma^\infty_{ham}(T\,T^*M), [\cdot,\cdot]\Big)$. The action $T^*(\tau)$ is then defined by letting $T^*(\tau)(X) = \vec{h}_{\tau(X)}$.) In this case, the obvious choice of a momentum map is given by letting $\mu(x,z)(X) = h_{\theta(X)}(x,z)$, i.e. $\mu(x,z) = \langle z, \theta(X)(x)\rangle$, for $(x,z) \in T^*M$. We use μ^τ to denote this canonical momentum map, so μ^τ is an L^*-valued function on T^*M. Then (V) implies

(VI) *if M is a smooth manifold, $H : T^*M \to \mathbb{R}$ is a locally Lipschitz function, and $\tau : L \to \Gamma^\infty(TM)$ is a smooth action of a Lie algebra L on M such that the induced symplectic action $T^*(\tau)$ is an infinitesimal group of symmetries of H, then the momentum map μ^τ is constant along all trajectories of $\vec{H}$.*

Statements (I), (II), (III), (IV), (V) and (VI) are versions of Noether's Theorem.

3 MINIMIZING TRAJECTORIES OF HAMILTONIAN FAMILIES

We now present a form of Noether's Theorem for minimizing trajectories of a family of Hamiltonian functions. This will be the crucial ingredient needed for the statement and proof —in Section 4— of the control-theoretic analogue of Noether's Theorem.

If N is a symplectic manifold, a *Hamiltonian family* on N is a parametrized family $\mathcal{H} = \{H_u : u \in U\}$ of continuous real-valued functions on N. We call $\mathcal{H}$ *locally Lipschitz*, or *of class* C^k, if each H_u is locally Lipschitz, or of class C^k. If $\mathcal{H}$ is a Hamiltonian family of class C^1, then we define set-valued maps $\vec{\mathcal{H}}$, $\vec{\mathcal{H}}^{min}$, that assign to each $x \in N$ the subsets $\vec{\mathcal{H}}(x)$, $\vec{\mathcal{H}}^{min}(x)$, of T_xM given by

$$\vec{\mathcal{H}}(x) = \{\vec{H}_u(x) : u \in U\}\,. \tag{3.1}$$

$$\vec{\mathcal{H}}^{min}(x) = \{\vec{H}_u(x) : u \in U\,,\, H_u(x) = \min\{H_v(x) : v \in U\}\}\,. \tag{3.2}$$

A *trajectory* of a Hamiltonian family $\mathcal{H}$ of class C^1 is a locally absolutely continuous map $\gamma : I \to N$, defined on an interval $I \subseteq \mathbb{R}$, such that $\dot{\gamma}(t) \in \vec{\mathcal{H}}(\gamma(t))$ for almost all $t \in I$. A trajectory $\gamma : I \to N$ of $\mathcal{H}$ is *minimizing* if $\dot{\gamma}(t) \in \vec{\mathcal{H}}^{min}(\gamma(t))$ for almost every $t \in I$. We use $\mathcal{T}(\mathcal{H})$, $\mathcal{T}^{min}(\mathcal{H})$ to denote, respectively, the class of all trajectories of $\mathcal{H}$, and that of all minimizing trajectories of $\mathcal{H}$.

If $\mathcal{H} = \{H_u : u \in U\}$ and $\tilde{\mathcal{H}} = \{\tilde{H}_u : u \in \tilde{U}\}$ are two Hamiltonian families on N, we write $\mathcal{H} \subseteq \tilde{\mathcal{H}}$ if every function H_u also occurs in $\tilde{\mathcal{H}}$, i.e. if for every $u \in U$ there exists a $\tilde{u} \in \tilde{U}$ such that $H_u = \tilde{H}_{\tilde{u}}$. If V is an open subset of N, $\mathcal{H}$ is a Hamiltonian family on N, and $\rho : V \to N$ is a continuous map, then the *pullback* $\rho^*\mathcal{H}$ of $\mathcal{H}$ by ρ is the family $\{H_u \circ \rho : u \in U\}$. The *restriction* $\mathcal{H}\lceil V$ of $\mathcal{H}$ to V is the pullback of $\mathcal{H}$ by the inclusion map from V to N. A smooth diffeomorphism ρ from an open subset V_1 of N to another open subset V_2 is a *symmetry* of a Hamiltonian family $\mathcal{H}$ if $\rho^*\mathcal{H} \subseteq \mathcal{H}\lceil V_1$ and $(\rho^{-1})^*\mathcal{H} \subseteq \mathcal{H}\lceil V_2$.

An *infinitesimal symmetry* of a Hamiltonian family $\mathcal{H}$ on N is a smooth locally Hamiltonian vector field X on N such that the diffeomorphisms e^{tX} are symmetries of $\mathcal{H}$ for all $t \in \mathbb{R}$.

