EXTRACTA MATHEMATICAE Vol. **18**, Núm. 2, 129 – 151 (2003)

A Panorama of Geometrical Optimal Control Theory

M. Delgado-Téllez, A. Ibort

Dpto. de Matemáticas, Unidad Asociada de Matemáticas UCIIIM-CSIC, *Universidad Carlos* III *de Madrid,* 28911-*Leganés, Madrid, Spain*

AMS *Subject Class.* (2000): 49J15, 34A09, 34K35 *Received February* 14, 2002

1. Introduction

Control theory is a young branch of mathematics that has developed mostly in the realm of engineering problems. It is splitted in two major branches; control theory of problems described by partial differential equations where control are exercized either by boundary terms and/or inhomogeneous terms and where the objective functionals are mostly quadratic forms; and control theory of problems described by parameter dependent ordinary differential equations. In this case it is more frequent to deal with non-linear systems and non-quadratic objective functionals [49]. In spite that control theory can be consider part of the general theory of differential equations, the problems that inspires it and some of the results obtained so far, have configured a theory with a strong and definite personality that is already offering interesting returns to its ancestors. For instance, the geometrization of nonlinear affine-input control theory problems by introducing Lie-geometrical methods into its analysis started already by R. Brockett [9] is inspiring classical Riemannian geometry and creating what is called today subriemannian geometry.

In any case, the breadth of mathematics involved in modern control theory is so promising that mathematicians, and in particular "applied" mathematicians, should be aware of their developments, having the total certainty that they will find in this theory an important source of inspiration for their research.

In this review article we present a panorama of modern geometrical optimal control theory for dynamical systems. Optimal control theory for dynamical systems is perhaps, inside the vast subject of control theory, one of the oldest and most developed parts. Optimal control theory sinks directly to the sources of mechanics and the calculus of variations, and reaches its most profound results by the hand of Pontriaguine's maximum principle [49]. Rather recently, a battery of geometrical methods has been developed to address from a new perspective some old problems in the theory of control. Starting with the already mentioned pioneering work by R.W. Brockett [9], [10], to the most recent developments in the field, as reported for instance by H.J. Sussmann [52], and the work of young mathematicians that are conducting their research work in this direction [17], [42], [47], [18], etc., we are witnessing a flourishing of geometrical ideas at the foundations of control theory.

In these notes we will report on some problems in optimal control theory where, we believe, geometrical ideas will play a relevant role towards its understanding and eventual solution. Moreover, geometrically inspired ideas will provide alternative treatments to previously considered approaches. Concretely, we will discuss the problem of cheap and singular optimal control and its links to singular perturbations from a new perspective inspired by traditional constraints theory. We will also present the problem of implicit optimal control and we will analize the various singular behaviours arising in such situations. We will discuss also the problem of the integrability of optimal control problems and the notion of symmetry and, finally, we will present a few remarks on the problem of feedback linearizability and normal forms. We will not, however, touch such interesting problems as controllability or observability, that have already deserved a lot of attention, and where Lie methods have proved to be essential. Finally, we will leave the full discussion and details of the material presented here to the more technical papers quoted or announced in the list of references at the end.

2. Singular optimal control

2.1. Regular and singular optimal control

We will consider the problem of finding C^1-piecewise smooth curves $\gamma(t) = (x(t), u(t))$ with fixed endpoints in state space, $x(t_0) = x_0$ and $x(T) = x_T$, satisfying the control equation

$$\dot{x}^i(t) = f^i(x(t), u(t)), \tag{1}$$

and minimizing the objective functional

$$S(\gamma) = \int_{t_0}^{T} L(x(t), u(t))dt. \tag{2}$$

It is well-known that the solution of such problem is provided by Pontriaguine's maximum principle; this is, the curve $\gamma(t) = (x(t), u(t))$ is an optimal trajectory if there exists a lifting of $x(t)$ to the coestate space, $(x(t), p(t))$ satisfying Hamilton's equations

$$\dot{x}^i = \frac{\partial H}{\partial p_i}, \qquad \dot{p}_i = -\frac{\partial H}{\partial x^i}, \tag{3}$$

and such that

$$H(x(t), p(t), u(t)) = \max_v H(x(t), p(t), v), \qquad a.e. \quad t \in [t_0, T], \tag{4}$$

where H denotes the Hamiltonian function (Pontriaguine's Hamiltonian),

$$H(x, p, u) = p_i f^i(x, u) - L(x, u). \tag{5}$$

It is clear that a necessary condition for the maximum condition (4) is that the function defined by $\varphi_a = \frac{\partial H}{\partial u^a}$ will vanish. Hence, the trajectories solution of the optimal control problem will lie in the submanifold

$$M_1 = \{(x, p, u) \mid \varphi_a(x, p, u) = 0\},$$

of the total space M with coordinates (x, p, u).

The constraint functions $\varphi_a = 0$, called in what follows zeroth order constraints, will define implicitely a function

$$u^a = \psi^a(x, p) \tag{6}$$

whenever the matrix W_{ab} defined by

$$W_{ab} = \frac{\partial \varphi_a}{\partial u^b}$$

will be invertible. Thus, under these circumstances, on the submanifold M_1 there will be defined an optimal feedback function ψ^a given by eq. (6), i.e., whenever

$$\det W_{ab} = \det\left(p_i \frac{\partial^2 f^i}{\partial u^a \partial u^b} - \frac{\partial^2 L}{\partial u^a \partial u^b}\right) \neq 0. \tag{7}$$

If such condition is satisfied at a given point m we will say that the optimal control problem is regular at m. If $\det W_{ab} = 0$ at m we will say that the optimal control problem is singular at m. Generically, the singular points for a given optimal control problem will lie in a union of submanifolds of M_1, the singular locus of the problem. At regular points we will substitute (6) at 3 obtaining a first order system of differential equations that possess locally a unique solution for each initial data.

If we consider the particular example of control systems affine in the control variables, i.e.,

$$\dot{x}^i = A^i(x) + B^i_a(x)u^a$$

the regularity condition eq. (7) becomes simply

$$\det W_{ab} = \det \frac{\partial^2 L}{\partial u^a \partial u^b} \neq 0,$$

that coincides with the well-known regularity condition for Lagrangians in the ordinary calculus of variations, where the control equation eq. (1) is $\dot{x}^i = u^i$.

