Control and Cybernetics

vol. **38** (2009) No. 4B

Curvatures of single-input control systems[*][†]

by

Bronisław Jakubczyk

Institute of Mathematics, Polish Academy of Sciences
00-956 Warsaw, Śniadeckich 8, Poland
e-mail: b.jakubczyk@impan.gov.pl

Abstract: We introduce invariants of control-affine systems which we call curvatures. They are defined by the drift and the control distribution, given by the system. The curvatures allow us to analyse the variational equation along a given trajectory, as well as existence of conjugate points.

Keywords: control system, curvature, invariants, variational equation, conjugate point.

1. Introduction

In the paper we will study certain invariants, called curvatures, associated to a control-affine system with scalar control:

$$\Sigma: \quad \dot{x} = f(x) + ug(x), \qquad u \in \mathbb{R}, \quad x \in X.$$

We assume that the state space X is an open subset of $\mathbb{R}^n$ or, more generally, a differentiable manifold of class C^∞, of dimension n. The vector fields f and g are assumed smooth (of class C^∞) or sufficiently many times differentiable.

In our approach we will analyse the behaviour of Σ around a given trajectory or a family of trajectories. For technical simplicity it is convenient to assume that the trajectory or trajectories are trajectories of the drift f (see Remark 1 for a more general case). Therefore, we will treat the zero control $u^*(t) \equiv 0$ as a distinguished one and the drift f as given. On the other hand, the term ug will play a role of a "correctional control" or perturbation. Changing g for βg, where $\beta(x)$ is a nonvanishing function, will only reparametrise the control and will have no effect on the properties of Σ.

[*]The results of this article were presented at the conference "50 Years of Optimal Control", September 2008, Będlewo, Poland. Research supported by Polish Ministry of Research and Higher Education, grant N201 039 32/2703.

[†]Submitted: September 2009; Accepted: November 2009.

From our point of view one can represent Σ by the pair

$$(f, \mathcal{V}),$$

where f is the drift and $\mathcal{V}$ is the distribution of tangent lines defined by g,

$$\mathcal{V}(x) = \mathbb{R}g(x).$$

The curvatures of Σ, introduced in Section 3, will be invariants associated to the pair $(f, \mathcal{V})$, that is, they will not depend on the choice of the generator g of the distribution $\mathcal{V}$. On the other hand, changing f for $f + \alpha(x)g$ (i.e., adding feedback) will change the curvatures.

REMARK 1 If, instead of zero control, we distinguish a control $u^*(x)$ then we can replace the drift f with the new one,

$$\hat{f}(x) = f(x) + u^*(x)g(x),$$

and analyse the new dynamic pair $(\hat{f}, \mathcal{V})$. In order for the further constructions to work we have to assume that $u^*(x)$ is smooth. Similarly, if the distinguished control is $u^*(t)$, then we add $t = x^0$ as the new coordinate of the system, with the new system equation $\dot{x}^0 = 1$, thus modifying the system vector fields to $\bar{f}$ and $\bar{g}$. Then we replace the earlier drift with $\hat{f} = \bar{f} + u^*(x^0)\bar{g}$. Again, $u^*(\cdot)$ is required to be smooth.

There are two canonical problems, where our invariants are applicable. One is an analysis of the linearized version of Σ along a given trajectory. The second one is the variational equation (the Jacobi equation) along an extremal of an integral functional or, more generally, of an optimal control problem. In this case the drift f is replaced with the vector field on the tangent or cotangent bundle (or on a submanifold of the tangent/cotangent bundle), which defines the extremals of the problem, and $\mathcal{V}$ is the vertical distribution of the bundle. We briefly indicate possibilities of such applications in Sections 5, 6, and 8.

The curvatures seem to be an appropriate tool to determine if a given optimal control problem has conjugate points, the problem which has been treated using other methods by Bonnard and Kupka (1993), Bonnard and Chyba (2003), by Agrachev and other researchers. The case of scalar control, analysed in detail in the two references, seems treatable with our approach with additional advantage of omitting the difficulty of computing the normal form needed there.

The name "curvatures" is justified by the fact that, in the special case of f being the geodesic spray of a surface endowed with a Riemann metric, our curvature is a single scalar and coincides with the curvature appearing in the classical Jacobi equation. Similarly, if f is the Hamiltonian vector field defining the extremals of an optimal problem on $X = M^2$ satisfying certain regularity conditions, then the curvature coincides with the one defined in Agrachev and Sachkov (2004).

The approach can be extended to the vector-control case. In fact, the curvatures for general dynamic pair were used in Kryński (2008) as partial invariants in the equivalence problem for dynamic pairs. They are applicable for dynamic pairs defined by systems of ordinary differential equations, as special cases of dynamic pairs with $\operatorname{rank} \mathcal{V} > 1$ (Kryński, 2008; Jakubczyk and Kryński, 2009).

We introduce the curvatures in Section 3 and provide explicit formulas for computing them in Section 4. Next we write the variational equation (linearized equation along a trajectory) in terms of the curvatures (Section 5, Theorem 2). We introduce a notion of conjugate point, corresponding to needle variations. We prove that there are no conjugate points, if the curvatures are negative along the trajectory (Section 6). Finally, in Sections 7 and 8 we give an example how the curvatures can be used in order to establish if a given extremal of Σ, for a time-minimal problem, has conjugate points and how to estimate their position (we use results from Bonnard and Chyba, 2003).

