ESAIM: COCV 31 (2025) 7
https://doi.org/10.1051/cocv/2024079

ESAIM: Control, Optimisation and Calculus of Variations
www.esaim-cocv.org

CONJUGATE POINTS OF DYNAMIC PAIRS AND CONTROL SYSTEMS

B. Jakubczyk* and W. Kryński

Abstract. We study the geometry of dynamic pairs $(X, \mathcal{V})$ on a manifold M, where X is a vector field and $\mathcal{V}$ is a distribution on M, both satisfying a regularity condition. Special cases are pairs defined by systems of second order ODEs, geodesic sprays in Riemannian, Finslerian and Lagranian geometries, semi-Hamiltonian systems and control-affine systems. Analogs of conjugate points from the calculus of variations are defined for the pair $(X, \mathcal{V})$. The main results give estimates for the position of conjugate points in terms of a curvature operator, analogously to the Cartan–Hadamard and Bonet–Myers theorems. Contrary to classical cases, no metric is given a priori, the distribution $\mathcal{V}$ may be nonintegrable and the curvature operator is defined in terms of $(X, \mathcal{V})$.

Mathematics Subject Classification. 93C10, 34C14, 53A55.

Received October 27, 2023. Accepted October 27, 2024.

1. Introduction

In the present paper we consider *dynamic pairs* $(X, \mathcal{V})$, where X is a vector field on a manifold M and $\mathcal{V}$ is a distribution on M (a field of tangent m-planes), both satisfying a natural regularity condition. Such pairs are defined by geodesic sprays in (pseudo)-Riemannian and Finsler geometry, in Lagrangian and Hamiltonian mechanics, in spray geometry, and in geometry of systems of second order ODEs. More generally, they appear in geometric descriptions of control-affine systems. General theory of dynamic pairs was developed in [1] where a number of geometric objects invariantly assigned to dynamic pairs were introduced. These include a curvature operator, a covariant derivative and invariant metrics.

Our goal here is to apply these notions in order to study the infinitesimal behaviour of trajectories of X, in vicinity of a given one, relative to the distribution $\mathcal{V}$. To this aim we extend the classical notion of conjugate points from the variational calculus to our setting. Here a basic role is played by the aforementioned curvature operator, denoted K. This operator is analogous to the curvature operator which appears in the classical Jacobi equation, often called Jacobi endomorphism. Fiberwise, it is an endomorphism $K_x \colon \mathcal{V}(x) \to \mathcal{V}(x)$. All data are assumed to be C^∞-smooth.

Our formalism includes, as particular cases, similar questions deeply studied in Riemann and Finsler geometry and in the geometry of Hamiltonian vector fields on symplectic manifolds in presence of Lagrangian foliations with particular attention put on applications in sub-Riemannian geometry (see [2–8]). The theory of the Jacobi vector fields is also well established in the context of systems of second-order differential equations (see

Keywords and phrases: Dynamic pairs, control systems, geodesic flows, conjugate points, curvature, Jacobi fields.

Institute of Mathematics, Polish Academy of Sciences, ul. Śniadeckich 8, 00-656 Warszawa, Poland.

* Corresponding author: b.jakubczyk@impan.pl

[9–14]). In all the mentioned works distribution $\mathcal{V}$ is integrable. Furthermore, there is a canonical metric in the Riemannian, Finslerian and Hamiltonian settings and the operator K is symmetric with respect to the metric. In our case there is no canonical metric and the curvature K is not necessarily symmetric. This makes the problem harder to deal with.

Our motivating example comes from Geometric Control Theory. The most often considered class of non-linear systems are control-affine systems of the form

$$\dot{x} = X(x) + \sum_{i=0}^{m} u_i Y_i(x), \tag{1.1}$$

where $X, Y_1, \ldots, Y_m$ are vector fields on a manifold M, $x \in M$ is the state of the system, and $u_1, \ldots, u_m$ are components of the control. Such a system defines a dynamic pair $(X, \mathcal{V})$ where

$$\mathcal{V} = \operatorname{span}\{Y_1, \ldots, Y_m\}.$$

We shall assume that $X(x) \neq 0$ and $\operatorname{rk}\operatorname{span}\{Y_1, \ldots, Y_m, [X, Y_1], \ldots, [X, Y_m]\} = 2m$, where $[\cdot, \cdot]$ denotes Lie bracket. These assumptions are often met in applications. Vector field X, usually called the *drift* of the system, corresponds to the free motion of the system with no controls. In this regard, our results can be interpreted in terms of stability of the flow generated by X with respect to perturbations tangent to $\mathcal{V}$. In Section 3 we introduce the concept of conjugate points (times) and provide estimates for the first conjugate time. As a by-product of our results we also obtain new theorems on the location of conjugate points for general second-order ODEs. In this context our results strengthen a result in [12] concerning a specific case where the curvature has a one-dimensional invariant eigenspace.

Another motivation comes from optimal control theory where the problem of location of conjugate points is crucial for determining how long from a starting point a given extremal is optimal, analogously to Riemannian and Finslerian geometries. Even if in our study we do not have any optimality criterion, the formalism and results in Section 4 are related to the rich research in this field, especially to the works [2, 4–7, 15]. Especially relevant are the general comparison theorems of Barilari and Rizzi [7] on the location of conjugate points in sub-Riemannian problems (see Rem. 4.4 in Sect. 4 for more details).

In Section 2 we recall definitions and basic facts concerning the dynamic pairs studied in a more general setting in [1]. New results are contained in Sections 3 and 4. In particular, in Section 3.1 the notion of conjugate points is given and main Theorems 3.3 and 3.6 are stated. Further, we consider Jacobi vector fields in Section 3.2. They are used in Section 3.4 in proofs of the main results. Section 4 is devoted to (semi)-Hamiltonian version of dynamic pairs where, additionally to X and $\mathcal{V}$, a skew-symmetric 2-form σ on M is given (possibly, non-closed). Finally, Section 5 is devoted to special cases. In particular we consider geodesic sprays, mechanical control systems (see [16, 17]), and certain classes of nonlinear systems of second-order ODEs.

2. Dynamic pairs, normal frames and curvature

In this section we mainly recall basic notions and constructions introduced before in [1] and we formulate a number of results used later in the paper. Note that it is shown in [1] that in the case of the second-order systems of ordinary differential equations our constructions reinterpret the standard approach that goes back to [18–20]. A similar approach in this context is also used in [10–14].

2.1. Dynamic pairs

Let M be a smooth differentiable manifold. A *dynamic pair* on M is a pair $(X, \mathcal{V})$, where X is a vector field on M and $\mathcal{V} \subset TM$ is a smooth distribution on M which is a field of tangent planes of dimension $m = \operatorname{rk}\mathcal{V}$.

The vector fields $V \in \mathcal{V}$, together with the Lie brackets $[X, V]$ span a new distribution $\mathcal{D} \subset TM$ denoted

$$\mathcal{D} = \mathcal{V} + [X, \mathcal{V}].$$

Above and later on we often write $V \in \mathcal{V}$, instead of $V \in \mathrm{Sec}(\mathcal{V})$, meaning that V is a section of $\mathcal{V}$, usually a local section.

Definition 2.1. A dynamic pair $(X, \mathcal{V})$ is called *regular* on $U \subset M$ if it satisfies three conditions:

$$X(x) \neq 0, \quad x \in U, \tag{$R1$}$$

$$\mathrm{rk}\, \mathcal{D} = 2m \tag{$R2$}$$

(*i.e.* the rank of $\mathcal{D}$ is maximal possible), and

$$[X, \mathcal{D}] \subset \mathcal{D}, \tag{I}$$

i.e. $\mathcal{D}$ is invariant under the local flow of X. Conditions (R1) and (R2) will be called *regularity conditions*, while (I) will be called *invariance condition*.

Remark 2.2. We will usually identify M with U, taking it sufficiently small so that (R1), (R2) and (I) are satisfied on M. A sufficiently interesting case is that of $\mathcal{D} = TM$. In this case $\dim M = 2m$ and the invariance condition (I) is automatically satisfied. Another natural case is $\dim M = 2m + 1$ and $TM = \mathcal{D} \oplus \mathrm{span}\{X\}$. In this case most results of the paper also hold under a weaker condition: $[X, \mathcal{D}] \subset \mathcal{D} \mod X$ which only requires that all computations are performed modulo X. However, for simplicity, we assume that (I) is always satisfied. Condition (R1) is not needed for the definition of the curvature below but it is essential in Proposition 2.6, where it implies the non-singularity of equation (2.4).

2.2. Curvature operator

If $V_1, \dots, V_m$ span $\mathcal{V}$ on M or an open subset $U \subset M$ (resp. along a trajectory c of X) then the tuple

$$\hat{V} = (V_1, \dots, V_m),$$

will be called a *frame* of $\mathcal{V}$ on U (resp. along c).

Note that condition (R2) implies that the vector fields V_j and $[X, V_j]$, $0 \le j \le m$, are pointwise linearly independent and they span $\mathcal{D}$. Moreover, it follows from (I) that $[X, [X, V_j]]$ are also in $\mathcal{D}$. Thus

$$[X, [X, V_j]] = \sum_i (H_0)^i_j V_i + \sum_i (H_1)^i_j [X, V_i],$$

where the functions $(H_0)^i_j$, $(H_1)^i_j$ form two $m \times m$ matrices H_0 and H_1. In matrix notation we have

$$[X, [X, \hat{V}]] = \hat{V} H_0 + [X, \hat{V}] H_1. \tag{2.1}$$

Definition 2.3. The curvature matrix in the frame $\hat{V}$, of the regular pair $(X, \mathcal{V})$, is the matrix

$$K = -H_0 + \frac{1}{2} X(H_1) - \frac{1}{4} H_1^2, \tag{2.2}$$

where $X(H_1)$ is the matrix of derivatives along X of the coefficients of H_1.

The following result proved in [1], Formula (11) and Proposition 2.5 implies that the matrix K defines, at each $x \in M$, a linear operator $K_x : \mathcal{V}(x) \to \mathcal{V}(x)$ which will be called *curvature operator* of $(X, \mathcal{V})$.

Proposition 2.4. *When the local frame $\hat{V}$ is changed for $\hat{V}' = \hat{V}G$, with an $m \times m$ invertible matrix G of smooth functions, the curvature matrix is transformed to*

$$K' = G^{-1}KG. \tag{2.3}$$

The curvature K can be defined in a more conceptual way. For this we need the following notion.

Definition 2.5. A local section V of $\mathcal{V}$ is called *normal* if

$$[X, [X, V]] \in \mathcal{V}.$$

A local frame $\hat{V} = (V_1, \ldots, V_m)$ of $\mathcal{V}$ on $U \subset M$ is called *normal frame* of $\mathcal{V}$ if all V_i are normal sections of $\mathcal{V}$. We also call $V_1, \ldots, V_m$ *normal generators* of $\mathcal{V}$.

Note that in [1] the normal sections satisfy a weaker condition: $[X, [X, V]] \in \mathcal{V} \mod X$. However, in the present article, we always assume that the invariance condition (I) holds, which implies the consistency of Definition 2.5 with the one given in [1] (compare Rem. 2.2 above).

Normal frames in $\mathcal{V}$ exist (see [1], Prop. 2.2) and have the following properties.

Proposition 2.6. *For a regular dynamic pair $(X, \mathcal{V})$ on M the following holds.*

(i) Given $x_0 \in M$ and a frame F in $\mathcal{V}(x_0)$, then along the (non-closed) trajectory c of X starting from x_0 there is a unique normal frame $\hat{V}$ of $\mathcal{V}$ which coincides with F at x_0. In particular, there is a normal frame in neighbourhood of x_0.

(ii) If $\hat{V}$ is a frame of $\mathcal{V}$ on an open $U \subset M$ and H_1 is defined via (2.1), then $\hat{W} = \hat{V}G$ is a normal frame of $\mathcal{V}$ on U if and only if

$$X(G) = -\frac{1}{2}H_1G. \tag{2.4}$$

In particular, if $\hat{V}$ is normal then $\hat{W} = \hat{V}G$ is normal if and only if $X(G) = 0$ and a vector field $Y = \sum f^iV_i$ is normal if and only if $X(f^i) = 0$.

Remark 2.7. Equation (2.4) implies that the function $d(t) = \det G(x(t))$ satisfies the differential equation $\dot{d} = ad$ along any trajectory $x(t)$ of X, where $a(t) = -\frac{1}{2}\operatorname{tr} H_1(x(t))$. Assuming $d(0) \neq 0$ we have $d(t) \neq 0$, thus (2.4) has a global, non-degenerate solution G along the whole trajectory $x(t)$.

