Geometric aspects of the Maximum Principle and lifts over a bundle map

B. Langerock

Department of Mathematical Physics and Astronomy, Ghent University, Krijgslaan 281 S9,B-9000 Gent Belgium

Abstract

A coordinate-free proof of the Maximum Principle is provided in the specific case of an optimal control problem with fixed time. Our treatment heavily relies on a special notion of variation of curves that consist of a concatenation of integral curves of time-dependent vector fields with unit time component, and on the use of a concept of lift over a bundle map. We further derive necessary and sufficient conditions for the existence of so-called abnormal extremals.

Key words: control theory, Maximum Principle, abnormal extremals, lifts over bundle maps. 49Kxx, 53Cxx.

1 Introduction and preliminary definitions

The results presented in this paper find their origin in some recent work on sub-Riemannian geometry [1], and are also strongly inspired by some ideas developed in the book by L.S. Pontryagin et al. [2]. The main purpose is to provide a comprehensive and coordinate-free proof of the Maximum Principle and, at the same time, to present a version of this principle that may be readily accessible to researchers studying the variational approach to dynamical systems subjected to nonholonomic constraints, also called Vakonomic dynamics. Applications of our results can be found, for instance, in sub-Riemannian geometry, where the problem of characterizing length-minimizing curves (see [1] and references therein) can be solved by means of the Maximum Principle. Also the construction of a Lagrangian and Hamiltonian dynamics on

Email address: bavo.langerock@rug.ac.be (B. Langerock).
URL: http://kraz.rug.ac.be/eng (B. Langerock).

Lie-algebroids (see, for instance, [3–5]) could be tackled using the formalism described in the present paper. This will discussed in a forthcoming paper.

For the present paper, we restrict ourselves to control problems satisfying strong smoothness conditions and we only consider optimal control problems with fixed time. The extension of our results to more general optimal control problems is currently under investigation.

We now first give some preliminary definitions and fix some notations. All manifolds considered in this paper are real, finite dimensional smooth manifolds without boundary, and by smooth we will always mean of class C^∞. The set of (real valued) smooth functions on a manifold B will be denoted by $C^\infty(B)$, the set of smooth vector fields by $\mathcal{X}(B)$ and the set of smooth one-forms by $\mathcal{X}^*(B)$. The set of all smooth (local or global) sections of an arbitrary fibre bundle $\tau : E \to B$ will be denoted by $\Gamma(\tau)$. A family $\mathcal{D}$ of vector fields on a manifold B is said to be *everywhere defined* if, given any point $x \in B$, there exists an element $X \in \mathcal{D}$ such that x is contained in the domain of X.

We now recall the concept of piecewise curve as introduced in [1]. First of all, by a *curve* in an arbitrary manifold B we shall always mean a smooth mapping $c : I \to B$, with $I \subset \mathbb{R}$ a closed interval, and such that c admits a smooth extension to an open interval containing I. A mapping $c : [a, b] \to B$ will be called a *piecewise curve* in B if there exists a finite subdivision $a_0 := a < a_1 < \ldots < a_{\ell-1} < a_\ell := b$ such that the following conditions are fulfilled:

(1) c is left continuous at each point a_i for $i = 1, \ldots, \ell$, i.e. $\lim_{t\to a_i^-} c(t)$ exists and equals $c(a_i)$;
(2) $\lim_{t\to a_i^+} c(t)$ is defined for all $i = 1, \ldots, \ell$ and $\lim_{t\to a_0^+} c(t) = c(a_0)$ (i.e. c is right continuous at $a_0 = a$);
(3) for each $i = 1, \ldots, \ell$, the mapping $c^i : [a_{i-1}, a_i] \to B$, defined by $c^i(t) = c(t)$ for $t \in]a_{i-1}, a_i]$ and $c^i(a_{i-1}) = \lim_{t\to a_{i-1}^+} c(t)$, is smooth (i.e. is a curve in B).

We will also say that the piecewise curve c is "induced by the smooth curves c^i". A piecewise curve which is continuous everywhere will simply be called a *continuous piecewise curve* and it corresponds to what is usually called a 'piecewise smooth curve' in the literature. For example, consider two smooth curves $\gamma^i : [a_{i-1}, a_i] \to B$ with $i = 1, 2$ such that $\gamma^1(a_1) = \gamma^2(a_1)$. According to the above definition, the curve $\gamma : [a_0, a_2] \to B$ defined by $\gamma(t) = \gamma^i(t)$ if $t \in]a_{i-1}, a_i]$ and $\gamma(a_0) = \gamma^1(a_0)$, is a continuous piecewise curve induced by γ^1, γ^2. On the other hand, the piecewise curve $\dot\gamma$, induced by $\dot\gamma^1, \dot\gamma^2$, provides an example of a piecewise curve which, in general, need not be continuous.

In this paper we will also encounter the notion of piecewise section of a bundle fibred over the real line, say $\pi : B \to \mathbb{R}$, the definition of which is similar to that definition of a piecewise curve. A smooth section $\sigma \in \Gamma(\pi)$, defined on a closed interval $I = [a, b]$, is always assumed to be the restriction of a smooth section of π defined on an open interval containing I. Clearly, any section of π determines a curve in B. On the other hand, if $\gamma : I \to B$ is a curve in B, then it will determine a section of π iff $\pi(\gamma(t)) = t$ for all $t \in I$. We say that $\sigma : I = [a, b] \to B$ is a *piecewise section* of π if σ is a piecewise curve in B and, in addition, $\pi(\sigma(t)) = t$ for all $t \in I$. Let $\sigma^i : [a_{i-1}, a_i] \to B$, with $i = 1, \ldots, \ell$ and $a_0 = a < a_1 < \ldots < a_\ell = b$, represent a finite number of curves that induce such a piecewise section σ. Then, the curves σ^i necessarily satisfy $\pi(\sigma^i(t)) = t$, which implies that they are smooth (local) sections of π. We then say that the smooth sections σ^i induce the piecewise section σ. A *continuous piecewise section* σ is a piecewise section $\sigma : I \to B$ such that, in addition, σ is a continuous mapping.

2 A geometric framework for control theory

We can now proceed towards the construction of a differential geometric setting for certain control problems. It should be emphasized that, although our formulation is not the most general one, if only for the rather strong smoothness conditions we impose, it occurs to us that there is a sufficiently large and relevant class of control problems that fit within the framework described below (see for instance [6] for a different approach).

Definition 2.1 *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, (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, $J^1\tau$ is the first jet bundle of $\tau : M \to \mathbb{R}$, with projections $\tau_1 : J^1\tau \to \mathbb{R}$ and $\tau_{1,0} : J^1\tau \to M$. The typical fibre of M plays the role of configuration space and will be denoted by Q. It follows from the definition that we have the following commutative diagram:

$$\begin{tikzcd} U \arrow[r, "\rho"] \arrow[d, "\nu"'] & J^1\tau \arrow[dl, "\tau_{1,0}"] \\ M \arrow[d, "\tau"'] & \\ \mathbb{R} & \end{tikzcd}$$

Let u denote a (local) section of $\tau \circ \nu$, i.e. $u : I \subseteq \mathbb{R} \to U$ with $\tau(\nu(u(t))) = t$. With u we can associate a section c of τ, called the *base section of* u and defined by $c = \nu \circ u$.

Definition 2.2 *A smooth section $u \in \Gamma(\tau \circ \nu)$ is said to be a* smooth control *if $\rho \circ u = j^1 c$, where c denotes the base section of u and $j^1 c$ its first jet extension. A smooth section $c \in \Gamma(\tau)$ is called a smooth controlled section if c is the base section of a smooth control u.*

Let (t, x^i, u^a) denote an adapted coordinate system on U (i.e. adapted to both fibrations τ and ν). The condition for $u \in \Gamma(\tau \circ \nu)$ to be a smooth control is expressed in coordinates as follows: putting $u(t) = (t, x^j(t), u^a(t))$ we must have that $\rho^i(t, x^j(t), u^a(t)) = \dot{x}^i(t)$ for all t. Note that these equations are in agreement with the definition of a control as given in [2, p 56], where $M = \mathbb{R} \times \mathbb{R}^n$ and U is an (open) subset of $M \times \mathbb{R}^k$.

Definition 2.3 *A* control *$u : I = [a, b] \to U$ is a piecewise section of $\tau \circ \nu$ such that u is induced by a finite number of smooth controls and, in addition, its projection $\nu \circ u$ is a continuous piecewise section of τ. A continuous piecewise section $c : I \to M$ of τ will be called be a* controlled section *if it is the base section of a control.*

In the following, we shall show that one can associate with any section of ν a vector field on M. These vector fields will generate controls in the sense that (segments of) their integral curves can be regarded as controlled sections of τ. Moreover, we shall see that also the converse holds: each controlled section appears to consist of a concatenation of integral curves of such vector fields. First, we shall specify what we precisely mean by a "concatenation of integral curves" of vector fields.

Let B denote an arbitrary manifold and consider a finite *ordered* set of, say, ℓ vector fields on B: $(X_\ell, \ldots, X_1)$, which need not all be different. Let $\{\phi^i_s\}$ denote the flow of X_i. The *composite flow* Φ induced by $(X_\ell, \ldots, X_1)$ is then defined as the mapping

$$\Phi : V \subset \mathbb{R}^\ell \times B \to B : ((t_\ell, \ldots, t_1), x) \mapsto \phi^\ell_{t_\ell} \circ \ldots \circ \phi^1_{t_1}(x)\,,$$

whose domain is a subset V of $\mathbb{R}^\ell \times B$. For brevity we shall write $\Phi_T(x)$ for $\Phi((t_\ell, \ldots, t_1), x)$, where $T := (t_\ell, \ldots, t_1)$. We shall sometimes refer to T as the *composite flow parameter*. Assume that $(t_1, x) \in \mathrm{Dom}(\phi^1)$ and that $(t_{i+1}, (\phi^i_{t_i} \circ \ldots \circ \phi^1_{t_1})(x)) \in \mathrm{Dom}(\phi^{i+1})$ for $i = 1, \ldots, \ell - 1$, then $((t_\ell, \ldots, t_1), x) \in \mathrm{Dom}(\Phi)$. It can be proven that $\mathrm{Dom}(\Phi)(= V)$ is an open set (which might be empty) and that for each $x \in B$, $T \mapsto \Phi_T(x)$ is a smooth mapping defined on an open neighborhood of $0 \in \mathbb{R}^\ell$. If we fix a value $T \in \mathbb{R}^\ell$ of the composite flow parameter, then $\Phi_T : B \to B$ determines a diffeomorphism defined on an open subset of B. We refer to [7] (Appendix 3) for further details on composite flows.

Fixing again some $T = (t_\ell, \ldots, t_1) \in pr_1(V) \subset \mathbb{R}^\ell$ (with pr_1 the projection of V onto $\mathbb{R}^\ell$), we can associate with any $x \in \mathrm{Dom}(\Phi_T)$ and with arbitrary $a_0 \in \mathbb{R}$, a continuous piecewise curve $\gamma : [a_0, a_0 + |t_1| + \ldots + |t_\ell|] \to B$ as follows: putting $a_i = a_0 + \sum_{j=1}^{i} |t_j|$ and $\mathrm{sgn}(t_i) := 0, +1, -1$ depending on whether $t_i = 0, t_i > 0, t_i < 0$, respectively, let

$$\gamma(t) = \begin{cases} \phi^1_{\mathrm{sgn}(t_1)(t-a_0)}(x) & \text{for } \; t \in [a_0, a_1] \\ \phi^2_{\mathrm{sgn}(t_2)(t-a_1)}(\phi^1_{t_1}(x)) & \text{for } \; t \in \,]a_1, a_2] \\ \ldots & \\ \phi^\ell_{\mathrm{sgn}(t_\ell)(t-a_{\ell-1})}(\ldots \phi^2_{t_2}(\phi^1_{t_1}(x))) & \text{for } \; t \in \,]a_{\ell-1}, a_\ell], \end{cases}$$

For $t \in]a_{i-1}, a_i[$ we then have $\dot\gamma(t) = \mathrm{sgn}(t_i) X^i(\gamma(t))$ and, hence, the restriction of γ to $]a_{i-1}, a_i[$ is an integral curve of X_i, resp. $-X_i$, for $t_i > 0$, resp. $t_i < 0$. Note that $\gamma(a_\ell) = \Phi_T(x)$, i.e. the endpoint of γ coincides with the image of x under the composite flow map Φ_T. If all $t_i \geq 0$, then we say that γ is a *concatenation of integral curves* through x associated with Φ (or, with the ordered set $(X_\ell, \ldots, X_1)$) and corresponding to the value T of the composite flow parameter. Indeed, we than have $\dot\gamma(t) = X_i(\gamma(t))$ for any $t \in]a_{i-1}, a_i]$.

Let us now return to the geometric control structure (τ, ν, ρ) and recall the definition of the total time derivative operator on the first jet bundle $J^1\tau$, denoted by $\mathbf{T} : J^1\tau \to TM$. This is the vector field along the projection $\tau_{1,0}$ defined by

$$\mathbf{T}(j^1_t c) = Tc\left(\left.\frac{\partial}{\partial t}\right|_t \right),$$

where $c \in \Gamma(\tau)$. Note that $\tau_M(\mathbf{T}(j^1_t c)) = c(t) = \tau_{1,0}(j^1_t c)$. Let σ be a section of ν, then $\rho \circ \sigma$ is a section of $\tau_{1,0}$ and composing it with the total time derivative, we obtain a mapping $\mathbf{T} \circ \rho \circ \sigma : M \to TM$, which is a smooth section of τ_M. The vector field $\mathbf{T} \circ \rho \circ \sigma$ is projectable with respect to τ, and its projection on $\mathbb{R}$ is given by $\frac{\partial}{\partial t}$, i.e. $T\tau \circ \mathbf{T} \circ \rho \circ \sigma = \frac{\partial}{\partial t} \circ \tau$. This implies that, if $\{\phi_s\}$ denotes the flow of $\mathbf{T} \circ \rho \circ \sigma$ and $\{\lambda_s\}$ the flow of $\frac{\partial}{\partial t}$ on $\mathbb{R}$ (i.e. $\lambda_s(t) = t + s$), then the equality $(\tau \circ \phi_s)(m) = \lambda_s(\tau(m))$ holds for any $m \in M$ and for all s in a neighborhood of 0 such that $\phi_s(m)$ is defined.

For a given $\sigma \in \Gamma(\nu)$, let $\{\phi_s\}$ again denote the flow of the vector field $\mathbf{T} \circ \rho \circ \sigma$ on M. Assume that $m \in \mathrm{Dom}(\phi_\epsilon)$, for some fixed $\epsilon > 0$ and let $\tau(m) = t_0$. Consider then the curve $\gamma(t) = \phi_{t-t_0}(m)$ in M, defined on $I = [t_0, t_0 + \epsilon]$. From the above we know that $\tau(\gamma(t)) = \lambda_{t-t_0}(t_0) = t$, implying that $\gamma : I \to M$ can be regarded as a section of τ. Moreover, γ is the base section of the section $u : I \to U$ defined by $u(t) = \sigma(\gamma(t))$. From the definition of u it easily follows that $\mathbf{T}(\rho(u(t))) = \dot{\gamma}(t) = \mathbf{T}(j^1_t\gamma)$, which is equivalent to $\rho(u(t)) = j^1_t\gamma$ and, hence, u is a smooth control. We may therefore conclude that (up to a reparameterization) any integral curve of $X = \mathbf{T} \circ \rho \circ \sigma$, with $\sigma \in \Gamma(\nu)$, determines a controlled section. Indeed, if $\gamma : [a, b] \to M$ is such an integral curve, with $m = \gamma(a)$, then the curve $\gamma' : [\tau(m), \tau(m) + b - a] \to M\,, t \mapsto \gamma(t - \tau(m) + a)$ is a reparametrization of γ, representing a controlled section of τ. From now on, it will always be tacitly assumed that the integral curves γ of a vector field of the form $\mathbf{T} \circ \rho \circ \sigma$ will be parameterized in this way.

