Jacobi fields in optimal control: one-dimensional variations

# Jacobi Fields in Optimal Control II: One-dimensional variations

A. Agrachev, I. Beschastnyi

November 21, 2018

### Abstract

In this second paper of the series we specify general theory developed in the first paper. Here we study the structure of Jacobi fields in the case of an analytic system and piece-wise analytic control. Moreover, we consider only 1-dimensional control variations. Jacobi fields are piece-wise analytic in this case but may have much more singularities than the control. We derive ODEs that these fields satisfy on the intervals of regularity and study behavior of the fields in a neighborhood of a singularity where the ODE becomes singular and the Jacobi fields may have jumps.

## Introduction

In this paper we continue the study of the second variation of optimal control problems. Our main technique are the so-called $\mathcal{L}$-derivatives that were introduced in [6, 1]. One can think of them as a rule that for a given critical point assigns to a certain space of variations a Lagrangian space in some symplectic space. In the first article we presented the theoretical basis and gave an algorithm how to compute an approximation of an $\mathcal{L}$-derivative with arbitrary good precision.

In this article we focus on examples and simpler characterization of Jacobi curves. To be more precise we study the following optimal control problem

$$\dot{q} = f(q, u), \qquad u \in U \subset \mathbb{R}, \qquad q \in M \tag{1}$$

$$q(0) = q_0, \qquad q(T) = q_T,$$

$$J_T[u] = \int_0^T L(u, q)dt \to \min.$$

Here $M$ is a $n$-dimensional manifold. For simplicity we assume that $f(u, q)$ and $L(u, q)$ are analytic in both variables, that $U$ is a polytope or a one-dimensional smooth manifold, that controls are $L^\infty$ functions and that the time $T$ is fixed. For simplicity and conceptual clarity we make the assumption that extremal control takes values in the vertices or one-dimensional edges of $U$.

We begin by recalling the necessary results from symplectic geometry in Section 1. This part is standard and can be found in several books like [16] or [17]. We present this section for reader's convenience and in order to fix the basic notations. Then we recall all the necessary results from the previous article [3] about the $\mathcal{L}$-derivatives and the construction of Jacobi curves for optimal control problems in Sections 2 and 3.

Using these techniques we then proceed to the study of singular and bang-bang cases in Sections 4 and 5, where we give efficient algorithms for characterizing Jacobi curves. These cases were already studied by many authors (see for example [12, 26, 25, 21, 23, 10]). These results can be recovered using the constructions from this paper and Morse-type theorems from our previous article [3].

In the second part we study extremals along which at a single point the Legendre condition becomes degenerate. In this case the Jacobi curve should be a solution of a singular Jacobi DE, but it can not be characterized as the usual boundary value problem, because we lose both existence and uniqueness of solutions. This kind of systems were previously studied in the classical calculus of variations, in particular by Morse himself and some of his students [19, 18] using functional-analytic technique or in [28] using differential geometry and fields of extremals. Some special examples with similar singularities were recently studied in [13]. We show how using the technique of $\mathcal{L}$-derivatives we can still characterize the desired Jacobi curve in a simplest singular example as boundary value problem of an ODE, but with conditions on the first k-jet of a solution (and not just the initial value).

In Section 6 we show that if we restrict ourselves only to one-dimensional variations, using a change of variables under some non-degeneracy conditions we can separate the dynamics of the system into a regular and singular part in an invariant symplectic subspace of a dimension at most four. In Section 7 we explain the idea heuristically when the dimension of the manifold on which the control system is defined is equal to one. In the remaining sections we make all those ideas rigorous. In Section 8, we find sufficient conditions for existence and non-existence of $\mathcal{L}$-derivatives and in Section 9 we characterize Jacobi curves as singular BVPs.

# 1 Linear symplectic geometry and Lagrangian Grassmanian

In this sections we recall basic facts from symplectic geometry and fix notations that we use in the article.