Even more, in the case of linear control problems with quadratic regulators (LQ systems), i.e., of the form,

$$\dot{x}^i = A^i_j x^j + B^i_a u^a, \tag{8}$$

with Lagrangian function,

$$L = \frac{1}{2} P_{ij} x^i x^j + Q_{ia} x^i u^a + \frac{1}{2} R_{ab} u^a u^b. \tag{9}$$

the regularity condition becomes simply the regularity of the matrix R_{ab}. The system will be singular if and only if the matrix R is singular. Notice that regularity does not requires that R will be definite, condition that will be needed for the extremal trajectories to be an actual minimun of the objective functional. In obvious matrix notation

$$\dot{x} = A \cdot x + B \cdot u, \tag{10}$$

and

$$L = \frac{1}{2} x^t \cdot P \cdot x + x^t \cdot Q \cdot u + \frac{1}{2} u^t \cdot R \cdot u. \tag{11}$$

Thus for regular LQ systems, the optimal feedback function (6) will be

$$u = R^{-1} p^t \cdot B - R^{-1} Q \cdot x,$$

where R^{-1} denotes the inverse of the matrix R_{ab}, $(R^{-1})^{ab}R_{bc} = \delta^a_c$. Then substituting in eq. (8) and the corresponding coestate equation, we obtain the linear system

$$\dot{x} = (A - BR^{-1}Q) \cdot x + BR^{-1}p^t \cdot B$$

and

$$\dot{p} = p^t \cdot (A + QR^{-1}B) + (P - QR^{-1}Q^t) \cdot x.$$

Moreover, introducing the partial feedback $p(t)^t = K(t)x(t)$, we obtain the linear system

$$\dot{x} = (A - BR^{-1}Q + BR^{-1}BK)x \tag{12}$$

and the matrix function K satisfies the matrix Ricatti equation,

$$\dot{K} = K[A - BR^{-1}Q] - K^t[A - QR^{-1}B] + KBR^{-1}BK - QR^{-1}Q^T + P.$$

Thus, solving this last equation in the interval $[t_0, T]$, and replacing in the equation for x, we will obtain a time-dependent system that we can eventually solve explicitely.

An important observation is in place here. Let us suppose that R is regular but has a small eigenvalue that we denote by ϵ. We can think alternatively that $R = R(\epsilon)$ is a one-parameter family of matrices such that $R(0)$ is not invertible. For instance $R(\epsilon) = \epsilon R_0$ where R_0 is regular. In such case, $R^{-1} = \epsilon^{-1}R_0^{-1}$, and eq. (12) becomes the singularly perturbed problem

$$\epsilon\dot{x} = (\epsilon A - BR_0^{-1}Q + BR_0^{-1}BK)x.$$

Thus, the analysis of a singular LQ systems leads naturally to a singularly perturbed linear system. This situation has been studied extensively in [16]. This singular perturbation is different to the often-studied singular perturbation problem that results when the state equation is singularly perturbed [44], [38], [39]. Generalized necessary and sufficient conditions for minimality of singular arcs are studied in [5].

The analysis of these systems has great interest. We must notice first that for a general singular problem, even an LQ system, the behavior of the matrix R near the singular set can be much more complicated that the naive example before. Then, the reduced system, i.e., $\epsilon \to 0$, will be in general quite involved. We have addressed such problem in the general setting by adapting ideas taken from the general theory of singular lagragians. The analysis proceeds as sketched below and is described in detail in [20], [21].

Thus, either if we have a singular system,

$$\det W_{ab} = 0,$$

or we have a regular family $W_{ab}(\epsilon)$, where $\epsilon = (\epsilon_1, \dots, \epsilon_r)$ is a multiparameter, with singular limit at $\epsilon \to 0$, we shall obtain a reduced system that will be defined recursively by imposing the existence of solutions compatible with the constraints. The natural way to express these conditions is by means of the geometric formulation that receives the name, among others, of the presymplectic constraint algorithm (PCA), originally developed in its primitive form by P.A.M. Dirac to quantize mechanical systems defined by singular lagrangians [24] and brought to maturity by the work of many people, ending with the formulation presented in [28]. The PCA will eventually produce a reduced submanifold M_∞ where the state and coestate equations will be well posed, i.e., its integral curves will lie in M_∞ (even if they will not be unique). Next step will be to match the resulting reduced solutions with the endpoints. To do that we will proceed by considering a singularly perturbed system, constructed step by step of the PCA, ending after this process with a "cheap" or nearly optimal control singularly perturbed system called the singular turnpike [20].

2.2. The geometrical setting of optimal control theory To proceed further with the programme sketched at the end of the last paragraph, we must clarify first the geometrical nature of the different objects involved in the previous discussion. All the discussion before was done locally, hence spurious identifications took place sometimes among various coordinates. We will assume that the state space of the system, or configuration space, is a smooth manifold P without boundary and with local coordinates x^i. Control variables will be local coordinates on an affine bundle over P. Such bundle structure for the control variables $\pi\colon C \to P$ means that the control variables are attached to points of the state space and transform under changes of coordinates in state space

$$\acute{x}^i = \varphi^i(x),$$

affinely, this is,

$$\acute{u}^a = \varphi^a_b(x) u^b + \rho^a(x).$$

Later on (see Section 5.1) we will return to discuss some consequences of this assumption.

Thus, an ordinary differential equation on P depending on the parameters u^a, is nothing but a vector field Γ along the projection map π, i.e., Γ is a smooth map $\Gamma\colon C \to TP$ such that $\tau_P \circ \Gamma = \pi$, where $\tau_P\colon TP \to P$ denotes the canonical projection. Locally the vector field Γ will be written as

$$\Gamma = f^i(x,u)\frac{\partial}{\partial x^i},$$

and could also be considered as a map from smooth functions on P to smooth functions on C with properties similar to a derivation (see [48], [46], [14], [15] for a thorough discussion on calculus along maps). Finally, the coestate space of the system is simply the contangent bundle $\pi_P\colon T^*P \to P$.