We now introduce the following everywhere defined family of vector fields on M:

$$\mathcal{D} = \{\mathbf{T} \circ \rho \circ \sigma \mid \forall \sigma \in \Gamma(\nu)\}. \tag{2.1}$$

Take some arbitrary sections $\sigma_i \in \Gamma(\nu), i = 1, \ldots, \ell$ and put $X_i := \mathbf{T} \circ \rho \circ \sigma_i \in \mathcal{D}$. Then, any concatenation $\gamma : [a_0, a_\ell] \to M$ of integral curves associated to the ordered set $(X_\ell, \ldots, X_1)$, corresponding to a value parameter $T \in \mathbb{R}^\ell_+$ of the composite flow parameter (where $\mathbb{R}^\ell_+ = \{(t_\ell, \ldots, t_1) \mid t_i \geq 0\}$) and such that $a_\ell = a_0 + t_1 + \ldots + t_\ell$, determines a piecewise controlled section if $\tau(\gamma(a_0)) = a_0$. Indeed, it is an easy exercise to see that the piecewise section u induced by $u_i(t) = \sigma_i(\gamma(t))$, for $t \in [a_{i-1}, a_i]$, controls γ (we are using here the notations of Section 1). In the following we prove that the converse also holds, i.e. the base section of any control can be regarded as a concatenation of integral curves of vector fields belonging to $\mathcal{D}$. We only prove the result for smooth controls; the proof for the more general case then easily follows.

Let $u : I \to U$ be a smooth control with base section $c := \nu \circ u$. First, assume that the image $u(I)$ is contained in the domain of an adapted coordinate chart V of U with coordinates (t, x^i, u^a). Consider a smooth extension $\tilde{u}$ of u, defined on an open interval $\tilde{I}$ containing I, i.e. $\tilde{u} : \tilde{I} \to U$ is a local section of $\tau \circ \nu$ with $\tilde{u}(t) = u(t)$ for all $t \in I$. Upon reducing $\tilde{I}$ if necessary, we may always assume that $\tilde{u}(\tilde{I}) \subset V$, and in terms of the adapted coordinates on V we can then write $\tilde{u}(t) = (t, x^i(t), \tilde{u}^a(t))$. We can now define a local section σ of ν on the open subset $V' = \nu(V) \cap \tau^{-1}(\tilde{I})$ of M as follows: $\sigma(t, x^i) = (t, x^i, \tilde{u}^a(t)), \forall (t, x^i) \in V'$. The map $\rho \circ \sigma$ determines a section of $\tau_{1,0}$ satisfying $\rho \circ \sigma(c(t)) = j^1c(t)$ for any $t \in I$. This implies that c is an integral curve of $\mathbf{T} \circ \rho \circ \sigma$. In case the image set $u(I)$ is not fully contained in an adapted coordinate chart, we can always cover the compact set $u(I)$ with a finite number of adapted coordinate charts and choose a subdivision of I such that the image of each subinterval is entirely contained in one of these coordinate charts. The construction above can then be carried out for the restriction of u to each of these subintervals, and it readily follows that the base section c is a concatenation of integral curves of vector fields in $\mathcal{D}$. As mentioned above, the extension of this proof to the case of general controls is straightforward. Summarizing, we have shown that the following property holds.

Proposition 2.4 *A continuous piecewise section $c : I \to M$ is a controlled section iff c is a concatenation of integral curves of vector fields in $\mathcal{D}$.*

With the family of vector fields $\mathcal{D}$ on M we can associate a 'quasi-order relation' R on M (i.e. a reflexive and transitive relation) as follows: R is the subset of $M \times M$ defined by $(m, n) \in R$ if there exists a control $u : [a, b] \to U$ such that $\nu(u(a)) = m$ and $\nu(u(b)) = n$ (we will say that 'the control u takes m to n'). For brevity we shall also denote $(m, n) \in R$ by $m \to n$, and if we want to indicate the control u explicitly, we will write $m \xrightarrow{u} n$. From Proposition 2.4 it follows that $m \to n$ iff there exists a composite flow Φ associated with an ordered set $(X_\ell, \dots, X_1)$, with $X_i \in \mathcal{D}$, such that $n = \Phi_T(m)$ for some $T \in \mathbb{R}^\ell_+$. For any $m \in M$, the subset $R_m \subset M$, defined by

$$R_m = \{n \in M \mid m \to n\},$$

is called the *set of reachable points from m*.

In the next section we will first show that a quasi-order relation can be associated to any everywhere defined family of vector fields on an arbitrary manifold, and that the notion of 'set of reachable points' can be introduced in this more general setting. We will then investigate some properties of a set of reachable points that will play an important role in the further treatment.

3 Some properties of the set of reachable points

Given an everywhere defined family of vector fields $\mathcal{D}$ on an arbitrary manifold B one can define a quasi-order relation R on B as follows: for $x, y \in B$ we put $(x, y) \in R$ if there exists a composite flow Φ, associated with an ordered set $(X_\ell, \dots, X_1)$, $X_i \in \mathcal{D}$, such that $\Phi_T(x) = y$ for some $T \in \mathbb{R}^\ell_+$. We then also write $x \xrightarrow{(\Phi,T)} y$ (or simply $x \to y$). As described in the previous section, the concatenation of integral curves through x, determined by Φ and T, is a continuous piecewise curve γ such that $\dot\gamma(t) \in \mathcal{D}$ for all t where the derivative exists. As in the previous section, we can then define the set of reachable points from $x \in B$ as the subset $R_x = \{y \in B \mid x \to y\}$. Note that $R_x \neq \emptyset$ for all $x \in B$, since $\mathcal{D}$ is assumed to be everywhere defined.

Let D denote the smallest generalized integrable distribution, generated by $\mathcal{D}$ (in the sense of H.J. Sussmann, see e.g. [7]) and let us denote the leaf of D

through a given point $x \in B$ by L_x. Recall that D_x is defined as the space spanned by all tangent vectors of the form $T\Phi_T(Y((\Phi_T)^{-1}(x)))$, for $Y \in \mathcal{D}$, Φ a composite flow associated with a (finite) ordered set of vector fields belonging to $\mathcal{D}$, and $T \in \mathbb{R}^\ell$ such that $x \in \mathrm{Im}(\Phi_T)$. Then it is a simple exercise to see that $R_x \subset L_x$ for any $x \in B$. If $\mathcal{D} = -\mathcal{D}$, then the relation R is symmetric. Indeed, if $\Phi_T(x) = y$, with Φ the composite flow determined by $(X_\ell, \ldots, X_1)$ and $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell_+$, then $(\Phi_T)^{-1}(y) = x$ and an elementary computation shows that $(\Phi_T)^{-1} = \Psi_{T^*}$, where Ψ is the composite flow corresponding to the ordered set $(-X_1, \ldots, -X_\ell)$ (where, by assumption, $-X_i \in \mathcal{D}$) and $T^* = (t_1, \ldots, t_\ell)$, i.e. we also have $y \to x$. In this case R determines an equivalence relation for which the equivalence classes are precisely the leafs of the foliation of the smallest integrable distribution D generated by $\mathcal{D}$, i.e. $R_x = L_x$ for any $x \in B$.

Remark 3.1 It should be emphasized here that the everywhere defined family of vector fields (2.1) associated to a control structure, can never be invariant under multiplication by -1 since, by construction, each vector field belonging to this $\mathcal{D}$ is of the form $\mathbf{T} \circ \rho \circ \sigma$ for some $\sigma \in \Gamma(\nu)$ and, therefore, projects onto the fixed vector field $\frac{\partial}{\partial t}$ on $\mathbb{R}$. Moreover, the relation $m \to n$ is an order relation (i.e. transitive, reflexive and not symmetric) since, if $m \to n$ then $\tau(m) \leq \tau(n)$ holds.

We will now investigate the local structure of the set of reachable points R_x for a given $x \in B$. For that purpose we will introduce a special class of variations of a concatenation of integral curves of vector fields in $\mathcal{D}$, connecting x with some $y \in R_x$, such that these variations will lead us from x to points in a neighborhood of y that also belong to R_x. The following description is merely intended to give a general intuitive idea of the kind of variation we have in mind. We will be more specific later on.

Consider the composite flow Φ corresponding to an ordered set of, say, ℓ vector fields in $\mathcal{D}$, and let $T \in \mathbb{R}^\ell_+$ be such that $\Phi_T(x) = y$. Let $\gamma : [a, b] \to B$ be the concatenation of integral curves induced by Φ and T, as constructed in the previous section, with $\gamma(a) = x$ and $\gamma(b) = y$. Roughly speaking, a *variation of* γ consists of a 1-parameter family of continuous piecewise curves $\gamma_\epsilon : [a, b] \to B$, where ϵ varies over an open interval containing 0, such that the following conditions are verified:

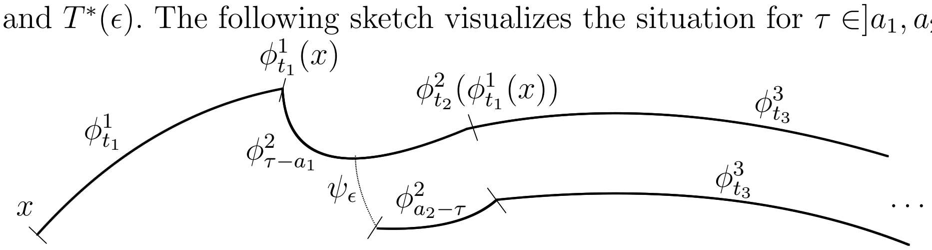
(1) $\gamma_0 = \gamma$;
(2) for all ϵ, $\gamma_\epsilon(a) = x$;
(3) for any $\epsilon \geq 0$ we have that γ_ϵ is a concatenation of integral curves of vector fields in $\mathcal{D}$;
(4) the map $\epsilon \mapsto \gamma_\epsilon(b)$ is a smooth curve through b.

The tangent vector to the curve $\epsilon \mapsto \gamma_\epsilon(b)$ at $\epsilon = 0$ is called the *tangent* *vector to the variation* γ_ϵ (note that $\gamma_0(b) = \gamma(b) = \Phi_T(x) = y$). Rather than considering all possible variations satisfying the above conditions, we will mainly deal with a specific class of variations, to be determined below, called *single variations*. It will be shown that the tangent vectors at y to these single variations generate a convex cone in D_y (where we recall that D refers to the smallest integrable distribution generated by $\mathcal{D}$) and, moreover, we will prove that each vector belonging to this cone is in fact a tangent vector to a variation. If we agree to call *dimension of a cone* the dimension of the linear space generated by all vectors belonging to the cone, then the main result of this section can be summarized as follows: if the dimension of the cone of tangent vectors at y to single variations equals the dimension of D_y, say d, then there exists a coordinate chart V on the leaf L_y, with $y \in V$ and coordinate functions denoted by $(x^1, \ldots, x^d)$, such that for any point $z \in V$ for which $x^i(z) \geq 0$ for all $i = 1, \ldots, d$, we have that $z \in R_x$.

Consider again a concatenation of integral curves $\gamma : [a, b] \to B$ associated with the composite flow $\Phi : V \subset \mathbb{R}^\ell \times B \to B$ of an ordered set of ℓ vector fields $(X_\ell, \ldots, X_1)$ in $\mathcal{D}$, and with a given value $T \in \mathbb{R}^\ell_+$ of the corresponding composite flow parameter, such that $\gamma(a) = x$ and $\gamma(b) = \Phi_T(x) = y$. We now proceed towards the construction of what will be called a single variation of γ. Let $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell_+$ and put $a_0 = a, a_\ell = b$ and $a_i = a_{i-1} + t_i$ for $i = 1, \ldots, \ell$. Choose an arbitrary point $\tau \in]a_0, a_\ell]$ and let Y be any vector field on B such that $\gamma(\tau)$ belongs to the domain of Y. To fix the ideas, let us assume that $a_{i-1} < \tau \leq a_i$. The flow of Y will be denoted by $\{\psi_s\}$ and, as before, $\{\phi^i_s\}$ denotes the flow of X_i. We can then consider the composite flow $\Phi^* : V' \subset \mathbb{R}^{\ell+2} \times B \to B$, associated with the ordered set of $\ell + 2$ vector fields $(X_\ell, \ldots, X_i, Y, X_i, \ldots, X_1)$. Next, define

$$\begin{aligned} T^* : \mathbb{R} \to \mathbb{R}^{\ell+2} &: \\ \epsilon \mapsto T^*(\epsilon) &= (t_\ell, \ldots, t_{i+1}, a_i - \tau, \epsilon, \tau - a_{i-1}, t_{i-1}, \ldots, t_1)\,. \end{aligned} \tag{3.1}$$

It is easily seen that there exists an open neighborhood $\tilde{I} \subset \mathbb{R}$ of 0, such that x is contained in the domain of the map $\Phi^*_{T^*(\epsilon)}$ for all $\epsilon \in \tilde{I}$. For each $\epsilon \in \tilde{I}$, let γ_ϵ denote the concatenation of integral curves through x corresponding to Φ^* and $T^*(\epsilon)$. The following sketch visualizes the situation for $\tau \in]a_1, a_2]$:



The tangent vector to the smooth curve $\epsilon \mapsto \gamma_\epsilon(b) = \Phi^*_{T^*(\epsilon)}(x)$ at $\epsilon = 0$ is then given by

$$\left.\frac{\partial}{\partial \epsilon}\right|_0 \Phi^*_{T^*(\epsilon)}(x) = T\Phi^{a_\ell}_\tau(Y(\gamma(\tau))) \in T_y B\,,$$

where, in order to simplify the notations, we have introduced the mapping $T\Phi^{a_\ell}_\tau : T_{\gamma(\tau)}B \to T_y B$, given by

$$T\Phi^{a_\ell}_\tau(v) = T\phi^\ell_{t_\ell} \circ T\phi^{\ell-1}_{t_{\ell-1}} \circ \ldots \circ T\phi^i_{a_i-\tau}(v)\,, \quad \forall v \in T_{\gamma(\tau)}B\,.$$

Assume now that $Y \in \mathcal{D}$. Then one can see that the 1-parameter family of continuous piecewise curves γ_ϵ satisfies the conditions proposed above for a variation of γ.

Next, suppose we take $Y = -X_i$ and $\tau \in]a_{i-1}, a_i]$ for some $i \in \{1, \ldots, \ell\}$, then for $\epsilon > 0$ (but sufficiently small) and for any $t \in]\tau, \tau + \epsilon]$, the tangent vector $\dot\gamma_\epsilon(t)$ to the concatenation of integral curves through x, induced by Φ^* and $T^*(\epsilon)$, in general will not be contained in $\mathcal{D}$ since $-X_i$ does not have to belong to $\mathcal{D}$. Consequently, if $-X_i \notin \mathcal{D}$, the γ_ϵ resulting from the choice $Y = -X_i$ is, strictly speaking, not a variation in the sense put forward above. However, we can easily remedy the situation by constructing a "reduced" composite flow as follows. Putting $\widehat{T}(\epsilon) = (t_\ell, \ldots, t_i - \epsilon, \ldots, t_1) \in R^\ell$, we see that for ϵ sufficiently small, $\Phi_{\widehat{T}(\epsilon)}$ is well-defined in a neighborhood of x and, moreover, since $\phi^i_{a_i-\tau} \circ \phi^i_{-\epsilon} \circ \phi^i_{\tau - a_{i-1}} = \phi^i_{t_i-\epsilon}$, it follows that $\Phi^*_{T^*(\epsilon)} = \Phi_{\widehat{T}(\epsilon)}$. The concatenation of integral curves determined by Φ and $\hat{T}(\epsilon)$ does verify the conditions for a variation of γ. The tangent vector at $\epsilon = 0$ to this "reduced" variation equals

$$\left.\frac{\partial}{\partial \epsilon}\right|_0 \Phi^*_{T^*(\epsilon)}(x) = \left.\frac{\partial}{\partial \epsilon}\right|_0 \Phi_{\widehat{T}(\epsilon)}(x) = -T\Phi^{a_\ell}_\tau(X_i(c(\tau))).$$

We have thus shown that if $\tau \in]a_{i-1}, a_i]$, a variation of the given γ is also determined by the ordered set $(X_\ell, \ldots, X_i, -X_i, X_i, \ldots, X_1)$.