Now, if (R1), (R2), (I) hold and $\hat{V} = (V_1, \ldots, V_m)$ is a normal frame in $\mathcal{V}$, there are functions K_i^j such that

$$[X, [X, V_i]] + \sum_j K_i^j V_j = 0, \tag{2.5}$$

or in the matrix form

$$[X, [X, \hat{V}]] + \hat{V}K = 0. \tag{2.6}$$

This means that $H_0 = -K$ and $H_1 = 0$ in equation (2.1). Thus $K = (K_i^j)$ coincides with the curvature matrix defined in Definition 2.3. The fact that the curvature matrix K is defined in a normal frame will have a special meaning for us. Thus we introduce

Definition 2.8. Let $c : [0, T] \to M$ be an integral curve of X. The curvature matrix

$$K(t) = (K_i^j(c(t))), \quad t \in [0, T],$$

expressed in a normal frame along c (equivalently, defined by Eq. (2.5) along c) will be called *normal curvature matrix* along c of the regular pair $(X, \mathcal{V})$.

2.3. Systems of second order ODEs

As mentioned before, basic examples of regular dynamic pairs are provided by systems of second order differential equations $\ddot{x} = F(t, x, \dot{x})$, $t \in \mathbb{R}$, $x \in \mathbb{R}^m$. Geometric invariants of such systems were deeply studied by many authors beginning with the works of Kosambi [20], Cartan [18], and Chern [19]. Our general constructions presented in the next sections coincide, when specified to this case, with works of [10, 13] (see also [12] and references therein). An invariant characterization of general dynamic pairs which correspond to systems of ODEs is given in [1], Theorem 4.5.

For systems of second order ODEs we take $M = J^1(\mathbb{R}, \mathbb{R}^m) \simeq \mathbb{R} \times \mathbb{R}^m \times \mathbb{R}^m$ which is the space of 1-jets of mappings $\mathbb{R} \to \mathbb{R}^m$ with natural coordinates t, x^i, y^i on M (y^i replacing $\dot{x}^i$). We have a natural vector field on M which is the total derivative

$$X = \partial_t + \sum_{i=1}^m y^i \partial_{x^i} + \sum_{i=1}^m F^i \partial_{y^i}.$$

As the distribution $\mathcal{V}$ we take span $\{\partial_{y^1}, \ldots, \partial_{y^m}\}$ (the tangent distribution to the foliation of fibers of the bundle of 1-jets of curves in $\mathbb{R}^m$). It is then easy to check that, for $V_j = \partial_{y^j}$, we have $[X, V_j] = -\partial_{x^j} - \sum_i F^i_{y^j} \partial_{y^i}$ and

$$\mathcal{D} = \text{span}\,\{\partial_{x^1}, \ldots, \partial_{x^m}, \partial_{y^1}, \ldots, \partial_{y^m}\}$$

Thus $(X, \mathcal{V})$ is a regular dynamic pair. Moreover, we have

$$[X, [X, V_j]] = \sum_i F^i_{y^j} \partial_{x^i} + \sum_{i,k} (-F^i_{ty^j} - y^k F^i_{x^k y^j} - F^k F^i_{y^k y^j} + F^i_{x^j} + F^k_{y^j} F^i_{y^k}) \partial_{y^i}.$$

Taking into account that the first sum and the last term in the second sum give together $-\sum_i F^i_{y^j} [X, V_i]$ we see that the matrices in (2.1) are equal to

$$H_0 = (-F^i_{ty^j} - y^k F^i_{x^k y^j} - F^k F^i_{y^k y^j} + F^i_{x^j}), \quad H_1 = (-F^i_{y^j}).$$

Thus the general formula (2.2) for curvature gives the matrix

$$K^i_j = -F^i_{x^j} - \frac{1}{4} F^i_{y^k} F^k_{y^j} + \frac{1}{2} F^k F^i_{y^k y^j} + \frac{1}{2} y^k F^i_{x^k y^j} + \frac{1}{2} F^i_{ty^j}, \tag{2.7}$$

or in more compact form

$$K = -F_x + \frac{1}{2} X(F_{\dot{x}}) - \frac{1}{4}(F_{\dot{x}})^2. \tag{2.8}$$

Formula (2.7) appeared for the first time in the aforementioned works [18–20] and, in recent literature, it is often referred to as the Jacobi curvature or Jacobi endomorphism (*cf.* [10]). Note that the trace-free part of K is referred to as the torsion in [21].

Remark 2.9. Notice, that the case of autonomous equations $\ddot{x} = F(x, \dot{x})$ can also be analysed on the simpler manifold $M = \mathbb{R}^m \times \mathbb{R}^m$, in which case we remove ∂_t from the total derivative. The remaining ingredients define a regular dynamic pair with the distribution $\mathcal{D}$ being the whole tangent bundle TM and the curvature given by (2.7), with $F_{ty} = 0$.

2.4. Invariant splitting and connection

Proposition 2.6 says that if both $\hat{V}$ and $\hat{W} = \hat{V}G$ are normal frames of $\mathcal{V}$ then $X(G) = 0$. Hence $[X, \hat{W}] = [X, \hat{V}]G$. Consequently, given a normal frame $\hat{V} = (V_1, \ldots, V_m)$, the distribution

$$\mathcal{H} = \operatorname{span}\{[X, V_1], \ldots, [X, V_m]\} = \operatorname{span}\{[X, \hat{V}]\}$$

does not depend on the choice of such frame. Clearly $\mathcal{H} \subset \mathcal{D}$, which follows from $\mathcal{V} \subset \mathcal{D}$ and $[X, \mathcal{D}] \subset \mathcal{D}$.

The distribution $\mathcal{H}$ will be called *horizontal* and vector fields in $\mathcal{H}$ will also be called *horizontal*. By Proposition 2.6 a vector field $Y = \sum f^i V_i$, expressed in a normal frame $(V_1, \ldots, V_m)$, is normal iff $X(f^i) = 0$ and then $[X, Y] = \sum f^i [X, V_i]$ is horizontal.

If $\hat{V} = (V_1, \ldots, V_m)$ is a general frame of $\mathcal{V}$, then $\mathcal{H}$ can be computed as follows (see [1], Prop. 2.5).

Proposition 2.10. *If $[X, [X, \hat{V}]] = [X, \hat{V}]H_1 + \hat{V}H_0$, for some $m \times m$ matrices H_0, H_1, then*

$$\mathcal{H} = \operatorname{span}\left\{ \operatorname{ad}_X \hat{V} - \frac{1}{2}\hat{V}H_1 \right\}.$$

Condition (R2) implies that the distribution $\mathcal{D}$ can be written as

$$\mathcal{D} = \mathcal{V} \oplus \mathcal{H}$$

which we call the *canonical split* of $\mathcal{D}$. This relation defines canonical projections

$$\pi_{\mathcal{V}} \colon \mathcal{D} \to \mathcal{V} \quad \text{and} \quad \pi_{\mathcal{H}} \colon \mathcal{D} \to \mathcal{H}.$$

Definition 2.11. A *covariant derivative* along X of vector fields in $\mathcal{V}$ is the operator $D_X \colon \operatorname{Sec}(\mathcal{V}) \to \operatorname{Sec}(\mathcal{V})$ given by

$$D_X V := \pi_{\mathcal{V}}[X, V], \quad V \in \mathcal{V}.$$

More generally, D_X extends to an operator acting on sections of $\mathcal{D}$ by formula

$$D_X Y := \pi_{\mathcal{V}}[X, \pi_{\mathcal{V}} Y] + \pi_{\mathcal{H}}[X, \pi_{\mathcal{H}} Y].$$

Note that D_X is also well defined as an operator $\operatorname{Sec}_c(\mathcal{V}) \to \operatorname{Sec}_c(\mathcal{V})$, or $\operatorname{Sec}_c(\mathcal{D}) \to \operatorname{Sec}_c(\mathcal{D})$, where Sec_c denotes sections along a given trajectory $c \colon [0, T] \to M$ of X. It satisfies the Leibniz rule

$$D_X(fY) = X(f)Y + f D_X Y.$$

and its definition implies the following characterization.

Proposition 2.12. *For vector fields $V \in \mathcal{V}$ and $H \in \mathcal{H}$ defined locally or along c we have:*

$$D_X V = 0 \iff V \text{ is normal},$$

$$D_X H = 0 \iff H = [X, W], \text{ where } W \in \mathcal{V} \text{ is normal}.$$

Remark 2.13. The canonical splitting in the context of second order systems of ODEs and the corresponding covariant differential D_X coincides with the analogous notions used in [10, 13]. Counterparts of these notions for higher order ODEs were introduced in [1].

Remark 2.14. The operators $A\colon \mathcal{V} \to \mathcal{H}$ and $B\colon \mathcal{H} \to \mathcal{V}$ defined by

$$A = \pi_{\mathcal{H}} \circ \mathrm{ad}_X, \quad B = \pi_{\mathcal{V}} \circ \mathrm{ad}_X$$

are vector bundle morphisms, canonically defined by $(X, \mathcal{V})$. This is a consequence of the fact that, any time we compose the operator ad_X defined on sections $Sec(\Delta)$ of a distribution $\Delta \subset TM$ with a projection π along Δ, we obtain a vector bundle morphism because $\pi(\mathrm{ad}_X(fY)) = \pi(X(f)Y + f[X,Y]) = f\pi([X,Y])$. In addition, it follows from (R2) that A is a local isomorphism of bundles, given on a normal frame by $AV_i = [X, V_i]$. The operators A and B have canonical extensions to $\mathcal{D}$ defined by $\hat{A} = \pi_{\mathcal{H}} \circ \mathrm{ad}_X \circ \pi_{\mathcal{V}}$ and $\hat{B} = \pi_{\mathcal{V}} \circ \mathrm{ad}_X \circ \pi_{\mathcal{H}}$. Using them we get the decomposition $\mathrm{ad}_X = D_X + \hat{A} + \hat{B}$.

Remark 2.15. The curvature operator K of a regular dynamic pair $(X, \mathcal{V})$ can be equivalently defined as $K = -B \circ A$ which follows from formula (2.5).

2.5. Invariant metrics

Any regular dynamic pair $(X, \mathcal{V})$ admits, along any trajectory of X (or locally), metrics which are invariant under covariant differentiation along X. We begin with defining such a metric on the distribution $\mathcal{V}$. Consider a normal frame $\hat{V} = (V_1, \ldots, V_m)$ in $\mathcal{V}$, along a trajectory $c\colon [0,T] \to M$ of X. Let g be the positive definite metric in $\mathcal{V}$ such that $V_1, \ldots, V_m$ are orthonormal. Equivalently, if $Y = \sum \phi^i V_i$, $Z = \sum \psi^i V_i$ are vector fields in $\mathcal{V}$, along c, the metric g is given by

$$g(Y,Z) = \sum_i \phi^i \psi^i.$$

By adding minus signs in the above sum we can define a metric with any prescribed signature. It follows from Proposition 2.12 that $D_X V_i = 0$, thus $D_X Y = \sum X(\phi^i) V_i$ and $D_X Z = \sum X(\psi^i) V_i$ and we see that the metric is invariant under the covariant derivative D_X, *i.e.*,

$$X(g(Y,Z)) = g(D_X Y, Z) + g(Y, D_X Z). \tag{2.9}$$

Any metric on $\mathcal{V}$ satisfying (2.9) will be called *invariant metric*. The above metric on $\mathcal{V}$ can be extended to $\mathcal{D} = \mathcal{V} \oplus \mathcal{H}$. Namely, given a normal frame $\hat{V} = (V_1, \ldots, V_m)$ in $\mathcal{V}$, we define g by declaring that the vector fields $V_1, \ldots, V_m$, $[X, V_1], \ldots, [X, V_m]$ are orthonormal. Then, for arbitrary vector fields in $\mathcal{D}$,

$$Y = \sum \phi^i V_i + \sum \widetilde{\phi}^i [X, V_i], \quad Z = \sum \psi^i V_i + \sum \widetilde{\psi}^i [X, V_i]$$

(all defined along c or locally) we have

$$g(Y,Z) = \sum_i (\phi^i \psi^i + \widetilde{\phi}^i \widetilde{\psi}^i).$$

By Proposition 2.12 we also have $D_X[X, V_i] = 0$, thus the invariance formula (2.9) holds for all $Y, Z \in \mathcal{D}$. Note that, since the normal frame along c is uniquely defined by its initial value at $c(0)$, by Proposition 2.6, the invariant metric along c is unique up to a linear transformation in $\mathcal{D}(c(0))$. Thus we have proved

Proposition 2.16. *Suppose we are given a regular dynamic pair $(X, \mathcal{V})$, an integral curve $c : [0,T] \to M$ of X, and a frame F in $\mathcal{V}(c(0))$. Then the following holds.*

(i) There is a unique invariant positive definite (respectively, of prescribed signature) metric g on $\mathcal{D}$ along c such that F is orthonormal with respect to g restricted to $\mathcal{V}(c(0))$. The normal frame which coincides with F at $c(0)$ is orthonormal with respect to g along c.