The first observation is that the Pontriguine's necessary conditions for extremal trajectories have a sound geometrical description as a presymplectic system. In fact, the total space of the system will be $M = T^*P \times_P C$. The previous notation means that M is the bundle over P with fibre M_x at $x \in P$, given by $T^*_xP \times C_x$. In M there is a canonical closed 2-form Ω defined as $\Omega = pr_1^*\omega_0$, where ω_0 is the canonical symplectic form on T^*P and $pr_1\colon M \to T^*P$ is the canonical projection. The coordinate expression for Ω is simply,

$$\Omega = dx^i \wedge dp_i.$$

The 2-form Ω is degenerate and its characteristic distribution is spanned by the "vertical" vectors $\partial/\partial u^a$. The Pontriaguine's hamiltonian function $H\colon M \to \mathbb{R}$, eq. (3), defines a presymplectic system (M,Ω,H), whose dynamical vector fields $\tilde{\Gamma}$ are the solutions of the dynamical equation

$$i_{\tilde{\Gamma}}\Omega = dH. \tag{13}$$

The relevance of the presymplectic system eq. (13) comes from the following result.

Theorem 1. *The curve $\gamma(t)$ is an extremal trajectory for the optimal control system* (1)-(2) *if there exists a lifting $\tilde{\gamma}(t)$ of $\gamma(t)$ to M which is an integral curve of a vector field defined by the dynamical equation* (13).

See [21] for a proof. Similar ideas have been discussed in [42].

2.3. The PCA and singular perturbation It is important to notice that there are vector fields satisfying the dynamical equation (13) only at points m of M such that $Z(H) = 0$ for all Z in the kernel of Ω_m. Thus, the dynamical equation (13) will have in principle solutions only on the submanifold M_1 defined by the equations (as we notice already by means of a different analysis),

$$\varphi_a^{(1)} = \frac{\partial H}{\partial u^a} = 0. \tag{14}$$

Now on M_1 there is at least a vector field $\tilde{\Gamma}$ satisfying (13). However, such vector fields are not, in general, tangent to M_1. A simple computation shows that if the system is regular, i.e., $\det W_{ab} \neq 0$ on M_1, then there is a unique dynamical vector field $\tilde{\Gamma}$ parallel to M_1. Such vector field will be obtained explicitely by using the optimal feedback condition (6). At regular points the manifold M_1 projects (locally) symplectomorphically onto T^*P, and the vector field $\tilde{\Gamma}$ projects to a Hamiltonian vector field on T^*P whose integral curves satisfy hamilton's equations for the hamiltonian function

$$h(x,p) = H(x,p,\psi(x,p)).$$

If the system is singular however, there are points of M_1 where the vector fields solution to (13) will not be tangent to M_1. Thus, their integral curves will exit the submanifold where the extremal trajectories must lie. Those points must be removed from the analysis leaving in this way a subset $M_2 \subset M_1$ where some vector field solution to (13) is tangent to M_1. In fact, such subset is defined by the conditions

$$\tilde{\Gamma}(\varphi_a^{(1)}) = 0, \quad \text{on} \quad M_1,$$

for all $\tilde{\Gamma}$ solution of (13). We will assume that the subset M_2 is a submanifold of M_1. We shall denote the functions defining locally M_2 on M_1 by $\varphi_b^{(2)}$ and we will call them secondary constraints. Clearly the argument goes on, and we will obtain a family of submanifolds defined recursively as follows.

$$M_{k+1} = \{\, x \in M_k \mid \tilde{\Gamma}(\varphi_b^{(k)})(x) = 0 \,\}, \quad k \geq 1.$$

Eventually the recursion will stop and $M_r = M_{r+1} = M_{r+2} = \cdots$ for certain finite r (we will say then that the system is noetherian). In this way we will obtain a stable submanifold

$$M_\infty = \bigcap_{k \geq 0} M_k$$

where we can try to solve the dynamical equation (13) and the integral curves of the corresponding vector fields will be the extremal singular arcs of the singular optimal control problem. Next subsection will show an example of such procedure.

Finally, we can complete our picture by singularly perturbing the vector fields $\tilde{\Gamma}$ and construct cheap optimal controls of the singular problem. Such singular perturbation is done iteratively step by step in the PCA (see [20] for details of the construction).

2.4. A SIMPLE EXAMPLE: THE CUSP SINGULAR OPTIMAL CONTROL PROBLEM We shall consider the optimal control problem on $P = \mathbf{R}$ with control space the 1 dimensional trivial bundle $C = P \times \mathbf{R}$, given by the equation

$$\dot{x} = u$$

and Lagrangian density

$$L(x, u) = -xu^2 - u^4.$$

Pontriaguine's Hamiltonian is given by

$$H(x, p, u) = pu + xu^2 + u^4$$

and the primary constraint $\varphi^{(1)} = \partial H/\partial u$ is given by

$$\varphi^{(1)}(x, p, u) = p + 2xu + 4u^3.$$

The manifold $M_1 = \{ \varphi^{(1)} = 0 \}$ is the well-known cusp singularity [4]. On M_1 the vector fields $\tilde{\Gamma}$ solution of eq. (13) will have the form

$$\tilde{\Gamma}_c = u\frac{\partial}{\partial x} - u^2\frac{\partial}{\partial p} + c\frac{\partial}{\partial u},$$

where c is an undetermined constant. The vector field $\tilde{\Gamma}_c$ will be tangent to M_1 iff $\tilde{\Gamma}_c(\varphi^{(1)} = 0$ on M_1. Thus, this condition implies that

$$u^2 + cW = 0$$

on M_1, with $W = \partial^2 H/\partial u^2 = 2x + 12u^2$. Thus, if we are at a regular point $W \neq 0$ and then $c = -u^2/2(x + 6u^2)$. However there are singular points given precisely by the equation $W = 0$. The singular locus Σ_1 of the equation will be defined then as

$$\Sigma_1 = \{ (x, p, u) \mid \varphi^{(1)}(x, p, u) = 0, W(x, p, u) = 0 \}.$$

The set Σ_1 corresponds clearly to the points where the surface M_1 does not project transversaly into the plane (x,p).