To conclude, if we are given a continuous piecewise curve $\gamma : [a, b] \to B$, with $\gamma(a) = x$, such that γ consists of a concatenation of integral curves determined by the composite flow Φ and composite flow parameter $T = (t_\ell, \ldots, t_1) \in \mathbb{R}^\ell_+$ of an ordered set of vector fields $(X_\ell, \ldots, X_1)$ belonging to $\mathcal{D}$, we introduce the following definition.

Definition 3.2 *A* single variation *of γ is a 1-parameter family of continuous piecewise curves $\gamma_\epsilon : [a, b] \to B$, passing through x, with $\gamma_0 = \gamma$, and such that for each ϵ the corresponding γ_ϵ is the continuous piecewise curve determined* *by the composite flow Φ^* and composite flow parameter $T^*(\epsilon)$ associated to an ordered set of vector fields of the form $(X_\ell, \ldots, X_i, Y, X_i, \ldots, X_1)$ for some $i \in \{1, \ldots, \ell\}$, with $Y \in \mathcal{D} \cup \{-X_i\}$ and where $T^*(\epsilon)$ is given by (3.1). (We will also briefly refer to γ_ϵ as 'the single variation determined by Φ^* and $T^*(\epsilon)$'.)*

For later use we introduce the shorthand notation: $\mathcal{D}_{-X} := \mathcal{D} \cup \{-X_i \,|\, i = 1, \ldots, \ell\}$. Whenever we consider a single variation determined by an ordered set $(X_\ell, \ldots, X_i, Y, X_i, \ldots, X_1)$ for some $Y \in \mathcal{D}_{-X}$, it will always be understood that $Y = -X_j$ can only occur if $i = j$.

Given a single variation γ_ϵ of γ, determined by a composite flow Φ^* and composite flow parameter $T^*(\epsilon)$, one can always obtain a 'new' variation by considering a suitable reparameterization $\epsilon(\epsilon')$. More precisely, let $\epsilon' \mapsto \epsilon(\epsilon')$ denote a smooth map satisfying $\epsilon(0) = 0$ and $\delta = \frac{d\epsilon}{d\epsilon'}(0) > 0$. Then it is not difficult to verify that Φ^* and $T^*(\epsilon(\epsilon'))$ also determine a variation since $\delta > 0$ implies that, in a neighborhood of 0, $\mathrm{sgn}(\epsilon) = \mathrm{sgn}(\epsilon')$. The tangent vector to the curve $\epsilon' \mapsto \Phi^*_{T^*(\epsilon(\epsilon'))}(m)$ at $\epsilon' = 0$ equals $\delta\; T\Phi^{a_\ell}_\tau(Y(\gamma(\tau)))$. From this one can easily derive that any positive multiple of a tangent vector to a single variation is again a tangent vector to a (not necessarily single) variation. Note that if $\delta Y \in \mathcal{D}_{-X}$, then $\delta\; T\Phi^{a_\ell}_\tau(Y(\gamma(\tau)))$ is again a tangent vector to a single variation. In general, however, if $Y \in \mathcal{D}_X$, the vector field δY need not be contained in $\mathcal{D}_{-X}$. All this naturally leads to the following definition.

Definition 3.3 *Let $y \in R_x$ and fix a composite flow Φ, corresponding to an ordered set $(X_\ell, \ldots, X_1)$ of vector fields in $\mathcal{D}$, such that $\Phi_T(x) = y$ for some $T \in \mathbb{R}^\ell_+$. The* variational cone *at y associated to Φ and T, is the cone $C_y R_x(\Phi, T)$ in $T_y B$ consisting of all finite linear combinations, with positive coefficients, of tangent vectors to single variations, i.e.*

$$C_y R_x(\Phi, T) = \{ \sum_{i=1}^{s} \delta^i T\Phi^{a_\ell}_{\tau^i}(Y_i(\gamma(\tau^i))) \mid Y_i \in \mathcal{D}_{-X}, \delta^i \geq 0, \\ \tau^i \in]a_0, a_\ell], s \in \mathbb{N}\}.$$

If no confusion can arise, we will often drop the explicit reference to Φ and T and simply denote the variational cone by $C_y R_x$. It is easily seen that $C_y R_x$ is a convex set. Indeed if $v, w \in C_y R_x$, then $(1-t)v + tw \in C_y R_x$, for any $t \in [0, 1]$. As a consequence of the next lemma it will be seen that any element of $C_y R_x(\Phi, T)$ can be regarded as a tangent vector to a variation of the continuous piecewise curve through x associated with Φ and T. First, we introduce an alternative notation for composite flows which will sometimes be more convenient, in particular when considering compositions of composite flows.

Let $(Z_\ell, \ldots, Z_1)$ denote an ordered family of vector fields on a manifold B, with composite flow Ψ. If $\{\psi^i_s\}$ represents the flow of Z_i for $i = 1, \ldots, \ell$, then it will turn out to be convenient to write $\psi^\ell \star \ldots \star \psi^1$ for the composite flow Ψ, whereby it is understood that $(\psi^\ell \star \ldots \star \psi^1)_T := \psi^\ell_{t_\ell} \circ \ldots \circ \psi^1_{t_1} = \Psi_T$ for any admissible $T = (t_\ell, \ldots, t_1)$. Using this notation, we are able to define the *composition* $\Psi_{(2)} \star \Psi_{(1)}$ of two composite flows $\Psi_{(2)}, \Psi_{(1)}$, with $\Psi_{(i)} = \psi^{\ell_i}_{(i)} \star \ldots \star \psi^1_{(i)}$ for $i = 1, 2$, as follows

$$\Psi_{(2)} \star \Psi_{(1)} = \psi^{\ell_2}_{(2)} \star \ldots \star \psi^1_{(2)} \star \psi^{\ell_1}_{(1)} \star \ldots \star \psi^1_{(1)}.$$

We now have the following result, the proof of which is quite technical. As before, we start from a given continuous piecewise curve $\gamma : [a_0, a_\ell] \to B$, with $\gamma(a_0) = x$, associated to the composite flow of an ordered set of ℓ vector fields $(X_\ell, \ldots, X_1)$ in $\mathcal{D}$, and a fixed value T of the composite flow parameter.

Lemma 3.4 *Consider any finite number of (say, s) tangent vectors to single variations of γ, namely $v_i = T\Phi^{a_\ell}_{\tau^i}(Y_i(\gamma(\tau^i)))$, with $Y_i \in \mathcal{D}_{-X}$ and $\tau^i \in]a_0, a_\ell]$ for $i = 1, \ldots, s$. Then, there exists a composite flow Φ^* associated to $\ell + 2s$ vector fields, and a smooth mapping $T^* : \mathbb{R}^s \to \mathbb{R}^{\ell+2s}, (\epsilon^1, \ldots, \epsilon^s) \mapsto T^*(\epsilon^1, \ldots, \epsilon^s)$ such that:*

(1) $\Phi^*_{T^*(0)} = \Phi_T$*;*
(2) x belongs to the domain of $\Phi^*_{T^*(\epsilon^1,\ldots,\epsilon^s)}$ for all $(\epsilon^1, \ldots, \epsilon^s)$ in some open neighborhood $I^{(s)}$ of $(0, \ldots, 0) \in \mathbb{R}^s$;
(3) for each fixed $(\epsilon^1, \ldots, \epsilon^s) \in I^{(s)}$, with $\epsilon^i > 0$ for all i, the tangent vector to the concatenation of integral curves through x determined by Φ^* and $T^*(\epsilon^1, \ldots, \epsilon^s)$ is everywhere contained in $\mathcal{D}$ (possibly after a 'reduction' of Φ^* in the sense described above) such that, in particular, $\Phi^*_{T^*(\epsilon^1,\ldots,\epsilon^s)}(x) \in R_x$;
(4) the tangent vector at $\epsilon = 0$ to the curve $\epsilon \mapsto \Phi^*_{T^*(\epsilon\delta^1,\ldots,\epsilon\delta^s)}(x)$ equals $\delta^i v_i$, for all $\delta^i \in \mathbb{R}$ (and where the curve is defined on a sufficiently small interval such that $(\epsilon\delta^1, \ldots, \epsilon\delta^s) \in I^{(s)}$).

Proof. Without loss of generality, we may assume that the instants τ^i are ordered in such a way that $\tau^1 \leq \tau^2 \ldots \leq \tau^s$. Moreover, whenever some of the successive τ^i coincide, the ordering should be such that from the corresponding vector fields Y_i, those that do not belong to $\mathcal{D}$ always precede those that do belong to $\mathcal{D}$. More precisely, assume $\tau^i = \ldots = \tau^j$ with $1 \leq i < j \leq s$, and let $\tau^i \in]a_{r-1}, a_r]$ for some $r \in \{1, \ldots, \ell\}$. Then we require that if $Y_k = -X_r$ for some $k \in \{i, \ldots, j\}$, and $-X_r \notin \mathcal{D}$, we have $k < k'$ for all those $k' \in \{i, \ldots, j\}$ for which $Y_{k'} \in \mathcal{D}$. Such an arrangement can always be achieved by simply taking a suitable permutation of the ordered set $(Y_i, \ldots Y_j)$, if necessary. Henceforth, we will always assume, for simplicity, that the Y_i's already appear in the correct ordering.

For $j = 1, \ldots, \ell$, let s_j denote the maximum of the set $\{i \mid \tau^i \in]a_{j-1}, a_j]\}$ and put $s_j = s_{j-1}$ if $\{i \mid \tau^i \in]a_{j-1}, a_j]\} = \emptyset$ and $s_0 = 0$. The number of τ^i's belonging to the j-th subinterval is then given by $n_j = s_j - s_{j-1}$. Let $\{\psi^i_s\}$ denote the flow of Y_i (and, as before, $\{\phi^j_s\}$ refers to the flow of X_j). Using the 'star' notation introduced above, we now consider for each $j \in \{1, \ldots, \ell\}$, the composite flow $\Phi^*_j : \mathbb{R}^{1+2n_j} \times B \to B$ defined by

$$\Phi^*_j = \begin{cases} \phi^j \star \psi^{s_j} \star \phi^j \star \psi^{s_j - 1} \star \ldots \star \phi^j \star \psi^{s_{j-1}+1} \star \phi^j & \text{if } \; n_j > 0, \\ \phi^j & \text{if } \; n_j = 0, \end{cases}$$

and a mapping $T^*_j : \mathbb{R}^{n_j} \mapsto \mathbb{R}^{1+2n_j}$ (where it is understood that if $n_j = 0$, then $T^*_j \in \mathbb{R}$):

$$T^*_j(\epsilon^{s_{j-1}+1}, \ldots, \epsilon^{s_j}) = \begin{cases} (a_j - \tau^{s_j}, \epsilon^{s_j}, \tau^{s_j} - \tau^{s_j - 1}, \epsilon^{s_j - 1}, \ldots, \\ \quad \tau^{s_{j-1}+2} - \tau^{s_{j-1}+1}, \epsilon^{s_{j-1}+1}, \tau^{s_{j-1}+1} - a_{j-1}) \\ \text{if } n_j > 0, \\ (a_j - a_{j-1}) \text{ if } n_j = 0. \end{cases}$$

Next, by Φ^* we denote the 'composition' of all the composite flows Φ^*_j, i.e. $\Phi^* = \Phi^*_\ell \star \ldots \star \Phi^*_1$. Then, Φ^* itself is a composite flow which can be evaluated at points of $\mathbb{R}^{\ell+2s} \times B$. If we define the mapping $T^* : \mathbb{R}^s \to \mathbb{R}^{\ell+2s}$ by

$$T^*(\epsilon^1, \ldots, \epsilon^s) = (T^*_\ell(\epsilon^{s_{\ell-1}+1}, \ldots, \epsilon^{s_\ell}), \ldots, T^*_1(\epsilon^1, \ldots, \epsilon^{s_1})),$$

then it is easily seen that $(T^*(0, \ldots, 0), x) \in \mathrm{Dom}(\Phi^*)$ and $y = \Phi^*_{T^*(0,\ldots,0)}(x))$. This implies, in particular, that there exists an open neighborhood $I^{(s)}$ of $(0, \ldots, 0) \in \mathbb{R}^s$ for which the map $(\epsilon^1, \ldots, \epsilon^s) \mapsto \Phi^*_{T^*(\epsilon^1,\ldots,\epsilon^s)}(x)$ is well defined and, hence, (2) holds. Note that $\Phi^*_{T^*(\epsilon^1,\ldots,\epsilon^s)}(x)$ can still be written as:

$$\Phi^*_{T^*(\epsilon^1,\ldots,\epsilon^s)}(x) = (\Phi^*_\ell)_{T^*_\ell(\epsilon^{s_{\ell-1}+1},\ldots,\epsilon^{s_\ell})} \circ \ldots \circ (\Phi^*_1)_{T^*_1(\epsilon^1,\ldots,\epsilon^{s_1})}(x).$$

For $s = 1$ the definitions of Φ^* and T^* coincide with those encountered in the construction of a single variation. For any $(\delta^1, \ldots, \delta^s) \in \mathbb{R}^s$ and ϵ varying over a sufficiently small interval centered at 0, such that the image of the map $\epsilon \mapsto (\epsilon\delta^1, \ldots, \epsilon\delta^s)$ is contained in $I^{(s)}$, a straightforward, but rather tedious computation shows that the tangent vector to the curve $\epsilon \mapsto \Phi^*_{T^*(\epsilon\delta^1,\ldots,\epsilon\delta^s)}(x)$, at $\epsilon = 0$, equals $\delta^i v_i$, proving (4). It is also easily seen that when putting $\epsilon^i = 0$ for all i, we obtain $\Phi^*_{T^*(0)} = \Phi_T$, proving (1).

The proof of (3) we will be provided for a particular, simplified case from which the idea for the general proof can then be easily deduced. Recall that we have chosen the ordering of the τ^i in such a way that, whenever we have a sequence $\tau^i, \ldots, \tau^j, (i < j)$ with $\tau^i = \tau^{i+1} = \cdots = \tau^j$, those vector fields Y_k which belong to the set $\{-X_1, \ldots, -X_\ell\}$ and which are not contained in $\mathcal{D}$, always appear before all the $Y_{k'} \in \mathcal{D}$ in the sequence $Y_i, \ldots, Y_j$. Consider now the particular case where $a_0 < \tau^1 = \tau^2 = \tau^3 < a_1 < \tau^4$, $Y_1 = -X_1 (\notin \mathcal{D})$ and $Y_2, Y_3 \in \mathcal{D}$. Then,

$$\begin{aligned}(\Phi_1^*)_{T_1^*(\epsilon^1,\epsilon^2,\epsilon^3)} &= \phi^1_{a_1-\tau^1} \circ \psi^3_{\epsilon^3} \circ \psi^2_{\epsilon^2} \circ \phi^1_{-\epsilon^1} \circ \phi^1_{\tau^1-a_0} \\ &= \phi^1_{a_1-\tau^1} \circ \psi^3_{\epsilon^3} \circ \psi^2_{\epsilon^2} \circ \phi^1_{\tau^1-a_0-\epsilon^1}\,.\end{aligned}$$

Therefore, we can define a new composite flow, associated with vector fields in $\mathcal{D}$, by putting $\widehat{\Phi}_1 = \phi^1 \star \psi^3 \star \psi^2 \star \phi^1$, and a new composite flow parameter $\widehat{T}_1(\epsilon^1, \epsilon^2, \epsilon^3) = (a_1 - \tau^3, \epsilon^3, \epsilon^2, \tau^1 - a_0 - \epsilon^1)$. Then $(\Phi_1)^*_{T_1^*(\epsilon^1,\epsilon^2,\epsilon^3)} = (\widehat{\Phi}_1)_{\widehat{T}_1(\epsilon^1,\epsilon^2,\epsilon^3)}$ and, for ϵ^1 sufficiently small, the components of $\widehat{T}_1(\epsilon^1, \epsilon^2, \epsilon^3)$ are positive, from which (3) readily follows for the 'reduced' composite flow $\widehat{\Phi}_1$ and the reduced composite flow parameter $\widehat{T}_1$. A similar reasoning can be applied to the general case, which completes the proof of the lemma. $\square$

The previous lemma implies, among others, that any $v \in C_y R_x(\Phi, T)$ can be regarded as a tangent vector to a variation of the continuous piecewise curve γ through x, determined by Φ and T. Indeed, by definition of the cone $C_y R_x(\Phi, T)$ we can always write v (in a non-unique way) as $v = \sum_{i=1}^s \delta^i v_i$ for a finite number of tangent vectors to single variations $v_i = T\Phi^{a_\ell}_{\tau^i} Y_i(\gamma(\tau^i))$, with $\delta^i > 0$. We can then associate to these v_i a composite flow Φ^*, and a composite flow parameter $T^*(\epsilon^1, \ldots, \epsilon^s)$, as in the above lemma. Then Φ^* and $\epsilon \mapsto T^*(\epsilon\delta^1, \ldots, \epsilon\delta^s)$ determine a one-parameter family of continuous piecewise curves satisfying the conditions for a variation of γ. Moreover, from the above lemma it follows that the tangent vector to the curve $\epsilon \mapsto \Phi^*_{T^*(\epsilon\delta^1,\ldots,\epsilon\delta^s)}(x)$ at $\epsilon = 0$ precisely equals v, which we wanted to demonstrate.