The following result is the analogue of Version (II) of Noether's Theorem for minimizing trajectories of Hamiltonian families:

Theorem 3.1. *Let N be a symplectic manifold and let $\mathcal{H} = \{H_u : u \in U\}$ be a Hamiltonian family of class C^1 on N. Let K be a smooth function on N such that $\vec{K}$ is an infinitesimal symmetry of $\mathcal{H}$. Then K is constant along every minimizing trajectory of $\mathcal{H}$.*

PROOF. Let $\gamma : I \to N$ be a minimizing trajectory of $\mathcal{H}$. Let $f(t) = K(\gamma(t))$. Then f is locally absolutely continuous, so our conclusion will follow if we show that the derivative of f vanishes almost everywhere. Let E be the set of those $t \in I$ such that $\dot{\gamma}(t)$ exists and belongs to $\vec{\mathcal{H}}^{min}(\gamma(t))$. Then E is of full measure in I. So it suffices to show that $f'(t) = 0$ for $t \in E$. Pick a $t \in E$. Let $p = \gamma(t)$. Pick $u \in U$ such that $\dot{\gamma}(t) = \vec{H}_u(p)$ and

$$H_u(p) = \min\{H_w(p) : w \in U\}\,. \tag{3.3}$$

Clearly, $f'(t) = (\vec{H}_u K)(p)$, i.e. $f'(t) = \{H_u, K\}(p)$. So we need to prove that $\{H_u, K\}(p) = 0$, and this will follow if we show that

$$\left.\frac{d}{ds}\right|_{s=0} H_u\left(\mathrm{e}^{s\vec{K}}(p)\right) = 0\,. \tag{3.4}$$

Now, since the map $\rho_s = \mathrm{e}^{s\vec{K}}$ is a symmetry of $\mathcal{H}$, there exists, for every s such that $\mathrm{e}^{s\vec{K}}(p)$ is defined —and in particular for every s in some interval $(-\delta, \delta)$, $\delta > 0$— a $v(s)$ such that $H_u \circ \rho_s = H_{v(s)}$ on the domain of ρ_s. This implies that $H_u(\rho_s(p)) = H_{v(s)}(p)$ for $s \in (-\delta, \delta)$. On the other hand, $H_{v(s)}(p) \geq H_u(p)$, because of (3.3). So $H_u(\rho_s(p)) \geq H_u(p)$ for $s \in (-\delta, \delta)$. Therefore the continuously differentiable function $(-\delta, \delta) \ni s \to H_u(\rho_s(p))$ has a minimum at $s = 0$. So (3.4) follows, and our proof is complete. □

Remark 3.1. An alternative proof of Theorem 3.1 is possible if $\mathcal{H}$ satisfies the additional requirement that the function $\mathcal{H}_{inf} : N \to [-\infty, \infty]$ given by

$$\mathcal{H}_{inf}(x) = \inf\{H_u(x) : u \in U\}$$

is everywhere finite and locally Lipschitz. Indeed, in that case it is easy to see that $\mathcal{H}_{inf}$ is constant along trajectories of $\vec{K}$, so Version (II) of Noether's Theorem implies that K is constant along every trajectory of $\vec{\mathcal{H}}_{inf}$. On the other hand, *every minimizing trajectory of $\mathcal{H}$ is a trajectory of $\vec{\mathcal{H}}_{inf}$.* (Indeed, suppose that $\gamma : I \to N$ and $\gamma \in \mathcal{T}^{min}(\mathcal{H})$. Let E be the set of those $t \in I$ such that $\dot{\gamma}(t)$ exists and is equal to $\vec{H}_u(\gamma(t))$ for some $u \in U$ for which $H_u(\gamma(t)) = \min\{H_w(\gamma(t)) : w \in U\}$. Then E is of full measure in I. Let $t \in E$, and write $p = \gamma(t)$. Then the function $x \to \psi(x) \stackrel{\text{def}}{=} H_u(x) - \mathcal{H}_{inf}(x)$ vanishes at $x = p$, and is everywhere nonnegative. Therefore ψ has a local minimum at p. Since ψ is locally Lipschitz, it follows from standard properties of the generalized gradient that $0 \in \partial\psi(p)$. Moreover, $\partial\psi(p) = \{dH_u(p) - v : v \in \partial\mathcal{H}_{inf}(p)\}$. So $dH_u(p) \in \partial\mathcal{H}_{inf}(p)$. Therefore $\vec{H}_u(p) \in \vec{\mathcal{H}}_{inf}(p)$. So $\dot{\gamma}(t) = \vec{H}_u(\gamma(t)) \in \vec{\mathcal{H}}_{inf}(\gamma(t))$. Since this is true for every $t \in E$, γ is a trajectory of $\vec{\mathcal{H}}_{inf}$, as stated.) Combining these two observations, we see that K is constant along every minimizing trajectory of $\mathcal{H}$. □

