PROCEEDINGS OF THE FOURTH INTERNATIONAL CONFERENCE ON
DYNAMICAL SYSTEMS AND DIFFERENTIAL EQUATIONS
May 24 – 27, 2002, Wilmington, NC, USA pp. **1–10**

OPTIMAL CONTROL PROBLEMS WITH VARIABLE ENDPOINTS

BAVO LANGEROCK

Abstract. In a previous paper [5] we have proven a geometric formulation of the maximum principle for non-autonomous optimal control problems with fixed endpoint conditions. In this paper we shall reconsider and extend some results from [5] in order to obtain the maximum principle for optimal control problems with variable endpoint conditions. We only consider the case where one of the endpoints may vary, whereas the other is kept fixed.

1. **Introduction and preliminary definitions.** The development of optimal control theory in a differential geometric setting has been carried out by for instance H.J. Sussmann in [14], where a coordinate-free formulation of the maximum principle is given. Many interesting problems that can be regarded as a control problems are encountered in differential geometry. For instance, the problem of characterising length minimising curves in sub-Riemannian geometry (see [4, 11, 12]) has become one of the standard examples in "geometric optimal control theory". Another field of applications can be found in the geometric formulation of Lagrangian systems subjected to nonholonomic constraints (see [10] and references therein). More recently, the formulation of Lagrangian systems on Lie-algebroids, which has been studied in [2, 8, 15], can also be treated as an optimal control problem.

In [5] we have given a proof of the coordinate-free maximum principle for (time-dependent) optimal control systems with fixed endpoint conditions, relying on the approach of L.S. Pontryagin et al. in [7]. As a side result of our approach, we were able to derive some necessary and sufficient conditions for the existence of what are called *(strictly) abnormal extremals* in sub-Riemannian geometry (for an example of a strictly abnormal extremal, we refer to [9]). In this paper, it is our goal to present an extension of the maximum principle for (time-dependent) optimal control problems with *variable endpoint* conditions.

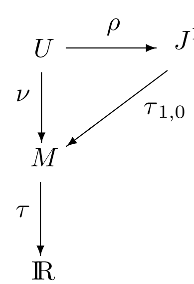
The outline of the paper is as follows. In the remainder of this section, we briefly describe the geometric setting in which we study control systems. The maximum principle and some intermediate results obtained in [5] are summarised in Section 2 without giving all the technical details. These results are indispensable in order to present a comprehensive treatment of control problems with variable endpoints in Section 3.

We now proceed towards the definition of a geometric control structure. It should be noted that we impose rather strong smoothness conditions. However, it occurs to us that there is a sufficiently large and relevant class of control problems that fits into the framework presented below (see [14] for a different approach).

1991 *Mathematics Subject Classification.* 49Kxx, 53Cxx.
Key words and phrases. optimal control theory, variable endpoints.

Definition 1. A *geometric control structure* is a triple (τ, ν, ρ) consisting of: (i) a fibre bundle $\tau : M \to \mathbb{R}$ over the real line, where M is called the event space and with typical fibre Q, which is referred to as the configuration manifold, (ii) a fibre bundle $\nu : U \to M$, called the control space, and (iii) a bundle morphism $\rho : U \to J^1\tau$ over the identity on M, such that $\tau_{1,0} \circ \rho = \nu$.

In the above definition, the first jet space of τ is denoted by $J^1\tau$ and the projections $J^1\tau \to \mathbb{R}$ and $J^1\tau \to M$ are denoted by τ_1 and $\tau_{1,0}$ respectively. The map ρ, which is called the anchor map, makes the following diagram commutative:



A section $u : I \to U$ of $\tau \circ \nu$ is called a *smooth control* if $\rho \circ u = j^1 c$, where $c = \nu \circ u$ is called the *smooth base section* of the control u.

In order to fix the ideas we will first elaborate on the notion of smooth controls. Fix a bundle adapted coordinate chart on M and let $(t, x^1, \ldots, x^n)$ (where $\dim Q = n$) denote the associated coordinate functions. In particular the projection τ maps $(t, x^1, \ldots, x^n)$ onto t. Similarly we consider an adapted coordinate chart of the control space U, with coordinate functions $(t, x^1, \ldots, x^n, u^1, \ldots, u^k)$ (with $\dim U = 1 + n + k$). A smooth control u, being a section of $\tau \circ \nu$, locally represented by n functions $x^i(t)$, $i = 1, \ldots, n$ and k functions $u^a(t)$, $a = 1, \ldots, k$, and has to satisfy, by definition, the following equation:

$$\dot{x}^i(t) = \rho^i(t, x^i(t), u^a(t)).$$

The above equations are easily recognised (see [7]) as the "law of motion" that occurs in standard control theory.

It turns out, however, that the class of smooth controls should be further extended to sections admitting (a finite number of) discontinuities in the form of certain 'jumps' in the fibres of ν, such that the corresponding base section is piecewise smooth (see also [7]). For instance, assume that $u_1 : [a, b] \to U$ and $u_2 : [b, c] \to U$ are two smooth controls with respective base sections c_1 and c_2, such that $c_1(b) = c_2(b)$. The composite control $u_2 \cdot u_1 : [a, c] \to U$ of u_1 and u_2 is defined by:

$$u_2 \cdot u_1(t) = \begin{cases} u_1(t) & t \in [a, b], \\ u_2(t) & t \in]b, c]. \end{cases}$$

It is readily seen that $u_2 \cdot u_1$ is (in general) discontinuous at $t = b$, however, the base $\nu \circ (u_2 \cdot u_1)$ is continuous. This definition can easily be extended to any finite number of smooth controls, yielding what we shall call in general a *control* (a detailed definition can be found in [5]). We say that a control $u : [a, b] \to U$ with base section c takes x to y if $c(a) = x$ and $c(b) = y$, with $x, y \in M$.