Note that $C_y R_x (= C_y R_x(\Phi, T))$ is entirely contained in D_y (with D, as before, the smallest generalized integrable distribution generated by $\mathcal{D}$). If the dimension of the cone $C_y R_x$ equals $d = \dim D_y$, then this is equivalent to saying that the the interior of the convex cone $C_y R_x$, with respect to the standard vector space topology on D_y, is not empty. Indeed, if we have d independent vectors $v^1, \ldots, v^d \in C_y R_x$, then the interior of the simplex in D_y, determined by the ordered set $(0, v^1, \ldots, v^d)$, is contained in $C_y R_x$. The converse is an immediate consequence of the fact that any (nonempty) open ball in a vector space spans the full space.

Before stating the main result of this section, we recall that L_y denotes the leaf of D passing through y (and, of course, $L_y = L_x$). From the theory of integrable distributions, we know that L_y is an immersed submanifold of B whose dimension equals the rank of D at y.

Theorem 3.5 *Assume that the dimension of the cone $C_y R_x$ equals the dimension d of D_y. Then there exists a coordinate chart V on the leaf L_y, with $y \in V$ and coordinate functions denoted by $(x^1, \ldots, x^d)$, such that for any point $z \in V$ for which $x^i(z) \geq 0$ for all $i = 1, \ldots, d$, we have that $z \in R_x$.*

Proof. By assumption, the linear space spanned by all elements of $C_y R_x$ equals D_y. We can therefore select a basis $\{v_1, \ldots, v_d\}$ of the linear space D_y, with $v_i \in C_y R_x$ for all i. By definition of $C_y R_x$, each v_i can then be written as a finite linear combination of tangent vectors to single variations, i.e.

$$v_i = \sum_{j=1}^{s_i} \delta^j_{(i)} v^{(i)}_j, \; i = 1, \ldots, d, \tag{3.2}$$

for some $\delta^j_{(i)} \in \mathbb{R}_+$, and where each $v^{(i)}_j$ is of the form $v^{(i)}_j = T\Phi^{a_\ell}_{\tau^j_{(i)}} Y^{(i)}_j(\gamma(\tau^j_{(i)}))$ for some $Y^{(i)}_j \in \mathcal{D}_{-X}$, $\tau^j_{(i)} \in]a_0, a_\ell]$. Although these decompositions are not uniquely determined, for the remainder of the proof we assume that for each of the given basis vectors v_i one particular decomposition has been singled out, i.e. we make a fixed choice for the $v^{(i)}_j$ and for the positive real numbers $\delta^j_{(i)}$ appearing in (3.2). In total we thus have $s = s_1 + \ldots + s_d$ tangent vectors to single variations $v^{(i)}_j$ which, however, need not all be different and/or linearly independent. For convenience, we introduce the following ordering: $(v^{(1)}_1, \ldots, v^{(1)}_{s_1}, v^{(2)}_1, \ldots, v^{(d)}_{s_d})$ and we denote an arbitrary element of this ordered set by w_α, with $\alpha = 1, \ldots s$ and such that $w_\alpha = v^{(1)}_\alpha$ for $\alpha = 1, \ldots s_1$, $w_\alpha = v^{(2)}_{\alpha - s_1}$ for $\alpha = s_1 + 1, \ldots, s_1 + s_2$, *etc.* ... . According to Lemma 3.4 we can associate to the s tangent vectors to single variations, w_α, a composite flow Φ^* and a map $T^* : \mathbb{R}^s \to \mathbb{R}^{\ell + 2s}$ such that

(1) $\Phi^*_{T^*(0)} = \Phi_T$,
(2) $\Phi^*_{T^*(\epsilon^1, \ldots, \epsilon^s)}(x) \in R_x$ if all $\epsilon^i \geq 0$,
(3) for any fixed $(\delta^1, \ldots, \delta^s) \in \mathbb{R}^s$, the tangent vector to the curve $\epsilon \mapsto \Phi^*_{T^*(\epsilon\delta^1, \ldots, \epsilon\delta^s)}(x)$ at $\epsilon = 0$ equals $\delta^\alpha w_\alpha$.

With the convention that $s_0 := 0$, we have for any $v \in D_y$ that

$$v = l^i v_i = \sum_{i=1}^{d} \sum_{j=1}^{s_i} l^i \delta^j_{(i)} w_{s_0 + \ldots + s_{i-1} + j} \in D_y \, .$$

Putting

$$(\delta^1, \dots, \delta^s) := (l^1 \delta^1_{(1)}, \dots, l^1 \delta^{s_1}_{(1)}, l^2 \delta^1_{(2)}, \dots, l^d \delta^{s_d}_{(d)}),$$

we can still write v as

$$v = \sum_{\alpha=1}^{s} \delta^\alpha w_\alpha \, .$$

Since the $\delta^j_{(i)}$ in (3.2) have been fixed, it follows that all the coefficients δ^α, appearing in this decomposition of v, are determined unambiguously. Therefore, the following mapping is well-defined:

$$\widetilde{T} : D_y \to \mathbb{R}^{\ell+2s}, \; v \mapsto \widetilde{T}(v) = T^*(\delta^1, \dots, \delta^s),$$

and, clearly, $\widetilde{T}$ is smooth.

From the properties of Φ^* and T^*, one can further deduce that, on a sufficiently small open neighborhood W of the origin in the linear space D_y, the mapping given by

$$f : W(\subset D_y) \to B, \, v \mapsto \Phi^*_{\widetilde{T}(v)}(x)$$

is well-defined and smooth. Moreover, by definition of Φ^*, we have that $f(0) = y$ and $\mathrm{Im}\, f \subset L_y$. Let $j : L_y \hookrightarrow B$ denote the natural inclusion and let us write $\widetilde{f}$ for f, regarded as a mapping from W into L_y, such that the following relation holds: $j \circ \widetilde{f} = f$. Since j is an immersion and f is smooth, it follows that $\widetilde{f} : W(\subset D_y) \to L_y$ is smooth. In view of the natural identification $T_0 D_y \cong D_y$, it is easily proven, using property (3) of Φ^* and T^*, that the tangent map of f at 0 satisfies, for any $v = \delta^\alpha w_\alpha \in D_y$,

$$T_0 f(v) = \left.\frac{d}{d\epsilon}\right|_0 f(\epsilon v) = \left.\frac{d}{d\epsilon}\right|_0 \Phi^*_{T^*(\epsilon\delta^1, \dots, \epsilon\delta^s)}(x) = \delta^\alpha w_\alpha = v \, .$$

This, in turn, implies that $T_0 \widetilde{f} : D_y \to T_y(L_y) \equiv D_y$ is the identity map and, hence, $\widetilde{f}$ induces a diffeomorphism from an open neighborhood $\widetilde{W} \subset W$ of $0 \in D_y$ onto a an open neighborhood V of y in L_y. Hence, to each point $z \in V$ there corresponds a unique $v \in \widetilde{W}$, with $\widetilde{f}(v) = z$ and, with respect to the basis $\{v^i : \; i = 1, \dots, d\}$ of D_y chosen above, we can write $v = l^i v_i$. The open set V then becomes the domain of a local coordinate chart on L_y, with coordinate functions x^i $(i = 1, \dots, d)$ defined by putting $x^i(z) = l^i$. Finally, from property (2) of Φ^* and T^* it follows that for those vectors $v = l^i v_i \in \widetilde{W}$

for which all $l^i \geq 0$, we have $z = f(v) \in R_x$ since, in this case, all the coefficients δ^a appearing in the decomposition $v = \delta^\alpha w_\alpha$ are also non-negative. This completes the proof of the theorem. $\square$

Observe that the coordinate vector fields on L_y corresponding to the special chart constructed in the previous theorem are such that (using the notations from the proof of the theorem) $\left.\dfrac{\partial}{\partial x^i}\right|_y = T_0\widetilde{f}(v^i) = v^i$. This observation will be of use in proving the following result, which is a straightforward consequence of Theorem 3.5.

Corollary 3.6 *Assume that C_yR_x has a non empty interior with respect to the topology of D_y (denoted by $\operatorname{int}(C_yR_x)$). Then, for any curve $\theta : [0,1] \to (L_x =)L_y$ with $\theta(0) = y$ and $0 \neq \dot\theta(0) \in \operatorname{int}(C_yR_x)$ there exists an $\epsilon > 0$ such that $\theta(t') \in R_x$ for $0 \leq t' \leq \epsilon$.*

Proof. As pointed out before, the fact that C_yR_x has nonempty interior implies that the 'dimension' of the cone equals that of D_y and so the previous theorem applies. One can always fix a basis v_i in D_y, with $v_i \in C_yR_x$, such that the $\dot\theta(0)$ is contained in the interior of the simplex spanned by $(0, v^1, \ldots, v^d)$. In particular, this means that $\dot\theta(0) = k^i v_i$ with all $k^i \in]0,1[$. Consider the coordinate chart $(x^1, \ldots, x^d)$ on L_y, in a neighborhood of y, associated with the basis $v^1, \ldots, v^d$ as constructed in Theorem 3.5. Note, in passing, that $x^i(y) = 0$ for all i. Now, since $\left.\frac{\partial}{\partial x^i}\right|_y = v_i$ for $i = 1, \ldots, d$, and putting $\theta^i = x^i \circ \theta$, we find that

$$\left.\frac{d}{dt'}\right|_0 \theta^i(t') = k^i, \text{ for } i = 1, \ldots, d.$$

This implies that for all $i = 1, \ldots, d$, $\dot\theta^i(0) > 0$ and hence, since $\theta^i(0) = 0$, $\theta^i(t') > 0$ for $0 \leq t' \leq \epsilon$ and ϵ sufficiently small, i.e. $x^i(\theta(t')) > 0$ for $i = 1, \ldots, d$. According to Theorem 3.5 this implies that $\theta(t') \in R_x$ for all $0 \leq t' \leq \epsilon$. $\square$

To close this section, we return to the framework of a geometric control structure.

Let (τ, ν, ρ) denote an arbitrary geometric control structure. It is easily seen that the previous definitions and results can be applied, in particular, to the everywhere defined family of vector fields $\mathcal{D} = \{\mathbf{T} \circ \rho \circ \sigma \mid \sigma \in \Gamma(\nu)\}$ on M. Consider a pair $(m, n) \in M \times M$ such that $m \xrightarrow{(\Phi,T)} n$ and let $C_nR_m(\Phi, T)$ denote the associated cone of variations. Since M is fibred over the real line, the kernel of the tangent map $T\tau$ defines a sub-bundle $V\tau = \ker T\tau$ of TM, called the vertical bundle to τ. We will now define a 'sub-cone' of C_nR_m which is vertical in the sense that it is contained in $V_n\tau$ and which satisfies $V_nR_m \subset C_nR_m$.

Definition 3.7 *The* vertical variational cone at n*, associated to* Φ *and* T*, is given by:*

$$V_nR_m(\Phi, T) = \{\sum_{i=1}^{s} \delta^i T\Phi^{a_\ell}_{\tau^i}(Y^i(c(\tau^i)) - \dot{c}(\tau^i)) \\ \mid \delta^i \geq 0, \tau^i \in]a_0, a_\ell], Y^i \in \mathcal{D}, i = 1, \ldots, s\}$$

As for the variational cone, we shall also sometimes simply write V_nR_m if there can be no confusion regarding the related Φ and T.

4 The cost coordinate and optimality

In this section we give a straightforward application of Corollary 3.6 leading to necessary conditions to be satisfied by an optimal control. We first specify how the notion of optimality of a control can be formulated within the present geometric framework.

Let (τ, ν, ρ) be an arbitrary geometric control structure (with $\tau : M \to \mathbb{R}$, $\nu : U \to M$, $\rho : U \to J^1\tau$, as in Definition 2.1) and let $L \in C^\infty(U)$ denote a function on the control bundle U. If $u : I = [a, b] \to U$ is a control, then the *cost* of u with respect to L is defined by

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

If we put $m = \nu(u(a))$ and $n = \nu(u(b))$, we have, with the notations from Section 2, that $m \xrightarrow{u} n$ and, in particular, $n \in R_m$. We say that the control u

is *optimal* if $\mathcal{J}(u) \leq \mathcal{J}(u')$ for any other control u' such that $m \xrightarrow{u'} n$. For the further discussion, it will be helpful to introduce the following notation:

$$\mathcal{J}_u^{(t_1,t_2)} = \int_{t_1}^{t_2} L(u(t))dt,$$

where $t_1, t_2 \in [a, b]$, with $t_1 \leq t_2$. Note that, in this notation, $\mathcal{J}(u) = \mathcal{J}_u^{(a,b)}$. The function L is sometimes referred to as the *cost function*.

Definition 4.1 *A* geometric optimal control structure (τ, ν, ρ, L) *consists of a geometric control structure* (τ, ν, ρ) *and a cost function* L.

We will now show that to every geometric optimal control structure (τ, ρ, ν, L) one can associate an *extended geometric control structure*, $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ in which the cost function is incorporated into the bundle map $\overline{\rho}$. For that purpose, we first introduce the product space $\overline{M} := M \times \mathbb{R}$, the points of which will be denoted by (m, J). For reasons to become clear later on, J will be called the *cost coordinate*. The fibration τ of M over $\mathbb{R}$ induces the fibration $\overline{\tau} : \overline{M} \to \mathbb{R}\,, (m, J) \mapsto \overline{\tau}(m, J) = \tau(m)$. Next, for the extended control bundle we take $\overline{U} = U \times \mathbb{R}$, with projection onto $\overline{M}$ given by $\overline{\nu}(u, J) = (\nu(u), J)$. Finally, we can define a bundle map $\overline{\rho} : \overline{U} \to J^1\overline{\tau}$ as follows: $\overline{\rho}(u, J) = (\rho(u), J, L(u))$, where we have used the canonical identification between $J^1\overline{\tau}$ and $J^1\tau \times \mathbb{R}^2$ obtained as follows: given any section $\overline{c}(t) = (c(t), J(t))$ of $\overline{\tau}$, we map $j^1_t\overline{c}$ onto $(j^1_t c, J(t), \dot{J}(t))$. Note that $\overline{\tau}_{1,0}(\overline{\rho}(u, J)) = \overline{\nu}(u, J)$ and, therefore, $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ is indeed a well-defined geometric control structure.