(ii) $\mathcal{V}$ and $\mathcal{H}$ are orthogonal with respect to g along c.

(iii) $g(AV, AW) = g(V, W)$ for all $V, W \in \mathcal{V}$ along c, where $A\colon \mathcal{V} \to \mathcal{H}$ is defined in Remark 2.14.

Given a metric g on $\mathcal{V}$, possibly not positive definite, we can define a *directional curvature* of a regular dynamic pair $(X, \mathcal{V})$, using its curvature operator K.

Definition 2.17. The *curvature in direction* $v \in \mathcal{V}$ of the triple $(X, \mathcal{V}, g)$ is

$$k(v) = \frac{g(Kv, v)}{g(v, v)}, \quad \text{for} \quad v \in \mathcal{V} \quad \text{such that} \quad g(v, v) \neq 0.$$

3. Conjugate points and Jacobi vector fields

In this section we introduce a definition of conjugate points for dynamic pairs and state our main results on existence and location of conjugate points in terms of the curvature operator of the dynamic pair. For proving the results we introduce Jacobi vector fields generalized to the current context and establish correspondence between conjugate points and properties of Jacobi fields (Sect. 3.2). The proofs follow in Section 3.4.

3.1. Conjugate points

Let $(X, \mathcal{V})$ be a regular dynamic pair. Consider an integral curve $c\colon [0, T] \to M$ of X, $c(0) = x_0$, which is not a closed orbit. Regularity condition (R2) implies that the space $\mathcal{V}(x_0)$ transported by the flow of X along the trajectory $c(t)$ is transversal to the space $\mathcal{V}(c(t))$ for t small enough. Points $c(t)$ where this is not true will be called conjugate to $c(0)$, analogously as in second order ODEs and Calculus of Variations.

Definition 3.1. Given c, a point $0 < t^* \leq T$ is called *conjugate time* and the corresponding point $c(t^*)$ *conjugate point* to x_0 on c if

$$\exp(t^* X)_* \mathcal{V}(x_0) \cap \mathcal{V}(c(t^*)) \neq \{0\}.$$

The dimension of the above intersection is called *the multiplicity* of t^* or $c(t^*)$.

The intuition behind the definition is as follows. Consider a smooth family $c_s\colon [0, T] \to M$ of integral curves of X, $s \in (-\epsilon, \epsilon)$, uniquely defined by the curve of initial points $\gamma(s) = c_s(0)$. Then $t^* > 0$ is conjugate time if there exists a curve of initial points $\gamma(s), \gamma(0) = x_0$, which for all s is tangent to $\mathcal{V}$ such that the curve of endpoints $s \to c_s(t^*)$ is tangent to the subspace $\mathcal{V}(c_0(t^*))$, *i.e.* $\frac{\mathrm{d}}{\mathrm{d}s} c_s|_{s=0}(t^*) \in \mathcal{V}(c_0(t^*))$. Informally, there exists a vector $v \in \mathcal{V}(c(0))$ such that an infinitesimally small perturbation of the initial condition $c(0)$ in the direction of v causes the perturbation of the trajectory $c(t)$ of X which, after time t^*, is infinitesimally small in a direction $w \in \mathcal{V}(c(t^*))$.

If $\mathcal{V}$ is the control distribution $\mathrm{span}\{Y_1, \ldots, Y_m\}$, defined by a control affine system (1.1), then this means that there is a control value u such that if a small control impulse in direction u is applied at time $t = 0$ then the perturbed trajectory arrives at time t^* to a point infinitesimally close to the reference trajectory in a direction which belongs to $\mathcal{V}(c(t^*))$, thus another impulse of control can "remove" the effect of the impulse at $t = 0$, placing the perturbed trajectory back to $c(t)$.

Remark 3.2. If distribution $\mathcal{V}$ is integrable then one can consider locally defined quotient manifold $N = M/\mathcal{V}$ that comes with a natural projection $\pi\colon M \to N$. Then any integral curve $t \mapsto c(t)$ of X defines a curve in N *via* projection $\tilde{c}(t) = \pi(c(t))$. If a family of curves $c_s(t)$, $s \in (-\epsilon, \epsilon)$, satisfies $\frac{\mathrm{d}}{\mathrm{d}s} c_s(0) \in \mathcal{V}$ for all s then $\tilde{c}_s(0) = \tilde{c}_0(0)$ for all s and the curve $s \to c_s(0)$ defines a point in N. Hence we arrive to the usual definition 'of conjugate points in this case.

Now we can formulate two results that generalize the classical Cartan–Hadamard and Bonnet–Myers theorems on conjugate points on Riemannian and Finslerian manifolds.

Theorem 3.3. *(i) Given a regular dynamic pair $(\mathcal{V}, X)$ and an invariant metric g, if there exists $\lambda \in \mathbb{R}$ such that the directional curvature satisfies $k(v) \leq \lambda$ on c or, more explicitely,*

$$g(Kv, v) \leq \lambda g(v, v), \quad \text{for} \quad v \in \mathcal{V}(c(t)), \quad t \in [0, T], \tag{3.1}$$

then there are no conjugate times in $(0, T]$, if $\lambda \leq 0$, and there are no conjugate times in the interval $(0, t_c)$, $t_c = \min\{T, \frac{\pi}{\sqrt{\lambda}}\}$, if $\lambda > 0$.

(ii) More generally, if the estimate (3.1) holds with the constant λ replaced by a continuous function $\lambda(t)$, $t \in [0, T]$, then there is no conjugate time smaller then the first zero in $(0, T]$ of the solution $\psi(t)$ of the Cauchy problem

$$\ddot{\psi} = -\lambda(t)\psi, \quad \psi(0) = 0, \quad \dot{\psi}(0) = 1. \tag{3.2}$$

In particular, there is no conjugate time in $(0, T]$ if $\psi(t) \neq 0$ for $t \in (0, T]$.

Remark 3.4. Note that the estimate (3.1) in statement (i) can be replaced by the equivalent estimate $\lambda^K(t) \leq \lambda$, $t \in [0, T]$, where $\lambda^K(t)$ denotes the maximal eigenvalue of the symmetric (with respect to the scalar product $g(t)$) part of $K(t)$. Similarly one can take $\lambda(t) = \lambda^K(t)$ in statement (ii). If the curvature $K(t)$ is diagonalizable and the corresponding eigenspaces are parallel with respect to D_X, then the invariant metric g can be chosen such that the eigenspaces are orthogonal with respect to g (compare Rem. 3.9 below). In this case $K(t)$ is symmetric and statement (ii) of Theorem 3.3 gives an optimal estimate.

This is not the case for a general invariant g, as shown in the example below. Nevertheless, one can search for the optimal estimate by minimizing λ satisfying (3.1) over the set of invariant metrics g (recall that invariant metrics along an integral curve of X form a finite-dimensional family as they are defined by a choice of an orthonormal frame of $\mathcal{V}$ at a fixed point, cf. Prop. 2.16). An example of such search is also given in the example.

Example 3.5. Consider a system given by the second order equations

$$\begin{pmatrix} \ddot{x}_1 \\ \ddot{x}_2 \end{pmatrix} = -\begin{pmatrix} 0 & 1 \\ 0 & \alpha \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Then, by (2.8), $K = \begin{pmatrix} 0 & 1 \\ 0 & \alpha \end{pmatrix}$ and a normal frame is given by vector fields ∂_{y_1}, ∂_{y_2}, where y_i is a coordinate on the space of 1-jets corresponding to $\dot{x}_i$. Taking positive definite g of the diagonal form

$$g = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \qquad a, b > 0,$$

one gets for $v = v_1\partial_{y_1} + v_2\partial_{y_2}$ the estimate

$$k(v) = \frac{g(Kv, v)}{g(v, v)} = \frac{av_1v_2 + b\alpha(v_2)^2}{a(v_1)^2 + b(v_2)^2} \leq \frac{1}{2}\sqrt{\frac{a}{b}} + \alpha.$$

The estimate is not optimal and can be improved by suitable choosing the parameters of g. Namely, taking $a > 0$ arbitrarily small we see that we can choose an invariant metric so that any $\lambda > \alpha$ can be placed in the estimate (3.1) in Theorem 3.3. This implies that there are no conjugate points on an interval of any prescribed length, if $\alpha = 0$, and there are no conjugate points in the interval $(0, \frac{\pi}{\sqrt{\alpha}})$ for $\alpha > 0$ (and this estimate is optimal).

To state the second result consider a non-closed integral curve $c : [0, T) \to M$ of X, with $T \in (0, \infty]$, and let $(X, \mathcal{V})$ be a regular dynamic pair defined in a neighbourhood of c.

Theorem 3.6. *If, along c, the curvature K is symmetric with respect to some, parallel with respect to D_X, positive definite metric g and*

$$\operatorname{tr} K \geq \kappa > 0 \quad \textit{along } c$$

then there is a conjugate point in the interval $(0, T^*]$, $T^* = \pi\sqrt{\frac{m}{\kappa}}$, provided that $T^* < T$.

Example 3.7. We will demonstrate that the symmetry of K in the above theorem is crucial. For this we will identify points (x, y) in $\mathbb{R}^2$ with complex numbers $z = x + iy \in \mathbb{C}$. Let $\lambda = a + ib$, where $a, b \in \mathbb{R}$ are given constants. Consider the second order equation

$$\ddot{z} = \lambda^2 z.$$

One can easily verify that the curve in $\mathbb{C}$ defined by $z^*(t) = A(t) + iB(t)$,

$$A(t) = \sinh(at)\cos(bt), \quad B(t) = \cosh(at)\sin(bt),$$

satisfies the equation and $z^*(0) = 0$. Explicitly, the equation reads

$$\begin{pmatrix} \ddot{x} \\ \ddot{y} \end{pmatrix} = \begin{pmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and we see that this is a second order system in normal form with the curvature matrix

$$K = -\begin{pmatrix} a^2 - b^2 & -2ab \\ 2ab & a^2 - b^2 \end{pmatrix}.$$

If (v, w) denote coordinates associated to (x, y) in the "tangent bundle" $T\mathbb{R}^2$, one can take the constant basis ∂_v, ∂_w as a normal basis of the vertical distribution. Then the usual Euclidean metric is a invariant metric. For $a = 0$ and $b > 0$ the curvature matrix is the identity matrix and the equations take the form $\ddot{x} = -x$, $\ddot{y} = -y$. Those are at the same time the Jacobi equations of the system. The double conjugate time is $t^* = \pi$.

When we perturbe such system by taking $b > 0$ and $a \neq 0$, $|a| < b$ then still $\operatorname{tr} K > 0$ but K is not symmetric. The theorem does not guaratee the existence of conjugate points and, in fact, there are no such points. Namely, the solution $z^*(t)$ is nonzero for $t > 0$ as both $A(t)$ and $B(t)$ do not have such common zeros. Moreover, a general nontrivial solution $z(t)$ satisfying $z(0) = 0$ is of the form $z(t) = z^*(t)w$, with $w \in \mathbb{C} \setminus \{0\}$, thus it is also nonvanishing fot $t > 0$.

The following result does use any metric and, thus, it does not require symmetry of K.

Theorem 3.8. *(i) If the curvature matrix $K_t = K(c(t))$, $t \in [0, T]$, has an invariant one-dimensional subspace $\mathcal{Y} \subset \mathcal{V}$ which is parallel with respect to D_X along c, i.e., $D_X V \in \mathcal{Y}$ for any section V of $\mathcal{Y}$, and the corresponding eigenvalue satisfies*

$$\lambda(t) \geq \kappa > 0, \quad t \in (0, T], \tag{3.3}$$

(equivalently, the normal curvature matrix K_t on c has a one-dimensional diagonal block with the eigenvalue satisfying (3.3)) then there is a conjugate time in $(0, \pi/\sqrt{\kappa}\,]$, provided that $\pi/\sqrt{\kappa} \leq T$.