We now introduce the notion of optimality. Assume that *a cost function* $L \in C^\infty(U)$ is given. For any control $u : [a,b] \to M$ we can define its *cost* $\mathcal{J}(u)$:

$$\mathcal{J}(u) = \int_a^b L(u(t))dt.$$

A control u taking x to y is said to be *optimal* if, given any other control u' taking x to y then, we have

$$\mathcal{J}(u) \leq \mathcal{J}(u').$$

The problem of finding the optimal controls taking a given point to another given point is called an *optimal control problem with fixed endpoint conditions.*

On the other hand, assume that two immersed submanifolds $i : S_i \to M$ and $j : S_f \to M$ are given. A control u taking a point $x \in i(S_i)$ to a point $y \in j(S_f)$ is said to be *optimal* if, given any other control u' taking $x' \in i(S_i)$ to $y' \in j(S_f)$, then $\mathcal{J}(u) \leq \mathcal{J}(u')$. The problem of finding the optimal controls taking a point in S_i to a point in S_f is called an *optimal control problem with variable endpoint conditions.*

The maximum principle gives necessary conditions to be satisfied by optimal controls. In Section 2 we present a version of the maximum principle, proven in [5], for optimal control problems with *fixed* endpoint conditions whereas in Section 3 we prove necessary conditions for optimal controls in the case of optimal control problems with *variable* endpoints, in the specific case where either $S_i = \{x\}$ or $S_f = \{y\}$ (i.e. one of the endpoints is kept fixed). The more general case is more involved and is left for future work. The case where one of the endpoints is kept fixed, admits an elegant and concise approach, which is worth to be treated separately.

2. **Optimal control problems with fixed endpoints.**

This section gives a quick review of some results obtained in [5], to which we refer for the proofs and further technical details.

2.1. **Controllability and the cone of variations.**

Consider the total time derivative $\mathbf{T}$, which is a vector field along $\tau_{1,0}$, i.e. $\mathbf{T} : J^1\tau \to TM : j^1_t c \mapsto T_t c(\partial_t)$ (where ∂_t is the standard vector field on $\mathbb{R}$ and $T_t c$ the tangent map of $c : \mathbb{R} \to M$ at t). Using this map, the everywhere defined family of vector fields $\mathcal{D} = \{\mathbf{T} \circ \rho \circ \sigma \mid \sigma \in \Gamma(\nu)\}$ on M, generated by (local) sections of ν, is well defined. The family $\mathcal{D}$ plays a crucial role in deriving the maximum principle, as will become clear from the following observation. In [5] it is proven that the base of any control is a concatenation of integral curves of vector fields in $\mathcal{D}$ and vice versa (see below). We first proceed with some elementary definitions associated with this family of vector fields (see also [6], where one can find a detailed treatment of the theory of generalised distributions).

A *composite flow of vector fields in* $\mathcal{D}$ is defined in the following way. Let $(X_\ell, \ldots, X_1)$ denote an ordered family of ℓ vector fields in $\mathcal{D}$ and let $\{\phi^i_t\}$ denote the flow of $X_i = \mathbf{T} \circ \rho \circ \sigma_i$. For $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell$, belonging to a suitable domain the composite flow of the X_i is defined by

$$\Phi_T(x) = \phi^\ell_{t_\ell} \circ \ldots \circ \phi^1_{t_1}(x),$$

where T is called the *composite flow parameter.* The concatenation γ through $x \in \tau^{-1}(a)$, associated with Φ and T, is a piecewise smooth curve defined as follows:

$\gamma : [a, a + |t_1| + \ldots + |t_\ell|] \to M$ with

$$\gamma(t) = \begin{cases} \phi^1_{\mathrm{sgn}(t_1)(t-a)}(x) & \text{for} \quad t \in [a, a_1] \\ \phi^2_{\mathrm{sgn}(t_2)(t-a_1)}(\phi^1_{t_1}(x)) & \text{for} \quad t \in]a_1, a_2] \\ \ldots & \\ \phi^\ell_{\mathrm{sgn}(t_\ell)(t-a_{\ell-1})}(\ldots(\phi^1_{t_1}(x))\ldots) & \text{for} \quad t \in]a_{\ell-1}, a_\ell], \end{cases}$$

where $a_i = a + \sum_{j=1}^{i} |t_j|$, $\mathrm{sgn}(t_i) = \frac{t_i}{|t_i|}$ for $t_i \neq 0$ and $\mathrm{sgn}(0) = 0$. Assume that all $t_i \geq 0$, then $\dot\gamma(t) = X_i(\gamma(t))$, and thus γ is an integral curve of X_i when restricted to $[a_{i-1}, a_i]$. Moreover, since every $X_i = \mathbf{T} \circ \rho \circ \sigma_i$ is τ-related to ∂_t on $\mathbb{R}$ we have that $\tau(\gamma(t)) = t$ and thus γ determines (at least on every smooth part) a section of τ. Define the sections $u_i(t) = \sigma_i(\phi_{t-a_i}(x))$ on $[a_{i-1}, a_i]$. Then it is easily seen that $u = u_\ell \cdot \ldots \cdot u_1$ is a control with base curve γ; we shall call u the control induced by the ordered family of sections $(\sigma_\ell, \ldots, \sigma_1)$. We can conclude that every concatenation of integral curves of vector fields of an ordered set of $\mathcal{D}$ is the base section of a control.