Next, we shall prove that any control defined on a geometric optimal control structure (τ, ν, ρ, L) induces a control on the extended structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$, and vice versa. Let $u : I = [a, b] \to U$ be a control related to (τ, ν, ρ, L), with $\nu(u(a)) = m$ and $\nu(u(b)) = n$. We shall construct a control $\overline{u}$ in the associated structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$ such that for any $J_0 \in \mathbb{R}$ we have $(m, J_0) \xrightarrow{\overline{u}} (n, J_0 + \mathcal{J}_u^{(a,b)})$. More precisely, define the map $\overline{u} : I \to \overline{U}$ by putting

$$\overline{u}(t) = (u(t), J_0 + \mathcal{J}_u^{(a,t)}).$$

It is easily seen that $\overline{u}$ determines a piecewise section of $\overline{\tau} \circ \overline{\nu}$ whose projection onto $\overline{M}$ is a continuous piecewise section. Furthermore, the first-order jet of the base section $\overline{\nu} \circ \overline{u}$ equals $j^1_t(\nu \circ u, J_0 + \mathcal{J}_u^{(a,t)}) = (j^1_t(\nu \circ u), J_0 + \mathcal{J}_u^{(a,t)}, L(u(t)))$. Since u is a control, we readily obtain the equality $\overline{\rho} \circ \overline{u} = j^1(\overline{\nu} \circ \overline{u})$, which implies that $\overline{u}$ is indeed a control. On the other hand, the projections of $\overline{u}(a)$ and $\overline{u}(b)$ onto $\overline{M}$ are given by (m, J_0) and $(n, J_0 + \mathcal{J}(u))$, respectively. It follows that $(m, J_0) \xrightarrow{\overline{u}} (n, J_0 + \mathcal{J}(u))$ for the extended geometric control problem (and for arbitrary $J_0 \in \mathbb{R}$).

Conversely, let $\overline{u} : [a,b] \to \overline{U}$, $t \mapsto \overline{u}(t) = (u(t), J(t))$ represent a control on the extended geometric control structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$. Then, if the base section is written as $(\overline{\nu} \circ \overline{u})(t) = \overline{c}(t) = (c(t), J(t))$ we can deduce from $\overline{\rho} \circ \overline{u} = j^1\overline{c}$ that $(\rho \circ u)(t) = j^1_t c$, i.e. $u : [a,b] \to U$ is a control. Moreover, the cost coordinate satisfies $\dot{J}(t) = L(u(t))$ and, hence,

$$J(t) = J(a) + \int_a^t L(u(t))dt = J(a) + \mathcal{J}_u^{(a,t)},$$

In particular, we have $J(b) = J(a) + \mathcal{J}_u^{(a,b)}$.

Summarizing the preceding discussion, we have proven the following result.

Proposition 4.2 *Let (τ, ν, ρ, L) denote a geometric optimal control structure. Then for any $m, n \in M$ and $J_m, J_n \in \mathbb{R}$, we have that $m \xrightarrow{u} n$ and $\mathcal{J}(u) = J_n - J_m$ for some control u iff $(m, J_m) \xrightarrow{(u,J)} (n, J_n)$ in the associated extended geometric control structure, where $J : [a,b] \to \mathbb{R}$ is given by $J(t) = J_m + \mathcal{J}_u^{(a,t)}$.*

Consider again an arbitrary geometric optimal control structure (τ, ν, ρ, L) and assume $m \xrightarrow{u} n$ for some control u. According to the previous proposition we then know that, for any $J_0 \in \mathbb{R}$, one can define an appropriate function $J(t)$ such that $(m, J_0) \xrightarrow{(u,J)} (n, J_0 + \mathcal{J}(u))$. Let $\overline{c} = \overline{\nu} \circ (u, J)$ be the base section of the control (u, J). On $\overline{M}$ we can then consider the variational cone $C_{(n,J_0+J(u))}R_{(m,J_0)}$, resp. the vertical variational cone $V_{(n,J_0+J(u))}R_{(m,J_0)}$, associated to a composite flow $\overline{\Phi}$ and composite flow parameter $\overline{T}$ determining the controlled section $\overline{c}$, with $\overline{\Phi}_{\overline{T}}(m, J_0) = (n, J_0 + \mathcal{J}(u))$. The proof of the following proposition relies on Corollary 3.6.

Proposition 4.3 *Let (τ, ν, ρ, L) denote a geometric optimal control structure and assume $m \xrightarrow{u} n$ for a control u which is optimal. Then the interior of $C_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}$ does not contain the tangent vector $-\frac{\partial}{\partial J}\Big|_{(n,J_0+\mathcal{J}(u))}$.*

Proof. Assume that $(-\partial/\partial J)_{(n,J_0+\mathcal{J}(u))} \in \mathrm{int}(C_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)})$. Consider the 'vertical' curve $\theta(t) = (n, J_0 + \mathcal{J}(u) - t)$ in $\overline{M}$, defined for $t \in [0,1]$, whose tangent vector at $t = 0$ precisely equals $(-\partial/\partial J)_{(n,J_0+\mathcal{J}(u))}$. From Corollary 3.6 it then follows that there exists an $\epsilon > 0$, sufficiently small, such that $\theta(t) \in R_{(m,J_0)}$ for $t \in [0, \epsilon]$. From this, one can deduce that there exists a control $\overline{u}'$ for which $(m, J_0) \xrightarrow{\overline{u}'} (n, J_0 + \mathcal{J}(u) - \epsilon)$. In view of previous considerations, this further implies that there exists a control u' on (τ, ν, ρ, L) such that $m \xrightarrow{u'} n$, with cost $\mathcal{J}(u') = \mathcal{J}(u) - \epsilon$, $\mathcal{J}(u') < \mathcal{J}(u)$. Since u was assumed to be optimal, this clearly leads to a contradiction. $\square$

Before proceeding, we first recall some properties and terminology regarding linear spaces and convex cones in a linear space. Let $\mathcal{V}$ be an arbitrary (finite dimensional) linear space. A hyperplane in $\mathcal{V}$ (i.e. a linear subspace of co-dimension one) can always be defined as the set of all vectors $v \in \mathcal{V}$ satisfying $\langle \eta, v \rangle = 0$ for some (non-zero) co-vector $\eta \in \mathcal{V}^*$. Such a hyperplane divides $\mathcal{V}$ into two 'half-spaces' which are given by the set of all v such that $\langle \eta, v \rangle \leq 0$, resp. $\langle \eta, v \rangle \geq 0$, and which are called the 'negative' half-space and the 'positive' half-space, respectively. If C is a convex cone in $\mathcal{V}$ which does not span the full space, then there always exists a hyperplane such that C is contained in one of the corresponding half-spaces.

If we now return to the situation described in the previous proposition, it follows from the above considerations that, under the conditions of Proposition 4.3, there exists a hyperplane in the tangent space $T_{(n,J_0+\mathcal{J}(u))}\overline{M}$ such that the variational cone $C_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}$ is contained in, say, the corresponding negative half-plane, whereas the vector $(-\partial/\partial J)_{(n,J_0+\mathcal{J}(u))}$ belongs to the positive half-plane. From the fact that the vertical variational cone $V_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}$ is a subset of $C_{(n,J_0+J(u))}R_{(m,J_0)}$, contained in the vertical subspace $V_{(n,J_0+\mathcal{J}(u))}\overline{\tau}$, the following result is a straightforward consequence of Proposition 4.3.

Corollary 4.4 *If $m \xrightarrow{u} n$ and if u is optimal, then there exists a hyperplane in $T_{(n,J_0+\mathcal{J}(u))}\overline{M}$, determined by some $\overline{\eta} \in V^*_{(n,J_0+J(u))}\overline{\tau}$ (the dual space of the vertical tangent space $V_{(n,J_0+\mathcal{J}(u))}\overline{\tau}$) such that*

(1) $\langle \overline{\eta}, -\frac{\partial}{\partial J}\Big|_{(n,J_0+\mathcal{J}(u))} \rangle \geq 0$, *and*
(2) $\langle \overline{\eta}, v \rangle \leq 0$ *for all* $v \in V_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}$.

In order to relate the previous result to a more familiar formulation of the necessary conditions for an optimal control, in terms of solutions of differential equations, we will need a minor generalization of the theory of connections over a bundle map as developed, for instance, in [8].

5 Lifts over bundle maps

For the sake of completeness, we first briefly recall the setting for defining *a lift over a bundle map.*

Consider a smooth manifold B and a fibre bundle $\nu : N \to B$, equipped with a bundle map $\Lambda : N \to TB$ fibred over the identity, as shown in the following commutative diagram.

$$\begin{array}{ccc} N & \xrightarrow{\Lambda} & TB \\ \nu \downarrow & & \downarrow \tau_B \\ B & \xrightarrow[\mathrm{id}_B]{} & B \end{array}$$

Note that, unlike the treatment in [8], we do not require N to be a vector bundle. Next, let $\pi : E \to B$ denote an arbitrary fibre bundle over B and consider the pull-back bundle $\pi^* N$. We can then define the following notion of lift.

Definition 5.1 *A* lift over Λ *is a bundle map* $h : \pi^* N \to TE$ *fibred over the identity on* E *such that the following diagram commutes:*

$$\begin{array}{ccc} \pi^* N & \xrightarrow{h} & TE \\ \tilde{\pi}_2 \downarrow & & \downarrow T\pi \\ N & \xrightarrow[\Lambda]{} & TB \end{array}$$

A lift h over Λ allows us to define the h*-lift* of a section s of ν. More precisely, the h-lift of $s \in \Gamma(\nu)$ is a section of τ_E defined by $s^h(e) = h(e, s(\pi(e)))$, for all $e \in E$. Note that s^h determines a vector field on E.

A Λ-*admissible curve* $c : I = [a, b] \to N$ is a smooth curve such that the base curve $\nu \circ c = \tilde{c}$ in B satisfies $\dot{\tilde{c}}(t) = \Lambda(c(t))$. If we assume that $\Lambda(n) \neq 0$ for all $n \in N$, then any Λ-admissible curve is a concatenation of integral curves of vector fields belonging to the family $\mathcal{D}' = \{\Lambda \circ s \mid s \in \Gamma(\nu)\}$. Indeed, let $c : I \to N$ denote a Λ-admissible curve, with base curve $\tilde{c}$. Then $\dot{\tilde{c}}(t) \neq 0$ for all t, i.e. $\tilde{c}$ is an immersion. Following an argument of S. Helgason (see [9, p 28]), one can prove that there exists a finite subdivision $\{I_i\}$ of I such that for the restriction of c to each of these subintervals I_i there exists a local section s_i of ν verifying $s_i(\tilde{c}(t)) = c(t)$ for all $t \in I_i$. It is easily seen that $\tilde{c}_{|I_i}$ is an integral curve of $\Lambda \circ s_i$.

Remark 5.2 We can apply all this to a geometric control structure (τ, ν, ρ), where we take $B = M, N = U, \Lambda = \mathbf{T} \circ \rho$. A control can then be equivalently characterized as a $(\mathbf{T} \circ \rho)$-admissible curve $u : I \to U$, with the additional constraint that it should be a section of $\tau \circ \nu$, i.e.$(\tau \circ \nu)(t) = t$ for all t. We also recover here the property that each $(\mathbf{T} \circ \rho)$-admissible curve is a concatenation of integral curves of vector fields in $\mathcal{D}$.

Assume now that the bundle E is a vector bundle and let Δ be the dilation vector field on E, with flow $\{\delta_t\}$. A lift h over Λ is then said to be *linear* if

$T\delta_t \circ h(n,e) = h(n, \delta_t(e))$ for any t. Consider bundle adapted coordinate charts on N and E, denoted by (x^i, n^α) and (x^i, e^A), respectively. In coordinates, h then reads

$$h(x^i, n^\alpha, e^A) = \Lambda^j(x^i, n^\alpha) \left.\frac{\partial}{\partial x^j}\right|_e + \Gamma^A(x^i, n^\alpha, e^A) \left.\frac{\partial}{\partial e^A}\right|_e,$$

and h is a linear lift iff $\Gamma^A(x^i, n^\alpha, e^A) = \Gamma^A_B(x^i, n^\alpha) e^B$. The functions Γ^A_B are called the *coefficients of* h. For the remainder of this section, we always take E to be a vector bundle (over B).

Given a linear lift h and a Λ-admissible curve $c : [a,b] \to N$, with base curve $\tilde{c}$, take any $e \in E$ such that $\pi(e) = \tilde{c}(a)$. We can then construct a curve c^h in E through e, called the *h-lift* of c, which is uniquely determined by the differential equation $h(c^h(t), c(t)) = \dot{c}^h(t)$, with initial condition $c^h(a) = e$ (see also [8]).

Next, we show that a linear lift h always induces a derivative operator ∇, acting on sections of π. Let $\pi_2 : V\pi \cong E \times_B E \to E$ denote the projection onto the second factor, then, in analogy with the case where N is a vector bundle and Λ a linear bundle map (see [8]), we can define a mapping $K : \Lambda^* TE \to E$ according to: $K(n, w) = \pi_2(w - h(\tau_E(w), n))$. Given any $x \in B$, $n \in N_x (= \nu^{-1}(x))$ and any local section $\psi \in \Gamma(\pi)$, defined on an open neighborhood of x, we put

$$\nabla_n \psi := K(n, T_x\psi(\Lambda(n))).$$

Clearly, $\nabla_n \psi \in E_x (= \pi^{-1}(x))$. The map ∇_n thus defined, is a derivative operator on $\Gamma(\pi)$ since, for arbitrary $f \in C^\infty(B)$, $\psi_1, \psi_2 \in \Gamma(\pi)$ (all at least defined on a neighborhood of x) we find that

$$\begin{aligned} &\nabla_n f\psi = \Lambda(n)(f)\psi(x) + f(x)\nabla_n \psi, \\ &\nabla_n(\psi_1 + \psi_2) = \nabla_n \psi_1 + \nabla_n \psi_2. \end{aligned}$$

An operator on $\Gamma(\pi)$ satisfying these properties is called a Λ-*derivative*. Given any section $s \in \Gamma(\nu)$, we can define the operator ∇_s on $\Gamma(\pi)$ by

$$\nabla_s \psi(x) := \nabla_{s(x)} \psi,$$

and, obviously, $\nabla_s \psi$ is again a section of π. It is easily seen that there is a one-to-one correspondence between Λ-derivatives and linear lifts over Λ. Using the above coordinate expression for h, we obtain that the Λ-derivative determined by h locally reads (for $n=(x^i,n^\alpha)\in N_x$)

$$(\nabla_n\psi)^A=\Lambda^j(x^i,n^\alpha)\frac{\partial\psi^A}{\partial x^j}(x^i)-\Gamma^A_B(x^i,n^\alpha)\psi^B(x^i).$$

It also follows that $\nabla_s\psi=0$ for $s\in\Gamma(\nu)$ and $\psi\in\Gamma(\pi)$ iff $T_x\psi(\Lambda(s(x)))=s^h(\psi(x))$ for all $x\in B$.

Similar to what we have in standard connection theory, a derivative operator can be constructed which acts on sections of π defined along the base curve of a Λ-admissible curve $c:I=[a,b]\to N$. Indeed, consider a curve in E, $\tilde\psi:I\to E$, such that $\pi\circ\tilde\psi=\nu\circ c(=\tilde c)$, then the Λ-derivative associated to the linear lift h and acting on $\tilde\psi$ equals

$$\nabla_c\tilde\psi(t):=K(c(t),\dot{\tilde\psi}(t)).$$

It is not difficult to prove that $\nabla_c\tilde\psi(t)=0$ for all t iff $\tilde\psi=c^h$. If $\nabla_c\tilde\psi\equiv 0$, we say that $\tilde\psi$ is *h-transported along* c and that $\tilde\psi(b)$ is the *h-transport* of $\tilde\psi(a)$ along c. We conclude by pointing out that any Λ-admissible curve c in N determines a linear map $c^b_a:E_{c(a)}\to E_{c(b)}$, called the *h-transport operator along* c, defined by $c^b_a(e)=\tilde\psi(b)$, where $\tilde\psi$ is the unique solution of the equation $\nabla_c\tilde\psi(t)=0$ with $\tilde\psi(a)=e$.