(ii) Moreover, if there are r such subspaces with the eigenvalues $\lambda_1(t),\dots,\lambda_r(t)$ then the set of conjugate points along c of the pair $(X,\mathcal{V})$ includes all zeros on $(0,T]$ of solutions of the equations

$$\ddot{y}_i = -\lambda_i(t)y_i, \tag{3.4}$$

$i=1,\dots,r$, with initial conditions $y_i(0)=0$, $\dot{y}_i(0)\neq 0$ (if t^* is a common zero of k of these equations then it is conjugate point of multiplicity k or larger).

Remark 3.9. The assumption on eigenspaces being parallel with respect to D_X is equivalent to the fact that the normal curvature matrix is diagonal in a normal frame. This follows from the following property of one-dimensional subspaces $\mathcal{Y}$ of $\mathcal{V}$ defined along c.

$\mathcal{Y}$ is parallel with respect to D_X if and only if it has a normal generator.

The property follows from Proposition 2.12. Namely, if V is a normal generator of $\mathcal{Y}$ then $D_XV=0$ and then arbitrary $Y=fV$ satisfies $D_X(Y)=X(f)V\in\mathcal{Y}$. Vice versa, if Y is a generator of $\mathcal{Y}\subset\mathcal{V}$ which is invariant under D_X then $D_XY=fY$ for some function f. Then, taking the generator $V=gY$ of $\mathcal{Y}$ and requiring that $X(g)+fg=0$, $g(0)=1$, gives that $D_XV=X(g)Y+gD_XY=(X(g)+fg)Y=0$ along c, thus V is normal.

Example 3.10. As an example consider a linear system of ODEs of the form

$$\ddot{x}=A(t)x,$$

where $x\in\mathbb{R}^m$ and A is a time dependent $m\times m$ matrix. Then, by (2.8), $K=-A$ and a normal frame is given by vector fields ∂_{y_i}, $i=1,\dots,m$, where (t,x_i,y_i) are standard coordinates on $J^1(\mathbb{R},\mathbb{R}^m)$, with y_i corresponding to $\dot{x}_i$. Consequently, any other normal basis is obtained by multiplying $(\partial_{y_1},\dots,\partial_{y_m})$ by a matrix B with constant coefficients. Therefore, if there is a normal section $t\mapsto V(t)$ of distribution $\mathcal{V}$ such that $V(t)$ is an eigenvector of the matrix $A(t)$, then $V(t)$ is of the form $V(t)=\sum_i^m v_i\partial_{y_i}$ where v_i are constant coefficients. It follows that the matrices $A(t)$ have the same eigenspace spanned by $V(0)$, regardless of t. Therefore, if $A(t)$ is diagonalizable by normal fields then, the vector fields are constant in t and the matrix $\frac{\mathrm{d}}{\mathrm{d}t}A(t)$ has necessarily the same eigenspaces as $A(t)$. Later, in Example 5.4 we will present 2-dimensional non-linear systems with diagonizable K such that the eigenspaces are parallel with respect to D_X.

3.2. Jacobi vector fields

Consider a regular dynamic pair $(X,\mathcal{V})$ on M. In order to prove main results on conjugate points we extend to our context the classical notion of the Jacobi field in $\mathcal{V}$ and define its counterpart in $\mathcal{D}=\mathcal{V}+[X,\mathcal{V}]$, called here the symmetry vector field.

Definition 3.11. A vector field $J\in\mathcal{V}$ is a *Jacobi field* of $(X,\mathcal{V})$ if

$$D_XD_XJ+KJ=0. \tag{3.5}$$

Definition 3.12. A vector field $S\in\mathcal{D}$ is a *symmetry vector field* of X if

$$[X,S]=0. \tag{3.6}$$

Both definitions make sense on an open subset of M, as well as on a single integral curve $c:[0,T]\to M$ of X. In the latter case we replace condition $[X,S]=0$ by

$$\exp(tX)_*S_0=S_t,\quad t\in[0,T], \tag{3.7}$$

where $S_t=S(c(t))$. Equation (3.5) is called the *Jacobi equation.*

Condition (3.6) implies that the flow $\exp(tX)$ of X transports trajectories of the symmetry field S onto its trajectories, preserving time parametrization. In particular, given an integral curve $\gamma(s)$ of S, the curve $\gamma_t(s) = \exp(tX)(\gamma(s))$ is again a trajectory of S. In studying conjugate points of $(X, \mathcal{V})$ we are particularly interested in curves γ passing through $x_0 = \gamma(0)$ which are tangent to $\mathcal{V}$ at x_0 and transporting them by the flow $\exp(tX)$.

Symmetry vector fields of X play the role of "total Jacobi fields" and are in one-to-one correspondence with Jacobi fields of $(X, \mathcal{V})$. To show this we use the vector bundle isomorphism $A : \mathcal{V} \to \mathcal{H}$, $AV = \pi_{\mathcal{H}}[X, V]$ (Rem. 2.14).

Proposition 3.13. *If $J \in \mathcal{V}$ is a Jacobi field of $(X, \mathcal{V})$ then the vector field*

$$S_J = -D_X J + AJ \tag{3.8}$$

is a symmetry field of X. Vice versa, if S is a symmetry of X then the vector field

$$J_S = A^{-1}\pi_{\mathcal{H}} S \tag{3.9}$$

is a Jacobi field of $(X, \mathcal{V})$.

Proof. Let J be a Jacobi field. Locally, we may write $J = \sum f^i V_i$, where $V_1, \ldots, V_m$ are normal generators of $\mathcal{V}$. Then $D_X V_i = 0$ and $D_X J = \sum X(f^i)V_i$. By the definition of A and the fact that $[X, V_i] \in \mathcal{H}$ (definition of $\mathcal{H}$) we have $AJ = \sum f^i[X, V_i]$. Thus

$$\begin{aligned}[X, S_J] &= -[X, \sum X(f^i)V_i] + [X, \sum f^i[X, V_i]] \\ &= -\sum X^2(f^i)V_i + \sum f^i[X,[X,V_i]] = -D_X D_X J - KJ = 0\end{aligned}$$

since $D_X D_X(\sum f^i V_i) = \sum X^2(f^i)V_i$ and $K(\sum f^i V_i) = \sum_{i,j} f^i K_i^j V_j = -\sum_i f^i[X,[X,V_i]]$.

In order to prove the second assertion let $V_1, \ldots, V_m$ be a local normal frame of $\mathcal{V}$. Given a symmetry field S, we may write

$$S = \sum_i g^i V_i + \sum_i f^i [X, V_i].$$

Then

$$J_S = A^{-1}\pi_{\mathcal{H}} S = A^{-1}\left(\sum f^i[X, V_i]\right) = \sum f^i V_i.$$

We will show that $J_S = \sum f^i V_i$ is a Jacobi field. The assumption $[X, S] = 0$ gives

$$0 = \sum (X(g^i)V_i + f^i[X,[X,V_i]]) + \sum (g^i + X(f^i))[X, V_i]$$

and, taking into account $[X,[X,V_i]] = \sum_j -K_i^j V_j$, we see that $[X, S] = 0$ is equivalent to

$$g^i + X(f^i) = 0, \qquad X(g^i) - \sum_j K_j^i f^j = 0. \tag{3.10}$$

Thus $X^2(f^i) = -\sum_j K^i_j f^j$ and we find that

$$D_X^2 \left(\sum f^i V_i\right) = \sum X^2(f^i)V_i = -\sum_{i,j} f^j K^i_j V_i = -\sum f^j K V_j = -K\left(\sum f^j V_j\right),$$

where K is the curvature operator defined in the normal basis $V_1, \dots, V_m$ by the matrix (K^i_j). We conclude that $D_X^2 J_S = -K J_S$ which means that J_S is a Jacobi field. □

Clearly, existence of a vector field S in a neighbourhood of c which satisfies condition (3.6) implies existence of a vector field S_t along c satisfying (3.7), and vice versa. According to its definition a point $t^* \in (0, T]$ is a conjugate time if there exists a nonzero $v = S_0 \in \mathcal{V}(x_0)$ such that $\exp(t^*X)_* S_0 \in \mathcal{V}(c(t^*))$ or, equivalently, there exists a nonzero symmetry field S_t along c such that

$$S_0 \in \mathcal{V}(c(0)) \quad \text{and} \quad S_{t^*} \in \mathcal{V}(c(t^*)). \tag{3.11}$$

Using Proposition 3.13 we see that existence of a symmetry field S_t along c is equivalent to existence of a Jacobi field J_t where $J_t = A^{-1}\pi_{\mathcal{H}} S_t$, in which case the above boundary conditions are equivalent to $J_0 = 0$ and $J_{t^*} = 0$. Additionally, the dimensions of the space of such symmetry vector fields and the corresponding Jacobi vector fields coincide. Thus we obtain

Proposition 3.14. *A point $t^* \in (0, T]$ is a conjugate time along an integral curve $c\colon [0, T] \to M$ of X iff there exists a nontrivial Jacobi field $J \in \mathcal{V}$ along c such that*

$$J(c(0)) = 0 \quad \text{and} \quad J(c(t^*)) = 0. \tag{3.12}$$

The multiplicity of a conjugate time equals to the number of linearly independent Jacobi fields satisfying (3.12).

3.3. Jacobi frames and Riccati equation

Existence and location of conjugate points can also be characterized using special Jacobi frames and a corresponding Riccati equation assigned to the regular dynamic pair $(X, \mathcal{V})$.

Definition 3.15. A tuple of Jacobi fields $\hat{J} = (J_1, \dots, J_m)$ defined along a trajectory $c : [0, T] \to M$ of X is a *Jacobi frame* on $(0, \tau)$, $\tau \le T$, if they are linearly independent at each point $c(t)$, $t \in (0, \tau)$, and a *special Jacobi frame* on $(0, \tau)$ if in addition they satisfy $J_i(0) = 0$.

Special Jacobi frames are defined by solutions of the Jacobi equation satisfying the initial conditions

$$J_i(0) = 0, \quad D_X J_i(0) = v_i, \quad i = 1, \dots, m, \tag{3.13}$$

where $v_1, \dots, v_m$ is a frame in $\mathcal{V}(c(0))$. It is easy to see that initial conditions (3.13) imposed on the solutions J_i of (3.6) imply that the tuple $\hat{J} = (J_1, \dots, J_m)$ evaluated at $c(t)$ is a frame in $\mathcal{V}(c(t))$ for t small enough. Vice versa, if $v_1, \dots, v_n$ were linearly dependent then also the corresponding Jacobi fields would be linearly dependent, with the same constants, thus they would not form a Jacobi frame. We will show that the first point where $\hat{J}$ ceases to be a frame is a conjugate point.

Consider an interval $(0, \tau)$ where the Jacobi fields $J_1, \dots, J_m$ defined by (3.13) form a frame along c. Then we may write

$$D_X J_i = \sum_j G^j_i J_j \tag{3.14}$$

where the matrix $G = (G_i^j)$ is well defined on $(0, \tau)$. The matrix G satisfies a Riccati differential equation. Namely, differentating the above equality with respect to t we get

$$D_X D_X J_i = \sum_j \dot{G}_i^j J_j + \sum_j G_i^j D_X J_j = \sum_j \dot{G}_i^j J_j + \sum_{j,k} G_i^j G_j^k J_k.$$

We can write $D_X D_X J_i = -\sum K_i^j J_j$, where (K_i^j) is the matrix of the curvature written in the Jacobi frame $\hat{J}$. Then the above equality becomes

$$\sum_j \left(\dot{G}_i^j + \sum_k G_i^k G_k^j + K_i^j \right) J_j = 0, \quad i = 1, \ldots, m,$$

and we deduce that the matrix G satisfies the *Riccati equation*

$$\dot{G} + G^2 + K = 0 \tag{3.15}$$

in the interval $(0, \tau)$ where $\hat{J}$ is a frame. Moreover, the formula (3.14) defining G can be written in the form $D_X \hat{J} = \hat{J} G$ (recall that we treat $\hat{V}$ and $D_X \hat{V}$ as row vectors of sections of $\mathcal{V}$). Thust we have $\hat{J} = D_X \hat{J} G^{-1}$ and, by the initial conditions (3.13),

$$\lim_{t \to 0^+} G^{-1}(t) = 0. \tag{3.16}$$

Proposition 3.16. *Given a trajectory $c\colon [0,T] \to M$ of X and a point $\tau \in (0, T]$, the following conditions are equivalent.*

(i) There is no nontrivial Jacobi field J on $[0, T]$ satisfying (3.12) for some positive $t^* \le \tau$.

(ii) There is a special Jacobi frame $\hat{J}$ on $(0, \tau)$.