On the other hand, if $u : [a, b] \to U$ is a smooth control, then its base section c is an immersion (i.e. $\dot c = \mathbf{T} \circ \rho \circ u \neq 0$). Therefore, one can find a finite subdivision of $[a, b]$ such that on every subinterval, the section u along c can be extended to a section σ of ν (see for instance [1]). It is then easily seen that the restriction of c to each subinterval is an integral curve of the vector field $\mathbf{T} \circ \rho \circ \sigma$ in $\mathcal{D}$. This implies that the base curve of a control is a concatenation of integral curves associated with vector fields in $\mathcal{D}$.

Consider the following relation on M: $x \to y$ iff there exists a composite flow Φ of vector fields in $\mathcal{D}$ such that $\Phi_T(x) = y$ for some $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell$ with $t_i \geq 0$ for all $i = 1, \ldots, \ell$. This relation is reflexive, transitive, but not symmetric since if $x \to y$ then $\tau(x) \leq \tau(y)$.

Using the notations introduced above, we can say that $x \to y$ iff there exists a control $u : [a, b] \to U$ with base curve c such that $c(a) = x$ and $c(b) = y$. This justifies the following definition: the set of *reachable points from* x, denoted by R_x, is the set of all $y \in M$ such that $x \to y$. In particular, if u is a control with base section c taking x to y, we shall write $x \xrightarrow{u} y$. It is always assumed that, if we have fixed a control u taking x to y, then we shall only consider an ordered family of vector fields $(X_\ell, \ldots, X_1)$, with composite flow Φ such that $\Phi_T(x) = y$ and such that u is induced by the ordered family $(\sigma_\ell, \ldots, \sigma_1)$ of sections of ν and, in addition, $X_i = \mathbf{T} \circ \rho \circ \sigma_i$.

Using this convention, assume that $x \xrightarrow{u} y$, consider a composite flow Φ and composite flow parameter $T \in \mathbb{R}^\ell$, and let γ denote the concatenation associated with Φ and T through x. Using the same notations from above, assume that $s \in]a_{i-1}, a_i]$. Then we define $T\Phi^b_s : T_{\gamma(s)}M \to T_yM$ by

$$T\Phi^b_s = T\phi^\ell_{t_\ell} \circ \ldots \circ T\phi^i_{a_i - s}.$$

Definition 2. The *cone of variations* C_yR_x at $y \in R_x$ is the convex cone in T_yM generated by the following set of tangent vectors in T_yM:

$$\{T\Phi^b_s(Y(\gamma(s))) \mid \forall s \in]a, b] \text{ and } Y \in \mathcal{D}\} \;\cup\; \{T\Phi^b_s(-\dot\gamma(s)) \mid \forall s \in]a, b]\},$$

i.e. C_yR_x consists of all finite linear combinations, with nonnegative coefficients, of tangent vectors in the above set.

In [5] we have proven the following theorem, which is fundamental for a proof of the maximum principle.

Theorem 1. *Given any curve $\theta : [0,1] \to M$ through y at $t = 0$ and whose tangent vector at $t = 0$ belongs to the interior of the cone C_yR_x, then there exists an $\epsilon > 0$ such that $\theta(t) \in R_x$ for all $t \in [0, \epsilon]$.*

In particular, the above theorem implies that the interior of the cone C_yR_x can be regarded as a "tangent cone" to the set of reachable points. Before proceedings we introduce the *vertical cone of variations* V_yR_x at $y \in R_x$, which is the convex cone generated by the set: $\{T\Phi^b_s(Y(\gamma(s))-\dot\gamma(s)) \mid \forall s \in]a,b] \text{ and } Y \in \mathcal{D}\}$. It is easily seen that V_yR_x is contained in C_yR_x. Many results will be formulated in terms of this vertical cone of variations. For simplicity, we always assume that C_yR_x has a nonempty interior.

2.2. **Optimality and the extended control structure.** In the following we concentrate on optimal control problems (with fixed endpoints). Assume that *a cost function* $L \in C^\infty(U)$ is given. One of the basic ideas in the book of L.S. Pontryagin et al. in [7], was to consider an 'extended geometric control structure' in which the cost function becomes part of the anchor map. More specifically, consider the manifolds $\overline{M} = M \times \mathbb{R}$, $\overline{U} = U \times \mathbb{R}$ and let $\overline{\tau} : \overline{M} \to \mathbb{R} : (m, J) \mapsto \tau(m)$, $\overline{\nu} : \overline{U} \to \overline{M} : (u, J) \mapsto (\nu(u), J)$ denote two projections, making $\overline{M}$ and $\overline{U}$ into bundles over respectively $\mathbb{R}$ and $\overline{M}$. The extended anchor map $\overline{\rho} : \overline{U} \to J^1\overline{\tau}$ is defined by $\overline{\rho}(u, J) = (\rho(u), J, L(u))$ (where we have used the standard identification of $J^1\overline{\tau}$ with $J^1\tau \times \mathbb{R}^2$). From now on we shall refer to $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ as the *extended geometric control structure*. It is instructive to see how the control structure (τ, ν, ρ) and the extended control structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ can be related. Consider a control u in the control structure (τ, ν, ρ) and define the following section of $\overline{\tau} \circ \overline{\nu}$: $\overline{u}(t) = (u(t), J(t))$ with

$$J(t) = J_0 + \int_a^t L(u(s))ds,$$

where $J_0 \in \mathbb{R}$ can be chosen arbitrary (note that $J(b) = J_0 + \mathcal{J}(u)$). It is easily seen that $\overline{u}$ is a control in the extended control structure: $\overline{\rho} \circ \overline{u} = j^1\overline{c}$ follows from

$$\rho \circ u = j^1c, \text{ and } \dot J(t) = L(u(t)).$$

By reversing the above arguments, one can prove that any control in the extended control structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ determines a control in the control structure (τ, ν, ρ).