Given a symplectic space $(\mathbb{R}^{2n}, \sigma)$, where $\sigma$ is a symplectic form we can always assume by the Darboux theorem that

$$\sigma(\lambda_1, \lambda_2) = \lambda_1^T J \lambda_2, \qquad \forall \lambda_i \in \mathbb{R}^{2n},$$

where $J$ is the standard complex structure

$$J = \begin{pmatrix} 0 & \mathrm{id}_n \\ -\mathrm{id}_n & 0 \end{pmatrix}.$$

In particular $J^2 = -\text{id}_{2n}$. Coordinates in which $\sigma$ has such a form are called *Darboux coordinates*. We use usual position-momenta notations in this case, i.e. we write $\lambda = (p, q) = (p_1, ..., p_n, q_1, ..., q_n)$.

In Darboux coordinates a Hamiltonian system with a perhaps time-dependent Hamiltonian $H : \mathbb{R}^{2n} \times \mathbb{R} \to \mathbb{R}$ is a system of ODEs

$$\dot{\lambda} = -J\nabla H(t, \lambda),$$

where $\nabla H$ is the $\mathbb{R}^{2n}$-gradient of $H$. In particular, if $H$ is quadratic of the form

$$H(t, \lambda) = p^T C(t) p - 2q^T A(t) p - q^T B(t) q,$$

where $B(t), C(t)$ are symmetric matrices, we obtain a linear Hamiltonian system

$$\frac{d}{dt} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} A(t) & B(t) \\ C(t) & -A^T(t) \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}. \tag{2}$$

Given $J$ we can define the *symplectic group* $\text{Sp}(2n)$ and the corresponding *symplectic algebra* $\text{sp}(2n)$ as

$$\text{Sp}(2n) = \left\{ M \in \text{Mat}(2n \times 2n, \mathbb{R}) \ : \ M^T J M = J \right\},$$

$$\text{sp}(2n) = \left\{ X \in \text{Mat}(2n \times 2n, \mathbb{R}) \ : \ X^T J + J X = 0 \right\}.$$

If we write down $X \in \text{sp}(2n)$ as block matrix, we will see that it has the same form as the matrix in the Hamiltonian system (2). Therefore we immediately can see that the flow $\Phi(t)$ of (2) is symplectic.

We define the *skew-orthogonal complement* of a linear subspace $\Gamma$ as

$$\Gamma^{\angle} = \{ \mu \in \mathbb{R}^{2n} \ : \ \sigma(\mu, \lambda) = 0, \ \forall \lambda \in \Gamma \}.$$

A subspace $\Gamma$ is called

– *isotropic* if $\Gamma \subset \Gamma^{\angle}$ or equivalently if $\sigma|_\Gamma = 0$,

– *Lagrangian* if $\Gamma = \Gamma^{\angle}$ or if equivalently $\sigma|_\Gamma = 0$ and $\dim \Gamma = n$,

– *coisotropic* if $\Gamma \supset \Gamma^{\angle}$,

– *symplectic* if $\dim(\Gamma \cap \Gamma^{\angle}) = 0$ or if equivalently $\sigma|_\Gamma$ is non-degenerate.

Since $\sigma$ is skew-symmetric, any one-dimensional direction $\mathbb{R}v$, $v \in \mathbb{R}^{2n}$ is isotropic. Two main examples of Lagrangian subspaces are the *horizontal subspace* $\Sigma$ and the *vertical subspace* $\Pi$ defined as

$$\Pi = \left\{ (p, q) \in \mathbb{R}^{2n} \ : \ q = 0 \right\},$$
$$\Sigma = \left\{ (p, q) \in \mathbb{R}^{2n} \ : \ p = 0 \right\}.$$

3

We can construct other examples as follows. Let $S = S^T$ be a symmetric matrix. Then

$$\Lambda_S = \{(p, Sp) \in \mathbb{R}^{2n} \ : \ p \in \mathbb{R}^n\}$$

is a Lagrangian subspace transversal to $\Sigma$. Conversely to any and Lagrangian subspace $\Lambda$ transversal to $\Sigma$ (we denote this by $\Lambda \pitchfork \Sigma$) we can associate a symmetric operator $S$ from $\Pi$ to $\Sigma$.