Moreover, if (ii) holds and $\hat{J}$ is any special Jacobi frame on $(0, \tau)$ then it holds:

(iii) The matrix G defined by (3.14) satisfies the Riccati equation (3.15) on $(0, \tau)$ and (3.16).

Proof. Equivalence of both conditions follows from the fact that any Jacobi vector field J along c which satisfies $J(0) = 0$ is a linear combination $J = \sum c_i J_i$ where c_i are determined by the initial condition $D_X J(0) = \sum c_i v_i$. The last statement follows from the above calculation showing that the matrix defined by (3.14) satsfies (3.15). □

Note that statement (iii) is independent of the choice of the special Jacobi frame since such frame is determined by the “initial velocity” $(v_1, \ldots, v_m)$ in condition (3.13). A linear change of this initial velocity frame tranforms the Jacobi frame $\hat{J}$ and the matrices K and G with the same invertible matrix with constant coefficients.

Corollary 3.17. *Any of the three conditions in the proposition implies that the pair $(X, \mathcal{V})$ does not have a conjugate point in $(0, T)$ on c.*

3.4. Proofs

With all earlier preparations, the proofs can be reduced to classical arguments. To prove Theorem 3.6, we will need the index functional introduced below by formula (3.19). Throughout this section, we will use a normal frame in $\mathcal{V}$ along c and the corresponding invariant metric g on $\mathcal{V}$ making the frame orthonormal (see the beginning of Section 2.5). The curvature matrix $K(t)$, called normal, is taken in such frame.

Proof of Theorem 3.3. We first prove statement (ii). Consider a Jacobi field $J \in \mathcal{V}$ along $c\colon [0,T] \to M$, with $J(c(0)) = 0$, $(D_X J)(c(0)) = v \neq 0$. By Proposition 3.14 it is enough to estimate the first zero of $J_t = J(c(t))$ in $(0,T]$. This is equivalent to estimating the first zero of the function

$$\varphi = \sqrt{g(J,J)}$$

on $c(t)$, $t \in (0,T]$. Differentiating and using the invariance of the metric we get

$$\dot{\varphi} = \frac{g(D_X J, J)}{\varphi},$$

and

$$\ddot{\varphi} = \frac{g(D_X D_X J, J)}{\varphi} + \frac{g(D_X J, D_X J)}{\varphi} - \frac{g^2(D_X J, J)}{\varphi^3}.$$

The Cauchy–Schwartz inequality $g^2(D_X J, J) \leq g(D_X J, D_X J)\, g(J,J)$ and $g(J,J) = \varphi^2$ imply that the last two terms in the above expression give a nonnegative number. This and the Jacobi equation for J imply that

$$\ddot{\varphi} \geq -\frac{g(KJ,J)}{\varphi} = -\frac{k(J)g(J,J)}{\varphi} = -k(J)\varphi$$

with $k(J)$ the sectional cuvature. From the assumption (3.1) we have $k(J_t) \leq \lambda(t)$ and we finally get

$$\ddot{\varphi} \geq -\lambda(t)\varphi. \tag{3.17}$$

Note that the initial conditions on J imply that $\varphi(0) = 0$ and $\dot{\varphi}(0) = \sqrt{g(v,v)} > 0$ (since, for $\psi = \varphi^2$, we have $\dot{\psi} = 2\dot{\varphi}\varphi = 2g(D_X J, J)$, thus $\psi(0) = \dot{\psi}(0) = 0$, and $\ddot{\psi} = g(D_X^2 J, J) + g(D_X J, D_X J)$, thus $\ddot{\psi}(0) = g(v,v)$). It follows that φ is positive on some interval $(0, t_z)$, where t_z is taken the first zero of φ in $(0,T]$, or $t_z = T$ if φ does not have such a zero. Now we can write, instead of (3.17), the equality

$$\ddot{\varphi} = a(t)\varphi \tag{3.18}$$

on $(0, t_z)$, where $a(t) = \ddot{\varphi}/\varphi \geq -\lambda(t)$. Applying the classical Sturm comparison theorem (see [22]) we find out that the first zero t^* of φ is larger or equal to the first zero of ψ solving the equation $\ddot{\psi} = -\lambda(t)\psi$ with $\psi(0) = 0$, $\dot{\psi}(0) = \dot{\varphi}(0) > 0$. This proves statement (ii).

Statement (i) follows from (ii). Namely, the first zero of the solution $\psi(t)$ of the Cauchy problem (3.2) is $t = \pi/\sqrt{\lambda}$, if $\lambda > 0$ (then $\psi(t) = a/\sqrt{\lambda}\sin(\sqrt{\lambda}t)$), or it does not exist if $\lambda \leq 0$. □

For the proof of Theorem 3.6 we need a lemma. Let g be a positive definite metric on $\mathcal{V}$ defined along c and parallel with respect to D_X. Given $0 < r < T$, we introduce the *index functional* along c defined on sections V of $\mathcal{V}$ along $\bar{c} = c|_{[0,r]}$ by

$$I(V,V) = \int_0^r (g(D_X V, D_X V) - g(KV,V))\mathrm{d}t. \tag{3.19}$$

Lemma 3.18. *Let K be symmetric with respect to g on c and let $0 < r < T$. Then there are no conjugate points in $(0,r]$ if and only if for any section V of $\mathcal{V}$ along c such that $V(c(0)) = 0 = V(c(r))$ we have*

$$I(V,V) \geq 0 \tag{3.20}$$

and equality holds only if $V \equiv 0$ on $[0,r]$.

The proof will follow from Lemma 3.20.

Remark 3.19. In the case of symmetric curvature operators an alternative approach to conjugate points often exploits analysis of the associated Riccati equation, see Section 3.3. This viewpoint is adapted in [7] where estimates for the blow-up times of the Riccati equation (see [23]) are crucial components of the proofs.

Proof of Theorem 3.6. Assume that there are no conjugate points in the interval $(0, T^*]$. We will show that this leads to contradiction with the definition of T^*. Consider the interval $I = [0, r]$ where $T^* < r < T$. Then there there are no such points in $(0, r]$ for some $r > T^*$ close enough to T. This follows from the fact that r is a conjugate time if and only if the m-tuple $(J_1, \ldots, J_m)$ of Jacobi fields satisfying the initial conditions $J_i(0) = 0$ and $D_X J_i(0) = e_i$, with $(e_1, \ldots, e_m)$ a basis in $\mathcal{V}(c(0))$, is linearly dependent at $c(r)$. Linear independence of the tuple at $c(T^*)$ implies their independence at $c(r)$ for r close o T^* which implies the above statement.

Let $(V_1, \ldots, V_m)$ be a normal frame along c defined by the initial conditions $V_i(c(0)) = e_i$. Assume that the basis $(e_1, \ldots, e_m)$ of $\mathcal{V}(c(0))$ is orthonormal with respect to g. Then the normal frame is orthonormal along c, by invariance of g and V_i with respect to the flow of X. Let φ be a smooth function on $[0, r]$ not vanishing everywhere and such that $\varphi(0) = \varphi(r) = 0$. Define sections $W_i = \varphi V_i$ of $\mathcal{V}$ along c. Then by Lemma 3.18 we have $I(W_i, W_i) > 0$, *i.e.*,

$$\int_0^r ((\dot\varphi)^2 g(V_i, V_i) - \varphi^2 g(KV_i, V_i))\mathrm{d}t = \int_0^r ((\dot\varphi)^2 - \varphi^2 g(KV_i, V_i))\mathrm{d}t > 0.$$

Summing up with respect to i we obtain

$$\int_0^r (m(\dot\varphi)^2 - \mathrm{tr}(K)\varphi^2)\mathrm{d}t > 0.$$

Since φ was any nontrivial smooth function vanishing at 0 and r we can take $\varphi = \sin(at)$, where $a = \pi/r$. The above inequality and $\operatorname{tr} K \geq \kappa$ imply

$$\int_0^r ma^2 \cos^2(at)\mathrm{d}t - \int_0^r \kappa \sin^2(at))\mathrm{d}t > 0.$$

Taking into account that the integrals of $\cos^2(aT)$ and $\sin^2(at)$ over the period $[0, r]$ are equal we deduce that $ma^2 > \kappa$. This implies that $r < \pi\sqrt{m/\kappa} = T^*$ and contradicts our supposition that $r > T^*$. This contradiction implies the assertion of the theorem. □

The next lemma, proved in the Appendix A, can also be deduced from results in Chapter 11 of [24].

Lemma 3.20. *Let (P, Q) be $m \times m$ matrices defined on the interval $[0, T)$ (possibly infinite) which satisfy the differential equations*

$$\dot P = Q, \quad \dot Q = -KP,$$

and the initial conditions $P(0) = 0$, $Q(0) = Id$, where K is a symmetric matrix depending continuously on t. If P is invertible on the interval $(0, r]$, for some $r < T$, then for any C^1-smooth $w\colon [0, r] \to \mathbb{R}^m$ satisfying $w(0) = 0 = w(r)$ we have

$$\int_0^r (\langle \dot w, \dot w\rangle - \langle Kw, w\rangle)\mathrm{d}t \geq 0. \tag{3.21}$$

Equality takes place iff $w \equiv 0$ (the angle brackets denote standard scalar product in $\mathbb{R}^m$).

Proof of Lemma 3.18. Proposition 3.14 allows us to reduce the lemma to Lemma 3.20 by the following argument. Let $\hat{J} = (J_1, \ldots, J_m)$ be a tuple of Jacobi fields satisfying initial conditions $J_i(0) = 0$ and $D_X J_i(0) = e_i$, where $(e_1, \ldots, e_m)$ is a basis in $\mathcal{V}(c(0))$. Let $(V_1, \ldots, V_m)$ be a normal frame along c orthonormal with respect to g which exists by Proposition 2.16. Expressing the Jacobi fields and the Jacobi frame in the normal frame used as a basis at $c(t)$ for all t we can replace $\hat{J}(c(t))$ by its matrix, denoted $P(t)$. Then $D_X\hat{J}(c(t))$ corresponds to $\dot{P}$ and $D_X^2\hat{J}(c(t))$ corresponds to $\ddot{P}(t)$. Moreover, the Jacobi equation $D_X^2\hat{J} + K\hat{J} = 0$ is equivalent to the equation $\ddot{P} + KP = 0$ when the curvature operator is replaced with its matrix in the normal frame denoted again by K. This equation can be replaced by the system

$$\dot{P} = Q, \quad \dot{Q} = -KP.$$

By Proposition 3.14 a point $t > 0$ is a conjugate time for the dynamic pair $(X, \mathcal{V})$ iff there is a nontrivial Jacobi field J such that $J(c(0)) = 0$ and $J(c(t)) = 0$. This is equivalent to the fact that the above Jacobi tuple $\hat{J}$ is linearly dependent at $c(t)$ and, in consequence, the matrix P is not invertible at t. In this way we see that the assumptions of the lemma imply the assumptions of Lemma 3.20. Namely, nonexistence of conjugate points in $[0, r]$ for the dynamic pair implies that the matrix P defined here is invertible on $[0, r]$. Using Lemma 3.20 we conclude that the integral functional in the lemma has property (3.21).

To conclude we should show that the property (3.21) in Lemma 3.20 implies the similar statement (3.20) on the functional (3.19) in Lemma 3.20. Writing $V(c(t)) = \sum_i w_i(t)V_i(c(t))$ and defining $w = (w_1, \ldots, w_m)$ we have $w(0) = 0 = w(r)$, as $V(0) = 0 = V(r)$. Since the normal frame was chosen orthonormal with respect to g we have

$$I(V, V) = \int_0^r (g(D_X V, D_X V) - g(KV, V))\mathrm{d}t = \int_0^r (\langle \dot{w}, \dot{w}\rangle - \langle Kw, w\rangle)\mathrm{d}t$$

and the inequality (3.21) in Lemma 3.20 implies the inequality (3.20) in Lemma 3.18. This completes the proof. □

Proof of Theorem 3.8. We first prove statement (ii). Note that existence along c of a one dimensional subspace of $\mathcal{V}$ parallel with respect to D_X means that there exists a nonvanishing section $V \in \mathcal{V}$ along c which is normal, *i.e.*, satisfies $D_X V = 0$. Existence of r such subspaces means that there are r normal, linearly independent along c sections of $\mathcal{V}$. These sections can be completed to a normal frame in $\mathcal{V}$ along c since such a frame is uniquely defined by a choice of a frame at the initial point $c(0)$ of c, by Proposition 2.6. The normal curvature matrix K_t along c has r one dimensional diagonal blocks with the eigenvalues $\lambda_i(t)$. The Jacobi equation in Definition 3.8 written in this frame has r single second order equations of the form (3.4) and an independent system of $m - r$ equations. Taking zero initial conditions for the values of all m variables and for all their derivatives at $t = 0$ with the exception of $\dot{y}_i(0) \neq 0$ for a single $1 \leq i \leq k$ we obtain r linearly independent (over $\mathbb{R}$) solutions of the Jacobi equation. If k of these solutions vanish at some $t^* \in (0, T]$ then Proposition 3.14 implies that t^* is a conjugate time of multiplicity at least k. This proves statement (ii).