The points such that $W=0$ do not have vectors solution of (13) tangent to M_1 unless $u=0$. If this is the case, then there are vectors tangent to M_1 given by $(0,0,c)$, c arbitrary. Notice that the singular subset $\Sigma_2 = \{\,(x,p,u) \mid \varphi^{(1)} = 0, W = 0, u = 0\,\}$ consists simply of the point $(0,0,0)$ and because the tangent space to M_1 at $(0,0,0)$ is the plane (x,u), the vector $(0,0,c)$ is tangent to M_1. In this case the final constraint submanifold is given by

$$M_\infty = \{\,(x,p,u) \mid \varphi^{(1)} = 0, W \neq 0\,\} \cup \{\,(0,0,0)\,\}.$$

3. Implicit Optimal Control

3.1. An elementary example of implicit optimal control An important class of problems in control theory have the form of the following simple example, $(x,y) \in \mathbf{R}^{n+m}$:

$$A \cdot \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = B \cdot \begin{pmatrix} x \\ y \end{pmatrix} + C \cdot u \tag{15}$$

where the matrix A is of rank n. If the matrix A is diagonalizable, in a new set of coordinates, denoted again by x, y, the previous equation takes the form,

$$\dot{x} = ax + by + cu \tag{16}$$
$$0 = cx + dy + du \tag{17}$$

Such differential-algebraic equations have deserved to be studied by themselves. See for instance [12], [7] for a discussion of nonlinear differential-algebraic equations. One of the natural ways to address such equations is by introducing a small parameter ϵ and consider the singularly perturbed system,

$$\dot{x} = ax + by + cu \tag{18}$$
$$\epsilon\dot{y} = cx + dy + du. \tag{19}$$

If the matrix A were not diagonalizable, the resulting system will be even more involved and it will not be so obvious how to introduce small parameters ϵ to render the problem a purely differential one. It could happen moreover, that the matrix A is of non constant coefficients and we will have to consider the structure of the bifurcation set where the rank of the matrix jumps. This situation will correspond to consider the first order approximation to a general implicit problem in control theory given by an implicit differential equation of the form

$$\Phi(\dot{x}, \dot{y}, x, y, u) = 0.$$

We shall address here not the most general implicit problem but the first order approximation to the general situation, i.e., systems of the form,

$$A(x) \cdot \dot{x} = f(x, u), \tag{20}$$

where x denotes the state variables on the state space P, the vector function f denotes generalized "forces" and live in an auxiliary "force" space F (that can be thought as a vector bundle over $F \to P$) and A is a matrix depending on the state variables x (that can be thought as a vector bundle map $A\colon TP \to F$ or as an F–valued 1–form on P). Thus the implicit optimal control problem will consists in finding the curves $\gamma(t) = (x(t), u(t))$ such that $x(t_0) = x_0$, $x(T) = x_T$,

$$A(x(t)) \cdot \dot{x}(t) = f(x(t), u(t)) \tag{21}$$

minimizing the objective functional eq. (2).

The strategy to solve this problem will be to extend Pontriaguine's maximum principle to this setting, and afterwards to proceed geometrically as in the non-implicit (or normal) optimal control problem discussed in Section 2.

Theorem 2. [22] *A C^1-piecewise smooth curve $\gamma(t) = (x(t), u(t))$ is a solution of the cuasilinear system*

$$A_i^a(x, u)\dot{x}^i = f^a(x, u)$$

and an extremal of the objective functional

$$S(x, u) = \int_{t_0}^{t_1} L(x, u) dt$$

with fixed endpoints $x(t_0) = x_0$; $x(T) = x_T$ if there exists a lifting $\tilde{\gamma} = (x(t), u(t), \zeta(t))$ of the curve γ to the dual bundle $F^* \to P$ of the force bundle F, satisfying the set of implicit differential equations:

1. $A_i^a(x, u)\dot{x}^i = \dfrac{\partial H}{\partial \xi_a}$

2. $A_i^a(x,u)\dot{\xi}_a + (F_A(x,u))_{ij}^a \dot{x}^j \xi_a + \xi_a \dot{u^\alpha} \dfrac{\partial A_i^a}{\partial u^\alpha} = -\dfrac{\partial H}{\partial x^i}$

3. $\xi_a \dfrac{\partial A_i^a}{\partial u^\alpha} \dot{x}^i = \dfrac{\partial H}{\partial u^\alpha}$

where H denotes the Hamiltonian function on $M = F^* \times_P C$ given by,

$$H(x,u,\zeta) = \zeta_a f^a(x,u) - L(x,u). \tag{22}$$

3.2. Presymplectic description of extremal trajectories It is a simple observation that the bundle F^* is equipped with a canonical 1-form θ_A defined by

$$\theta_A = \zeta_a A_i^a(x) dx^i$$

that can be thought also as the pull-back of the canonical Liouville 1-form on T^*P by the dual A^* of the map A. Then, the 2-form $d\theta_A$ induces a presymplectic 2-form Ω on M. Hence as in the case of normal optimal control, we have a presymplectic system (M, Ω, H) defined on M that provides a dynamical equation formally identical to (13). That such presymplectic system is useful to solve the implicit optimal control problem is the content of the following theorem.

Theorem 3. [22] *Obtaining smooth extremal trajectories for an implicit optimal control system* (21)-(2) *is equivalent to find the integral curves of the presymplectic system defined by the presymplectic 2–form $\Omega = pr_1^* d\theta_A$ on M and the Hamiltonian function H* (22)*, in the sense that solutions of the equations* (1)-(3) *of Thm.* 2*, are integral curves of vector fields $\tilde{\Gamma}$ satisfying* (13).

The presymplectic system defines implicit equations whose consistency and uniqueness must be carefully analyzed. We will reproduce the PCA algorithm described in the section above restricting eventually the space M to a subspace M_∞ where the solutions of (13) live. Such analysis follows closely the work in [29], [30], [31].

4. Optimal control with symmetries and integrability

An important aspect of the theory of mechanical systems is the existence of symmetries and in consequence, because of Noether's theorem, of constants of the motion that are very useful in its integration. Such ideas have played a crucial role in the development of celestial mechanics in the past and, mainly through its quantum mechanical counterpart, in modern physics in general. In the theory of control such ideas do not have played such a significant role so far partly because the strong interest in various issues that in a sense a contrary to integrability (controllability for instance). Nevertherless it was already noticed by R. Brockett in [8] the importance of control on Lie groups and homogeneous spaces. In fact, the systems on Lie groups have a close resemblance with the situation discussed by S. Lie [43] and that give rise to the so called nonlinear superposition principles which represents a generalized form of integrability as discussed for instance in [35].

Apart from these observations, little has been worked out around the problem of integrability on control theory. L.E. Faybusovich has discussed more recently the issue of explicit solvability for control problems [25], [26] and, even more recently, A. Bloch and P. Crouch offer in [6] a presentation of optimal control systems on coadjoint orbits related with reduction problems and integrability. The issue is gaining interest and some recent developments can be found in [18].