In particular, we have that if $x \xrightarrow{u} y$ with cost $\mathcal{J}(u)$, then $(x, J_0) \xrightarrow{\overline{u}} (y, J_0+\mathcal{J}(u))$ in the extended control structure (for arbitrary $J_0 \in \mathbb{R}$). And, vice versa, if

$$(x, J_0) \xrightarrow{\overline{u}} (y, J_1)$$

then there exists a control u such that $x \xrightarrow{u} y$ and $\mathcal{J}(u) = J_1-J_0$. These observations justify the choice for calling the coordinate J the *cost coordinate*.

Assume that we fix a control u taking x to y and the associated control $\overline{u}$ in the extended control structure taking $(x, 0)$ to $(y, \mathcal{J}(u))$ (i.e. we fix $J_0 = 0$). Similar to the construction in the control structure (τ, ν, ρ), we can associate with $\overline{u}$ a composite flow and composite flow parameter and, consequently, we can introduce the cone of variations $C_{(y,\mathcal{J}(u))}R_{(x,0)}$ in the extended control structure and apply Theorem 1, in the following sense. If u is optimal, then the tangent vector $-\partial_J$, corresponding to the cost coordinate, can not be contained in the interior of the cone of variations $C_{(y,\mathcal{J}(u))}R_{(x,0)}$. Indeed, using the notations from Theorem 1, the curve defined by $\theta(t) = (y, \mathcal{J}(u) - t)$ for $t \in [0, 1]$ satisfies $\dot\theta(0) = -\partial_J$. Then, if $-\partial_J$ is contained in the interior of $C_{(y,\mathcal{J}(u))}R_{(x,0)}$, there exists an $\epsilon > 0$ such that

$(x,0) \to (y, \mathcal{J}(u) - \epsilon)$ or, equivalently, there exists a control taking x to y with cost $\mathcal{J}(u) - \epsilon$, which is impossible since u is assumed to be optimal.

The condition that $-\partial_J$ is not contained in the interior of $C_{(y,\mathcal{J}(u))}R_{(x,0)}$ can be translated into a 'differential equation' and a maximality condition, which are known as the necessary conditions provided by the maximum principle. In the remainder of this section we explain how these conditions are obtained.

2.3. **Multipliers and the maximum principle.** Before proceeding we still have to introduce a few additional concepts. First, let $V\tau$ denote the bundle of vertical tangent vectors to τ, i.e. $V\tau = \{w \in TM \mid T\tau(w) = 0\}$, with dual bundle $V^*\tau$. Note that the vertical cone of variations V_yR_x is entirely contained in $V_y\tau$, justifying the denomination "vertical cone". Consider the fibred product $U \times_M V\tau$ of U and $V\tau$ over M, i.e. $(u,w) \in U \times_M V\tau$ if $u \in U$, $w \in V\tau$ and $\nu(u) = \tau_M(w)$, where $\tau_M : TM \to M$ denotes the tangent bundle projection. Let $\lambda \in \mathbb{R}$ and define a section σ_λ of the fibration $U \times_M T^*M \to U \times_M V^*\tau$ by $\sigma_\lambda(u,\eta) = (u,\alpha)$ where α is uniquely determined by the conditions

1. $\langle \alpha, \mathbf{T}(\rho(u))\rangle + \lambda L(u) = 0$, and
2. α projects onto η.

The map σ_λ is smooth, as is easily seen from the coordinate expression $\alpha = -(\rho^i(u)\eta_i + \lambda L(u))dt + \eta_i dx^i$. Using σ_λ we can pull-back the canonical symplectic two-form ω on T^*M to a closed two-form ω_λ on $U \times_M V^*\tau$: $\omega_\lambda = \sigma_\lambda^*(pr_2^*\omega)$ with $pr_2 : U \times_M T^*M \to T^*M$, the standard projection onto the second factor. Herewith, we can introduce the notion of a multiplier of a control.

Definition 3. Let u denote a control with base c. A pair $(\eta(t), \lambda)$, where $\eta(t)$ is a piecewise smooth section of $V\tau$ along c and $\lambda \in \mathbb{R}$, is called a multiplier of u if the following three properties are satisfied:

1. $i_{(\dot{u}(t),\dot{\eta}(t))}\omega_\lambda = 0$ (on every smooth part of the curve $(u(t), \eta(t))$),
2. for any fixed $t \in I$, the function $u' \mapsto \langle \sigma_\lambda(u(t), \eta(t)), \mathbf{T}(\rho(u'))\rangle + \lambda L(u')$, defined on $\nu^{-1}(c(t))$, attains a global maximum for $u' = u(t)$,
3. $(\eta(t), \lambda) \neq 0$ for any $t \in I = [a,b]$.