Remark 3.2. If $\mathcal{H}$ is a locally Lipschitz Hamiltonian family, then we can still define a set-valued map $\vec{\mathcal{H}}$, by letting

$$\vec{\mathcal{H}}(x) = \bigcup \{\vec{H}_u(x) : u \in U\}\,. \tag{3.5}$$

One can then define *trajectories*, and *minimizing trajectories*, exactly as in the C^1 case. However, *Theorem 3.1 need not be true if the Hamiltonians H_u are only locally Lipschitz, even if the minimized Hamiltonian $\mathcal{H}_{inf}$ is smooth.* For an example of this, let N be an arbitrary symplectic manifold, and let U be the set of all nonnegative locally Lipschitz real-valued functions on N. Then define $\mathcal{H} = \{H_u : u \in U\}$ by just letting $H_u = u$. It is clear that $\mathcal{H}_{inf} \equiv 0$, which is obviously smooth. Moreover, given any $\bar{x} \in N$ and any vector $v \in T_{\bar{x}}N$, it is easy to show that there exists $u \in U$ such that $u(\bar{x}) = 0$ and $v \in \vec{u}(\bar{x})$. (It suffices to construct a nonegative locally Lipschitz function $w : N \to \mathbb{R}$ that, for some chart $x \to \kappa(x)$ near $\bar{x}$ such that $\kappa(\bar{x}) = 0$, coincides with $||\kappa(x)||$ near $\bar{x}$. Then $\partial w(\bar{x})$ is a neigborhood of 0 in $T^*_{\bar{x}}N$, so $\vec{w}(\bar{x})$ is a neigborhood of 0 in $T_{\bar{x}}N$. It follows that, if $r > 0$ is sufficiently large, and we take $u = rw$, then $v \in \vec{u}(\bar{x})$.) This implies that every absolutely continuous curve in N is a minimizing trajectory of $\mathcal{H}$. On the other hand, if K is an arbitrary smooth function on N, then $\vec{K}$ is an infinitesimal symmetry of $\mathcal{H}$. So, if Theorem 3.1 was true for $\mathcal{H}$, it would follow that every smooth function on N is constant along every absolutely continuous curve on N, a fact that is obviously false if $\dim N > 0$. □

Remark 3.3. The preceding example shows, in particular, that when $\mathcal{H}$ is a locally Lipschitz family, it need not be true that all minimizing trajectories of $\mathcal{H}$ are trajectories of $\mathcal{H}_{inf}$, even when $\mathcal{H}_{inf}$ is smooth. This explains why the proof of Theorem 3.1 given in Remark 3.1 fails for locally Lipschitz families. More precisely, the step where continuous differentiability is crucial is the passage from "$0 \in \partial\psi(p)$" to "$\dot{\gamma}(t) \in \vec{\mathcal{H}}_{inf}(\gamma(t))$," which depends on the fact that $\vec{H}_{u(t)}(\gamma(t))$ consists of a single point. □

Naturally, Theorem 3.1 has several corollaries involving momentum maps. We summarize these in the following

Theorem 3.2. *Let N be a symplectic manifold and let $\mathcal{H} = \{H_u : u \in U\}$ be a Hamiltonian family of class C^1 on N. Let L be a Lie algebra, and let $\theta : L \to \Gamma^\infty_{ham,loc}(TN)$ be a smooth symplectic action of L on N such that every $\theta(X)$, $X \in L$, is an infinitesimal symmetry of $\mathcal{H}$. Then: (i) every $x \in N$ belongs to the domain of a local momentum map μ of θ, and (ii) every local momentum map μ is constant along all minimizing trajectories of $\mathcal{H}$ that are contained in* $\mathrm{Dom}(\mu)$. *In particular, if $N = T^*M$, where M is a manifold, and $\tau : L \to \Gamma^\infty(TM)$ is a smooth action of a Lie algebra L on M such that the induced symplectic action $T^*(\tau)$ is an infinitesimal group of symmetries of $\mathcal{H}$, then the momentum map μ^τ is constant along all minimizing trajectories of $\mathcal{H}$.* □

4 AN OPTIMAL CONTROL VERSION OF NOETHER'S THEOREM

We now apply Theorems 3.1 and 3.2 to optimal control problems.

First, we define a *vector field system* on a manifold M to be a parametrized family $f = \{f_u : u \in U\}$ of vector fields on M. We call such a family *locally Lipschitz*, or *of class* C^k, if every vector field f_u is locally Lipschitz, or of class C^k. The set U is called the *control space* of f. A *Lagrangian* on M is a family $f^0 = \{f^0_u : u \in U\}$ of continuous functions on M. Again, the set U is the *control space* of f^0, and we call f^0 *locally Lipschitz*, or *of class* C^k, if every function f^0_u is locally Lipschitz or of class C^k.