In these notes we address the problem of integrability of optimal control systems, from the point of view of its symmetries, searching for a generalized Noether's theorem that will associate any one-parameter group of symmetries with a conserved quantity.

4.1. Noether's theorem Let us consider a regular optimal control problem (C, Γ, L) as stated in Section 2. A symmetry transformation will be a map Φ of state and control variables that will transform optimal trajectories and controls into themselves [23]. More precisely, Φ will be of the form

$$\Phi(x, u) = (\varphi(x), \tilde{\varphi}(x)u),$$

i.e., Φ is a bundle map of the control vector bundle $C \to P$. It is not difficult to check that Φ maps optimal trajectories into optimal trajectories if it verifies

$$\varphi_* \circ \Gamma = \Gamma \circ \Phi,$$

and

$$\Phi^* L = L + \dot{h},$$

where $\dot{h} = \tilde{\Gamma}(h)$ and $\tilde{\Gamma}$ denotes the unique vector field extending Γ to the submanifold M_0 and satisfying the dynamical equation (13). Hence if we have a one-parameter group of symmetries Φ_s, then its infinitesimal generator X,

$$X \circ \Phi_t = \frac{d}{dt}\Phi_t,$$

has the form

$$X = \zeta^i(x)\frac{\partial}{\partial x^i} + u^a \zeta_a^b(x)\frac{\partial}{\partial u^b}.$$

Moreover,

$$\Phi_t^* L = L + \dot{h}_t,$$

and

$$X(L) = \frac{d}{dt}\dot{h}_t|_{t=0}.$$

We shall denote the gauge function in the r.h.s. of the previous equation by μ.

$$\mu = \frac{d}{dt}\dot{h}_t|_{t=0}.$$

Then, we have the following result:

Theorem 4. [23] *Let X be the infinitesimal generator of a one-parameter group of symmetries of the optimal control problem (C, Γ, L). Then the function*

$$J(x,p,u) = \zeta^i(x)p_i - \mu(x), \tag{23}$$

is a constant of the motion, i.e., if $(x(t), p(t), u(t))$ is an optimal trajectory, then

$$J(x(t), p(t), u(t)) = J(x(t_0), p(t_0), u(t_0)), \qquad \forall t \in [t_0, T]. \tag{24}$$

4.2. Particular cases The first application of the previous result, Thm. 4, comes when particularizing it to control systems of the form

$$\dot{x}^i = u^i.$$

Then, on M_0, we have

$$p_i = \frac{\partial L}{\partial u^i},$$

and the conserved quantity associated to any one-parameter group of symmetries, eq. (24), becomes

$$J(x,u) = \frac{\partial L}{\partial u^i}\zeta^i(x) - \mu(x),$$

which is the statement of the well-known Noether's theorem for Lagrangian mechanics.

If we have vertical symmetries, i.e., symmetries X such that $\zeta^i = 0$, then the conserved quantity is 0. This is not strange as vertical symmetries affect only to control variables, thus the system is insensitive to these changes.

5. Linearization and normal forms

5.1. Linearization and feedback linearization In the previous sections we have used as simple examples linear systems of the form eq. (10) Thus, if we are considering a nonlinear problem

$$\dot{x} = f(x, u) \tag{25}$$

we can face the question of whether there is a new local system of coordinates y^i in state space, and new controls v^a such that (25) transforms in the simpler linear form (10). Before determining the type of transformations that we will allow in state and control space, let us consider first the simpler problem of the linearization of a vector field X in the neighborhood of a critical point.

Let us then consider then a vector field $X = f^i(x)\partial/\partial x^i$ such that $X(m) = \mathbf{0}$. Then, we would like to know if the vector field is (locally) linear, that is, if there exists a set of coordinates y^i such that the vector field takes the form

$$X = A_i^j y^j \frac{\partial}{\partial y^i}.$$

We shall consider in what follows local coordinates such that $x(m) = 0$. The problem of the local linearizability of a vector field was already considered by H. Poincaré. If we expand f around $\mathbf{0}$, we obtain

$$\dot{x} = A \cdot x + \text{h.o.t.},$$

with

$$A_i^j = \frac{\partial f^j}{\partial x^i}(\mathbf{0}).$$

We shall denote by $\lambda_1, \ldots, \lambda_n$ the eigenvalues of A. We will say that they are resonant if there is collection of nonnegative integers $m_1, \ldots, m_n$, $\sum_k m_k \geq 2$, such that

$$\lambda_s = \langle m, \lambda \rangle.$$

The number $|m| = \sum_k m_k$ will be called the order of the resonance.

Theorem 5. [3] *If the eigenvalues of the matrix A are non resonant, then there is a formal change of coordinates $y = y(x)$ such that in the coordinates y, the vector field X becomes linear.*

Similar local results can be obtained for families of vector fields closing a semisimple Lie algebra with a common critical point [33], [32].

A geometrical characterization of (local and global) linearizability can be used to address the same question. Such aspects are discussed in depth in [13] (see [27] for some elementary applications). Global linearizability can be characterized by means of a global dilation vector field Δ. Then, X will be linearizable iff $[\Delta, X] = 0$. Locally, linearizability can be characterized by a SODE. The precise way to do this will be discussed elsewhere.

Now we arrive to the point about the linearizability of a general control system. If we think that the system (25) defines a vector field along the projection map of an affine bundle $\pi\colon C \to P$, then the coordinate transformations we consider will be of the form

$$y^i = \varphi^i(x) \tag{26}$$

$$v^a = \tilde{\varphi}(x)^a_b u^b + \rho^a(x). \tag{27}$$

This is exactly what has been called feedback linearization as was presented by R. Brockett [11]. If the system (25) is affine in the controls

$$\dot{x} = f(x) + g(x)u,$$

such problem has been solved in different ways [40], [36], [51]. See also [41] for a review of the problem and references therein.

5.2. Normal forms Similar questions to those addressed in the previous paragraphs, section 5.1, can be posed for optimal control systems. Now, the linear model for such systems will be an LQ systems as described in Section 2, eqs. (10-11). Hence we would like to know if a regular non-linear optimal control system can be brought locally to an LQ form.