Another concept that we will need is that of the 'dual of a cone'. Let C denote a convex cone in a vector space $\mathcal{V}$. The dual convex cone C^* in $\mathcal{V}^*$ is defined by

$$C^* = \{\alpha \in \mathcal{V}^* \mid \langle \alpha, v\rangle \leq 0,\ \forall v \in C\}.$$

A result from [7], which we shall take for granted here, tells us that, given two convex cones C and C', such that the interior of C has an empty intersection with C', then they can be *separated* in the sense that there exists an $\alpha \in C^*$ for which $\langle \alpha, C'\rangle \geq 0$. It is also instructive to note that $C^* = (\mathrm{cl}(C))^*$ and $C^{**} = \mathrm{cl}(C)$, where $\mathrm{cl}(C)$ denotes the closure of C (these results are taken from [3]).

We are now ready to state the following theorem, establishing a connection between the multipliers of a control and elements in the dual cone of $V_{(y,\mathcal{J}(u))}R_{(x,0)}$. We agree to write elements, say $\overline{\eta}$, of $V^*_{(y,\mathcal{J}(u))}\overline{\tau}$ as (η, λ), where $\eta \in V^*\tau$ and $\lambda \in \mathbb{R}$, using the identification of $V^*_{(y,\mathcal{J}(u))}\overline{\tau}$ with $V^*_y\tau \times \mathbb{R}$, i.e. $\overline{\eta} = \eta + \lambda dJ_{(y,\mathcal{J}(u))}$. A similar identification between $T^*_{(y,\mathcal{J}(u))}\overline{M}$ and $T^*_yM \times \mathbb{R}$ can be made valid.

Theorem 2. *Let (η, λ) denote a multiplier, then the covector in $V^*_{(y,\mathcal{J}(u))}\overline{\tau}$ defined by $(\eta(b), \lambda)$, is contained in the dual cone $(V_{(y,\mathcal{J}(u))}R_{(x,0)})^*$. On the other hand, if $(\eta_y, \lambda) \in (V_{(y,\mathcal{J}(u))}R_{(x,0)})^*$ then there exists a multiplier (η, λ) with $\eta(b) = \eta_y$.*

A straightforward corollary of this theorem is the *maximum principle*, formulated in the following way: *if a control u is optimal then there exists a multiplier (η, λ) with $\lambda \leq 0$.* Indeed, since $-\partial_J$ can not belong to the interior of $C_{(y,\mathcal{J}(u))}R_{(x,0)}$ if u is optimal, there must exist an element in $(C_{(y,\mathcal{J}(u))}R_{(x,0)})^*$, say (α_y, λ), such that $\langle(\alpha_y, \lambda), -\partial_J\rangle \geq 0$ or, more specifically, $\lambda \leq 0$. Since $V_{(y,\mathcal{J}(u))}R_{(x,0)} \subset C_{(y,\mathcal{J}(u))}R_{(x,0)}$, the restriction $(\eta_y, \lambda) \in V^*\overline{\tau}$ of (α_y, λ) to $V\overline{\tau}$ is contained in the dual cone of $V_{(y,\mathcal{J}(u))}R_{(x,0)}$. We conclude that there exists a $(\eta_y, \lambda) \in (V_{(y,\mathcal{J}(u))}R_{(x,0)})^*$ with $\lambda \leq 0$. Using Theorem 2, we have that there exists a multiplier (η, λ) such that $\lambda \leq 0$. This necessary condition is precisely the necessary condition of the maximum principle. The following theorem is a minor generalisation of a theorem proven in [5] which will be used in the following section.

Theorem 3. *Assume that u is a control taking x to y and let (η, λ) denote a multiplier. Fix a composite flow $\overline{\Phi}$ in the extended setting, such that $\overline{\Phi}_T((x,0)) = (y, \mathcal{J}(u))$ and with associated vertical cone of variations $V_{(y,\mathcal{J}(u))}R_{(x,0)}$. Then*

$$\eta(t) + \lambda dJ = (T\overline{\Phi}_t^b)^* \left(\eta(b) + \lambda dJ\right).$$

Moreover, if we define $\alpha(t) = \sigma_\lambda(u(t), \eta(t))$, then we also have:

$$\alpha(t) + \lambda dJ = (T\overline{\Phi}_t^b)^* \left(\alpha(b) + \lambda dJ\right).$$

Remark 1. Let (α_y, λ) denote an element in $(C_{(y,\mathcal{J}(u))}R_{(x,0)})^*$. By definition it is easily seen that both $\dot{\gamma}(b) + L(u)\partial_J$ and $-(\dot{\gamma}(b) + L(u)\partial_J)$ are elements in $C_{(y,\mathcal{J}(u))}R_{(x,0)}$. Since (α_y, λ) is an element in the dual cone, we have that

$$\langle(\alpha_y, \lambda), \pm\left(\dot{\gamma}(b) + L(u)\partial_J\right)\rangle \leq 0,$$

which is only possible if $\langle(\alpha_y, \lambda), \dot{\gamma}(b) + L(u)\partial_J\rangle = 0$. More specifically, since $\dot{\gamma}(b) = \mathbf{T}(\rho(u(b)))$, we obtain that $\alpha_y = \sigma_\lambda(u(b), \eta_y)$, where η_y is the restriction of α_y to elements in $V_y\tau$.