Statement (i) follows from statement (ii) and the Sturm comparison theorem. Namely, the theorem implies that if $\lambda(t)$ and the constant κ satisfy (3.3) then a nontrivial solution on $[0, T]$ of the scalar equation $\ddot{y} = \lambda(t)y$ satisfying $y(0) = 0$ has a zero between two consecutive zeros on $[0, T]$ of any nontrivial solution of the equation $\ddot{u} = \kappa u$. Taking the solution $u(t) = \sin(\sqrt{\kappa}\, t)$ we see that $u(t)$ has a zero in $(0, \pi/\sqrt{\kappa}\,]$. □

4. Semi-Hamiltonian systems

Classical results on conjugate points assume presence of a Riemannian or Finslerian metric on a manifold N and the vector field X is the geodesic spray on the tangent bundle $M = TN$ while the vertical distribution $\mathcal{V}$ is the distribution spanned by the vector fields tangent to the fibers of $TN \to N$. Another case where a canonical metric is given a priori is the case of X being a Hamiltonian vector field on the cotangent bundle T^*N and the distribution $\mathcal{V}$ is tangent to the fibers of $T^*N \to N$. In a more general setting such a metric is given if $\mathcal{V}$ is an integrable Lagrangian distribution on a symplectic manifold (M, σ), cf. [2, 25]. The metric on $\mathcal{V}$ is then given by $g(Y, Z) = \sigma([X, Y], Z)$. In our results below the closedness property of σ and integrability of $\mathcal{V}$ is not needed.

Definition 4.1. Let σ be a non-degenerate, skew-symmetric 2-form on a manifold M of even dimension. We will call a vector field X on M *semi-Hamiltonian* with respect to σ if its flow preserves σ, *i.e.* $\mathcal{L}_X \sigma = 0$, where $\mathcal{L}_X \sigma$ is the Lie derivative of σ with respect to X.

More generally, if $\mathcal{D}$ is a distribution on M of even rank and σ is a non-degenerate, skew-symmetric 2-form on $\mathcal{D}$ then X will be called *semi-Hamiltonian* (with respect to $(\mathcal{D}, \sigma)$) if $\mathcal{D}$ and σ are invariant under the flow of X:

$$[X, \mathcal{D}] \subset \mathcal{D}, \tag{I}$$

$$X(\sigma(Y, Z)) = \sigma([X, Y], Z) + \sigma(Y, [X, Z]), \quad \forall\ Y, Z \in \mathcal{D}. \tag{S}$$

Clearly, if σ is closed and $\mathcal{D} = TM$ then we get the classical definition of Hamiltonian vector field. If $\operatorname{rk} \mathcal{D} = 2m$ then we will call a rank m sub-distribution $\mathcal{V} \subset \mathcal{D}$ *Lagrangian* (with respect to σ) if the field of subspaces $\mathcal{V}(x) \subset \mathcal{D}(x)$ satisfies

$$\sigma_x(v, w) = 0, \quad \forall\ \ v, w \in \mathcal{V}(x), \quad x \in M. \tag{L}$$

Definition 4.2. A quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ where X is semi-Hamiltonian with respect to $(\mathcal{D}, \sigma)$, the sub-distribution $\mathcal{V} \subset \mathcal{D}$ is Lagrangian with respect to σ, and the regularity conditions (R1), (R2) hold, will be called *regular semi-Hamiltonian quadruple.*

Any semi-Hamiltonian quadruple defines a bilinear form on $\mathcal{V}$ by the formula

$$g(V, W) = \sigma([X, V], W), \quad V, W \in \mathcal{V}, \tag{4.1}$$

which is tensorial since $\mathcal{V}$ is Lagrangian. A regular semi-Hamiltonian quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ has the following properties (all data are assumed to be C^∞-smooth).

Proposition 4.3. *If $V_1, \ldots, V_m$ form a normal frame in $\mathcal{V}$ defined on an open subset $U \subset M$ then the following statements hold on U.*
(i) The subdistribution $\mathcal{H} = \operatorname{span}\{[X, V_1], \ldots, [X, V_m]\}$ of $\mathcal{D}$ is Lagrangian with respect to σ.
(ii) The matrix (g_{ij}) given by

$$g_{ij}(x) = \sigma([X, V_i], V_j)(x), \quad i, j = 1, \ldots, m,$$

is symmetric, nondegenerate, and defines a pseudo-Riemannian metric g on $\mathcal{V}$.
(iii) The metric g is invariant under covariant differentiation along X, i.e.,

$$X(g(Y, Z)) = g(D_X Y, Z) + g(Y, D_X Z). \tag{4.2}$$

(iv) The matrix K (Def. 2.3) is symmetric and defines a selfadjoint (relative to g) operator $K\colon \mathcal{V} \to \mathcal{V}$.

Remark 4.4. It can be deduced (*e.g.* exploiting structural equations from [2], Props. 7 and 8) that the curvature operator K used here coincides, in the Hamiltonian setting, with the curvature introduced in the work of Agrachev and coworkers, cf. [2, 4–6], once our regularity assumptions are satisfied. The approach in the above mentioned works starts from an optimality problem and goes through lifting of the system data along an extremal to a curve in a Lagrangian grassmannian. *Via* symplectic geometry arguments it is shown that the curve contains essential information on the optimality problem and the curvature is one of its invariants. So defined curvature is more geometric but difficult to compute explicitely, in general (explicit formulae given in special cases coincide with ours). The general way of defining the curvature in the regular sub-Riemannian case [7, 15] uses a construction of a normal frame which, implicitly, requires solving certain differential equations. Our formula (2.3) and more general formulae in [1], when adopted to the Hamiltonian setting of regular sub-Riemannian problems, may suggest a more direct way of computing the curvature. This requires further research which, potentially, may contribute to better understanding of dynamical properties of semi-Hamiltonian systems.

The following systems, when satisfy (R1) and (R2), provide examples of regular semi-Hamiltonian quadruples.

Example 4.5. Hamiltonian vector field on the cotangent bundle. The classical example of a Hamiltonian system on the phase space $M = T^*Q$ fits into above framework. Namely, given a function $H : T^*Q \to \mathbb{R}$ (Hamiltonian) on the cotangent bundle of a (configuration) manifold Q and the canonical symplectic structure σ on T^*Q, take $X = \vec{H}$ the Hamiltonian vector field on T^*Q defined by H and let $\mathcal{V}$ be the vertical distribution on the bundle $\pi : T^*Q \to Q$. The tuple $(\vec{H}, \mathcal{V}, \mathcal{D}, \sigma)$, where $\mathcal{D} = T(T^*Q)$, forms a regular Hamiltonian quadruple in the region where the Hamiltonian H is regular, *i.e.*, its derivative along the fibers is is nonzero and the square matrix of its second derivatives along the fibers is nondegenerate. The curvature operator K is then the classical Jacobi endomorphism of the Hamiltonian equations defined by $\vec{H}$.

A regular semi-Hamiltonian quadruple is obtained if we restrict the considerations to a level submanifold $M_c = \{H = c\} \subset T^*Q$ of the Hamiltonian H and take the vector field X equal to $\vec{H}$ restricted to M_c, $X = \vec{H}|_{M_c}$. The distribution $\mathcal{V}$ on M_c is defined as the vertical distribution of the cotangent bundle intersected with the tangent space to M_c: $\mathcal{V}(x) = T(T^*_{\pi(x)}Q) \cap T_xM_c$, for $x \in M_c$. Then, if $\dim Q = m + 1$, we have $\dim \mathcal{V}(x) = m$, $\dim M_c = 2m + 1$ and assuming regularity of such pair $(X, \mathcal{V})$ makes sense (typical examples are regular). In this case all statements of Proposition 4.3 hold true for the canonical symplectic form σ on T^*Q replaced with the skew-symmetric form $\hat{\sigma} = \sigma|_{M_c}$ restricted to the corresponding distribution $\mathcal{D}$ on M_c.

Example 4.6. Time dependent semi-Hamiltonian vector field. The previous example can be included into the following more general one. Consider a time dependent semi-Hamiltonian vector field $\tilde{X}(t, x)$ on a manifold $(N, \tilde{\sigma})$, $\dim N = 2m$, with nondegenerate skew-symmetric 2-form $\tilde{\sigma}$. Take:

$$M = \mathbb{R} \times N, \quad X = \frac{\partial}{\partial t} + \tilde{X}, \quad \mathcal{D} = \ker \mathrm{d}t,$$

where t is the canonical coordinate on $\mathbb{R}$, extended to M with the canonical projection $\mathbb{R} \times N \to \mathbb{R}$, the vector field $\tilde{X}$ is tangent to the fibers $\{t\} \times N$ and identical with the original $\tilde{X}$ on N. The 2-form σ is uniquely defined on M by the form $\tilde{\sigma}$ and the requirements: $\sigma|_{\{t\}\times N} = \tilde{\sigma}$ and $\sigma(\partial/\partial t, \cdot) = 0$. Take $\mathcal{V}(x) \subset \mathcal{D}(x)$ to be a Lagrangian subspace of $\tilde{\sigma}$ in $\mathcal{D}(x)$, for any x, so that $x \mapsto \mathcal{V}(x)$ is a smooth distribution. If $\mathcal{V} + [X, \mathcal{V}] = \mathcal{D}$ then so defined quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ is regular semi-Hamiltonian.

Example 4.7. Contact vector field on a contact manifold. Let (M, ω) be a manifold of dimension $2m+1$ with a given 1-form ω on M satisfying $\omega \wedge (\mathrm{d}\omega)^m(x) \neq 0$ for all x, called contact form. The contact form defines the rank $2m$ distribution $\mathcal{D} = \ker \omega$ called contact distribution. Take X to be a nonvanishing contact vector field on M, *i.e.*, a nonvanishing vector field satisfying $[X, \mathcal{D}] \subset \mathcal{D}$, and let

$$\sigma = \mathrm{d}\omega|_{\mathcal{D}}.$$

Finally, consider a Lagrangian sub-distribution $\mathcal{V} \subset \mathcal{D}$, $\mathrm{rk}\,\mathcal{V} = m$, satisfying by definition $\sigma(Y, Z) = 0$ for $Y, Z \in \mathcal{V}$. Assume that the triple $(X, \mathcal{V}, \mathcal{D})$ satisfies the regularity condition (R2). Then the curvature matrix K and covariant differentiation operator D_X on sections of $\mathcal{D}$ are well defined (Sects. 2.2 and 2.4). Together with σ we obtain the regular semi-Hamiltonian quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ on M.

A special case arises when X is the Reeb vector field canonically defined by ω using the formulae $\omega(X) = 1$ and $\mathrm{d}\omega(X, \cdot) = 0$. In this setting, let g be defined on $\mathcal{V}$ by Proposition 4.3 and then on $\mathcal{D}$ as an extension of g on

$\mathcal{V}$ as in Section 2.5 using the isomorphism $A\colon \mathcal{V} \to \mathcal{H}$ and declaring orthogonality of $\mathcal{V}$ and $\mathcal{H}$. This viewpoint puts the structure within the general framework of contact sub-pseudo-Riemannian geometry, where the Reeb vector field plays a crucial role, see [26–29]. In particular, one can expect that the curvature operator K of the triple $(X, \mathcal{V}, \mathcal{D})$ contributes into invariants of the pair $(\mathcal{D}, g)$, which were derived in [26] (see also [27]) for the case of Riemannian signature and in [28] for the general case. Precisely, by fixing an orthonormal frame $(V_1, \dots, V_n)$ in $\mathcal{V}$, one can construct an orthonormal frame $(X_1, \dots, X_{2n})$ of $\mathcal{D}$ as

$$X_i = V_i, \quad X_{n+i} = [X, V_i] \qquad \text{for} \;\; i = 1, \dots, n.$$

Additionally, we denote $X_0 := X$ and introduce the structural functions as

$$[X_i, X_j] = \sum_{l=0}^{2n} c_{ij}^l X_l.$$

Then, exploiting the fact that V_i are normal, one recognizes that $K_i^j = c_{0n+i}^j$, which in turn implies that the curvature operator is part of the invariant bilinear form h defined as the Lie derivative of g (considered on $\mathcal{D}$) with respect to X (*i.e.*, $h = \frac{1}{2}\mathcal{L}_X g$), as follows from explicit formula in terms of the structural functions [28], page 446.