6 The control lift and control derivative

Let (τ,ν,ρ) denote a geometric control structure. Consider the first-order jet bundle $J^1\nu$ of the bundle $\nu:U\to M$, with associated projections $\nu_1:J^1\nu\to M$, $\nu_{1,0}:J^1\nu\to U$. Recall that for any two local sections σ and σ' of ν, defined on a neighborhood of a point $m\in M$, we have that $j^1_m\sigma=j^1_m\sigma'\in J^1\nu$ iff $\sigma(m)=\sigma'(m)$ and $T_m\sigma=T_m\sigma'$ (as linear maps from T_mM into $T_{\sigma(m)}U$). Bearing this in mind, it is easily seen that the following mapping is well-defined:

$$\Lambda:J^1\nu\to TU,\ j^1_m\sigma\mapsto\Lambda(j^1_m\sigma)=T_m\sigma((\mathbf{T}\circ\rho)(\sigma(m))).\tag{6.1}$$

Moreover, Λ is a bundle map over the identity on U. In terms of appropriate bundle coordinates (t,x^i,u^a) on U and (t,x^i,u^a,u^a_t,u^a_i) on $J^1\nu$, Λ reads

$$\Lambda(t,x^i,u^a,u^a_t,u^a_i)=(t,x^i,u^a,1,\rho^j(t,x^i,u^a),u^b_t+\rho^j(t,x^i,u^a)u^b_j)).$$

We now consider the fibred product bundle $U\times_M V\tau$, with projections $p_1:U\times_M V\tau\to U$, $(u,v)\mapsto p_1(u,v)=u$ and $p_2:U\times_M V\tau\to V\tau$, $(u,v)\mapsto v$, whereby $\nu \circ p_1 = \tau_M \circ p_2$. Observing that $p_1 : U \times_M V\tau \to U$ is a vector bundle over U, we can apply the theory from the previous section to the case where $B = U$, $N = J^1\nu$, $E = U \times_M V\tau$ and Λ is given by (6.1). It will be seen that, within this setting, Λ-admissible curves are closely related to controls. For that purpose, we need the following straightforward extension of the definition of Λ-admissible curve to the class of piecewise curves: a piecewise curve ψ in $J^1\nu$ is said to be Λ-admissible if it is induced by (i.e. consists of a concatenation of) a finite number of smooth Λ-admissible curves.

In the sequel, we always assume that a piecewise Λ-admissible curve ψ in $J^1\nu$ has a continuous projection onto M and is parameterized such that $\tau(\nu_1(\psi(t))) = t$, i.e. such that ψ is a section of $\tau \circ \nu_1$. (Note that this is not a restriction since, given any Λ-admissible curve $\psi : [a, b] \to J^1\nu$, with $t_a = \nu_1(\psi(a))$, we can consider a reparametrization of ψ according to $\psi' : [t_a, t_a + b - a] \to J^1\nu$, $t \mapsto \psi'(t) = \psi(t - t_a + a)$. Then, ψ' is still Λ-admissible and, moreover, satisfies $\tau(\nu_1(\psi'(t))) = t$.)

Lemma 6.1 *The projection onto U of any Λ-admissible curve in $J^1\nu$ is a smooth control, and any control in U can be obtained as the projection of a piecewise Λ-admissible curve.*

Proof. We first prove that the projection $u = \nu_{1,0} \circ \psi$ of a Λ-admissible curve $\psi : [a, b] \to J^1\nu$ is a smooth control. By definition, we have $\dot{u}(t) = \Lambda(\psi(t))$. From $T\nu \circ \Lambda = \mathbf{T} \circ \rho \circ \nu_{1,0}$, it follows that $\dot{c}(t) = (\mathbf{T} \circ \rho)(u(t))$, where $c = \nu_1(\psi(t)) = \nu \circ \nu_{1,0}(\psi(t))$. This shows that the smooth curve u is $(\mathbf{T} \circ \rho)$-admissible, i.e. it is a smooth control.

On the other hand, assume that $u : [a_0, a_\ell] \to U$ is a control, with base curve $c = \nu \circ u$. We then know that c can be written as a concatenation of integral curves, induced by the composite flow Φ of an ordered set $(\mathbf{T} \circ \rho \circ \sigma_\ell, \ldots, \mathbf{T} \circ \rho \circ \sigma_1)$ for some $\sigma_i \in \Gamma(\nu)$, with composite flow parameter $T = (a_\ell - a_{\ell-1}, \ldots, a_1 - a_0)$. Furthermore, $u(t) = \sigma_i(c(t))$ for any $t \in]a_{i-1}, a_i]$. Putting $\psi_i(t) = j^1\sigma_i(c(t))$ for all $t \in [a_{i-1}, a_i]$ and $i = 1, \ldots, \ell$, we obtain that for any $t \in]a_{i-1}, a_i]$ the equality

$$\Lambda(\psi_i(t)) = T_{c(t)}\sigma_i(\dot{c}(t)) = \left.\frac{d}{dt}\right|_t (\sigma_i(c(t))) = \dot{u}(t),$$

holds. Therefore, according to the definition above, the piecewise curve $\psi : [a_0, a_\ell] \to J^1\nu$, induced by the smooth curves $\psi_i : [a_{i-1}, a_i] \to J^1\nu$, is a piecewise Λ-admissible curve, which completes the proof of the lemma. □

In the following we shall frequently make use of the natural identification $T(U \times_M V\tau) \cong TU \times_{TM} T(V\tau)$, without mentioning it explicitly. We further denote by $\mathfrak{s} : TTM \to TTM$ the canonical involution on TTM. The latter is characterized by the relations $T\tau_M \circ \mathfrak{s} = \tau_{TM}$ and $\tau_{TM} \circ \mathfrak{s} = T\tau_M$.

Remark 6.2 Recall that, given an arbitrary manifold B with local coordinates (x^i), and denoting the natural bundle coordinates on TB and TTB by (x^i, v^i) and $(x^i, v^i, \dot{x}^i, \dot{v}^i)$, respectively, then the canonical involution $\mathfrak{s}$ on TTB reads $\mathfrak{s}\,(x^i, v^i, \dot{x}^i, \dot{v}^i) = (x^i, \dot{x}^i, v^i, \dot{v}^i)$.

For a geometric control structure (τ, ν, ρ), with bundle map Λ given by (6.1), we have the following property.

Proposition 6.3 *The map $h^c : \nu^*_{1,0}(U \times_M V\tau) \to T(U \times_M V\tau)$, defined by*

$$h^c(j^1_m\sigma, (\sigma(m), v)) = (\Lambda(j^1_m\sigma), \mathfrak{s}\,(T(\mathbf{T} \circ \rho)(T_m\sigma(v)))),$$

for any $m \in M$, $\sigma \in \Gamma(\nu)$ and $v \in V_m\tau (\subset T_mM)$, is a linear lift over Λ.

Proof. We first verify that h^c indeed takes values in $T(U \times_M V\tau)$. For that purpose, consider bundle adapted coordinates (t, x^i, u^a) and (t, x^i, v^j) on U and $V\tau$, respectively. Take $m = (t, x^i) \in M$, $v = (t, x^i, v^j) \in V_m\tau$ and $j^1_m\sigma = (t, x^i, u^a, u^a_t, u^a_i) \in J^1\nu$, then:

$$\begin{aligned}\mathfrak{s}\,(T(\mathbf{T} \circ \rho)(T_m\sigma(v))) = &\left.\frac{\partial}{\partial t}\right|_v + \rho^i(t, x^j, u^a)\left.\frac{\partial}{\partial x^i}\right|_v \\ &+ \left(v^i \frac{\partial \rho^k}{\partial x^i}(t, x^j, u^a) + v^i u^b_i \frac{\partial \rho^k}{\partial u^b}(t, x^j, u^a)\right)\left.\frac{\partial}{\partial v^k}\right|_v .\end{aligned}$$

From this expression one can read that $\mathfrak{s}\,(T(\mathbf{T} \circ \rho)(T_m\sigma(X))) \in T_v(V\tau)(\subset T_vTM)$. Next, using the properties of the canonical involution operator $\mathfrak{s}$, and taking into account (6.1), it is easily seen that $T\nu(\Lambda(j^1_m\sigma)) = T\tau_M(\mathfrak{s}\,(T(\mathbf{T} \circ \rho)(T_m\sigma(v)))) = (\mathbf{T} \circ \rho)(\sigma(m)) \in T_mM$ which proves indeed that $\mathrm{Im}\, h^c \subset T(U \times_M V\tau)$.

From its definition it readily follows that h^c is a bundle map fibred over the identity on $U \times_M V\tau$, and we have that $Tp_1(h^c(j^1_m\sigma, (\sigma(m), X))) = \Lambda(j^1_m\sigma)$. This already guaranties that h^c is a lift over Λ in the sense of Definition 5.1. From the above coordinate expression we can also deduce that the $\partial/\partial v^k$-components Γ^k of h^c are linear in the fibre coordinates v^i of the vector bundle p_1. More precisely, we have $\Gamma^k(t, x^j, v^j, u^a, u^a_j) = \Gamma^k_i(t, x^j, u^a, u^a_j)v^i$ with

$$\Gamma^k_i(t, x^i, u^a, u^a_i) = \frac{\partial \rho^k}{\partial x^i}(t, x^j, u^a) + u^b_i \frac{\partial \rho^k}{\partial u^b}(t, x^j, u^a).$$

This shows, in particular, that h^c is a linear lift. $\square$

Note that the 'coefficients' Γ_i^k of h^c do not depend on the coordinates u_t^a of $J^1\nu$. In a remark at the end of this section we will return to this point in more detail.

Let us denote the Λ-derivative corresponding to h^c by D and let $\mathcal{V}(\nu)$ denote the set of τ-vertical vector fields along ν, i.e.

$$\mathcal{V}(\nu) = \{Z : U \to V\tau \,|\, \tau_M(Z(u)) = \nu(u), \text{ for all } u \in U\}.$$

Note that, in view of the relation $\nu \circ p_1 = \tau_M \circ p_2$, we have $\mathcal{V}(\nu) \cong \Gamma(p_1)$.

Proposition 6.4 *Given any $j_m^1\sigma \in J^1\nu$ and $Z \in \mathcal{V}(\nu)$, then $D_{j_m^1\sigma}Z \in V_m\tau$ and*

$$D_{j_m^1\sigma}Z = [\mathbf{T} \circ \rho \circ \sigma, Z \circ \sigma](m),$$

(where the square brackets on the right-hand side denote the ordinary Lie bracket of vector fields on M).

Proof. Recalling the coordinate expression of a Λ-derivative (cf. Section 5), we obtain, with a slight abuse of notation,

$$(D_{j_m^1\sigma}Z)^i = (\frac{\partial Z^i}{\partial t} + \rho^j \frac{\partial Z^i}{\partial x^j} + (u_t^a + \rho^j u_j^a)\frac{\partial Z^i}{\partial u^a} - \Gamma_j^i Z^j)_m.$$

The result then easily follows upon substituting $u_j^a = \frac{\partial \sigma^a}{\partial x^j}$ and $u_t^a = \frac{\partial \sigma^a}{\partial t}$ in the right-hand side, and comparing this with the coordinate expression of the Lie bracket $[\mathbf{T} \circ \rho \circ \sigma, Z \circ \sigma](m)$. $\square$

We shall now derive an explicit expression for the h^c-transport operator ψ_a^b determined by a Λ-admissible curve $\psi : I = [a, b] \to J^1\nu$. We first consider the case where ψ takes the special form $\psi(t) = j^1\sigma(c(t))$ for some curve $c : [a, b] \to M$ and a section $\sigma \in \Gamma(\nu)$. Note that such a ψ is Λ-admissible iff $u(t) := \sigma(c(t))$ is a smooth control, which still implies that in terms of the flow $\{\phi_s\}$ of the vector field $\mathbf{T} \circ \rho \circ \sigma$, we have $c(t) = \phi_{t-a}(c(a))$.

Lemma 6.5 *Let $\psi : [a, b] \to J^1\nu$, $t \mapsto \psi(t) = j^1\sigma(c(t))$ be a Λ-admissible curve, and let $\{\phi_s\}$ denote the flow of $\mathbf{T} \circ \rho \circ \sigma$. Then the h^c-transport operator $\psi_a^b : V_{c(a)}\tau \to V_{c(b)}\tau$ along ψ is given by $\psi_a^b = T\phi_{b-a}$.*

Proof. Representing the flow of $\frac{\partial}{\partial t}$ on $\mathbb{R}$ by $\{\lambda_s\}$, it immediately follows from $\tau \circ \phi_s = \lambda_s \circ \tau$ that, for any $v \in V\tau$, the vector $T\phi_s(v)$ also belongs to $V\tau$. In particular, we have $T\phi_{b-a}(v) \in V_{c(b)}\tau$.

Next, take $v_0 \in V_{c(a)}$ and let $X(t)$ denote the section of $V\tau$ along $c(t)$ which is uniquely determined by the conditions $D_\psi(u, X)(t) = 0$ and $X(a) = v_0$. This is still equivalent to

$$\left.\frac{d}{dt}\right|_t X(t) = \mathfrak{s}\left(T(\mathbf{T} \circ \rho)T_{c(t)}\sigma(X(t))\right). \tag{6.2}$$

Since $\mathfrak{s}\left(T(\mathbf{T} \circ \rho)T_{c(t)}\sigma(X(t))\right) = (\mathbf{T} \circ \rho \circ \sigma)^c(X(t))$, where $(\mathbf{T} \circ \rho \circ \sigma)^c$ denotes the complete lift of the vector field $\mathbf{T} \circ \rho \circ \sigma$ to TM, (6.2) tells us that $X(t)$ is an integral curve of $(\mathbf{T} \circ \rho \circ \sigma)^c$, passing through v_0. By construction of the complete lift of a vector field, the flow of $(\mathbf{T} \circ \rho \circ \sigma)^c$ is given by $\{T\phi_s\}$ and, therefore, $X(t) = T\phi_{t-a}(X(a))$. The result then follows immediately from the definition of the h^c-transport operator along ψ. $\square$

Next, we consider the case where $\psi : [a, b] \to J^1\nu$ is a piecewise Λ-admissible curve whose projection $c = \nu_1 \circ \psi$ onto M is continuous. Recall, in particular, that $u(t) := \nu_{1,0}(\psi(t))$ is a control (see Lemma 6.1). For the sequel we will need an extension of the definition of the Λ-derivative corresponding to h^c to piecewise curves. For that purpose, let $X : [a, b] \to V\tau$ be a continuous piecewise curve projecting onto the base curve $c(t)$ of ψ. Note, in particular, that (u, X) represents a piecewise section of $U \times_M V\tau$ along c. From the definition of piecewise curves it can be deduced that one can always find a sufficiently fine subdivision $a_0 = a < a_1 \ldots < a_\ell = b$ of the given interval $[a, b]$ such that ψ can be written as a concatenation of ℓ smooth Λ-admissible curves $\psi_i : [a_{i-1}, a_i] \to J^1\nu$ and X as a concatenation of ℓ smooth curves $X_i : [a_{i-1}, a_i] \to V\tau$. For the piecewise Λ-admissible curve ψ we now define the Λ-derivative D_ψ, acting on the piecewise section (u, X), as follows:

$$D_\psi(u, X)(t) := D_{\psi_i}(u, X_i)(t) \quad \text{for all } t \in]a_{i-1}, a_i],\ i = 1, \ldots, \ell,$$

and

$$D_\psi(u, X)(a_0) = D_{\psi_1}(u, X_1)(a_0).$$

It is easily seen in coordinates, for instance, that the mapping $D_\psi(u, X) : [a, b] \to U \times_M V\tau$ is indeed well defined. Given any $v_0 \in V_{c(a)}\tau$, one can readily verify that there exists a unique continuous piecewise curve $X(t)$ in $V\tau$ such that $D_\psi(u, X)(t) = 0$ for all $t \in [a, b]$, with $X(a) = v_0$. This implies that one may introduce a (composite) h^c-transport operator ψ_a^b along the piecewise

Λ-admissible curve ψ as follows: $\psi_a^b = (\psi_\ell)_{a_{\ell-1}}^{a_\ell} \circ \ldots \circ (\psi_1)_{a_0}^{a_1}$, where $(\psi_i)_{a_{i-1}}^{a_i}$ represents the h^c-transport operator along the smooth Λ-admissible curve ψ_i, as defined in the previous section. If, for a given ψ (and the corresponding control u), $X(t)$ solves the equation $D_\psi(u, X)(t) = 0$, it then follows from the definition that $\psi_a^b(X(a)) = X(b)$.