3. **Optimal control problems with variable endpoints.** Recall the definition of an optimal control problem with variable endpoints from Section 1 and consider two immersed submanifolds $i : S_i \to M$ and $j : S_f \to M$. We only treat the specific case where either $S_i = \{x\}$ or $S_f = \{y\}$. We would like to mention explicitly that, in this section, a control u taking a point $x \in i(S_i)$ to a point $y \in j(S_f)$ is said to be optimal if, given any other control u' taking $x' \in i(S_i)$ to $y' \in j(S_f)$, then $\mathcal{J}(u) \leq \mathcal{J}(u')$. Note that this notion of optimality is stronger then the notion of optimality from the previous section. Consequently, we can already conclude that if u is optimal among all controls taking points from S_i to S_f, then there exists a multiplier (η, λ) with $\lambda \leq 0$. In this section we shall construct two more conditions (one for the case where $S_i = \{x\}$ and one if $S_f = \{y\}$), which are known as the *transversality conditions* (see [7]). The first of these two transversality conditions (i.e. for $S_i = \{x\}$) is easily derived in the next section. The second condition ($S_f = \{y\}$) takes some more effort. In the sequel, we use the notations introduced in the previous section.

3.1. **The transversality condition at the endpoint ($\mathbf{S_i = \{x\}}$).** Assume that u is optimal. We consider the following tangent vector in $T_{(y,\mathcal{J}(u))}\overline{M}$: $\overline{w} = -\partial_J + Tj(v)$, where $v \in T_yS_f$ can be chosen arbitrarily. We shall now prove by contradiction, that $\overline{w}$ is not contained in the interior of $C_{(y,\mathcal{J}(u))}R_{(x,0)}$.

Assume that $\overline{w}$ is in the interior of $C_{(y,\mathcal{J}(u))}R_{(x,0)}$. Fix any curve in S_f, say $\theta : [0,1] \to S_f$, satisfying the condition that the tangent vector at $t = 0$ equals v, i.e. $\dot{\theta}(0) = v$. We now define a curve $\overline{\theta} : [0, 1] \to \overline{M}$ in the extended control structure and apply Theorem 1. More specifically, let $\overline{\theta}(t) = (j(\theta(t)), \mathcal{J}(u) - t) \in \overline{M}$. It is easily seen that $\overline{\theta}(0) = (y, \mathcal{J}(u))$, and that $\dot{\overline{\theta}}(0) = \overline{w} = -\partial_J + Tj(v)$.

Using Theorem 1, we know that there exists an $\epsilon > 0$ such that $(j(\theta(\epsilon)), \mathcal{J}(u) - \epsilon) \in R_{(x,0)}$. In particular, we have that there exists a control taking x to $j(\theta(\epsilon)) \in S_f$ with cost $\mathcal{J}(u) - \epsilon$. However, since u is assumed to be optimal, this is not possible. The same construction can be carried out for any positive multiple of $\overline{w}$, implying that the convex cone, generated by all vectors $\overline{w} = -\partial_J + Tj(v)$ for $v \in Tj(T_yS_f)$, can be separated from the cone of variations $C_{(y,\mathcal{J}(u))}R_{(x,0)}$. Therefore, there exists a hyperplane in $(\alpha_y, \lambda) \in (C_{(y,\mathcal{J}(u))}R_{(x,0)})^*$ such that $\langle(\alpha_y, \lambda), -\partial_J + Tj(v)\rangle \geq 0$, for all $v \in T_yS_f$. In particular, putting $v = 0$, we have $\langle(\alpha_y, \lambda), -\partial_J\rangle \geq 0$, or $\lambda \leq 0$. Following the same arguments as in the preceding section, we may conclude that there exists a multiplier (η, λ) with $\lambda \leq 0$, where $\eta(b)$ equals the restriction of α_y to $V\tau$. In Remark 1 we have stated that $\eta(b)$ satisfies $\sigma_\lambda(u(b), \eta(b)) = \alpha_y$.

Take an arbitrary vector in the tangent space to S_f, i.e. $v \in T_yS_f$. Then $\lambda \leq \langle\alpha_y, v\rangle$ holds by definition of (α_y, λ). This equations is also valid for $-v$, implying that $\lambda \leq -\langle\alpha_y, v\rangle$.

Assume $\lambda = 0$, then $\alpha_y \in (Tj(T_yS_f))^0$ (where we use $\mathcal{V}^0$ denotes the annihilator space in $\mathcal{W}^*$ of a linear subspace $\mathcal{V}$ of $\mathcal{W}$). On the other hand, if $\lambda < 0$, then we obtain $1 \geq \langle\alpha_y, \lambda^{-1}v\rangle$ and $1 \geq -\langle\alpha_y, \lambda^{-1}v\rangle$. Now, since this equation holds for any multiple of v, we find again that $\alpha_y \in (Tj(T_yS_f))^0$. We conclude that if u is optimal, *there exists a multiplier* (η, λ) *with* $\lambda \leq 0$ *and* $\sigma_\lambda(u(b), \eta(b)) \in (Tj(T_yS_f))^0$.

3.2. **The transversality condition at the starting point ($\mathbf{S_f = \{y\}}$).** In order to prove a similar result for the initial submanifold S_i, we are obliged to construct a new control structure (τ', ν, ρ'), which we shall call the *inverse control structure.*

Consider the following bundle $\tau' : M \to \mathbb{R} : x \mapsto -\tau(x)$. And consider the map $\xi : \Gamma(\tau) \to \Gamma(\tau')$ defined as follows, if $c :]a, b[\to M$ is contained in $\Gamma(\tau)$ then $\xi(c) :]-b, -a[\to M, t' \mapsto c(-t')$. It is easily seen that ξ is invertible, and that it induces a bundle morphism between $J^1\tau$ and $J^1\tau'$, which will be denoted by the same letter for the sake of simplicity, and is defined by

$$\xi(j^1_t c) = j^1_{-t}(\xi(c)).$$

It is easily seen that $\xi : J^1\tau \to J^1\tau'$ is fibred over the identity on M and over the mapping $-1 : \mathbb{R} \to \mathbb{R} : t \mapsto -t$, making the following diagram commutative.