Given the metric g and the curvature operator K, recall that the directional curvature is defined in Section 3.1 as $k(v) = g(Kv, v)/g(v, v)$, for $v \in \mathcal{V}$ such that $g(v, v) \neq 0$. Theorems 3.3 and 3.6 have the following counterparts in the semi-Hamiltonian setting.

Theorem 4.8. *Consider a regular semi-Hamiltonian quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ on M and let $c : [0, T] \to M$ be an integral curve of X. We assume that:*

(A) the metric (4.1) is strictly definite (positive or negative) at all points of c;

(B) the directional curvature satisfies $k(v) \leq \lambda$ for all $v \in \mathcal{V}(x) \setminus \{0\}$ and all x in c, where $\lambda \in \mathbb{R}$ is a constant.

Then there are no conjugate times in $(0, T]$, if $\lambda \leq 0$, and there are no conjugate times in the interval $(0, t_c)$, $t_c = \min\{T, \frac{\pi}{\sqrt{\lambda}}\}$, if $\lambda > 0$.

Proof. The proof is the same as the proof of Theorem 3.3, with the following modification. Instead of using an "artificial" metric introduced there we use the canonical metric $g(Y, Z) = \sigma([X, Y], Z)$ on $\mathcal{V}$ introduced above. Assumption (A) guarantees that it is definite and we may assume that it is positive definite (otherwise we replace it with $-g$). The metric is invariant under covariant differentiation with respect to D_X, as Proposition 4.3 says. Introducing the function $\varphi = \sqrt{g(J, J)}$ along a geodesic, where J is a Jacobi field, we can proceed in exactly the same way as in the proof of Theorem 3.3, thus proving Theorem 4.8. □

Theorem 4.9. *Consider a regular semi-Hamiltonian quadruple $(X, \mathcal{V}, \mathcal{D}, \sigma)$ on M and let $c : [0, T) \to M$ be an integral curve of X. We assume that (A) is satisfied and* $\operatorname{tr} K \geq \kappa > 0$*, along c. Then there is a conjugate point in $(0, T^*]$ if $T^* = \pi \frac{\operatorname{rk} \mathcal{V}}{\sqrt{\kappa}} < T$.*

Proof. The theorem follows from Theorem 3.6 and Proposition 4.3. Namely, for the metric g in Theorem 3.6 we take the metric from Proposition 4.3, if it is positive definite (otherwise we take $-g$). By the proposition the curvature K is symmetric with respect to such g and the assertion follows from Theorem 3.6. □

Proof of Proposition 4.3. To prove statement (ii) take $V_i, V_j \in \mathcal{V}$. Since $\mathcal{V}$ is Lagrangian, we have $\sigma(V_i, V_j) = 0$. Lie differentiating this equality with respect to X and using condition (S) gives

$$0 = \sigma([X, V_i], V_j) + \sigma(V_i, [X, V_j])$$

and proves the symmetry of g in (ii), since σ is antisymmetric.

Differentiating twice gives

$$0 = \sigma(\operatorname{ad}_X^2 V_i, V_j) + 2\sigma([X, V_i], [X, V_j]) + \sigma(V_i, \operatorname{ad}_X^2 V_j).$$

The side terms are zero which follows from (2.5) and the fact that $\mathcal{V}$ is Lagrangian. Thus $\sigma([X, V_i], [X, V_j]) = 0$, which shows (i).

Differentiating three times yields

$$0 = \sigma([X, \operatorname{ad}_X^2 V_i], V_j) + 3\sigma(\operatorname{ad}_X^2 V_i, [X, V_j]) + 3\sigma([X, V_i], \operatorname{ad}_X^2 V_j) + \sigma(V_i, [X, \operatorname{ad}_X^2 V_j])$$

and, applying (2.5) and the summation convention,

$$\begin{aligned} 0 &= \sigma([X, K_i^s V_s], V_j) + 3\sigma(K_i^s V_s, [X, V_j]) \\ &\quad + 3\sigma([X, V_i], K_j^s V_s) + \sigma(V_i, [X, K_j^s V_s]) \\ &= \sigma(K_i^s [X, V_s], V_j) + 3\sigma(K_i^s V_s, [X, V_j]) \\ &\quad + 3\sigma([X, V_i], K_j^s V_s) + \sigma(V_i, K_j^s [X, V_s]), \end{aligned}$$

where in the second equality we use the fact that $\sigma(X(K_i^s)V_s, V_j) = 0$, as $\mathcal{V}$ is Lagrangian. Using antisymmetry of σ and the definition of g_{ij} we get

$$0 = K_i^s g_{sj} - 3K_i^s g_{js} + 3K_j^s g_{is} - K_j^s g_{si} = -2K_i^s g_{sj} + 2K_j^s g_{is},$$

since g is symmetric. This proves (iv).

Since $\mathcal{V} = \operatorname{span}\{V_1, \ldots, V_m\}$ and $\mathcal{H}^1 = \operatorname{span}\{[X, V_1], \ldots, [X, V_m]\}$ are transversal Lagrangian subspaces in TN and σ is non-degenerate on TN, it follows that the matrix g in (ii) is non-degenerate and defines a pseudo-Riemannian metric on $\mathcal{V}$.

Finally, for proving statement (iii) it is enough to show that the coefficients g_{ij} are constant along c. Using invariance of σ with respect to X (condition (S)) we find that $X(g_{ij}) = \sigma([X[X, V_i]], V_j) + \sigma([X, V_i], [X, V_j]) = 0$ since V_i and V_j are normal, thus $[X[X, V_i]]$ and V_j are in the Lagrange distribution $\mathcal{V}$ and $[X, V_i]$, $[X, V_j]$ are in the Lagrange distribution $\mathcal{H}$. □

5. Special cases

Example 5.1. Geodesic spray. Let (N, g) be a Riemannian or pseudo-Riemannian manifold. Take $M = TN$ to be the tangent bundle of N, $\mathcal{V}$ the distribution tangent to the fibers of $\pi : TN \to N$, and X the geodesic spray corresponding to metric g. In local coordinates $X = \sum_i y^i \partial_{x^i} + \sum_{i,j,k} \Gamma^i_{jk}(x) y^j y^k \partial_{y^i}$ and $\mathcal{V} = \operatorname{span}\{\partial_{y^i}\}$. Clearly, $(X, \mathcal{V})$ is a regular pair. It can be shown (see [1], Sect. 2.3) that $K(x, y)$, $(x, y) \in M = TN$ satisfies the equality $K(x, y)v = R_x(v, y)y$, where R_x is the Riemann curvature tensor at $x \in N$ and v, y are considered as vectors in T_xN. The same holds in the case of Finsler manifolds and, more generally, spray manifolds in which case the coefficients Γ^i_{jk} are functions of (x, y), instead of x only. Our theorems imply then the corresponding versions of Cartan–Hadamard and Bonet–Myers theorems on Riemannian, Finsler and spray manifolds, respectively.

Example 5.2. Fully actuated control-affine systems. In Geometric Control Theory the most often considered class of non-linear systems are control-affine systems,

$$\dot{x} = f(x) + \sum_{i=0}^{m} u_i g_i(x), \tag{5.1}$$

where $f, g_1, \dots, g_m$ are vector fields on a manifold M, $x \in M$ is the state of the system, and $u_1, \dots, u_m$ are components of the control. Such a system defines a dynamic pair,

$$X = f, \quad \mathcal{V} = \operatorname{span}\{g_1, \dots, g_m\}.$$

The regularity conditions (R1), (R2) and the invariance condition (I) are often satisfied for systems found in applications. Assuming that $\dim M = 2m$, conditions (R1), (R2) mean that $f(x) \neq 0$ and $\operatorname{span}\{g_1, \dots, g_m, [f, g_1], \dots, [f, g_m]\} = TM = \mathcal{D}$.

These assumptions are satisfied for so called fully actuated mechanical systems used in Robotics. For such a system, $M = TQ$ is the velocity phase space, where Q is the configuration manifold. Fully actuated means that there is a separate control for each degree of freedom in Q (for each joint of the robot). The system is usually described (see [16] and also [17, 30]) by the Euler-Lagrange equations corresponding to a specified Lagrangian with added external forces (the external forces include controlled forces applied at the joints). The Lagrangian has the form

$$L(q, \dot{q}) = \frac{1}{2} g(\dot{q}, \dot{q}) - P(q),$$

where $q \in Q$ is the generalized position, the first term in L is the kinetic energy (mathematically, a Riemann metric g on Q) and P is the potential energy. The controlled equations of motion can be written, in coordinates q^i, $v^i = \dot{q}^i$, as

$$\begin{aligned} \dot{q}^i &= v^i, \\ \dot{v}^i &= -\sum_{j,k} \Gamma^i_{j,k}(q) v^j v^k - \sum_s g^{is} P_{q^s}(q) + \sum_s g^{is} \left(F_{0,s}(q, v) + \sum_j u^j F_{j,s}(q, v) \right). \end{aligned} \tag{5.2}$$

Here Γ^i_{jk} are coefficients of the Levi-Civita connection of the metric g and (g^{ts}) is the inverse matrix to the matrix of metric coefficients (g_{ij}) of g. The second term represents the potential force where P_{q^s} are partial derivatives of the potential energy P. The coefficients $F_{0,s}$ describe the external uncontrollable forces, while $F_{j,s}$ are coefficients of a controlled external force $F = \sum_j u_j F_j$. A toy example may be a children's swing of mass m, where Q is a circle, the Lagrangian is $L = \frac{1}{2} m (r\dot{\theta})^2 - c \cos(\theta + \pi)$ (with $\theta \in Q$ the angle and c a constant c) and an additional uncontrollable external force is given by the friction and air resistance. A controlled force may be generated by movements of a child on the swing.

The terms on the right-hand-side not containing controls define a vector field X on $M = TQ$, while the vector fields $\hat{F}_j = \sum_i \hat{F}^i_j \partial_{v^i}$ on M, with $\hat{F}^i_j = \sum_s g^{is} F_{j,s}$, define control vector fields g_j in (5.1) and span the distribution $\mathcal{V}$. The system (5.2) is called *fully actuated* if $\mathcal{V} = \operatorname{span}\{\partial_{v^1}, \dots, \partial_{v^m}\}$. This is the same distribution as the one defined by a second order ODE. Thus we have the following

Claim. *The fully actuated system* (5.2) *defines the same dynamic pair (with $\mathcal{V}$ as above) as the one assigned to the second order ODE obtained from* (5.2) *by choosing controls equal to zero, i.e., the ODE given by the vector field on $M = TQ$*

$$X = \sum_i v^i \partial_{q^i} + \left(-\sum_{jk} \Gamma^i_{j,k}(q) v^j v^k + \sum_s g^{is} \left(-P_{q^s}(q) + F_{0,s}(q, v) \right) \right) \partial_{v^i}.$$

Consider a special case where the metric coefficients are constant (the kinetic energy depends only on the components of the velocities and not on the positions) and the uncontrolled external force is independent of the velocity. Then $\Gamma^i_{j,k} = 0$ and

$$X = \sum_i v^i \partial_{q^i} + \sum_s g^{is}(-P_{q^s}(q) + F_{0,s}(q))\partial_{v^i}.$$

Choosing $V_j = \partial_{v^j}$ we see that $[X, V_j] = -\partial_{q^j}$ and $[X, [X, V_j]] = -\sum K^i_j V_j$, where

$$K^i_j = \sum_s g^{is} \left(-\frac{\partial^2 P}{\partial q^s \partial q^j} + \frac{\partial F_{0,j}}{\partial q^s} \right).$$

This means that the frame $V_j = \partial_{v^j}$ is a normal frame of $\mathcal{V}$ and $K = (K^i_j)$ is a normal curvature matrix, which can be used in Theorem 3.3.

Since the product of two symmetric matrices may be nonsymmetric, even if the coefficients $F_{0,s}$ are zero, the above curvature matrix can be nonsymmetric. This is due to the fact that the Euclidean metric corresponding to the chosen coordinates is not natural in this case. The natural invariant metric is given by the coefficients (g_{ij}) (now assumed constant) of the kinetic energy in the Lagrangian L. The matrix with coefficients $\sum_i g_{ki} K^i_j = -P_{q^k q^j}$ is symmetric which means that the curvature K is symmetric with respect to the metric g when the external force F_0 is absent. This means that in this case both Theorems 3.3 and 3.3 are applicable with the metric g.