We call the set of all Lagrangian planes *Lagrangian Grassmanian* and denote it by $L(n)$. It is a manifold, whose atlas is given by $\Lambda^{\pitchfork}$, which are the sets of Lagrangian planes transversal to $\Lambda \in L(n)$. Coordinate charts are maps from $\Lambda^{\pitchfork}$ to the space of symmetric matrices constructed like above. Throughout this paper we use another representation of a Lagrangian plane $\Lambda \in L(n)$ as a span of $n$ independent vectors $v_i$. It is clear that such a representation is not unique. We can replace $v_i$ by any linear span of the same vectors as long as they remain independent. This means that in general we need to quotient a natural $\mathrm{GL}(n)$ action. We can arrange $v_i$ in a single $n \times 2n$ matrix and we write

$$\Lambda = \begin{bmatrix} v_1 & ... & v_n \end{bmatrix},$$

where the square brackets indicate the equivalence class under the $\mathrm{GL}(n)$ action. We denote this action by

$$g \begin{bmatrix} v_1 & ... & v_n \end{bmatrix} := \begin{bmatrix} gv_1 & ... & gv_n \end{bmatrix}, \qquad g \in \mathrm{GL}(n).$$

For example, if $\Lambda \in \Sigma^{\pitchfork}$ we can write

$$\Lambda = \begin{bmatrix} \mathrm{id}_n \\ S \end{bmatrix},$$

where $S$ is a symmetric matrix like in the example above. Or we can assume that $v_i$ form an orthonormal basis of $\Lambda$ in $\mathbb{R}^{2n}$. Then

$$\Lambda = \begin{bmatrix} X \\ Y \end{bmatrix},$$

where $X^T X + Y^T Y = \mathrm{id}_n$ (orthonormality property) and $X^T Y - Y^T X = 0$ (Lagrangian property) are satisfied. A matrix $X + iY$ that satisfies these properties is unitary and the converse is true as well. We can choose $v_i$ in such a way up to a $\mathrm{O}(n)$-action, which is given by

$$\begin{pmatrix} O & 0 \\ 0 & O \end{pmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}, \qquad O \in \mathrm{O}(n).$$

This gives the usual identification of $L(n) \simeq \mathrm{U}(n)/\mathrm{O}(n)$.

We will use this idea many times when we will consider the singular case, so at this point it makes sense to consider a simple example that will be useful for us later.

**Example 1.1.** Suppose that we would like to find a simple representation of a Lagrangian plane $\Lambda \in L(2)$ knowing that $\dim(\Lambda \cap \Sigma) = 1$. Then it must be of the form

$$\Lambda = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} x_1 & 0 \\ y_1 & 0 \\ z_1 & z_2 \\ w_1 & w_2 \end{bmatrix}, \qquad x_1^2 + y_1^2 \neq 0.$$

We can assume that $v_1$ and $v_2$ are orthonormal. We then apply a rotation $O \in O(n)$, so that $y$ component of $v_1$ becomes zero. Then

$$\Lambda = \begin{bmatrix} Ov_1 & Ov_2 \end{bmatrix} = \begin{bmatrix} \sqrt{x_1^2 + y_1^2} & 0 \\ 0 & 0 \\ \tilde{z}_1 & \tilde{z}_2 \\ \tilde{w}_1 & \tilde{w}_2 \end{bmatrix},$$

but since $\Lambda$ is Lagrangian we must have $\tilde{z}_2 = 0$. Changing the basis we then find

$$\Lambda = \begin{bmatrix} \dfrac{Ov_1 - (\tilde{w}_1/\tilde{w}_2)Ov_2}{\sqrt{x_1^2 + y_1^2}} & \dfrac{Ov_2}{\tilde{w}_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ z & 0 \\ 0 & 1 \end{bmatrix}.$$