2. Notation and basic assumptions

Let f and g be arbitrary smooth vector fields on X. In coordinates

$$f(x) = \sum_j f^j(x) \frac{\partial}{\partial x^j}, \qquad g(x) = \sum_j g^j(x) \frac{\partial}{\partial x^j}.$$

Recall that the *Lie bracket* of f and g is the commutator

$$[f, g] = \sum_{j,k} \left(f_j \frac{\partial g^k}{\partial x_j} - g_j \frac{\partial f^k}{\partial x_j} \right) \frac{\partial}{\partial x^k},$$

which is a new vector field. We denote:

$$\operatorname{ad}_f g = [f, g], \quad \operatorname{ad}_f^2 g = [f, [f, g]] \,, \ldots\ldots, \operatorname{ad}_f^{r+1} g = [f, \operatorname{ad}_f^r g].$$

Given a smooth function $\beta : X \to \mathbb{R}$, we denote its Lie derivative along f by

$$L_f(\beta) = f(\beta) := \sum_j f^j \frac{\partial \beta}{\partial x^j}.$$

Let now f and g be vector fields defining Σ.

Assumptions on Σ: $\exists\, r \geq 1$ such that, pointwise,

$$(A1) \qquad g,\ \operatorname{ad}_f g,\ \ldots,\ \operatorname{ad}_f^r g \quad \text{are linearly independent and} \quad f \neq 0,$$

$$(A2) \qquad ad_f^{r+1} g \ = \ h_0\, g + h_1 \operatorname{ad}_f g + \cdots + h_r \operatorname{ad}_f^r g,$$

for some (unique) functions $h_0, \ldots, h_r$.

In particular, (A1) and (A2) are satisfied (with $r = n - 1$) if

$$(A) \quad g,\ \operatorname{ad}_f g,\ \ldots,\ \operatorname{ad}_f^{n-1} g \quad \text{are linearly independent and} \quad f \neq 0.$$

Alternative assumptions: $\exists\ r \geq 1$ such that, pointwise,

$$(A1') \quad f,\ g,\ \mathrm{ad}_f g,\ \dots,\ \mathrm{ad}_f^r g \quad \text{are linearly independent}$$

$$(A2') \qquad ad_f^{r+1} g \ = \ h_0\, g + h_1\, \mathrm{ad}_f g + \cdots + h_r\, \mathrm{ad}_f^r g + \bar{h}\, f,$$

for some (unique) functions $h_0, \dots, h_r, \bar{h}$.

In particular, (A1') and (A2') are satisfied (with $r = n-2$) if

$$(A') \quad f,\ g,\ \mathrm{ad}_f g,\ \dots,\ \mathrm{ad}_f^{n-2} g \quad \text{are linearly independent.}$$

The domain of validity of the above assumptions will be made precise in the statements of our results. They will be needed along a single trajectory of the vector field f, only, if the system properties are analysed along such a trajectory. In particular, $f \neq 0$ will mean that $f(x) \neq 0$ along a given trajectory or a neighbourhood of a given point. The functions $h_0, \dots, h_r$ will be used to compute the curvatures.

3. Definition of curvatures

Our definition is based on the following fact, saying that the function h_r in (A2) and (A2') can be annihilated, by replacing g with some $\bar{g} = \beta g$.

PROPOSITION 1 *(a) If a pair of vector fields (f, g) satisfies (A1) and (A2), or (A1') and (A2'), in a neighbourhood W of a point $x_0 \in X$ such that $f(x_0) \neq 0$, (respectively, along a given trajectory $\gamma : [t_0, t_1] \to X$ of f), then there exists a smooth, nonvanishing function β, defined in a neighbourhood $V \subset W$ of x_0 (respectively, along γ), such that, with $\bar{g} = \beta g$, we have*

$$ad_f^{r+1} \bar{g} \ = \ \bar{h}_0\, \bar{g} + \bar{h}_1\, \mathrm{ad}_f \bar{g} + \cdots + \bar{h}_{r-1}\, \mathrm{ad}_f^{r-1} \bar{g} \qquad (N)$$

in V (respectively, along γ), where in the case of (A1') and (A2') the equality (N) holds modulo f.

(b) If both g and $\bar{g}$ have the property (N), then β satisfies $L_f(\beta) = 0$, i.e., β is constant on trajectories (respectively on the trajectory γ) of f.

(c) The functions $\bar{h}_0, \dots, \bar{h}_{r-1}$ are independent of the choice of β.

The above proposition allows us to define the curvatures using the functions $\bar{h}_0, \dots, \bar{h}_{r-1}$, uniquely defined and independent of the choice of the generator g.

DEFINITION 1 *Assume (A1) and (A2). The functions*

$$k_0 = (-1)^r \bar{h}_0,\ k_1 = (-1)^{r-1} \bar{h}_1,\ \dots,\ k_{r-1} = -\bar{h}_{r-1}$$

are called curvatures *of Σ (or of $(f, \mathcal{V})$). In other words,* curvatures *are unique functions k_i defined, using the generator $\bar{g}$ in Proposition 1, by the equality*

$$ad_f^{r+1} \bar{g} \ = \sum_{i=0}^{r-1} (-1)^{r-i} k_i\, \mathrm{ad}_f^i \bar{g}. \qquad (C)$$

A vector field $\bar{g}$ *satisfying (N), or (C), is called* normal generator *(of* $\mathcal{V}$*) and (N) is called* normality condition. *The same definitions apply under the assumptions (A1'), (A2'). In that case the equalities (N) and (C) are understood modulo* f.

The alternating sign in the definition of k_i is chosen for simplicity of their geometric interpretation (see Section 6). If $r = 1$ then k_0 plays the role analogous to Gauss curvature in the Jacobi equation, which is of the form

$$\dot{v}_0 = -k_0 v_1, \quad \dot{v}_1 = v_0.$$

The sign of k_0 determines if v_0 and v_1 have oscillatory or non-oscillatory behaviour. To see the geometric role of the curvatures the reader may go directly to Sections 5, 6 and 7.