We shall prove below that a piecewise Λ-admissible curve ψ, with $\nu_1 \circ \psi = c$, can always be considered as being induced by smooth Λ-admissible curves $\psi^i : I_i = [a_{i-1}, a_i] \to J^1\nu$ of the form $\psi_i(t) = j^1\sigma_i(c(t))$, for some local section σ_i of ν. Using this property we then know from above that $(\psi_i)_{a_{i-1}}^{a_i} = T\phi^i_{a_i - a_{i-1}}$, with $\{\phi^i_s\}$ the flow of $\mathbf{T} \circ \rho \circ \sigma_i$. Denoting the composite flow of the ordered set $(\mathbf{T} \circ \rho \circ \sigma_\ell, \ldots, \mathbf{T} \circ \rho \circ \sigma_1)$ by Φ and using the shorthand notation introduced in Section 2, we find that the (composite) h^c-transport operator ψ_a^b is given by

$$\psi_a^b = T\Phi_a^b.$$

Indeed, a straightforward computation gives:

$$\begin{aligned}
\psi_a^b(X(a)) &:= (\psi_\ell)_{a_{\ell-1}}^b \circ \ldots \circ (\psi_1)_a^{a_1}(X(a)) \\
&= T\phi^\ell_{a_\ell - a_{\ell-1}} \circ \ldots \circ T\phi^1_{a_1 - a}(X(a)) \\
&= T\Phi_a^b(X(a)) = X(b).
\end{aligned}$$

In order to prove that any (piecewise) Λ-admissible curve can be written as a concatenation of smooth Λ-admissible curves of the form $j^1\sigma \circ c$, we shall prove that any smooth Λ-admissible curve ψ whose image is entirely contained in a coordinate chart, is of that form. The general result then follows by a similar argument as the one applied in Section 2 (when proving that the base curve of any control is a concatenation of integral curves of vector fields in $\mathcal{D}$). So, assume ψ can be written in coordinates as $\psi(t) = (t, x^i(t), u^a(t), u^a_t(t), u^a_i(t))$ for all $t \in I = [a, b]$. Since ψ is Λ-admissible, we then have that

$$\dot{u}^a(t) = u^a_t(t) + u^a_i(t)\dot{x}^i(t) \quad \text{and} \quad \dot{x}^i(t) = \rho^i(t, x^i(t), u^a(t)).$$

Consider now a smooth extension $\tilde{\psi}(t) = (t, \tilde{x}^i(t), \tilde{u}^a(t), \tilde{u}^a_t(t), \tilde{u}^a_i(t))$ of ψ, defined on an open interval $\tilde{I}$ containing I, such that Im $\tilde{\psi}$ is still contained in the same coordinate chart, with $\tilde{\psi}(t) = \psi(t)$ for all $t \in I$. Next, we can construct a local section σ of ν, defined on $\tau^{-1}(\tilde{I})$, as follows: $\sigma(t, x) = (t, x, \sigma^a(t, x))$, with $\sigma^a(t, x) = \tilde{u}^a(t) + \tilde{u}^a_i(t)(x^i - \tilde{x}^i(t))$. For each fixed $t \in I$ we find that

$$\sigma^a(t, x^i(t)) = u^a(t)$$

$$\frac{\partial \sigma^a}{\partial t}(t, x^i(t)) = \dot{u}^a(t) - u^a_i(t)\dot{x}^i(t) = u^a_t(t),$$
$$\frac{\partial \sigma^a}{\partial x^i}(t, x^i(t)) = u^a_i(t),$$

and, hence, we have that $j^1\sigma(t, x(t)) = \psi(t)$ for all $t \in I$, which is precisely what we wanted to prove.

We have seen that, given a piecewise Λ-admissible curve ψ in $J^1\nu$, with continuous piecewise base curve $c = \nu_1 \circ \psi$ and corresponding control $u = \nu_{1,0} \circ \psi$, we can regard the equation $D_\psi(u, X)(t) = 0$ as a differential equation for the component X of the curve (u, X) in $U \times_M V\tau$ that is h^c-transported along ψ. Returning to the given geometric control structure (τ, ρ, ν), we shall now explain the role of the h^c-transport operator in determining the vertical variational cone associated to a composite flow Φ and composite flow parameter T induced by an ordered set of vector fields of the form $(\mathbf{T} \circ \rho \circ \sigma)$, for some $\sigma \in \Gamma(\nu)$.

Given any control $u : [a, b] \to U$, with base curve $c = \nu \circ u$. In Section 2 we have seen that c is induced by the composite flow Φ of an ordered set of vector fields belonging to the family $\mathcal{D}$ given by (2.1), say $(\mathbf{T} \circ \rho \circ \sigma_\ell, \ldots, \mathbf{T} \circ \rho \circ \sigma_1)$, where $\sigma_i \in \Gamma(\nu)$, and let the composite flow parameter be $T = (a_\ell - a_{\ell-1}, \ldots, a_1 - a_0)$, with $a = a_0 < a_1 < \ldots < b = a_\ell$. If we put $c(a_0) = m$ and $c(a_\ell) = m'$, then the vertical variational cone $V_{m'}R_m(\Phi, T)$ is completely determined by the piecewise Λ-admissible curve ψ in $J^1\nu$ that is induced by the smooth curves $\psi_i(t) = j^1\sigma_i(c(t))$. Indeed, it follows from Definition 3.7 and from the above analysis, that any element of $V_{m'}R_m(\Phi, T)$ can be written as a linear combination of h^c-transported vertical tangent vectors along ψ, i.e.

$$V_{m'}R_m(\Phi, T) = \left\{ \sum_i \delta^i \psi^b_{\tau^i}(Y_i(c(\tau^i)) - \dot{c}(\tau^i)) \mid Y_i \in \mathcal{D}, \delta^i \geq 0, \tau^i \in]a, b] \right\}.$$

Roughly speaking, one can say that the (piecewise) Λ-admissible curve ψ corresponding to the control u, contains sufficient information regarding the sections σ_i in order to determine the vertical variational cone $V_{m'}R_m$. From now on we shall therefore write $V_{m'}R_m(\psi)$ if we want to emphasize that the vertical variational cone can be generated by the h^c-transport operator along the (piecewise) Λ-admissible curve ψ.

For later use we will need an extension of the action of the Λ-derivative D to 'vertical' forms, belonging to the dual of $\mathcal{V}(\nu)$. Consider the fibred product bundle $U \times_M V^*\tau$ with corresponding projections $p_1^* : U \times_M V^*\tau \to U$, $p_2^* : U \times_M V^*\tau \to V^*\tau$, such that $\nu \circ p_1^* = \tau_M^* \circ V^*\tau$. Here $\tau_M^* : V^*\tau \to M$ denotes the dual bundle of $V\tau \to M$. The dual module of $\mathcal{V}(\nu)$ is then given by the set

$$\mathcal{V}^*(\nu) = \{\eta : U \to V^*\tau \,|\, \tau^*_M(\eta(u) = \nu(u) \text{ for all } u \in U\}.$$

Obviously, we have $\mathcal{V}^*(\nu) \cong \Gamma(p_1^*)$. Given $\eta \in \mathcal{V}^*(\nu)$ and $Z \in \mathcal{V}(\nu)$, the natural pairing $\mathcal{V}(\nu)$, $\langle \eta, Z\rangle$ defines a function on U. In particular, for $j^1_m\sigma \in J^1\nu$, with $\sigma(m) = u$, we note that $p_2^*(\eta(u))$ and $p_2(Z(u))$ belong to the dual linear spaces $V^*_m\tau$ and $V_m\tau$, respectively. By requiring that for any fixed $\eta \in \mathcal{V}^*(\nu)$, the relation

$$\langle D_{j^1_m\sigma}\eta, Z(u)\rangle = \Lambda(j^1_m\sigma)(\langle \eta, Z\rangle) - \langle \eta(u), D_{j^1_m\sigma}Z\rangle, \tag{6.3}$$

should hold for all $Z \in \mathcal{V}(\nu)$, the element $D_{j^1_m\sigma}\eta \in (p_1^*)^{-1}(u)(\cong V^*_m\tau)$ is uniquely determined.

Consider a piecewise Λ-admissible curve $\psi : [a, b] \to J^1\nu$ with continuous piecewise projection $c = \nu_1 \circ \psi$ on M and corresponding control $u = \nu_{1,0} \circ \psi : [a, b] \to U$. Take a continuous piecewise section $\overline{\eta}(t)$ of $V^*\tau$ along $c(t)$ such that $(u, \overline{\eta})(t)$ defines a section of p_1^* along the curve $u(t)$. We then have the following property.

Lemma 6.6 *$D_\psi(u, \overline{\eta})(t) = 0$ iff $\overline{\eta}(t) = ((\psi^t_a)^{-1})^*(\overline{\eta}(a))$ for all $t \in [a, b]$.*

Proof. Fix some $t_0 \in I$ and take an arbitrary $X_0 \in V_{c(t_0)}\tau$. Using the h^c-transport operator along ψ, we can then construct a continuous piecewise section $X(t)$ of $V\tau$ along $c(t)$ by $X(t) = \psi^t_a((\psi^{t_0}_a)^{-1}(X_0))$. Note that $X(t_0) = X_0$. Then, with (6.3) we obtain

$$\begin{aligned}\langle D_\psi(u, \overline{\eta})(t_0), (u(t_0), X(t_0))\rangle = {} & \left.\frac{d}{dt}\right|_{t_0} \langle \overline{\eta}(t), X(t)\rangle \\ & -\langle (u(t_0), \overline{\eta}(t_0)), D_\psi(u, X)(t_0)\rangle.\end{aligned}$$

Now it follows from the definitions that both terms on the right-hand vanish separately if we take $\overline{\eta}(t) = ((\psi^t_a)^{-1})^*(\eta(a))$. Indeed, with this choice we have $\langle \overline{\eta}(t), X(t)\rangle \equiv \langle \eta(a), (\psi^{t_0}_a)^{-1}(X_0)\rangle = \text{const.}$, and $D_\psi(u, X)(t_0) = 0$ in view of the definition of $X(t)$. The remainder of the proof then follows from the uniqueness of solutions of a system of ordinary differential equations with given initial conditions. $\square$

The Λ-derivative D will play a crucial role in the proof of the Maximum Principle in the next section. In the following remark we briefly explain how some of the basic ideas in the treatment of the Maximum Principle in [2] can be related to our work.

Remark 6.7 The discussion of the Maximum Principle can be developed for controls that verify the weaker assumption of being measurable and bounded, instead of (piecewise) smooth (see, for instance, L.S. Pontryagin et al. [2]). Using local coordinate expressions, we will roughly sketch how the smoothness conditions we have imposed on controls can also be relaxed within our framework. The local expressions for the equation $D_\psi(u, X)(t) = 0$ reads

$$\dot{X}^k(t) = \left(\frac{\partial \rho^k}{\partial x^i}(t, c^j(t), u^a(t)) + u_i^b(t)\frac{\partial \rho^k}{\partial u^b}(t, c^j(t), u^a(t))\right) X^i(t).$$

The condition that the functions $u^a(t)$ and $u_i^a(t)$ be measurable and bounded, suffices to obtain a solution of this equation and, subsequently, to introduce a suitable notion of transport operator. This observation can be translated into our geometric framework as follows. Consider the set $V^1\nu := \cup_{m\in M}\{T_m\sigma_{|V\tau} : V_m\tau \to T_{\sigma(m)}U \mid \sigma \in \Gamma(\nu)\}$. It can be proven by standard arguments that $V^1\nu$ is an affine bundle over U, with coordinates (t, x^i, u^a, u_i^a) (see, for instance, [10]). Note that there exists a natural projection $\mu : J^1\nu \to V^1\nu$, locally expressed by $(t, x^i, u^a, u_i^a, u_t^a) \mapsto (t, x^i, u^a, u_i^a)$. From the fact that the coefficients Γ_i^k of h^c do not depend on the u_t^a (see the proof of Proposition 6.3) it easily follows that the Λ-derivative D_ψ only depends on $\mu \circ \psi$. Now, since ψ was assumed to be Λ-admissible, i.e. $\Lambda(\psi) = \dot{u}$, the smoothness condition on u could not be relaxed. However, the curve $\tilde{\psi} = \mu \circ \psi$ does not have to satisfy this condition, implying that the smoothness condition can be relaxed without losing the notion of derivative acting on sections of $V\tau$ along c. We can therefore conclude that, in order to define a vertical cone of variations associated with a measurable and bounded control u, we must fix a curve $\tilde{\psi}$ in $V^1\nu$. If one works in a coordinate chart, a natural choice of $\tilde{\psi}$ is the curve $\tilde{\psi}(t) = (t, c^i(t), u^a(t), u_i^a(t))$ with $u_i^a(t) = 0$. The equations of the derivative associated with $\tilde{\psi}$ then reduce to $\dot{X}^k(t) = \frac{\partial \rho^k}{\partial x^i}(t, c^j(t), u^a(t))X^i(t)$. These equations are precisely the "variational equations" introduced in [2, p79]. By fixing the coordinate chart, one can fix the section $\sigma^a(t, x) = u^a(t)$ and the curve $\tilde{\psi}(t) = (t, c^j(t), u^a(t), 0)$, implying that, respectively a fixed vertical cone of variations and a fixed derivative associated with $\tilde{\psi}$ can be defined. This essentially establishes the link between our approach and the one followed by L.S. Pontryagin et al..

7 The Maximum Principle and extremal controls

We will now derive the Maximum Principle by combining the tools developed in Section 6 and the necessary conditions for optimal controls derived in Section 4.

Let (τ, ν, ρ, L) denote an arbitrary geometric optimal control structure, with extended geometric control structure $(\overline{\tau}, \overline{\nu}, \overline{\rho})$. In view of the structure of the bundle $\overline{\tau} : \overline{M}(= M \times \mathbb{R}) \to \mathbb{R}$, $(m, J) \mapsto \tau(m)$, it is easily seen that the bundle of vertical tangent vectors $V\overline{\tau}$ is isomorphic to $V\tau \times \mathbb{R}^2$. Similarly, the bundle $V^*\overline{\tau}$ can be identified with $V^*\tau \times \mathbb{R}^2$. In particular, given a point $(m, J) \in \overline{M}(= M \times \mathbb{R})$, a co-vector $\overline{\eta} \in V^*_{(m,J)}\overline{\tau}$ can always be represented by a pair (η_m, η_J) for some $\eta_m \in V^*_m\tau$ and $\eta_J \in \mathbb{R}$.

Before proceeding, we still have to introduce a few additional concepts. First, we recall that the dual of a convex cone C in a vector space $\mathcal{V}$ is defined by the set $C^* = \{\alpha \in \mathcal{V}^* \mid \langle \alpha, v\rangle \leq 0,\ \forall v \in C\}$. A general result that will be used later on, tells that $C^* = (\mathrm{cl}(C))^*$ and $(C^*)^* = \mathrm{cl}(C)$, where cl denotes the closure of C in $\mathcal{V}$ (see e.g. [11] for a proof). Finally, for any $v \in \mathcal{V}$, the half-ray through 0 en v, i.e. $\{w \mid w = rv,\ \forall r \geq 0\}$, will be called the 'cone generated by v', and denoted $C(v)$.