An *optimal control system* is a 4-tuple $\mathcal{P} = (M, U, f, f^0)$, where M is a smooth manifold, U is a set, f is a vector field system on M with control space U, and f^0 is a Lagrangian on M with control space U. A *control* for $\mathcal{P}$ is a U-valued map $t \to \eta(t)$, defined on a compact subinterval I of $\mathbb{R}$. To a control η we associate a time-varying vector field $M \times I \ni (x,t) \to f^\eta(x,t) \in T_xM$ by letting $f^\eta(x,t) = f_{\eta(t)}(x)$, and a function $M \times I \ni (x,t) \to f^{0,\eta}(x,t) \in \mathbb{R}$ given by $f^{0,\eta}(x,t) = f^0_{\eta(t)}(x)$. A *trajectory* for a control $\eta : I \to U$ is an absolutely continuous curve $\gamma : I \to M$ such that $\dot\gamma(t) = f^\eta(\gamma(t), t)$ for almost all $t \in I$. If $I = [a,b]$, then we write $t_-(\gamma) \stackrel{\text{def}}{=} a$, $t_+(\gamma) \stackrel{\text{def}}{=} b$, $x_-(\gamma) \stackrel{\text{def}}{=} \gamma(a)$, $x_+(\gamma) \stackrel{\text{def}}{=} \gamma(b)$, and we refer to $t_-(\gamma)$, $t_+(\gamma)$, $x_-(\gamma)$, $x_+(\gamma)$, respectively, as the *initial time*, *terminal time*, *initial state* and *terminal state* of γ. A *controlled trajectory* of $\mathcal{P}$ is a pair (γ, η) such that η is a control and γ is a trajectory for η.

For a control η, we define $F^\eta(x,t) = (f^\eta(x,t), f^{0,\eta}(x,t))$. A control $\eta : I \to U$ is *$\mathcal{P}$-admissible* if the map F^η satisfies the following *Lipschitz-Carathéodory conditions*: (i) $F^\eta(x,t)$ is Lebesgue measurable with respect to t for each fixed x, and locally Lipschitz with respect to x for each fixed t, (ii) for each compact subset K of the domain $\mathrm{Dom}(\kappa)$ of a coordinate chart $\kappa = (x^1, \ldots, x^n)$ of M there exists an integrable function $\varphi : I \to \mathbb{R}$ such that, if we let $f^\eta(x,t) = \sum_{j=1}^n f^{j,\kappa,\eta}(x,t)\frac{\partial}{\partial x^j}$, for $(x,t) \in \mathrm{Dom}(\kappa) \times I$, and write $f^{0,\kappa,\eta}(x,t) = f^{0,\eta}(x,t)$, then the inequalities $|f^{j,\kappa,\eta}(x,t)| \le \varphi(t)$ and $|f^{j,\kappa,\eta}(x,t) - f^{j,\kappa,\eta}(y,t)| \le \varphi(t)||\kappa(x) - \kappa(y)||$ hold for $(x,y,t) \in K \times K \times I$, $j = 0, \ldots, n$.

An *admissible controlled trajectory* of $\mathcal{P}$ is a controlled trajectory (γ, η) such that η is admissible. We use $ACT(\mathcal{P})$ to denote the set of all controlled admissible trajectories of $\mathcal{P}$. The *cost* $C(\gamma,\eta)$ of a $(\gamma,\eta) \in ACT(\mathcal{P})$ is given by

$$C(\gamma,\eta) \stackrel{\text{def}}{=} \int_{t_-(\gamma)}^{t_+(\gamma)} f^{0,\eta}(\gamma(t),t)dt\,. \tag{4.1}$$

A $(\gamma,\eta) \in ACT(\mathcal{P})$ is *optimal* if $C(\gamma,\eta) \le C(\hat\gamma,\hat\eta)$ for all $(\hat\gamma,\hat\eta) \in ACT(\mathcal{P})$ such that $x_-(\hat\gamma) = x_-(\gamma)$ and $x_+(\hat\gamma) = x_+(\gamma)$, and *fixed-time optimal* if $C(\gamma,\eta) \le C(\hat\gamma,\hat\eta)$ for all $(\hat\gamma,\hat\eta) \in ACT(\mathcal{P})$ that satisfy $x_-(\hat\gamma) = x_-(\gamma)$, $x_+(\hat\gamma) = x_+(\gamma)$ and have the same duration as (γ,η), i.e. are such that $t_+(\hat\gamma) - t_-(\hat\gamma) = t_+(\gamma) - t_-(\gamma)$.

To an optimal control system $\mathcal{P} = (M, U, f, f^0)$ we associate, for each real number c, a Hamiltonian family $\mathcal{H}^c(\mathcal{P})$ on T^*M, by letting $\mathcal{H}^c(\mathcal{P}) = \{H_u^{c,\mathcal{P}} : u \in U\}$, where

$$H_u^{c,\mathcal{P}}(x,z) = \langle z, f_u(x)\rangle + cf_u^0(x) \quad \text{for} \quad x \in M\,,\ z \in T^*M\,. \tag{4.2}$$

Then it is clear that, if f and f^0 are locally Lipschitz, or of class C^k, the same is true of the Hamiltonian families $\mathcal{H}^c(\mathcal{P})$. If $(\gamma,\eta) \in ACT(\mathcal{P})$, an *extremal adjoint vector*, or *multiplier*, along (γ,η) is a pair (ζ,c) such that $c \in \mathbb{R}$, $c \geq 0$, and the following three conditions are satisfied:

(EAV1) ζ is an absoutely continuous field of covectors along γ, i.e. a map that assigns to each $t \in [t_-(\gamma), t_+(\gamma)]$ a covector $\zeta(t) \in T^*_{\gamma(t)}M$;

(EAV2) the curve $t \to \Xi(t) = (\gamma(t), \zeta(t))$ satisfies the "adjoint equation"

$$\dot{\Xi}(t) \in \vec{H}^{c,\mathcal{P}}_{\eta(t)}\Big(\Xi(t)\Big) \tag{4.3}$$

and the "Hamiltonian minimization condition"

$$H^{c,\mathcal{P}}_{\eta(t)}(\Xi(t)) = \min\Big\{H_u^{c,\mathcal{P}}(\Xi(t)) : u \in U\Big\} = \text{constant} \tag{4.4}$$

for almost all $t \in I$;

(EAV3) $(\zeta(t), c) \neq (0,0)$ for all $t \in I$.

A triple (γ,ζ,η) such that $(\gamma,\eta) \in ACT(\mathcal{P})$ and (ζ,c) is an extremal adjoint vector along (γ,η) will be called a *controlled c-biextremal* of $\mathcal{P}$. If (γ,ζ,η) is a controlled c-biextremal for some c, then we will say that (γ,ζ,η) is a *controlled biextremal.* A $(\gamma,\eta) \in ACT(\mathcal{P})$ will be called a *controlled extremal* if there exists an extremal adjoint vector along (γ,η), i.e. if there exist ζ such that (γ,ζ,η) is a controlled biextremal. A curve $\Xi = (\gamma,\zeta)$ in T^*M is a *biextremal* if there exist η such that (γ,ζ,η) is a controlled biextremal.

The Pontryagin Maximum Principle (cf. [1], [2], [3], [5], [6], [8], [13]) says that (a) if $(\gamma,\eta) \in ACT(\mathcal{P})$ is fixed-time optimal, then it is a controlled extremal, and (b) if in addition (γ,η) is optimal, then the multiplier (ζ,c) can be chosen so that the constant of (4.4) vanishes.

It is clear that, if (γ,ζ,η) is a controlled c-biextremal, and $r > 0$, then $(\gamma, r\zeta, \eta)$ is a controlled rc-biextremal. In particular, we can restate the first part of the Maximum Principle by restricting c to be either 0 or 1: if $(\gamma,\eta) \in ACT(\mathcal{P})$ is fixed-time optimal, then there exist c, ζ such that (γ,ζ,η) is a controlled c-biextremal and $c = 0$ or $c = 1$.

If (γ,ζ,η) is a controlled c-biextremal, then (γ,ζ) is a minimizing trajectory of the Hamiltonian family $\mathcal{H}^c(\mathcal{P})$. This means that Theorem 3.1 is applicable in any situation where we can produce interesting infinitesimal groups of symmetries of $\mathcal{H}^c(\mathcal{P})$. We now explain how such symmetries of the Hamiltonian families $\mathcal{H}^c(\mathcal{P})$ associated to an optimal control system $\mathcal{P}$ arise from symmetries of $\mathcal{P}$ itself.

A *symmetry* of an optimal control system $\mathcal{P} = (M, U, f, f^0)$ is a diffeomorphism $\sigma : V_1 \to V_2$, where V_1, V_2 are open subsets of M, such that for every $u \in U$ there exist $u_1, u_2 \in U$ for which (i) $d\sigma(x)(f_u(x)) = f_{u_1}(\sigma(x))$ and $d\sigma(x)(f_{u_2}(x)) = f_u(\sigma(x))$ for all $x \in V_1$, and (ii) $f^0_{u_1} \circ \sigma = f^0_u$ and $f^0_u \circ \sigma = f^0_{u_2}$. An *infinitesimal group of symmetries* of $\mathcal{P}$ is a smooth action $\tau : L \to \Gamma^\infty(TM)$ on M of a Lie algebra L, such that every diffeomorphism $\mathrm{e}^{t\tau(X)}$, $X \in L$, is a symmetry of $\mathcal{P}$.

To an infinitesimal group of symmetries $\tau : L \to \Gamma^\infty(TM)$ of an optimal control system $\mathcal{P} = (M, U, f, f^0)$ we associate the *momentum map* $\mu^\tau : T^*M \to L^*$, given by $\mu^\tau(x, z)(X) = \langle z, \tau(X)(x)\rangle$.