Proof of Proposition 1. (a) Assume that (A1) and (A2) hold. We will use the Leibniz property of Lie bracket,

$$[f, \beta g] = L_f(\beta) g + \beta [f, g],$$

and its iterated consequence

$$\mathrm{ad}_f^i(\beta g) = \sum_{j=0}^{i} \binom{i}{j} L_f^{i-j}(\beta) \mathrm{ad}_f^j g. \tag{1}$$

In particular, we have

$$\mathrm{ad}_f^{r+1}(\beta g) = \beta \mathrm{ad}_f^{r+1} g + (r+1) L_f(\beta) \mathrm{ad}_f^r g \qquad \mathrm{mod}\ g,\ \mathrm{ad}_f g, \ldots, \mathrm{ad}_f^{r-1} g.$$

Since $\mathrm{ad}_f^{r+1} g = h_r \mathrm{ad}_f^r g$, modulo $g, \mathrm{ad}_f g, \ldots, \mathrm{ad}_f^{r-1} g$, we find that

$$\mathrm{ad}_f^{r+1}(\beta g) = (\beta h_r + (r+1) L_f(\beta)) \mathrm{ad}_f^r g \qquad \mathrm{mod}\ g,\ \mathrm{ad}_f g, \ldots, \mathrm{ad}_f^{r-1} g.$$

Therefore, in order to get $\bar{h}_r = 0$ we should have

$$\beta h_r + (r+1) L_f(\beta) = 0. \tag{2}$$

This is a linear differential equation for β. It has a nonvanishing smooth solution β along any trajectory γ of f, and in a neighbourhood of a given point x_0 such that $f(x_0) \neq 0$. (If the trajectory evolves along a closed orbit, then the function $\beta(t)$ may have different values at points $t_1 \neq t_2$ such that $\gamma(t_1) = \gamma(t_2)$.)

If (A1') and (A2') hold, then the proof is the same, except that we consider the equations involving $\mathrm{ad}_f^{r+1}(\beta g)$ modulo $f, g, \mathrm{ad}_f g, \ldots, \mathrm{ad}_f^{r-1} g$.

(b) If also g satisfies (N) then $h_r = 0$. It follows from (2) that $L_f(\beta) = 0$.

(c) This statement follows from (b). Indeed, the only freedom of changing the normal generator $\bar{g}$ in (N) is to multiply it by β such that $L_f(\beta) = 0$. Then $\mathrm{ad}_f^i(\beta \bar{g}) = \beta \mathrm{ad}_f^i \bar{g}$, for any i, and the functions $\bar{h}_i$, defined by (N) with $\bar{g}$ replaced by $\tilde{g} = \beta \bar{g}$ are the same as those defined by $\bar{g}$. This ends the proof. ■

4. Computing curvatures

Before we discuss some problems involving curvatures, we address the question how they can be computed. One obvious way is to use the definition, that is to find them using formula (C). This requires finding a normal generator $\bar{g} = \beta g$ which is to be determined from the differential equation (2), i.e., from

$$L_f(\beta) = h\beta, \quad \text{where} \quad h = -\frac{1}{r+1} h_r. \tag{3}$$

Finding a normal generator, or solving the differential equation (3) for a normalizing function β, may not be an easy task. Computing β along a given trajectory γ of f is always possible, at least numerically. However, this may not suffice for computing the Lie brackets of $\bar{g} = \beta g$ and f.

In order to present explicit formulas for the curvatures we introduce a vector notation. Let

$$H = (h_0, h_1, \ldots, h_r),$$

where h_i are given by formula (A2) or (A2'). The curvatures are arranged into the row vector

$$K = ((-1)^r k_0, (-1)^{r-1} k_1, (-1)^{r-2} k_2, \ldots, -k_{r-1}, 0),$$

where we add the last zero component, for simplicity of further formulas.

Consider the differential operators of order 1,

$$D = L_f + h, \quad \bar{D} = L_f - h,$$

where $L_f = \sum f^j \frac{\partial}{\partial x^j}$ and h is the operator of multiplication by the function $h = -(r+1)^{-1} h_r$. We introduce the coefficients (functions)

$$L^i_j = \binom{i}{j} D^{i-j}(1), \quad \bar{L}^i_j = \binom{i}{j} \bar{D}^{i-j}(1), \qquad \text{if} \quad i \geq j,$$

where 1 is the constant function, equal to 1, and D^k, $\bar{D}^k$ denote the k-th powers of D and $\bar{D}$. We put $L^i_j = 0 = \bar{L}^i_j$, if $i < j$. Note that $L^i_i = 1 = \bar{L}^i_i$. If $h_r = 0 = h$ then $D(1) = 0 = \bar{D}(1)$, and then $L^i_j = 0 = \bar{L}^i_j$, if $i \neq j$.

Let $L^{r+1}_\bullet$ denote the row vector

$$L^{r+1}_\bullet = (L^{r+1}_0, L^{r+1}_1, \ldots, L^{r+1}_r)$$

and let L and $\bar{L}$ be the $(r+1) \times (r+1)$ matrices

$$L = (L^i_j)^r_{i,j=0}, \quad \bar{L} = (\bar{L}^i_j)^r_{i,j=0}.$$

If $h_r = 0$, then $L^{r+1}_\bullet = 0$ and $L = \bar{L} = I$ - the identity matrix.