Given an isotropic subspace $\Gamma$ and a Lagrangian plane $\Lambda$, we can construct a new Lagrangian plane $\Lambda^\Gamma$, which is a Lagrangian plane that contains $\Gamma$ and the dimension of $\Lambda \cap \Lambda^\Gamma$ is maximal. It is defined as

$$\Lambda^\Gamma = (\Lambda \cap \Gamma^\angle) + \Gamma = (\Lambda + \Gamma) \cap \Gamma^\angle.$$

If $\Gamma = \mathbb{R}X$ for some vector $X \in \mathbb{R}^{2n}$ we will simply write $\Lambda^X$ instead of $\Lambda^{\mathbb{R}X}$.

Let us have a look at another example that will be useful in future.

**Example 1.2.** Assume that $X = \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix}^T$ and we would like to construct $\Lambda^X$ for $\Lambda \in L(2)$. We have that

$$\Lambda = \begin{bmatrix} v_1 & v_2 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \\ z_1 & z_2 \\ w_1 & w_2 \end{bmatrix}.$$

Subspace $X^\angle$ consists of vectors $v \in \mathbb{R}^4$ whose $z$-component is zero. So assume first that $z_1 = z_2 = 0$. Then $\sigma(X, v_1) = \sigma(X, v_2) = 0$. But since it is a Lagrangian subspace, it means that $X \in \Lambda$ and by definition $\Lambda^X = \Lambda$. Thus we can take $v_1 = X$. In this case we obtain

$$\Lambda^X = \begin{bmatrix} X & v_2 \end{bmatrix} = \begin{bmatrix} X & v_2 - x_2 X \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix}.$$

5

Suppose that $X \notin \Lambda$. Then $z_1^2 + z_2^2 \neq 0$ and as a result $\sigma(X, z_1 v_1 + z_2 v_2) \neq 0$, but $\sigma(X, z_1 v_2 - z_2 v_1) = 0$. So $\Lambda \cap X^{\angle} = z_1 v_2 - z_2 v_1$ and by definition

$$
\Lambda^X = \begin{bmatrix} X & z_1 v_2 - z_2 v_1 \end{bmatrix} = \begin{bmatrix} 1 & z_1 x_2 - z_2 x_1 \\ 0 & z_1 y_2 - z_2 y_1 \\ 0 & 0 \\ 0 & z_1 w_2 - z_2 w_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & z_1 y_2 - z_2 y_1 \\ 0 & 0 \\ 0 & z_1 w_2 - z_2 w_1 \end{bmatrix}.
$$

Our main objects of study are going to be curves in the Lagrangian Grassmanian. The curves will always come from a flow $\Phi(t)$ of a linear Hamiltonian system (2). We will simply take a point $\Lambda$ and consider a curve $\Lambda(t) = \Phi(t)\Lambda$. More generally a linear Hamiltonian system induces a dynamical system on $L(n)$. We can write down an ODE for that system using local charts. Indeed, let $(p(t), q(t))$ be a solution of (2) and $S(t)$ be a curve of symmetric matrices that correspond to $\Lambda(t)$. Then $q(t) = S(t)p(t)$ and we differentiate this expression. This way we obtain a Riccati equation of the form

$$
\dot{S} + SA + A^T S + SBS - C = 0. \tag{3}
$$

Since a coordinate chart $\Sigma^{\pitchfork}$ is dense in $L(n)$ the opposite is also true: a Riccati equation of the form above gives rise to a Hamiltonian system and a well defined flow on $L(n)$. In order to write down a Riccati equation in a different chart we can apply a symplectic transformation to the corresponding Hamiltonian, s.t. a given Lagrangian plane $\Lambda$ is mapped to $\Sigma$. Then we simply insert the new expressions for $A, B$ and $C$ in (3).