$$\begin{array}{ccc} J^1\tau & \xrightarrow{\ \xi\ } & J^1\tau' \\ {\scriptstyle \tau_{1,0}}\downarrow & & \downarrow{\scriptstyle \tau'_{1,0}} \\ M & \longrightarrow & M \\ {\scriptstyle \tau}\downarrow & & \downarrow{\scriptstyle \tau'} \\ \mathbb{R} & \xrightarrow{\ -1\ } & \mathbb{R} \end{array}$$

The total time derivative $\mathbf{T}' : J^1\tau' \to TM$ is related to $\mathbf{T} : J^1\tau \to TM$ as follows: $\mathbf{T}'(\xi(j^1_t c)) = -\mathbf{T}(j^1_t c)$. Define a new anchor map $\rho' : U \to J^1\tau' : u \mapsto \xi \circ \rho(u)$. We shall now further investigate the relation between the geometric control structures (τ', ν, ρ') and (τ, ν, ρ), and in particular between the order relations induced by the families of vector fields $\mathcal{D}'$ and $\mathcal{D}$, respectively.

Assume that $\sigma \in \Gamma(\nu)$. The vector field $\mathbf{T}' \circ \rho' \circ \sigma$ equals $\mathbf{T} \circ \rho \circ \sigma$ up to a minus sign. Indeed, since $\mathbf{T}' \circ \xi = -\mathbf{T}$, we obtain that

$$\mathbf{T}' \circ \rho' \circ \sigma = \mathbf{T}' \circ \xi \circ \rho \circ \sigma = -\mathbf{T} \circ \rho \circ \sigma.$$

Therefore, we have $\mathcal{D}' = -\mathcal{D}$. Let us write the order relation induced by $\mathcal{D}'$ as $\rightarrow'$ and, as before the order relation induced by $\mathcal{D}$ as $\rightarrow$. We then have that, if $x \rightarrow y$, then $y \rightarrow' x$, and vice versa. Indeed, assume that $\Phi_T(x) = y$, where Φ is a composite flow associated with the ordered family $(X_\ell, \ldots, X_1)$ of vector fields in $\mathcal{D}$ and $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell_+$. It is now easily seen that $(\Phi_T)^{-1}$ is precisely Ψ_{T^*}, where Ψ is the composite flow associated with $(-X_1, \ldots, -X_\ell)$ of vector fields in $\mathcal{D}'$ and $T^* = (t_1, \ldots, t_\ell)$. Thus $\Psi_{T^*}(y) = x$ or $y \rightarrow' x$. Moreover, if we consider the cone of variations $C_x R'_y$ in $T_x M$ with respect to the inverse control setting, we obtain that $C_x R'_y = -(T\Phi^b_a)^{-1}(C_y R_x)$. We will now reformulate Theorem 1 in terms of the inverse control structure (τ', ν, ρ') leading to a new theorem in the control structure (τ, ν, ρ).

Let $\theta : [0, 1] \rightarrow M$ denote an arbitrary curve through $x \in M$ at $t = 0$ such that $\dot{\theta}(0)$ is contained in the interior of $C_x R'_y$. Using Theorem 1 we obtain that there exists an $\epsilon > 0$ such that $\gamma(t) \in R'_y$ for $0 \leq t \leq \epsilon$. This brings us to the following theorem, which is merely a reformulation of Theorem 1 applied to (τ', ν, ρ').

Theorem 4. *Let $\theta : [0, 1] \rightarrow M$ with $\gamma(0) = x$. If $-T\Phi^b_a(\dot{\theta}(0))$ is contained in the interior of $C_y R_x$, then there exists an $\epsilon > 0$ such that $\theta(t) \rightarrow y$ for all $0 \leq t \leq \epsilon$.*

We shall now use this result. Assume that $x \xrightarrow{u} y$, such that u is optimal. Let $v \in T_x S_i$ and let $\theta : [0, 1] \rightarrow S_i$ denote a curve with $\dot{\theta}(0) = v$. Let $\overline{\Phi}$ denote a composite flow in the extended setting such that $\overline{\Phi}_T((x, 0)) = (y, \mathcal{J}(u))$ and consider the curve $\overline{\theta} : [0, 1] \rightarrow \overline{M}$ defined by $\overline{\theta}(t) = (i(\theta(t)), t)$, with $\dot{\overline{\theta}}(0) = Ti(v) + \partial_J = \overline{w}$. Assume that $-T\overline{\Phi}^b_a(\overline{w})$ is contained in the interior of the cone of variations $C_{(y,\mathcal{J}(u))} R_{(x,0)}$ in the extended setting. Then, from Theorem 4, it is easily seen that $(i(\theta(\epsilon)), \epsilon) \rightarrow (y, \mathcal{J}(u))$ for some $\epsilon > 0$. This implies that there exists a control u' with $i(\theta(\epsilon)) \xrightarrow{u'} y$ with cost $\mathcal{J}(u') = \mathcal{J}(u) - \epsilon$. Since $\theta(\epsilon) \in S_i$ and since we assumed u to be optimal, this is impossible.