THEOREM 1 *The vector of curvatures, well defined under the assumptions (A1) and (A2), or (A1') and (A2'), is given by the formula*

$$K = (H + L_{\bullet}^{r+1})\,\bar{L}. \tag{K}$$

Note that the above formula is well defined along a single trajectory of f. For illustration we note that, in the simplest case of $r = 1$, we get from (K)

COROLLARY 1 *If $r = 1$ then $K = (-k_0, 0)$, where the curvature k_0 is*

$$k_0 = -h_0 + \frac{1}{2}L_f(h_1) - \frac{1}{4}\,h_1^2. \tag{4}$$

Proof. We compute

$$D(1) = h, \quad D^2(1) = D(h) = L_f(h) + h^2, \quad \bar{D}(1) = -h,$$

$$L_{\bullet}^2 = (L_0^2, L_1^2) = (L_f(h) + h^2, 2h),$$

and

$$\bar{L} = \begin{pmatrix} 1 & 0 \\ -h & 1 \end{pmatrix}.$$

Thus, $H + L_{\bullet}^2 = (h_0 + L_f(h) + h^2, h_1 + 2h)$ and we find

$$K = (H+L_{\bullet}^2)\bar{L} = (h_0+L_f(h)+h^2-hh_1-2h^2, 0) = (h_0+L_f(h)-h^2-hh_1, 0).$$

Taking into account that $h = -\frac{1}{2}h_1$ we get $K = (-k_0, 0)$, with k_0 as in (4). ∎

In order to prove the theorem we will use the following simple lemmata.

LEMMA 1 *If (3) holds then*

$$(L_f)^i(\beta) = \beta D^i(1), \quad (L_f)^i(\beta^{-1}) = \beta^{-1}\bar{D}^i(1). \tag{5}$$

Proof. Let α and β be smooth functions. Then (5) are special cases (with $\alpha = 1$) of the general formulas

$$(L_f)^i(\alpha\beta) = \beta D^i(\alpha), \quad (L_f)^i(\alpha\beta^{-1}) = \beta^{-1}\bar{D}^i(\alpha).$$

For $i = 1$ the former formula follows from the Leibnitz rule and from (3). For general i it is proved by induction:

$$(L_f)^{i+1}(\alpha\beta) = L_f(L_f)^i(\alpha\beta) = L_f(\beta D^i(\alpha)) = \beta(h+L_f)D^i(\alpha) = \beta D^{i+1}(\alpha),$$

where in the third equality we use (3). The latter formula is proved in the same way. ∎

Denote

$$V^0 = g,\ V^1 = \mathrm{ad}_f g, \ldots, V^r = \mathrm{ad}_f^r g,\ V^{r+1} = \mathrm{ad}_f^{r+1} g,$$

and let $V = (V^0, \ldots, V^r)^T$ (we treat V as a column vector). Similarly, we denote $\bar V^0 = \bar g,\ \bar V^1 = \mathrm{ad}_f \bar g, \ldots, \bar V^r = \mathrm{ad}_f^r \bar g,\ \bar V^{r+1} = \mathrm{ad}_f^{r+1} \bar g$ and $\bar V = (\bar V^0, \ldots, \bar V^r)^T$.

Lemma 2 *If $\bar g = \beta g$ and (3) holds, then*

$$\bar V = \beta L V, \quad V = \beta^{-1} \bar L \bar V \quad \text{and} \tag{6}$$

$$\bar V^{r+1} = \beta(H + L_\bullet^{r+1}) V. \tag{7}$$

Proof. It follows from the iterative Leibniz rule (1) in Section 3 that

$$\bar V^i = \sum_{j=0}^{i} \binom{i}{j} L_f^{i-j}(\beta)\ V^j.$$

Together with (5), this implies the first formula in (6). The second formula is proved analogously. Finally, to prove (7) note that the above formula applied for $i = r+1$ gives

$$\bar V^{r+1} = ad_f^{r+1}(\beta g) = \beta \mathrm{ad}_f^{r+1} g + \sum_{j=0}^{r} \binom{r+1}{j} L_f^{r+1-j}(\beta) V^j$$

$$= \beta V^{r+1} + \beta \sum_{j=0}^{r} L_j^{r+1} V^j = \beta H V + \beta L_\bullet^{r+1} V.$$

■

From (6) and the linear independence of the vectors in V and $\bar V$ we get

Corollary 2 *The matrices L and $\bar L$ are mutually inverse, i.e., $\bar L = L^{-1}$.*

Proof of Theorem 1. Consider a normal generator $\bar g = \beta g$. By its definition, the function β satisfies the relation (3). It follows from (C) that

$$ad_f^{r+1} \bar g = K \bar V = \sum_{j=0}^{r-1} (-1)^{r-i} k_i \bar V^i.$$

Using the first formula in Lemma 2 we find that $K\bar V = \beta K L V$, thus

$$ad_f^{r+1} \bar g = \beta K L V.$$

Using the second formula we get

$$ad_f^{r+1} \bar g = \beta(H + L_\bullet^{r+1}) V.$$

Both equalities yield

$$KLV = (H + L_\bullet^{r+1}) V.$$

As the vector fields $V^0, \ldots, V^r$ in $V = (V^0, \ldots, V^r)^T$ are linearly independent, this implies $KL = (H + L_\bullet^{r+1})$ and $K = (H + L_\bullet^{r+1}) \bar L$ (since $L \bar L = I$), which ends the proof. ■

5. Variational equation

Consider a system

$$\Sigma: \quad \dot{x} = f(x) + ug(x),$$

and a control $u^* : I = [t_0, t_1] \to \mathbb{R}$ such that the corresponding trajectory $x^* : I \to X$ of Σ is well defined. Recall that a *needle variation* of $u^*(\cdot)$ is a control

$$u(t) = \begin{cases} \bar{u} & \text{if } t \in (\tau - \epsilon, \tau] \\ u^*(t) & \text{otherwise}, \end{cases}$$

where $\tau \in (t_0, t_1]$ is given, $\bar{u} \in \mathbb{R}$ is a given value and $\epsilon > 0$ is small enough. The corresponding trajectory $x(t, \epsilon)$ is called a *needle variation* of $x^*(\cdot)$ at $t = \tau$.

For simplicity, we will assume that $0 \in (t_0, t_1)$, $\tau = 0$ and $u^*(t) \equiv 0$ (the general case can be reduced to this one if the control is smooth). Thus, the variation of the trajectory will be nontrivial on the interval $[0, T]$, where we take $T = t_1$.