Remark 5.3. The example shows the advantage of considering the dynamic pair $(X, \mathcal{V})$ instead of the vector fields $f = X$ and $g_1, \dots, g_m$ defined by the system (5.1). The freedom in the choice of the generators $V_j = \partial_{v^j}$ of the distribution $\mathcal{V}$ made the calculation of the curvature matrix much simpler than if we used the generators $g_1 = F_j, \dots, g_m = F_m$ defined by the equations (5.2).

Example 5.4. A class of ordinary differential equations. Consider a pair of second-order ordinary differential equations:

$$\ddot{x}_i = F_i(t, x, \dot{x}), \qquad i = 1, 2.$$

If the corresponding normal curvature matrix K is diagonalizable, then, depending on the signs of the eigenvalues, Theorems 3.3 and 3.6 can be utilized to obtain estimates for conjugate points. However, identifying a general class of such systems is a challenging task. In Example 3.10 we considered the simplest case of linear systems. Here, we study a specific class of non-linear systems illustrating this concept. Namely, let

$$\ddot{x}_1 = F(t, x_1, \dot{x}_1), \qquad \ddot{x}_2 = \frac{\dot{x}_2}{\dot{x}_1 - x_2}(F(t, x_1, \dot{x}_1) - 2\dot{x}_2)$$

where F is arbitrary smooth function of 3 variables. Recall that $\mathcal{V} = \text{span}\{\partial_{y_1}, \partial_{y_2}\}$, where y_i a coordinate function on $J^1(\mathbb{R}, \mathbb{R}^2)$ corresponding to $\dot{x}_i$. Then, the curvature operator is given in the basis $\partial_{y_1}, \partial_{y_2}$ by

$$K = \begin{pmatrix} \chi_1 & 0 \\ -\frac{y_2(\chi_1 + \chi_2)}{y_1 - x_2} & \chi_2 \end{pmatrix}$$

where $\chi_1 = -\partial_{x_1} F + \frac{1}{2} X(\partial_{y_1} F) - \frac{1}{4}(\partial_{y_1} F)^2$ is the curvature for the scalar ODE $\ddot{x}_1 = F(t, x_1, \dot{x}_1)$, and $\chi_2 = \frac{1}{2}\frac{X(F)}{y_1 - x_2} + \frac{3}{4}\frac{F(2y_2 - F)}{(y_1 - x_2)^2}$ (recall that X denotes here the total derivative, as always in the context of ODEs). Clearly, K is not-symmetric but it is diagonalizable and the corresponding eigen-vectors are

$$V_1 = \partial_{y_2} \quad \text{and} \quad V_2 = \partial_{y_1} + \frac{y_2}{y_1 - x_2}\partial_{y_2}.$$

It turns out that the eigenspaces spanned by V_1 and V_2 are invariant with respect to D_X. Indeed, applying Proposition 2.10 and Definition 2.11 with H_1 given in Section 2.3 we compute that

$$D_X V_1 = -\frac{1}{2}\left(\frac{F-4y_2}{y_1-x_2}\right)V_1 \qquad \text{and} \qquad D_X V_2 = -\frac{1}{2}(\partial_{y_1}F)V_2.$$

Hence, it follows that V_i, $i=1,2$, can be rescaled to get a normal frame, with $\mathrm{diag}(\chi_1,\chi_2)$ as the normal curvature matrix. Note that the two eigen-values χ_1 and χ_2 can be of different signs, depending on the choice of function F.

If $F=0$ then $K=0$. We point out that in this case the system has been found before, as the system whose solution curves are so-called chains of the flat para-CR structure in dimension 3, see [31]. We also point out that the 4-dimensional solution space of this system is equipped with a well defined conformal metric of split signature, which was deeply studied and referred to as the dancing metric in [32, 33].

Acknowledgements

We are grateful to the Reviewers for their helpful remarks and suggestions concerning the first version of the manuscript. W. K. was partially supported by the grant 2019/34/E/ST1/00188 from the National Science Centre, Poland.

References

[1] B. Jakubczyk and W. Kryński, Vector fields with distributions and invariants of ODE's. *J. Geom. Mech.* **5** (2013) 85–129.

[2] A. Agrachev, The curvature and hyperbolicity of Hamiltonian systems. *Proc. Steklov Math. Inst.* **256** (2007) 26–46.

[3] A. Agrachev and I. Beschastny, Jacobi fields in optimal control: morse and Maslov indices. *Nonlinear Anal.* **214** (2022) 112608.

[4] A. Agrachev and R. Gamkrelidze, Feedback–invariant optimal control theory and differential geometry. I. Regular extremals. *J. Dyn. Control Syst.* **3** (1997) 343–389.

[5] A. Agrachev and R. Gamkrelidze, Feedback–invariant optimal control theory and differential geometry. II. Jacobi curves for singular extremals. *J. Dyn. Control Syst.* **4** (1998) 583–604.

[6] A. Agrachev and I. Zelenko, Geometry of Jacobi curves II. *J. Dyn. Control Syst.* **8** (2002) 167–215.

[7] D. Barilari and L. Rizzi, Comparison theorems for conjugate points in sub-Riemannian geometry. *ESAIM: COCV* **22** (2016) 439–472.

[8] N. Caroff and H. Frankowska, Conjugate points and shocks in nonlinear optimal control. *Trans. Am. Math. Soc.* **348** (1996) 3133–3153.

[9] J.F. Carinena, I. Gheorghiu and E. Martínez, Jacobi fields for second-order differential equations on Lie algebroids, in *Dynamical Systems and Differential Equations, Proceedings of the 10th AIMS International Conference*, edited by M. de León, *et al.* Madrid, Spain (2015) 213–222.

[10] J.F. Carinena and E. Martinez, Generalized Jacobi equation and inverse problem in classical mechanics, in *Integral Systems, Solid State Physics and Theory of Phase Transitions*, edited by V.V. Dondonov and V. Manko. Nova Science (1992).

[11] M. Crampin, E. Martinez and W. Sarlet, Linear connections for systems of second-order ordinary differential equations. *Ann. Inst. Henri Poincare* **65** (1996) 223–249.

[12] S. Hajdú and T. Mestdag, Conjugate points for systems of second-order ordinary differential equations. *Int. J. Geom. Methods Mod. Phys.* **17** (2020).

[13] M. Jerie and G.E. Prince, Jacobi fields and linear connections for arbitrary second-order ODEs. *J. Geom. Phys.* 43 (2002) 351–370.

[14] S.V. Sabau, Some remarks on Jacobi stability. *Nonlinear Anal.* **63** (2005) e143–e153.

[15] A. Agrachev, D. Barilari and L. Rizzi, Curvature: a variational approach. *Mem. AMS* **256** (2007).

[16] F. Bullo and A. Lewis, Geometric Control of Mechanical Systems, Texts in Applied Mathematics (TAM), Vol. 49. Springer (2005).

[17] S. Ricardo and W. Respondek, When is a control system mechanical? *J. Geom. Mech.* **2** (2010) 265–302.

[18] E. Cartan, Observations sur le mémoir précédent. *Math. Zeitschrift* **37** (1933) 619–622.
[19] S.-S. Chern, Sur la géométrie d'un systéme d'équations différentialles du second ordre. *Bull. Sci. Math.* **63** (1939) 206–212.
[20] D.D. Kosambi, Parallelism and path-space. *Math. Zeitsch.* **37** (1933) 608–618.
[21] D.A. Grossman, Torsion-free path geometries and integrable second order ODE systems. *Selecta Math. (N.S.)* **6** (2000) 399–442.
[22] J. Diaz, J.B. McLaughlin and R. Joyce, Sturm comparison theorems for ordinary and partial differential equations. *Bull. Am. Math. Soc.* **75** (1969) 335–339.
[23] H. Abou-Kandil, G. Freiling, V. Ionescu and G. Jank, Matrix Riccati Equations in Control and Systems Theory. Birkhauser Basel (2003).
[24] P. Hartman, Ordinary Differential Equations. John Wiley and Sons, New York (1964).
[25] F.C. Chittaro, An estimate for the entropy of Hamiltonian flows. *J. Dyn. Control Syst.* **13** (2007) 55–67.
[26] A. Agrachev, Exponential mappings for contact sub-Riemannian structures. *J. Dynam. Control Syst.* **2** (1996) 321–356.
[27] E. Falbel and C. Gorodski, Sub-Riemannian homogeneous spaces in dimensions 3 and 4. *Geom. Dedicata* **62** (1996) 227–252.
[28] M. Grochowski and W. Kryński, Invariants of sub-pseudo-Riemannian structures and Einstein-Weyl geometry. *Radon Ser. Computat. Appl. Math.* **18** (2016) 434–453.
[29] M. Grochowski and W. Kryński, On contact sub-pseudo-Riemannian isometries. *ESAIM: COCV* **23** (2017) 1751–1765.
[30] M. Nowicki and W. Respondek, A mechanical feedback classification of linear mechanical control systems, Special issue "Advances in robot motion and control". *Appl. Sci.* **11** (2021) 10669.
[31] W. Kryński and O. Makhmali, Lewy curves in para-CR geometry. arXiv:2406.04798 (2024).
[32] G. Bor, L. Lamoneda and P. Nurowski, The dancing metric, G2-symmetry and projective rolling. *Trans. Am. Math. Soc.* **370** (2018) 4433–4481.
[33] M. Dunajski, Twistor theory of dancing paths. *SIGMA* **18** (2022) 027.

Appendix A.

Proof of Lemma 3.20. Since P is invertible on $I = (0, r]$ we may define u by $w = Pu$ and then $u(0) = 0 = u(r)$ and $\dot{w} = \dot{P}u + P\dot{u}$. We can decompose $L = \langle \dot{w}, \dot{w} \rangle - \langle Kw, w \rangle$ as

$$L = \langle P\dot{u}, P\dot{u} \rangle + L_1 + L_2$$

where

$$L_1 = \langle \dot{P}u, \dot{P}u \rangle + \langle \dot{P}u, P\dot{u} \rangle, \quad L_2 = \langle P\dot{u}, \dot{P}u \rangle - \langle KPu, Pu \rangle.$$

We will prove that $\int_I (L_1+L_2)\mathrm{d}t = 0$. This equality implies the inequality

$$\int_I L\,\mathrm{d}t = \int_0^r \langle P\dot{u}, P\dot{u}\rangle \mathrm{d}t \geq 0$$

and, by invertibility of P and $u(0)=0$, we will have $\int_0^r L\mathrm{d}t = 0$ iff $u\equiv 0$ iff $w \equiv 0$. This will conclude the proof.

To prove $\int_I (L_1+L_2)\mathrm{d}t = 0$ note that $\dot{P} = Q$ implies

$$\int_I L_1 \mathrm{d}t = \int_I \langle Qu, (\dot{Pu})\rangle \mathrm{d}t = -\int_I \langle \dot{Q}u, Pu\rangle + \langle Q\dot{u}, Pu\rangle \mathrm{d}t, \tag{A.1}$$

where the latter equality follows from integration by parts and $u(0)=0=u(r)$. Using also $\dot{Q} = -KP$ we get

$$\int_I L_2 \mathrm{d}t = \int_I \langle Qu, P\dot{u}\rangle + \langle \dot{Q}u, Pu\rangle \mathrm{d}t. \tag{A.2}$$

We claim that $\langle Qu, P\dot{u}\rangle = \langle Q\dot{u}, Pu\rangle$ and consequently

$$\int_I \langle Qu, P\dot{u}\rangle - \langle Q\dot{u}, Pu\rangle \mathrm{d}t = 0. \tag{A.3}$$

This follows from

$$\langle Qu, P\dot{u}\rangle = \langle P^T Q u, \dot{u}\rangle, \quad \langle Q\dot{u}, Pu\rangle = \langle P^T Q\dot{u}, u\rangle$$

and the symmetry of $P^T Q$ which is a consequence of

$$\frac{\mathrm{d}}{\mathrm{d}t} P^T Q = \dot{P}^T Q + P^T \dot{Q} = Q^T Q - P^T K P.$$

Here the right hand side is symmetric, by the assumption $K = K^T$, and the initial condition $P(0)=0$ implies symmetric $(P^T Q)(0) = 0$, thus the solution $(P^T Q)(t)$ is symmetric. From (A.1), (A.2) and (A.3) we see that $\int_I (L_1+L_2)\mathrm{d}t = 0$, which concludes the proof. □