Another concept that we will need, is that of a 'multiplier of a control'. For that purpose, we first construct a 1-parameter family of closed two-forms on $U \times_M V^*\tau$. Let $\tilde{\omega}$ be the closed two-form on the fibred product $U \times_M T^*M$, obtained by pulling back the canonical symplectic form on T^*M by the projection $U \times_M T^*M \to T^*M$. Next, for any real number λ we can define a section H_λ of the fibration $U \times_M T^*M \to U \times_M V^*\tau$ in the following way. Take $u \in U_m, \eta \in V^*_m\tau$ and put $H_\lambda(u, \eta) = (u, \alpha)$, where $\alpha \in T^*_mM$ is uniquely determined by the conditions $\langle \alpha, \mathbf{T}(\rho(u))\rangle + \lambda L(u) = 0$ and α projects onto η. The mapping H_λ is smooth, as can be easily seen from the following coordinate expression: putting $u = (t, x^i, u^a)$ and $\eta = p_i\, dx^i_{|m}$, a straightforward computation gives

$$H_\lambda(t, x^i, u^a, p_i) = \Big(t, x^i, u^a, -\rho^i(t, x^i, u^a)p_i - \lambda L(t, x^i, u^a), p_i\Big)\,.$$

We can now use H_λ to pull-back the closed two-form $\tilde{\omega}$ to a closed two form on $U \times_M V^*\tau$, which will be denoted by ω_λ. Herewith, we can now introduce the following definition of a multiplier.

Definition 7.1 *Given a control $u : [a, b] \to U$, a pair (η, λ) consisting of a continuous piecewise section η of $V^*\tau$ along $c = \nu \circ u$ and a real number λ, is called a* multiplier *of u if the following conditions 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) given any $t_0 \in [a,b]$, and putting $H_\lambda(u(t_0),\eta(t_0)) = (u(t_0),\alpha_0)$, the function $u' \mapsto \langle \alpha_0, \mathbf{T}(\rho(u'))\rangle + \lambda L(u')$, defined on $\nu^{-1}(c(t_0))$, attains a global maximum for $u' = u(t_0)$,

(3) $(\eta(t),\lambda) \neq (0,0)$ for all $t \in [a,b]$.

Returning to the geometric optimal control structure (τ,ν,ρ,L), let $\overline{u}(t) = (u(t), J_0 + \mathcal{J}_u^{(a,t)})$ represent a control in the extended geometric control setting, defined on an interval $[a,b]$. As before, c will denote the base curve of u in M (cf. Section 4), and we put $c(a) = m, c(b) = n$. The bundle map (6.1) associated to the extended geometric control structure will be written as $\overline{\Lambda}$. Given an arbitrary piecewise $\overline{\Lambda}$-admissible curve $\overline{\psi}$ in $J^1\overline{\nu}$ projecting onto $\overline{u}$, we will prove in the following theorem that the dual of the vertical variational cone $V_{(n,J_0+J(u))}R_{(m,J_0)}(\overline{\psi})$ only depends on u.

Theorem 7.2 *Let $\overline{\eta}_0 = (\eta_0,\lambda_0) \in V^*_{(n,J_0+\mathcal{J}(u))}\overline{\tau}$, with $\overline{\eta}_0 \neq 0$. Then we have that $\overline{\eta}_0 \in \left(V_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}(\overline{\psi})\right)^*$ if and only if there exists a section η of $V^*\tau$ along c, with $\eta(b) = \eta_0$, such that the pair (η,λ_0) is a multiplier of u.*

Proof. We prove that any $\overline{\eta}_0 \neq 0$ in the dual of the vertical variational cone determines a multiplier for u. The converse property will then simply follow by reversing the arguments.

Let $\overline{\eta}(t)$ denote the unique continuous piecewise curve in $V^*\overline{\tau}$ satisfying the equation $\overline{D}_{\overline{\psi}}(\overline{u},\overline{\eta})(t) = 0$, with $\overline{\eta}(b) = \overline{\eta}_0$. This implies that $\overline{\eta}(t) = (\psi_t^b)^*(\overline{\eta}_0)$. We can write $\overline{\eta}(t)$ as $\overline{\eta}(t) = (\eta(t),(J_0 + \mathcal{J}_u^{(a,t)},\eta_J(t)))$, where $\eta(t)$, resp. $(J_0 + \mathcal{J}_u^{(a,t)},\eta_J(t))$ are curves in $V^*\tau$, resp. $\mathbb{R}^2$, such that $\eta(b) = \eta_0$ and $\eta_J(b) = \lambda_0$. We will now prove that (η,λ_0) is a multiplier of u.

First of all, it is easily seen that condition (3) of Definition 7.1 holds. In order to prove that (1) and (2) of the definition hold, take an arbitrary $t_0 \in]a,b]$ and $u' \in U_{c(t_0)}$ arbitrary. Then, we find that

$$\overline{\psi}_{t_0}^b\left(\mathbf{T}(\overline{\rho}(u', J_0 + \mathcal{J}_u^{(a,t_0)})) - \mathbf{T}(\overline{\rho}(\overline{u}(t_0)))\right) \in V_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}(\overline{\psi}).$$

By contracting this tangent vector with $\overline{\eta}_0 \in \left(V_{(n,J_0+\mathcal{J}(u))}R_{(m,J_0)}(\overline{\psi})\right)^*$, and taking into account the definition of the dual of a cone, we obtain the following inequality:

$$\langle \eta(t_0), \mathbf{T}(\rho(u')) - \mathbf{T}(\rho(u(t_0)))\rangle + \eta_J(t_0)(L(u') - L(u(t_0))) \leq 0. \qquad (7.1)$$

This holds for any $t_0 \in]a,b]$ and any $u' \in U_{c(t_0)}$. Note that this inequality is also valid for $t_0 = a$. It suffices to consider a local trivialization of U and to interpret the left-hand side of the above inequality as a function of t_0, which is clearly continuous in a neighborhood of a. In particular, we deduce from the above that the function

$$u' \mapsto \langle \eta(t_0), \mathbf{T}(\rho(u')) - \mathbf{T}(\rho(u(t_0))) \rangle + \eta_J(t_0)(L(u') - L(u(t_0))),$$

defined on $U_{c(t_0)}$, admits a global maximum at $u' = u(t_0)$. In local coordinates this means, in particular, that we have:

$$\eta_i(t_0) \frac{\partial \rho^i}{\partial u^a}(u(t_0)) + \eta_J(t_0) \frac{\partial L}{\partial u^a}(u(t_0)) = 0, \tag{7.2}$$

and this holds for all $t_0 \in [a, b]$. These relations are used in the following to prove that the function $\eta_J(t)$ is constant and that $\eta(t)$ satisfies condition (1) of Definition 7.1. The coefficients of the linear $\overline{\Lambda}$-lift $\overline{h}^c$ are related to the coefficients of h^c in the following way (using a slight abuse of notation):

$$\begin{aligned}
\overline{\Gamma}^i_j &= \Gamma^i_j, & \overline{\Gamma}^i_J &= \frac{\partial \rho^i}{\partial u^a} u^a_J, \\
\overline{\Gamma}^J_i &= \frac{\partial L}{\partial x^i} + \frac{\partial L}{\partial u^a} u^a_i, & \overline{\Gamma}^J_J &= \frac{\partial L}{\partial u^a} u^a_J.
\end{aligned}$$

Herewith, the differential equations for $\eta(t)$ and $\eta_J(t)$ become, on every smooth part of $\overline{\eta}$:

$$\begin{aligned}
\dot{\eta}_J(t) &= -\overline{\Gamma}^i_J \eta_i(t) - \overline{\Gamma}^J_J \eta_J(t), \\
\dot{\eta}_i(t) &= -\Gamma^j_i \eta_j(t) - \overline{\Gamma}^J_i \eta_J(t).
\end{aligned}$$

Taking into account the relations (7.2), which hold for all values of $t_0 \in [a, b]$, it is easily seen that $\dot{\eta}_J(t) = 0$ and, hence, η_J is a constant function, with $\eta_J(t) \equiv \lambda_0$. Moreover, the functions $\eta_i(t)$ satisfy:

$$\dot{\eta}_i(t) = -\frac{\partial \rho^j}{\partial x^i} \eta_j(t) - \lambda_0 \frac{\partial L}{\partial x^i}.$$

Putting, in local coordinates, $h_{\lambda_0}(u, \eta) = \rho^i(u)\eta_i + \lambda_0 L(u)$, the two-form ω_{λ_0} reads: $\omega_{\lambda_0} = -dh_{\lambda_0} \wedge dt + dp_i \wedge dx^i$. After some tedious, but straightforward calculations it follows that the condition $i_{(\dot{u}(t), \dot{\eta}(t))} \omega_{\lambda_0} = 0$ is equivalently to

$$\begin{aligned}
\dot{c}^i(t) &= \frac{\partial h_{\lambda_0}}{\partial p_i}(u(t), \eta(t)) = \rho^i(u(t)), \\
0 &= \frac{\partial h_{\lambda_0}}{\partial u^a}(u(t), \eta(t)) = \frac{\partial \rho^i}{\partial u^a}(u(t))\eta_i(t) + \lambda_0 \frac{\partial L}{\partial u^a}(u(t)),
\end{aligned}$$

$$\dot{\eta}_i(t) = -\frac{\partial h_{\lambda_0}}{\partial x^i}(u(t),\eta(t)) = -\frac{\partial \rho^j}{\partial x^i}(u(t))\eta_j(t) - \lambda_0 \frac{\partial L}{\partial x^i}(u(t)),$$
$$\left.\frac{d}{dt}\right|_t (h_{\lambda_0}(u(t),\eta(t))) = \frac{\partial h_{\lambda_0}}{\partial t}(u(t),\eta(t)),$$

and it is easily seen that the curve $\eta(t)$ defined above, satisfies these equations. This shows that condition (3) of Definition 7.1 is satisfied. It finally remains to prove that also the second condition for a multiplier holds.

Consider the section $H_{\lambda_0} : U \times_M V^*\tau \to U \times_M T^*M$ and let us write for any $t_0 \in [a,b]$, $H_{\lambda_0}(u(t_0),\eta(t_0)) = (u(t_0),\alpha(t_0))$. Substituting this into (7.1), and recalling that $\eta_J(t) \equiv \lambda_0$, we obtain:

$$\langle \alpha(t_0), \mathbf{T}(\rho(u'))\rangle + \lambda_0 L(u') \,(= \langle \alpha(t_0), \mathbf{T}(\rho(u(t_0)))\rangle + \lambda L(u(t_0))) \leq 0,$$

proving that (2) is satisfied. This completes the proof that $(\eta(t), \eta_J = \lambda_0)$ is indeed a multiplier. $\Box$

As a consequence of the above theorem, the dual of the vertical variational cone, in the extended setting, only depends on the control u and, hence, this is also true for the closure of this cone. Moreover, as an interesting side result we obtain that the closure of the vertical variational cone $V_n R_m$ also depends on u only. Indeed, using the same techniques as in the above theorem it is easily seen that every multiplier with $\lambda = 0$, determines an element of the dual cone of $V_n R_m$, and vice versa. To simplify the notations we put $\overline{m} = (m, J_0)$ and $\overline{n} = (n, J_0 + \mathcal{J}(u))$. Recall Corollary 4.4, which is reformulated in the following way and leads us to a more familiar version of the maximum principle.

Corollary 7.3 *Assume that $m \xrightarrow{u} n$ and that u is optimal. Then there exists a multiplier (η,λ) with $\lambda \leq 0$.*

The following definitions are well known from the literature.

Definition 7.4 *A control u, with $m \xrightarrow{u} n$ is called an extremal if there exists a multiplier $(\eta(t),\lambda)$ for which $\lambda \leq 0$. An extremal is called normal, resp. abnormal, if there exists a multiplier $(\eta(t),\lambda)$ for which $\lambda < 0$, resp. $\lambda = 0$.*

An extremal is thus equivalently defined as a control for which the closed cone $\mathrm{cl}(V_{\overline{n}} R_{\overline{m}})$ does not contain $-\partial/\partial J$ in its interior. Note that an extremal can be simultaneously abnormal and normal. We say that an extremal is *strictly abnormal* if it is abnormal but not normal. The following proposition gives necessary and sufficient conditions for a control to be an abnormal extremal or a strictly abnormal extremal.

Proposition 7.5 *A control is an abnormal extremal iff* $\mathrm{cl}(V_n R_m) \neq V_n \tau$. *A* *control is a strictly abnormal extremal iff* $-\frac{\partial}{\partial J}$ *is in the border of* $\mathrm{cl}(V_{\overline{n}}R_{\overline{m}})$.

Proof. The first statement follows from the fact that every element in the dual cone $(\mathrm{cl}(V_nR_m))^*$ corresponds to a multiplier with $\lambda = 0$ (see above).

A control is strictly abnormal iff every element $\overline{\eta}_0$ in $(\mathrm{cl}(V_{\overline{n}}R_{\overline{m}}))^*$ satisfies $(\overline{\eta}_0)_J \geq 0$ (by definition). Using the definition of the dual cone and the fact that $C^{**} = \mathrm{cl}(C)$ for an arbitrary convex cone C, we obtain that $-\frac{\partial}{\partial J}$ is contained in $\mathrm{cl}(V_{\overline{n}}R_{\overline{m}})$. On the other hand, since u is an extremal we know that $-\frac{\partial}{\partial J}$ is not contained in the interior of the cone $\mathrm{cl}(V_{\overline{n}}R_{\overline{m}})$. $\square$

It should be noted that the condition $V_nR_m \neq V_n\tau$ does *not* depend on the cost function L. This justifies the notion of an abnormal extremal: u satisfies the necessary conditions for being a optimal control with respect to the cost L, however these conditions do not depend on L. The above result can be intuitively interpreted as follows: a control u is an abnormal extremal iff the family of vector fields $\mathcal{D}$ does not supply enough "vertical" variations to the control u. In the case of strictly abnormal extremals the maximum principle fails in the sense that Corollary 3.6 only gives information on those vectors lying in the interior of a variational cone, and not on those belonging to the boundary.

Acknowledgements

This work has been 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 and to J. Cortés and A. Ibort for many useful suggestions.

References

[1] B. Langerock, A connection theoretic approach to sub-riemannian geometry, J. Geom. Phys. To appear.

[2] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamklelidze and E.F. Mishchenko, The Mathematical Theory of Optimal Processes, Wiley, Interscience, 1962.

[3] J.F. Cariñena and E. Martínez, Lie algebroid generalization of geometric mechanics, in: P. U. J. Kubarski, R. Wolak (Eds.), Lie algebroids and related topics in differential geometry, Vol. 54 of Banach Center Publications, Warzawa, 2001, pp. 201–215.

[4] E. Martínez, Lagrangian mechanics on lie algebroids, Acta. Appl. Math. 67 (2001) 295–320.

[5] A. Weinstein, Lagrangian mechanics and groupoids, in: P. K. W.F. Shadwick, T. Ratiu (Eds.), Mechanics Day, 1995, pp. 207–232.

[6] H. Sussmann, An introduction to the coordinate-free maximum principle, in: B. Jakubczyk, W. Respondek (Eds.), Geometry of Feedback and Optimal Control, Marcel Dekker, New York, 1997, pp. 463–557.

[7] P. Libermann, C.-M. Marle, Symplectic Geometry and Analytical Mechanics, Reidel, Dortrecht, 1987.

[8] F. Cantrijn, B. Langerock, Generalised connections over a vector bundle map, Diff. Geom. Appl. To appear (math.DG/0201274).

[9] S. Helgason, Differential geometry, Lie groups, and symmetric spaces, Academic Press, New York, 1978.

[10] D. Saunders, The geometry of jet bundles, Cambridge University Press, Cambrige, 1979.

[11] G. Köthe, Topological vector spaces I, Springer-Verlag, Berlin, 1969.