Given a trajectory $\gamma : t \mapsto x(t)$, $t \in I$, of

$$\dot{x} = f(x), \quad x(0) = x_0, \tag{8}$$

and its needle variation $x(t, \epsilon)$ at $t = 0$, it is well known that the corresponding infinitesimal variation $\delta x(t)$ is propagating so that it satisfies the linearized equation

$$\dot{v}(t) = \frac{\partial f}{\partial x}(x(t))v(t), \quad v(0) = ag(x_0), \tag{9}$$

where $a = \bar{u} - u^*(\tau)$ and we denote

$$v(t) = \delta x(t) = \frac{\partial x}{\partial \epsilon}(t, 0).$$

If $f^t = \exp(tf)$ denotes the flow of f, we can write

$$v(t) = Df^t(x_0)v(0). \tag{10}$$

Lemma 3 *If (A1) and (A2) are satisfied, then*

$$v(t) = \sum_0^r v_i(t)(-1)^{i+1}\mathrm{ad}_f^i g(f^t(x_0)), \tag{11}$$

where v_i are suitable coefficients (the factor $(-1)^{i+1}$ is introduced for simplicity of variational equations in Theorem 2). The same is true under the assumptions (A1') and (A2'), in which case (11) holds modulo $f(f^t(x_0))$.

Proof. By our assumption, f does not vanish. Therefore, locally, there are coordinates $x^1, \dots, x^n$ such that $f = \partial/\partial x^1$. Denote $t = x^1$, then $f = \partial/\partial t$. In these coordinates the flow f^t of f is of the form $(0, x^2, \dots, x^n) \mapsto (t, x^2, \dots, x^n)$ and $Df^t(x_0) = Id$. Thus, $v(t) = const = v(0)$.

Moreover, if $g = \sum_j g^j \partial/\partial x^j$, then

$$\mathrm{ad}_f^i g = \sum_j \frac{\partial^i g^j}{\partial t^i} \frac{\partial}{\partial x^j}.$$

Write g as the vector $g = (g^1, \dots, g^n)$, then $ad_f^i g$ is represented by the vector

$$g^{(i)} := \left(\frac{\partial^i g^1}{\partial t^i}, \dots, \frac{\partial^i g^n}{\partial t^i} \right).$$

Our assumption (A2) takes the form

$$g^{(r+1)} = h_0 g^{(0)} + h_1 g^{(1)} + \cdots + h_r g^{(r)}, \tag{12}$$

along a considered trajectory on $[0, T]$. Since the coefficients $h_0, \dots h_r$ are smooth functions of t, this equation has a solution, uniquely determined by the initial values $g(0) = p_0, g^{(1)}(0) = p_1, \dots, g^{(r)}(0) = p_r$. Moreover, we have

$$g(t) \in \mathrm{span}\{p_0, \dots, p_r\}.$$

This follows from existence and uniqueness of solutions of linear ordinary differential equations. Namely, choosing a new basis in $\mathbb{R}^n$ such that $p_0, \dots, p_r$ are its first $r + 1$ elements, we see that (12) reduces to the subspace spanned by $p_0, \dots, p_r$ and has a solution in this subspace. On the other hand, by uniqueness, the original equation has the same solution, thus $g(t)$ lies in this subspace.

The same argument works backwards. This means that, for a given $t \in [0, T]$, we have $g(0) \in \mathrm{span}\{g(t), g^{(1)}(t), \dots, g^{(r)}(t)\}$. This gives (11) as, due to our notation, $g^{(i)}(t) = ad_f^i g(f^t(x_0))$ and, by the definition of the variation, we have $v(t) = v(0) = ag(0)$, where a is a constant. ∎

In general, the curvatures are not complete invariants of the pair $(f, \mathcal{V})$ (see Jakubczyk and Kryński, 2009) but they are complete for linear non-autonomous equations, in particular, for the variational equation along a given trajectory. This is implied by the following result:

THEOREM 2 *Under the assumptions (A1) and (A2) (or (A1') and (A2')), if the generator g is normal, the coefficients $v_0, \dots, v_r$ in (11) of the infinitesimal variation $v(t)$ satisfy the differential equations*

$$\begin{aligned} \dot{v}_0 &= -k_0 v_r \\ \dot{v}_i &= v_{i-1} - k_i v_r, \qquad i = 1, \dots, r-1, \\ \dot{v}_r &= v_{r-1} \end{aligned} \tag{13}$$

and initial conditions $v_0(0) = a$, $v_2(0) = \cdots = v_r(0) = 0$, where $k_i = k_i(f^t(x_0))$ are the curvatures along the trajectory $f^t(x_0)$ of f.

Proof. We assume that (A1), (A2) are fulfilled. (In the case of (A1') (A2') the same considerations work, provided we consider further equalities modulo f.) Then $\mathrm{ad}_f^{r+1}g = h_0 g + h_1 \mathrm{ad}_f g + \cdots + h_r \mathrm{ad}_f^r g$. If the generator g is normal then, according to our definition of the curvatures k_i, $h_i = (-1)^{r-i}k_i$, $i = 0, \ldots, r-1$, and $h_r = k_r = 0$. We can then write

$$\mathrm{ad}_f^{r+1}g = \sum_{i=0}^{r}(-1)^{r-i}k_i \mathrm{ad}_f^i g. \tag{14}$$

From the equality (11) we get

$$\begin{aligned} v(0) &= (Df^t(x_0))^{-1}v(t) = \sum_0^r (-1)^{i+1}v_i(t)(Df^t(x_0))^{-1}\mathrm{ad}_f^i g(f^t(x_0)) \\ &= \sum_0^r (-1)^{i+1}v_i(t)f_*^{-t}(\mathrm{ad}_f^i g)(x_0). \end{aligned}$$