We are now ready to state and prove the *control theory version of Noether's Theorem*:

Theorem 4.1. *Assume that $\mathcal{P} = (M, U, f, f^0)$ is an optimal control system of class C^1, L is a Lie algebra, and $\tau : L \to \Gamma^\infty(TM)$ is an infinitesimal group of symmetries of $\mathcal{P}$. Let Ξ be a biextremal of $\mathcal{P}$. Then the function $\mu^\tau : T^*M \to L^*$ is constant along Ξ, that is, all the scalar Hamiltonian functions $(x, z) \to \langle z, \tau(X)(x)\rangle$, $X \in L$, are constant along Ξ.*

PROOF. Let $\Xi = (\gamma, \zeta)$, and assume that η, c are such that (γ, ζ, η) is a controlled c-biextremal. Then the adjoint equation (4.3) and the minimization condition (4.4) say that Ξ is a minimizing trajectory of $\mathcal{H}^c(\mathcal{P})$. On the other hand, it is easy to see that $T^*(\tau)$ is a infinitesimal group of symmetries of $\mathcal{H}^c(\mathcal{P})$. So we can apply Theorem 3.2, and our conclusion follows. □

5 SHORTEST PATHS WITH A CURVATURE BOUND

As an illustration of the use of the Maximum Principle in conjunction with the control theory version of Noether's Theorem, we outline some of the basic steps of our recently obtained solution of the three-dimensional version of the "Markov-Dubins-Reeds-Shepp problem," posed by Markov in [7], and solved completely in 1957 by L. Dubins in [4] for the two-dimensional case. (Cf. also [9] for the solution of a related problem, and [11] for a detailed treatment of both problems using optimal control. The three-dimensional problem was open until May 1992.)

The problem is that of characterizing the shortest curves $t \to x(t) \in \mathbb{R}^3$ of class C^1 that are parametrized by arc length, satisfy a curvature bound $||\ddot{x}|| \leq 1$, and go from a given initial position and velocity to a given terminal position and velocity. Precisely, let $\mathcal{A}_0$ be the class of all curves $t \to x(t)$ that are defined on some compact interval, and satisfy (i) $x(\cdot)$ is of class C^1, (ii) $||\dot{x}(t)|| = 1$ for all t, (iii) $\dot{x}$ is absolutely continuous, and (iv) $||\ddot{x}(t)|| \leq 1$ for all t. For $\bar{x}$, $\hat{x}$ in $\mathbb{R}^3$, and $\bar{y}$, $\hat{y}$ unit vectors in $\mathbb{R}^3$, let $\mathcal{A}_0(\bar{x}, \bar{y}, \hat{x}, \hat{y})$ be the set of all those $x(\cdot) \in \mathcal{A}_0$ that start at $\bar{x}$ with velocity vector $\bar{y}$ and end at $\hat{x}$ with velocity vector $\hat{y}$. Let $\mathcal{M}_0(\bar{x}, \bar{y}, \hat{x}, \hat{y})$ be the set of those curves in $\mathcal{A}_0(\bar{x}, \bar{y}, \hat{x}, \hat{y})$ that are

of minimum length (i.e. shorter than any other curve in $\mathcal{A}_0(\bar{x},\bar{y},\hat{x},\hat{y})$). Let $\mathcal{M}_0$ be the union of the sets $\mathcal{M}_0(\bar{x},\bar{y},\hat{x},\hat{y})$ over all pairs of initial and terminal conditions. The problem is to characterize $\mathcal{M}_0$.

It is more convenient to introduce the velocity y as a new variable, taking values in the unit sphere S^2 in $\mathbb{R}^3$, and to consider the class $\mathcal{A}$ of arcs of the form $t \to (x(t), \dot{x}(t))$, where $x(\cdot) \in \mathcal{A}_0$. Precisely, we let $M = \mathbb{R}^3 \times S^2$, and define $\mathcal{A}$ as the class of all arcs $t \to (x(t), y(t))$ such that x is of class C^1, y is absolutely continuous, $\dot{x} \equiv y$, and $||\dot{y}|| \leq 1$ almost everywhere. Clearly, x is then parametrized by arc-length since, by definition, y takes values in S^2. So the length of the arc x is the same as the *time*, i.e. the length of the interval $\mathrm{Dom}(x)$. We let $\mathcal{M}$ be the class of all arcs $(x,y) \in \mathcal{A}$ that minimize time among all arcs with the same initial and terminal conditions.

We can formulate this as a control problem of the kind considered in this paper by writing the dynamical equations as

$$\dot{x} = y \ , \ \ \dot{y} = y \times w \ , \tag{5.1}$$

where the control w is restricted to taking values in $\mathbb{B}^3$, the closed unit ball in $\mathbb{R}^3$. In other words, the control space U is $\mathbb{B}^3$, and the vector fields f_w, for $w \in U$, are given by $f_w(x,y) = (y, y \times w)$. (Notice that $y \times w$ is tangent to S^2, so $f_w(x,y)$ is tangent to M.) The Lagrangian $f^0 = \{f^0_w : w \in U\}$ is just given by $f^0_w \equiv 1$. The admissible controls then turn out to be the measurable functions $t \to w(t) \in \mathbb{B}^3$.