More precisely, let us suppose that we perform a change of variables in the state, coestate and control variables, $\Phi(y, \xi, v) = (x, \zeta, u)$, of the particular form

$$x^i = \varphi^i(y) \tag{28}$$

$$\zeta_i = \frac{\partial y^j}{\partial x^i}\xi_j \tag{29}$$

$$u^a = \tilde{\varphi}^a_b(y) v^b + \rho^a(y) \tag{30}$$

and that the resulting system is of LQ type. That is,

$$K(y,\xi,v) = H\circ\Phi(y,\xi,v) = \xi^t\cdot A\cdot y + \xi^t\cdot B\cdot v - \frac{1}{2}y^t\cdot P\cdot y - y^t\cdot Q\cdot v - \frac{1}{2}\xi^t\cdot R\cdot\xi.$$

Then, the linear equations of motion will have the form

$$\dot{y} = A\cdot y + B\cdot v \tag{31}$$

$$\dot{\xi} = -A^t\cdot\xi + P\cdot y + Q^t\cdot v \tag{32}$$

$$0 = B^t\cdot\xi - Q^t\cdot y - R\cdot v \tag{33}$$

Thus, the optimal feedback condition becomes,

$$v = C\cdot\xi + D\cdot y,$$

with $C = R^{-1}B$, $D = -R^{-1}Q$, and we will obtain, after substituting it, the linear hamiltonian equations,

$$\dot{y} = (A + BD)\cdot y + BC\cdot\xi \tag{34}$$

$$\dot{\xi} = -(A^t - Q^tC)\cdot\xi + (P + Q^tD)\cdot y. \tag{35}$$

Thus, the system has to be feedback linearizable and simultaneously, the Lagrangian L has to become quadratic. Notice that the point $(y,v) = (\mathbf{0},\mathbf{0})$ must be a critical point for the vector field $f(x,u)$. We will assume that $\Phi(\mathbf{0}) = \mathbf{0}$. Now, in order for L to be reducible to its quadratic part, we need,

$$L(\mathbf{0},\mathbf{0}) = 0, \qquad \frac{\partial L}{\partial x}(\mathbf{0},\mathbf{0}) = 0, \qquad \frac{\partial L}{\partial u}(\mathbf{0},\mathbf{0}) = 0.$$

Thus,

$$f(x,u) = A\cdot x + B\cdot u + \text{h.o.t.}, \qquad L(x,u) = \frac{1}{2}x^t\cdot P\cdot x + x^t\cdot Q\cdot u + \frac{1}{2}u^t\cdot R\cdot u + \text{h.o.t.},$$

with

$$A = \frac{\partial f}{\partial x}(\mathbf{0},\mathbf{0}), \quad B = \frac{\partial f}{\partial u}(\mathbf{0},\mathbf{0}), \quad P = \frac{\partial^2 L}{\partial x\partial x}(\mathbf{0},\mathbf{0}),$$
$$Q = \frac{\partial^2 L}{\partial x\partial u}(\mathbf{0},\mathbf{0}), \quad R = \frac{\partial^2 L}{\partial u\partial u}(\mathbf{0},\mathbf{0}).$$

The first important observation is that in general this problem has no solution. The reason for this is that because of the regularity condition, the optimal feedback control $u^a = \psi(x,\zeta)$, where ζ_i denotes the coestate variables, allows to reduce the problem to a Hamiltonian system with equations

$$\dot{x}^i = f^i(x,\psi(x,\zeta)) \tag{36}$$

$$\dot{\zeta}_i = -\zeta_j \frac{\partial f^j}{\partial x^i}(x,\psi(x,\zeta)) + \frac{\partial L}{\partial x^i}(x,\psi(x,\zeta)). \tag{37}$$

The Hamiltonian function of such system becomes,

$$h(x,\zeta) = \zeta_i f^i(x,\psi(x,\zeta)) - L(x,\psi(x,\zeta)).$$

The conclusion of the argument is based on the following (non-completely trivial) observation. The transformed hamiltonian function $k = k(y,\xi)$ of hamiltonian $h = h(x,\zeta)$ corresponding to a canonical map $(\varphi^*)^{-1}$ obtained by lifting a map φ on state variables, coincides with the Hamiltonian function obtained by first transforming the function H by means of the transformation Φ on the total space M with coordinates (x,ζ,u), and then restricting to a hamiltonian function by using the transformed feedback optimal condition. Thus, if a regular optimal control problem could be reduced to an LQ system by a change of coordinates of the form (36), then, there will exits a canonical change of coordinates such that the hamiltonian h will be linearized, but it is well-known that a Hamiltonian vector field cannot be linearized in general (not even formally). H. Poincaré noticed the impossibility of removing fourth order terms in the Taylor expansion by formal changes of coordinates.

Thus, we have arrived to the first negative result concerning the linearizability of optimal control problems. However, regarding the stability of equilibrium points, it is important to understand the local structure of such points. We know that the stability of equilibrium points for Hamiltonian systems is neutral. It is possible to reduce formally h to normal forms, called Poincaré-Birkhoff normal forms, in the neighborhood of an stable point. In fact, at an stable point, the quadratic part of the hamiltonian can be written in the form

$$h(q,p) = \sum \omega_k (p_k^2 + q_k^2),$$

where (q,p) are canonical coordinates and the frequences ω_k are, generically, all different. Then, if the frequences ω_k have no resonances of order less than N, there exists a formal change of canonical coordinates that brings the Hamiltonian h to the form

$$k(Q,P) = \sum_{|k|\leq N} a_k \tau^k,$$

up to terms of order $N+1$, where k denotes the multiindex $k = (k_1, k_2, \dots, k_r)$, $|k| = k_1 + k_2 + \cdots + k_r$, $\tau_i = \frac{1}{2}(P_i^2 + Q_i^2)$ and $\tau^k = \tau_1^{k_1} \cdots \tau_r^{k_r}$ [2].