Differentiating both sides with respect to t and taking into account the relation $d/dt(f_*^{-t}h) = f_*^{-t}(\mathrm{ad}_f(h))$, which holds for any vector field h, we get

$$\begin{aligned} 0 &= \sum_0^r (-1)^{i+1}\dot v_i(t)f_*^{-t}(\mathrm{ad}_f^i g)(x_0) + \sum_0^r (-1)^{i+1}v_i(t)f_*^{-t}(\mathrm{ad}_f^{i+1}g)(x_0) \\ &= \sum_0^r (-1)^{i+1}\dot v_i(t)f_*^{-t}(\mathrm{ad}_f^i g)(x_0) + \sum_0^{r-1} (-1)^{i+1}v_i(t)f_*^{-t}(\mathrm{ad}_f^{i+1}g)(x_0) \\ &\quad +(-1)^{r+1}v_r(t)f_*^{-t}(\mathrm{ad}_f^{r+1}g)(x_0). \end{aligned}$$

Using the expression (14) we get

$$(-1)^{r+1}v_r(t)f_*^{-t}(\mathrm{ad}_f^{r+1}g)(x_0) = \sum_{i=0}^{r}(-1)^{i+1}v_r(t)k_i f_*^{-t}(\mathrm{ad}_f^i g)(x_0),$$

thus

$$0 = \sum_0^r (-1)^{i+1}(\dot v_i(t) - v_{i-1}(t) + v_r(t)k_i)f_*^{-t}(\mathrm{ad}_f^i g)(x_0),$$

where we put $v_{-1} = 0$ and $k_r = 0$. Thus, since $\mathrm{ad}_f^i g$, $i = 0, \ldots, r$ are linearly independent, and so are $f_*^{-t}(ad_f^i g)$, we obtain the system of equations

$$\dot v_i = v_{i-1} - k_i v_r, \quad i = 0, \ldots, r.$$

This system coincides with (13), as we have chosen $v_{-1} = 0$ and $k_r = 0$. ∎

The fact that, in the special basis $(-1)^i \mathrm{ad}_f^{i-1} g(f^t(x_0))$, where g is normal, the coefficients of the variation $v(t) = \delta x(t)$ satisfy the special system of linear differential equations (13), with curvatures $k_i(t)$ as the only nontrivial coefficients, gives hopes that a lot can be said about the behaviour of the variation $v(t)$ in terms of the curvatures. The following section contains a simple example.

6. Absence of conjugate points

Consider again the system

$$\Sigma : \quad \dot{x} = f(x) + ug(x), \qquad u \in \mathbb{R},\ x \in X.$$

We introduce a notion of conjugate points of Σ along a trajectory $\gamma : I \to X$ of the vector field f, $I = [t_0, t_1]$. Assume that (A1) and (A2) hold along γ and

$$r = \max\{k\ :\ g, \mathrm{ad}_f g, \ldots, ad_f^k g \ \text{ are linearly independent along } \gamma\}.$$

Definition 2 *Two points $x_1 = \gamma(t')$ and $x_2 = \gamma(t'')$, with $t', t'' \in I$, $t' < t''$, are called* conjugate points *of Σ if, for any needle variation at $\tau = t' \in I$ of $u^*(t) \equiv 0$, the corresponding infinitesimal perturbation $\delta x(t)$ of γ satisfies:*
$$\delta x(t'') \in \mathrm{span}\{\ g(\gamma(t'')),\ (\mathrm{ad}_f g)(\gamma(t'')),\ \ldots,\ (\mathrm{ad}_f^{r-1} g)(\gamma(t''))\ \}.$$

Note that we can take $\tau = t_0$, since the trajectory γ of f is well defined on $(t_0 - \epsilon, t_1]$ and the corresponding needle variation at $\tau = t_0$ is well defined.

In further considerations we assume, without losing generality, that $I = [0, T]$. Then, a point $x_c = x(t_c)$ is called *conjugate* if it is conjugate to $x_0 = \gamma(0)$, and t_c is called *conjugate time.*

Theorem 3 *(a) If Σ satisfies (A1) and (A2) along $\gamma : [0, T] \to X$ then there is a constant $\delta > 0$ such that Σ has no conjugate points on $\gamma|_{[0,\delta)}$.*
(b) If, in addition, the curvatures $k_0, k_1, \ldots, k_{r-1}$ are nonpositive *along γ, then Σ has no conjugate points on γ.*

Remark 2 The same holds for trajectories corresponding to arbitrary smooth control $u^*(t)$, or $u^*(x)$, if the assumptions (A1), (A2) and definitions of curvatures are suitably modified, see Remark 1.

Proof. Theorem 3 follows from Theorem 2, the ensuing lemma and from Remark 3 at the end of this section. ∎

Lemma 4 *Consider the Cauchy problem*

$$\begin{aligned}
\dot{v}_0 &= b_0 v_r \\
\dot{v}_1 &= v_0 + b_1 v_r \\
\vdots\ &\ \vdots\ \ \vdots \\
\dot{v}_r &= v_{n-1} + b_r v_r
\end{aligned}$$

with continuous coefficients $b_i(t)$, $i = 0, \ldots, r$, and initial conditions

$$v_0(0) = a \neq 0, \quad v_1(0) = \cdots = v_r(0) = 0.$$

The following statements hold:

(a) The components v_i of the solution are (i+1)-times differentiable at zero and

$$v_i(t) = \frac{a}{i!}t^i + o(t^i), \quad i = 0, \dots, r. \tag{15}$$

(b) If the coefficients $b_0(t), b_1(t), \dots, b_r(t)$ are nonnegative, then also

$$v_i(t) \neq 0, \text{ and } \quad \operatorname{sgn}(v_i(t)) = \operatorname{sgn}(a), \text{ for all } t > 0 \text{ and all } 0 \leq i \leq r. \tag{16}$$

(c) If the nonzero initial component is changed for $v_k(0) = a \neq 0$ and $v_i(0) = 0$, for $i \neq k$, then statement (a) holds, with assertion (15) changed for

$$v_i(t) = \frac{a}{(i-k)!}t^{i-k} + o(t^{i-k}), \quad i = k, \dots, r, \tag{17}$$

and statement (b) holds, with (16) satisfied for $i = k, \dots, r$ and $t > 0$ small enough.