It is easy to see that the class of all trajectories of this system is exactly $\mathcal{A}$. To characterize the minimum-time trajectories, we use the Maximum Principle. The Hamiltonian is

$$H = \langle \lambda, y \rangle + \langle \mu, y \times w \rangle \ , \tag{5.2}$$

where we write a general covector z in $\mathbb{R}^3 \times S^2$ as the pullback of a covector $\tilde{z}$ in $\mathbb{R}^6$, i.e. in $\mathbb{R}^3 \times \mathbb{R}^3$, and write $\tilde{z}$ as (λ, μ), where λ and μ belong to $\mathbb{R}^3$. (Strictly speaking, we should have written $\langle \lambda, y \rangle + \langle \mu, y \times w \rangle + c$, where c is a constant, but c has no effect on the adjoint equation or the minimization condition, so we omit it. Notice, however, that c has to be ≥ 0, and $H + c$ must vanish, so H has to be ≤ 0.)

To apply the Hamiltonian minimization condition we first use the cross-product identity $\langle A, B \times C \rangle = \langle B, C \times A \rangle$ to rewrite H as $H = \langle \lambda, y \rangle + \langle w, \mu \times y \rangle$. It is then clear that the minimization of H is achieved by letting

$$w = -\frac{\mu \times y}{||\mu \times y||} \ \text{ if } \ \mu \times y \neq 0 \ . \tag{5.3}$$

If $\mu \times y = 0$, then any value of w in $\mathbb{B}^3$ is minimizing. We can capture all these possibilities by rewriting the minimization condition in the form

$$||\mu \times y|| w = -\mu \times y \ . \tag{5.4}$$

At this point, we have —at least as long as $\mu \times y \neq 0$— a closed, smooth system of ordinary differential equations for x, y, λ, and μ, consisting of our original dynamical equations, plus the adjoint equations, plus the formula for w. We could try to analyze this system directly and, in particular, try to understand how the solutions in the "smooth" region where $\mu \times y \neq 0$ connect up with those in $\mu \times y = 0$.

It turns out that it is much simpler to use Noether's Theorem to reduce the dimension of the system by finding conserved quantities. Our system is clearly invariant under the 6-dimensional group of rigid motions of $\mathbb{R}^3$. Consider first the action of the translations. For $v \in \mathbb{R}^3$, let τ_v be the translation $x \to x + v$. Then τ_v acts on the state and control variables of our system via $\tau_v(x, y, w) = (x+v, y, w)$. The infinitesimal generators of the action of the translations on M are the vector fields X_v given by $X_v(x, y) = (v, 0)$. The Hamiltonian function corresponding to X_v is $h_{X_v} = \langle \lambda, v \rangle$. By Noether's Theorem, this function is constant along biextremals. Since this is true for every vector v, it follows that the vector λ itself has to be constant.

The constancy of λ could have been ascertained directly from the adjoint equations, by observing that the Hamiltonian does not contain x explicitly, and therefore the adjoint equation for λ says that $\dot{\lambda} = 0$. A much more interesting conservation law is derived by using rotational invariance. A rotation matrix $R \in SO(3)$ acts on $M \times \mathbb{B}^3$ via $R(x, y, w) = (Rx, Ry, Rw)$. The infinitesimal generators are the skew-symmetric matrices $A \in so(3)$. To each such matrix there corresponds a vector field Y_A on M, given by $Y_A(x, y) = (Ax, Ay)$. The Hamiltonian h_{Y_A} corresponding to Y_A is given by $h_{Y_A} = \langle \lambda, Ax \rangle + \langle \mu, Ay \rangle$. So this quantity has to be conserved for every A. Now recall that the skew-symmetric transformations or $\mathbb{R}^3$ are exactly the maps of the form $u \to v \times u$, where $v \in \mathbb{R}^3$. So what we have shown is that the expression $\langle \lambda, v \times x \rangle + \langle \mu, v \times y \rangle$ has to be conserved for every vector v. Using $\langle A, B \times C \rangle = \langle B, C \times A \rangle$ again, we can rewrite this expression as $\langle v, x \times \lambda \rangle + \langle v, y \times \mu \rangle$. Since this has to be constant for every v, we conclude that the vector

$$V = x \times \lambda + y \times \mu \tag{5.5}$$

is conserved. Moreover, we can also conclude that λ and V cannot both vanish. Indeed, if $\lambda = 0$ and $V = 0$, then it would follow that $y \times \mu = 0$. But this would mean that the pullback of μ to S^2 (i.e. the component of μ orthogonal to y) vanishes, so the pullback of (λ, μ) vanishes as a covector in M, contradicting the nontriviality condition of the Maximum Principle.