We can summarize the previous discussion as follows:

THEOREM 6. *Let $\dot{x}^i = f^i(x,u)$ denote a regular optimal control system with objective functional defined by a Lagrangian function $L = L(x,u)$. We shall assume that $f^i(\mathbf{0},\mathbf{0}) = 0$, $L(\mathbf{0},\mathbf{0}) = 0$ and $dL(\mathbf{0},\mathbf{0}) = \mathbf{0}$. Let $h(x,\zeta) = H(x,\zeta,u(x,\zeta))$ be the hamiltonian function defined by the optimal feedback functions $u^a = \psi^a(x,\zeta)$. The linear approximation of the hamiltonian vector field defined by the hamiltonian h, is defined by a quadratic hamiltonian function h_0,*

$$h_0(x,\zeta) = \frac{1}{2}\zeta L\zeta + \zeta M x + \frac{1}{2} x N x.$$

with

$$L = -W^{-1}\frac{\partial f}{\partial u}\frac{\partial f}{\partial u} + \frac{\partial^2 L}{\partial u \partial u}\left(\frac{\partial \psi}{\partial \zeta}\right)^2 + \frac{\partial L}{\partial u}\frac{\partial^2 \psi}{\partial \zeta^2} \tag{38}$$

$$M = \frac{\partial f}{\partial x} - W^{-1}\frac{\partial f}{\partial u}\left(\zeta\frac{\partial^2 f}{\partial x \partial u} - \frac{\partial^2 L}{\partial x \partial u}\right) + \frac{\partial^2 L}{\partial x \partial x} + \frac{\partial^2 L}{\partial x \partial u}\frac{\partial \psi}{\partial x} + \frac{\partial L}{\partial u}\frac{\partial^2 \psi}{\partial x^2} \tag{39}$$

$$N = \frac{\partial^2 L}{\partial x \partial x} + \frac{\partial^2 L}{\partial x \partial u}\frac{\partial \psi}{\partial x} + \frac{\partial L}{\partial u}\frac{\partial^2 \psi}{\partial x^2} \tag{40}$$

Then, if the linear approximation is estable, i.e., all eigenvalues of the hamiltonian matrix

$$T = J\left(\begin{array}{c|c} L & M \\ \hline M^t & N \end{array}\right),$$

are imaginary and different, and do not have resonances of order smaller than N, the hamiltonian h can be brought to a Birkhoff normal form of degree N by a formal canonical change of canonical coordinates.

Several remarks are in order here. First, we must notice that the canonical change of coordinates that we will eventually find as a result of the previous theorem, will mix the state and coestate variables, but will not destroy the structure of state/control variables.

The linear approximation of the optimal control problem given by the hamiltonian h_0, can be brought to a normal form by linear symplectic changes of coordinates. The family of normal forms were obtained by Williamson [54] and can also be consulted in Appendix 6 of [2]. It will be important to understand the existence of approximate feedback linearizations for nonstable configurations.

6. Outlook

From the previous discussion, it is clear that geometrical ideas are providing a powerful insight in old and new problems in the theory of control and optimal control. In this paper we have reviewed some aspects of classical problems in optimal control theory such as the construction of singular arcs. We have also addressed recent problems like the integrability of optimal control problems and we have posed new problems like the construction of normal forms for optimal control. There are many other problems which are waiting for a geometrical perspective, like the inverse problem that could be analyzed along the lines suggested in [34].

As a consequence of all this we hope to have convinced the reader that geometrical ideas, combined with new analytical and numerical tools will lead to new applications and solutions to important problems in this domain of applied mathematics. In particular the use of numerical methods with a geometrical background will prove to be particularly stimulating. In this sense the adaptation of symplectic integrators to the optimal control setting [50], or Veselov discretization methods [53], [45] will certainly provide an outburst of activity in the field. Such ideas will be discussed in forthcoming papers.

Acknowledgements

Partial financial support for this work was provided by MCYT project BFM-2001/2272.

References

[1] Agrachev, A.A. GamKrelidze, R.V., Simplectic geometry for optimal control, in "Nonlinear Controllability and Optimal Control" Ed. H.J. Sussmann, Marcel & Dekker Inc., N.Y., 1990, 263 – 267.

[2] Arnol'd, V.I., "Les Méthodes Mathématiques de la Mécanique Classique", Éditions Mir., Moscou, 1976.

[3] Arnol'd, V.I., "Geometrical Methods in the Theory of Ordinary Differential Equations", Springer-Verlag, Berlin, 1983.

[4] Arnol'd, V.I., Varchenko, A., Goussein-Zadé, S., "Singularités des Applications Différentiables", Vol. 1. Éditions Mir, Moscou, 1986.

[5] Bell, D.J., Jacobson, D.H., "Singular Optimal Control Problems", Academic Press, London, 1975.

[6] Bloch, A.M., Crouch, P.E., Optimal control, optimization and analytical mechanics, in "Mathematical Control Theory", J. Baillieul and J.C. Willems, eds., Springer-Verlag, 1999, 265 – 321.

[7] Brenan, K.E., Campbell, S.L., Petzold, L.R., "Numerical Solutions of Initial Value Problems in Differential-Algebraic Equations", North-Holland, Amsterdam, 1989.

[8] Brockett, R.W., Systems theory on group manifolds and coset spaces, *SIAM J. Control Optim.,* **10** (1970), 265 – 284.

[9] Brockett, R.W., Lie theory and control systems defined on spheres, *SIAM J. Appl. Math.,* **25** (1973), 213 – 225.

[10] Brockett, R.W., Lie algebras and Lie groups in control theory, in "Geometrical Methods in System Theory", D.Q. Mayne, R. W. Brockett eds., Reidel, Drodrechet, 1973, 43 – 82.

[11] Brockett, R.W., Feedback invariants for nonlinear systems, in "Proceedings of the International Congress of Mathematicians", Helsinki, 1978, 1357 – 1368. S.L. . . (1980); .

[12] Campbell, S.L., "Singular Systems of Differential Equations", Pitman, San Francisco, 1980; ibid. "Singular Systems of Differential Equations II", Pitman, San Francisco, 1982.

[13] Cariñena, J.F., Ibort, A., Marmo, G., Morandi, G., Dynamical systems, (2000) in preparation.

[14] Cariñena, J.F., López, C., Martínez E., A new approach to the converse of Noether's theorem, *J. Phys. A: Math. & Gen.,* **22** (1989), 4777 – 4486.

[15] Cariñena, J.F., López, C., Martínez E., Sections along a map applied to higher-order Lagrangian Mechanics, *Acta Apli. Mathematicae,* **25** (1991), 127 – 151.

[16] Clements, D.J., Anderson, B.D.O., "Singular Optimal Control: The Linear-Quadratic Problem", Springer-Verlag, Berlin, 1978.

[17] Cortés, J., Martínez, S., Optimal control for nonholonomic systems with symmetry, in "Proc. IEEE Conf. Decision & Control", Sidney, Australia (2000), 5216 – 18.