Proof. (a) Given a smooth function f, defined in a neighbourhood of zero, we denote by $\operatorname{ord} f$ the order at $0 \in \mathbb{R}$, that is – the smallest order of a nonvanishing derivative of f at zero, and $\operatorname{ord} f = 0$ if $f(0) \neq 0$ ($\operatorname{ord} f = \infty$ if such derivative does not exist). Put $v_{-1} = 0$, then our equations take the iterative form

$$\dot{v}_i = v_{i-1} + b_i v_r, \quad i = 0, \dots, r. \tag{18}$$

Note that, due to nontrivial initial conditions and the form of the equations (18), we can not have $\operatorname{ord} v_r = \infty$. From (18) it follows that, for $i = 1, \dots, r$,

$$\operatorname{ord} v_{i-1} = \operatorname{ord}(\dot{v}_i - b_i v_r) = \operatorname{ord} v_i - 1, \text{ if } \operatorname{ord} v_r \geq \operatorname{ord} v_i \text{ and } v_i(0) = 0,$$

which gives the sequence of implications

$$\operatorname{ord} v_i \leq \operatorname{ord} v_r \implies \operatorname{ord} v_{i-1} = \operatorname{ord} v_i - 1 < \operatorname{ord} v_r. \tag{19}$$

Applying inductively these implications for $i = r, r-1, \dots, 1$ we find out that

$$\operatorname{ord} v_{i-1} = \operatorname{ord} v_i - 1, \text{ and } \operatorname{ord} v_{i-1} < \operatorname{ord} v_r \quad i = 1, \dots r. \tag{20}$$

Together with the initial condition $v_0(0) = a \neq 0$ this yields

$$\operatorname{ord} v_i = i, \quad i = 0, \dots r.$$

Denote by $v_i^*(t)$ the lowest order Taylor terms of $v_i(t)$. Since $\operatorname{ord} v_{i-1} < \operatorname{ord} v_r$, (18) imply the equations

$$\dot{v}_i^* = v_{i-1}^*, \quad i = 1, \dots, r.$$

Taking into account that $v_0^* = v_0(0) = a$ we find that $v_i^*(t) = a(i!)^{-1}t^i$, which proves the formula (15).

To show statement (b) assume $v_0(0) = a > 0$. It follows from statement (a) that all $v_i(t)$ are positive, for small t. Once this holds for small t, the equations (18) with nonnegative k_i imply that all the derivatives are nonnegative, thus $v_i(t) > 0$ for all $t > 0$. Namely, suppose this is not true and $t^* > 0$ is the infimum of t such that at least one component v_j vanishes at t^*. Then, since $v_j(t) > 0$ for $t < t^*$, the derivative $\dot{v}_j(t)$ must be negative at some moments $t < t^*$. This is impossible, since on the right hand side of (18) all components are nonnegative. The proof in the case of a negative is analogous.

The proof of statement (c) is similar but, proving (19) and (20), one should proceed with the induction argument taking $i = r, r-1, \ldots, k-1$, and then use the initial condition $v_k(0) = a$. If $a > 0$, then $v_i(t) > 0$, for $i = k, \ldots, r$ and $t > 0$ small enough. ■

Remark 3 Note that our definition (11) of the coefficients v_i and Theorem 2 imply that the conjugate time t_c is the time where v_r vanishes, i.e., $v_r(t_c) = 0$.

7. Conjugate points in the classic case

For $r = 1$ the system (13) takes the form

$$\dot{v}_0 = -k_0 v_1, \quad \dot{v}_1 = v_0.$$

Putting $y = v_1$ gives the second order equation

$$\ddot{y} = -Ky,$$

where $K = k_0$. It is well known that y has oscillatory behaviour if K is positive. One can estimate positions of zeros of $y(t)$ using the following classical result, which is a special case of the Sturm comparison theorem (see Hartman, 1964).

If two copies $\ddot{y}_j = -K_j y_j$, $j = 1, 2$, of the above equation satisfy

$$K_2(t) \geq K_1(t), \quad t \in [a, b], \quad \text{and} \quad \frac{y_1'(a)}{y_1(a)} \geq \frac{y_2'(a)}{y_2(a)}$$

and the function $y_1(t)$ has a zero in $(a, b]$, then $y_2(t)$ also has a zero in $(a, b]$. If one of the assumed equalities is strict, then $y_2(t)$ has a zero in $[a, b)$.

Above, we put $y_j'(a)/y_j(a) = +\infty$, if $y_j(a) = 0$, thus the assumed inequality is satisfied if $y_1(0) = 0 = y_2(0)$. The proof uses a new variable θ, satisfying $\cot\theta = y'/y$. Differentiation of this equality and elimination of y'' using $y'' = -Ky$ gives

$$\dot{\theta} = \cos^2\theta + K\sin^2\theta =: f(\theta).$$

Zeros of y correspond to the values $\theta = n\pi$. The inequality $K_2 \geq K_1$ implies that the right-hand sides of two copies of the above equation, corresponding to K_1 and K_2, satisfy $f_2(\theta) \geq f_1(\theta)$. This and $\theta_2(a) \geq \theta_1(a)$ imply that their solutions satisfy $\theta_2(t) \geq \theta_1(t)$ and the result follows. In particular, assuming $K_1 = K = const > 0$ and $y_1(a) = 0$, $y_1'(a) = c \neq 0$ we have $y_1(t) = cK^{-1/2}\sin(\sqrt{K}(t-a))$. Thus, if $K_2(t) \geq K$ then both assumed inequalities hold and $\ddot{y}_2 = -K_2 y_2$ has a root in $(a, b]$, where $b = a + \pi\sqrt{K}$.