Remark 5.1. Notice at this point the importance of formulating the Maximum Principle on a manifold. If we had taken the state space of our problem to be $\mathbb{R}^6$, then the nontriviality condition would just have said that $(\lambda, \mu) \neq (0, 0)$. It is easy to see that for any trajectory of our system one can find minimizing adjoint vectors that satisfy this weaker nontriviality property. So *if the problem is formulated in* $\mathbb{R}^6$ *rather than in* $\mathbb{R}^3 \times S^2$, *then every trajectory is an extremal and the Maximum Principle gives no information whatsoever.* □

If $\lambda = 0$, then $\mu \times y$ must be constant and nonzero. But, in view of (5.3), this implies that w is constant. Since y is orthogonal to w, we see that the velocity y is perpendicular to a fixed nonzero vector. This implies that the arc $x(\cdot)$ is entirely contained in a plane P in $\mathbb{R}^3$. It follows in particular that $x(\cdot)$ minimizes length among all the arcs in $\mathcal{A}$ that have the same initial and terminal conditions as $x(\cdot)$, and are contained in P. So $x(\cdot)$ is a solution of the two-dimensional version of our problem. The solution of the problem in dimension 2 was obtained by Dubins, who showed that every optimal trajectory is a concatenation of at most three pieces, at most one of which is a straight line segment, while the others are portions of circles of radius one. Moreover, in the three circles case, the intermediate circle has to have length $\geq \pi$. Since the solution of the plane problem is known, we will concentrate here on *nonplanar* optimal trajectories, i.e. trajectories that are not contained in a plane. For such a trajectory, $\lambda \neq 0$, as shown above.

It is interesting to consider the scalar conserved quantity

$$C \stackrel{\text{def}}{=} \langle \lambda, V \rangle \ . \tag{5.6}$$

Clearly, $C = \langle \lambda, y \times \mu \rangle$, since λ is orthogonal to $x \times \lambda$. Using 5.4, we find

$$C = \langle \lambda, w \rangle \, ||\mu \times y|| \ . \tag{5.7}$$

In other words, *the conserved quantity C is the product of two expressions, which need not be individually conserved.* This product representation has important consequences for the structure of the optimal trajectories. Notice that when $C \neq 0$ then $||\mu \times y||$ can never vanish, so we can write the equation for the trajectory in the form

$$\dot{x} = y \ , \ \dot{y} = y \times \left(\frac{V - x \times \lambda}{||V - x \times \lambda||} \right) \ , \tag{5.8}$$

with full assurance that the denominator never vanishes. The solutions of this equation are obviously smooth. A more detailed analysis shows that the curves that satisfy (5.8) are in fact parametrized by arc-length and have curvature 1, and their torsion satisfies a third-order differential equation without arbitrary constants. (In particular, there is a four-parameter family of such curves through every point.) The helixes (i.e. curves with constant torsion) are included in this family.

If $C = 0$, then nonsmooth curves can occur. The mechanism that permits non-smoothness is the product structure described above. The function $\langle \lambda, w \rangle$ can vanish identically while $\mu \times y \neq 0$, but if $\mu \times y$ ever becomes zero then it is possible for $\langle \lambda, w \rangle$ to become nonzero, as long as $\mu \times y$ stays equal to 0. A detailed analysis of this situation shows that in this case one obtains concatenations of circles and straight lines as in the plane case. All this is studied in detail in [12].

REFERENCES

1 F. Albrecht, *Topics in Control Theory*, Springer-Verlag, Berlin-Heidelberg-New York, 1968.

2 L. D. Berkovitz, *Optimal Control Theory*, Springer-Verlag, New York, 1974.

3 F. H. Clarke, *Optimization and Nonsmooth Analysis,* Wiley Interscience, New York, 1983.

4 L.E. Dubins, "On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal positions and tangents," *Amer. J. Math.* 79 (1957), pp. 497-516.

5 B. Kaskosz and S. Lojasiewicz Jr., "A Maximum Principle for generalized control systems," *Nonlinear Anal. TMA* **9** (1985), pp. 109-130.

6 E. B. Lee and L. Markus, *Foundations of Optimal Control Theory*, J. Wiley, New York, 1967.

7 A. A. Markov, "Some examples of the solution of a special kind of problem in greatest and least quantities," (in Russian) *Soobshch. Karkovsk. Mat. Obshch.* 1 (1887), pp. 250-276.

8 L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mischenko, *The Mathematical Theory of Optimal Processes* (translated by K. N. Trigoroff, L. W. Neustadt, editor), J. Wiley, 1962.

9 J.A. Reeds and L.A. Shepp, "Optimal paths for a car that goes both forwards and backwards," *Pacific J. Math.* 145 (1990), pp. 367-393.

10 H. J. Sussmann, "Orbits of families of vector fields and integrability of distributions," *Trans. Amer. Math. Society.* **180** (1973), pp. 171-188.

11 H. J. Sussmann and G. Tang, "Shortest paths for the Reeds-Shepp car: a worked out example of the use of geometric techniques in nonlinear optimal control," Rutgers Center for Systems and Control (Sycon) Report 91-10, September 1991. To appear in *SIAM J. Control.*

12 H. J. Sussmann, "Shortest paths with a Prescribed Bound on the Curvature: the three-dimensional Case," in preparation.

13 H. J. Sussmann, " An introduction to the coordinate-free Maximum Principle," in preparation.