The further analysis is basically the same as the one in Section 3.1. Again we consider the convex cone generated by all tangent vectors of the form $-T\overline{\Phi}^b_a(\partial_J + Ti(v))$, where $v \in T_x S_i$ is arbitrary. Using a result from above, this cone can be separated from the cone of variations $C_{(y,\mathcal{J}(u))} R_{(x,0)}$. This implies the existence of an element (α_y, λ) in the dual cone of $C_{(y,\mathcal{J}(u))} R_{(x,0)}$ such that, in addition,

$$\langle (\alpha_y, \lambda), -T\overline{\Phi}^b_a(\partial_J + Ti(v)) \rangle \geq 0.$$

Consider the restriction η_y of α_y to the set of vertical vectors $V\tau$ (see Remark 1). Since (η_y, λ) is an element of the dual to the vertical cone of variations, we have that there exists a multiplier (η, λ) (with $\eta(b) = \eta_y$ and $\alpha_y = \sigma_\lambda(u(b), \eta_y)$) such that

$$\langle \sigma_\lambda(u(b), \eta(b)) + \lambda dJ, -T\overline{\Phi}^b_a(\partial_J + Ti(v)) \rangle \geq 0,$$

for all $v \in T_x S_i$. If we use Theorem 3, then

$$\langle \sigma_\lambda(u(a), \eta(a)) + \lambda dJ, -\partial_J - Ti(v) \rangle \geq 0.$$

For $v = 0$, we find that $\lambda \leq 0$, which is the standard necessary condition from the maximum principle. If v can be chosen arbitrarily we have $\langle \alpha_y, Ti(v) \rangle \leq -\lambda$. If $\lambda =$

0, then $\langle \alpha_y, Ti(v) \leq 0$ and $-\langle \alpha_y, Ti(v)\rangle \leq 0$, implying that $\alpha_y \in (Ti(T_xS_i))^0$. If, on the other hand, $\lambda < 0$, then $\langle \alpha_y, Ti(\lambda^{-1}v)\rangle \geq 1$ and $-\langle \alpha_y, Ti(\lambda^{-1}v)\rangle \geq 1$. Since this inequality holds for any multiple of v, we obtain once more that $\alpha_y \in (Ti(T_xS_i))^0$. The results from Sections 3.1 and 3.2 can now be combined in the following theorem, which is a geometric version of the maximum principle for optimal control problem with variable endpoint conditions, provided either S_i or S_f reduces to a point.

Theorem 5. *Assume that u is a control taking x to y. If u is optimal among all controls with initial point in S_i and final point in S_f then there exists an extremal (η, λ) such that*

1. $\lambda \leq 0$,
2. $\sigma_\lambda(u(a), \eta(a)) \in (Ti(T_xS_i))^0$, *if* $S_f = \{y\}$, *or*
   $\sigma_\lambda(u(b), \eta(b)) \in (Tj(T_yS_f))^0$, *if* $S_i = \{x\}$.

It is easily seen that this theorem agrees with the results obtained in [7, p 48].

Acknowledgements. This research is supported by a grant from the "Bijzonder onderzoeksfonds" of Ghent University. I am indebted to F. Cantrijn for the many discussions and the careful reading of this paper.

REFERENCES

[1] S. Helgason, *Differential geometry, lie groups, and symmetric spaces*, Academic Press, New York, 1978.
[2] J.F. Cariñena and E. Martínez, *Lie algebroid generalization of geometric mechanics*, Lie algebroids and related topics in differential geometry (Warzawa) (P. Urbański J. Kubarski and R. Wolak, eds.), Banach Center Publications, vol. 54, 2001, pp. 201–215.
[3] G. Köthe, *Topological vector spaces I*, Springer-Verlag, Berlin, 1969.
[4] B. Langerock, *A connection theoretic approach to sub-riemannian geometry*, J. Geom. Phys. (2002), to appear.
[5] ———, *Geometric aspects of the maximum principle and lifts over a bundle map*, Acta. Appl. Math. (2003), to appear.
[6] P. Libermann and C.-M. Marle, *Symplectic geometry and analytical mechanics*, Reidel, Dortrecht, 1987.
[7] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamklelidze and E.F. Mishchenko, *The mathematical theory of optimal processes*, Wiley, Interscience, 1962.
[8] E. Martínez, *Lagrangian mechanics on lie algebroids*, Acta. Appl. Math. **67** (2001), 295–320.
[9] R. Montgomery, *Abnormal minimizers*, Siam Journal on control and optimization **32** (1994), 1605–1620.
[10] M. de León S. Martíez, J. Cortés, *Symmetries in vakonomic dynamics: applications to optimal control*, J. Geom. Phys. **38** (2001), 343–365.
[11] R.S. Strichartz, *sub-Riemannian geometry*, J. Diff. Geom. **24** (1986), 221–263.
[12] ———, *Corrections to "sub-Riemannian geometry"*, J. Diff. Geom. **30** (1989), 595–596.
[13] H.J. Sussmann, *Orbits of families of vector fields and integrability of distributions*, Trans. Amer. Math. Soc. **180** (1973), 171–188.
[14] ———, *An introduction to the coordinate-free maximum principle*, Geometry of Feedback and Optimal Control (New York) (B. Jakubczyk and W. Respondek, eds.), Marcel Dekker, 1997, pp. 463–557.
[15] A. Weinstein, *Lagrangian mechanics and groupoids*, Mechanics Day (P.S. Krishnaprasad, W.F. Shadwick and T.S. Ratiu, ed.), 1995, pp. 207–232.

Department of Mathematical Physics and Astronomy, Ghent University, Krijgslaan 281 S9, B-9000 Gent Belgium
E-mail address: bavo.langerock@rug.ac.be
URL: http://maphyast.rug.ac.be