8. Conjugate points for a time-optimal problem

We illustrate the use of our curvatures for estimating existence and position of conjugate times for a system in $\mathbb{R}^3$,

$$\Sigma: \quad \dot{x} = f(x) + ug(x), \quad |u| \leq M.$$

We admit $M = +\infty$, then $u(t) \in \mathbb{R}$. We make use of an analysis presented in Bonnard and Chyba (2003), where the conjugate time is defined in $\mathbb{R}^n$ in a different way, using normal forms of Σ and eigenfunctions of certain self-adjoint differential operators.

Define the functions on $\mathbb{R}^3$,

$$D = \det(g, [f,g], [g,[f,g]]), \quad D_1 = \det(g, [f,g], [f,[f,g]]),$$

and the vector field

$$S = f - \frac{D_1}{D} g.$$

It is well known that a trajectory γ of S is a singular extremal of the time-optimal problem for Σ, if the following assumptions hold:

(H0) γ is contained in the region of $x \in \mathbb{R}^3$ where $D(x) \neq 0$ and $f(x)$, $g(x)$ are linearly independent.

(H1) Along γ we have

$$[S,[S,g]](\gamma(t)) \in span\{g(\gamma(t)), [f,g](\gamma(t))\}.$$

Since S is smooth in the region where $D(x) \neq 0$ and it follows from (H0) that γ lies in this region, we may assume that S is smooth in the domain of consideration. Moreover, from the formula for S we see that $[S,g] = [f,g]$ modulo g. Therefore, it follows from (H1) that there are smooth functions $h_0(t)$ and $h_1(t)$ such that

$$[S,[S,g]](\gamma(t)) = h_0(t)g(\gamma(t)) + h_1(t)[S,g](\gamma(t)).$$

We see that our assumptions (A1), (A2) are satisfied, with $r = 1$, for the control system $\dot{x} = S(x) + ug(x)$. Thus, the curvature $k_0(t)$ is well defined along γ, with the formula given in Section 4,

$$k_0(t) = -h_0(t) + \frac{1}{2}h_1'(t) - \frac{1}{4}(h_1(t))^2. \tag{21}$$

The variational equations in Theorem 2 take the form

$$\dot{v}_0 = -k_0 v_1, \quad \dot{v}_1 = v_0, \tag{22}$$

or $\ddot{y} = -k_0 y$, where $y = v_1$. The latter equation plays the role of the Jacobi equation for the time-optimal problem.

Consider a time-minimal problem for the system Σ, with a starting point $x(0) = \gamma(0)$ and let $\gamma(t)$ be a fixed singular extremal (i.e. satisfying the necessary conditions of the Pontryagin Maximum Principle and corresponding to control $u^*(t) \in (-M, M)$). Assume, in addition, that γ satisfies (H0) and (H1). From the considerations ending the preceding section and the results in Bonnard and Chyba (2003) we deduce the following:

PROPOSITION 2 *Let $t_c > 0$ denote the first zero of v_1 of the solution of (22) with initial conditions $v_0(0) = 1$, $v_1(0) = 0$ (if such t_c does not exist, we put $t_c = +\infty$). Then, γ is a time minimal trajectory of Σ in a C^0 neigbourhood of γ for the fixed end-point problem. γ ceases to be time-optimal for $t > t_c$.*

The proof is a consequence of Lemma 21 in Bonnard and Chyba (2003) and the fact that the variational equation there can be replaced with our equation (22). Note that our approach gives an explicit formula (21) for computing the curvature k_0 in the variational equation. In particular, by denoting $T_K = \min\{\pi K^{-1/2}, T\}$ we have the following corollaries.

(i) If $k_0(t) \leq 0$ along $\gamma : [0,T] \to \mathbb{R}^3$, then γ is time-optimal on $[0,t]$ in a C^0 neighbourhood of γ, for all $0 < t \leq T$.

(ii) If $k_0(t) \leq K = const > 0$ along γ, then γ is optimal on $[0,t]$ in a C^0 neighbourhood of γ, for $t < T_K$.

(iii) If $k_0(t) \geq K = const > 0$ along γ, then γ is not optimal on $[0,t]$, in any C^0 neighbourhood of γ, for $\pi K^{-1/2} < t \leq T$.

The statements follow from the Sturm comparison theorem. In the first and second cases it implies that v_1 has no zeros on $[0,T]$ (respectively, on $[0,T_K)$), and in the third case it implies that it has a zero in the interval $[0,T_K]$, see the preceding section.

References

AGRACHEV A.A. and SACHKOV, YU.L. (2004) *Control Theory from the Geometric Viewpoint.* Springer-Verlag, New York.

BONNARD, B. and CHYBA, M. (2003) *Singular Trajectories and their Role in Control Theory. Mathématiques & Applications* **40**, Springer-Verlag, New York.

BONNARD, B. and KUPKA, I. (1993) Théorie des sigularités de l'application entrée/sortie et optimalité des trajectoires singulières dans le problème du temps minimal. *Forum Math.* **5**, 11-159.

Hartman, P. (1964) *Ordinary Differential Equations.* John Wiley & Sons, New York.

Jakubczyk, B. and Kryński, W. (2009) Vector fields with distributions and geometry of ODE's (preprint).

Kryński W. (2008) Equivalence problems for tangent distributions and ordinary differential equations. PhD Thesis, Institute of Mathematics, Polish Academy of Sciences, Warszawa (in Polish).