[18] Cortés Monforte, J., "Geometric, Control and Numerical Aspects of Nonholonomic Systems", Ph. Thesis, to be published in Lect. Notes in Maths., 2003.

[19] Crouch, P.E.. van der Schaft, A.J., "Variational and Hamiltonian Control Systems", Lect. Notes in Control and Inf. Theory, Vol. 101, Springer-Verlag, 1987.

[20] Delgado-Téllez, M., Ibort, A., "On the Geometry and Topology of Singular Optimal Control Problems and their Solutions", Proceedings of the fourth international conference on dynamical systems and differential equations, Wilmington 2002.

[21] Delgado-Téllez, M., Ibort, A., Singular Optimal Control: A geometric approach, in preparation (2002).

[22] Delgado-Téllez, M., Ibort, A., Implicit optimal control: a maximum principle, in preparation (2002).

[23] Delgado-Téllez, M., Ibort, A., Optimal control with symmetries and integability, in preparation (2002).

[24] Dirac, P.A.M., Generalized hamiltonian dynamics, 392 – ???; *Canad. J. Math.,* **2** (1950), 129 – 148.

[25] Faybusovich, L.E., Explicitly solvable optimal control problems, *Int. J. Control,* **48** (1988), 235 – 250.

[26] Faybusovich, L.E., Hamiltonian structure of dynamical systems which solve linear programming problems, *Physica D,* **53** (1991), 217 – 232.

[27] Gaeta, G., Marmo, G., Non-perturbative linearization of dynamical systems, *J. Phys. A: Math. & Gen.,* **29** (1996), 5035 – 5048.

[28] Gotay, M.J., Nester, J.M., Presymplectic manifolds, geometric constraint theory and the Dirac-Bergmann theory of constraints, *Ann. Inst. Henri Poincaré, Ser. A,* **30** (1979).

[29] Gracia, X., Pons, J.M., Constrained systems: a unified approach, *Int. J. Theor. Phys.,* **30** (1991), 511 – 516.

[30] Gracia, X., Pons, J.M., A generalized framework for constrained systems, *Diff. Geom. Appl.,* **2** (1992), 223 – 247.

[31] Gracia, X., Muñoz-Lecanda, M., Román-Roy, N., Singular systems: their origins, general features and non-numerical solutions, in "IFAC conference system structure and control. Ed. M. Gugliemi-IFAC", (1995), 37 – 42.

[32] Guillemin, V., Sternberg, S., Remarks on a paper by Hermann, *Trans. Am. Math. Soc.,* **130** (1968), 110 – 114.

[33] Hermann, R., The formal linearization of a semisimple Lie algebra of vector fields, *Trans. Am. Math. Soc.,* **130** (1968), 105 – 110.

[34] Ibort, L.A., Marín-Solano, J., On the inverse problem of the calculus of variations for a class of coupled dynamical systems, *Inverse Problems,* **7** (1991), 1 – 13.

[35] Ibort, A., Marmo, G., A new look at completely integrable systems and double Lie groups, *Contemporary Maths.,* **219** (1998), 154 – 172.

[36] Jakubczyk, B., Respondek, W., On linearization of control systems, *Bull. Acad. Pol., Sci. Ser. Sci. Math.,* **28** (1980), 517 – 522.

[37] Jurdjevic, V., "Geometric Control Theory", Cambridge Univ. Press, Cambridge, 1996.

[38] Kokotovic, P.V., Applications of singular perturbation techniques to control problems, *SIAM Review,* **26** (1984), 501 – 550.

[39] Kokotovic, P.V., Khalil, H.K., O'Reilly, J., "Singular Perturbation Methods in Control", Academic Press, London, 1986.

[40] Korobov, V.I., A general approach to the solution of the problem of synthesizing bounded controls in a control problem, *Math. USSR Sb.,* **37** (1979), 535 – 539.

[41] Krener, A.J., Feedback linearization, in "Mathematical Control Theory", J. Baillieul and J.C. Willems eds., Springer Verlag, (1999), 66 – 98.

[42] López, C., Martínez, E., Subfinslerian metrics associated to an optimal control problem, *J. Control Optim.,* **39** (2000), 798 – 811.

[43] Lie, S., "Vorlesungen über Differenialgleichungen mit bekannten infinitesimalen Transformationen. Bearbit und herausgegeben von Dr. G. Scheffers", Liepzig, B.G. Teubner, 1981.

[44] R.E. O'Malley, R.E., Jr, Singular perturbations and optimal control, in Lect. Notes in Maths., Vol. 680, Springer-Verlag, Berlin (1978), 170 – 218.

[45] Marsden, J.E., Patrick, G.W., Shkoller, S., "Multisymplectic Geometry, Variational Integrators, and Nonlinear PDEs", math.DG/9807080, 1998.

[46] Martínez, E., Geometría de ecuaciones diferenciales aplicada a la Mecánica, Publ. Sem. García Galdeano, Sec 2, 36, 1989.

[47] Martínez Díaz, S., "Geometric Methods in Nonlinear Control Theory with Applications to Dynamic Robotic Systems", Ph. D. dissertation, Univ. Carlos III de Madrid, 2002.

[48] Pidello, G., Tulczyjew, W., Derivations of differential forms and jet bundles, *Ann. Math. Pure et Appl.,* **147** (1987), 249 – 265.

[49] Pontryaguin, L.S., Boltyanskii, V.G., GamKrelidze, R.V., Mishchenko, E.R., "The Mathematical Theory of Optimal Processes", Interscience, N.Y., 1962.

[50] Sanz-Serna, J., Calvo, M., "Numerical Hamiltonian Problems", Chapman and Hall, London 1994.

[51] Su, R., On the linear equivalents of control systems, *Syst. Control Lett.,* **2** (1982), 48 – 52.

[52] Sussmann, H.J., Geometry and Optimal Control, in "Mathematical Control Theory", J. Baillieul and J.C. Willems eds., Springer Verlag, (1999), 140 – 198.

[53] Veselov, A.P., Integrable Lagrangian correspondences and the factorization of matrix polynomials, *Funct. Anal. Pril.,* **25** (1991), 38 – 49.

[54] Williamson, J., On the algebraic problem concerning the normal forms of linear dynamical systems, *Amer. J. Maths.,* **58** (1936), 141 – 163.