The last ingredient that we need is the *Maslov index*, which is a topological symplectic invariant of curves in the Lagrangian Grassmanian. Maslov index is the same for different curves in the same homotopy class, thus it does not change under small perturbations. We are going to give a simple coordinate definition, but one should remember that this number has many equivalent invariant definitions (see for example [14]).

Let $\Lambda : [0, 1] \to L(n)$ be a continuous curve. We assume that the curve $\Lambda(t)$ lies completely in some coordinate chart, i.e. there exists a plane $\Delta \in L(n)$, s.t. $\Lambda(t) \in \Delta^{\pitchfork}$ for all $t \in [0, 1]$. We call such a curve *simple*. Let $\Pi \in \Delta^{\pitchfork}$, and assume that end-points $\Lambda(0)$ and $\Lambda(1)$ are transversal to $\Pi$. We can then associate to $\Lambda(0), \Lambda(1)$ symmetric matrices $S_0, S_1$ that correspond to symmetric operators from $\Pi$ to $\Delta$. Maslov index of $\Lambda(t)$ with respect to $\Pi$ is defined as

$$
\mathrm{Mi}_\Pi \, \Lambda(t) = \frac{1}{2} \left( \mathrm{sign}\, S_1 - \mathrm{sign}\, S_0 \right).
$$

One can check that this definition does not depend on the choice of $\Delta$. To define the Maslov index for a general curves $\Lambda : [0, 1] \to L(n)$ one should split it in a number of simple arcs $\Lambda_i(t)$, s.t. each of them lies in its own coordinate chart. Then Maslov index of the whole curve is defined as a sum of the corresponding indices of simple arcs:

$$
\mathrm{Mi}_\Pi \, \Lambda(t) = \sum_{i=1}^{N} \mathrm{Mi}_\Pi \, \Lambda_i(t).
$$

6

From this definition we can easily see that the index of a simple curve depends only on the relative position of its end-points and remains the same if we perturb the curve as long as the transversality conditions are preserved. Using this one can prove that Maslov index of a curve is a homotopy invariant.

We have two very important properties of the Maslov index related to a change of the reference plane.

**Theorem 1.1** ([7]). *If a curve $\Lambda(\tau) \subset L(n)$ is closed, then its Maslov index does not depend on the choice of the reference plane, i.e.*

$$\mathrm{Mi}_{\Delta_1}\Lambda(\tau) = \mathrm{Mi}_{\Delta_2}\Lambda(\tau) = \mathrm{Mi}\,\Lambda(\tau), \qquad \forall \Delta_i \in L(n).$$

*If it is not closed we have the following estimate*

$$|\mathrm{Mi}_{\Delta_1}\Lambda(\tau) - \mathrm{Mi}_{\Delta_2}\Lambda(\tau)| \leq n.$$

There exist several other equivalent definitions of the Maslov index [16, 14]. Usually one defines it as an intersection index, but this definition is rather long and has many subtleties. Nevertheless it allows to prove easily the following theorem that we will use in Section 8.

**Theorem 1.2** ([7]). *Let $H(t)$ be a quadratic non-autonomous Hamiltonian and let $\Phi(t)$ be a flow of the corresponding Hamiltonian system. Fix two transversal Lagrangian planes $\Delta$ and $\Lambda$ and assume that $\Delta \pitchfork \Phi(T)\Lambda$. If $H(t)|_\Delta \geq 0$ for $t \in [0, T]$, then*

$$\mathrm{Mi}_\Delta \Phi(t)\Lambda = \sum_{t \in [0,T]} \dim\left(\Delta \cap \Phi(t)\Lambda\right).$$

*If $H(t)|_\Delta \leq 0$, then the same formula holds with a minus sign in front of the sum.*

These two theorems will be our main tool in proving oscillation results in Section 8.

## 2  $\mathcal{L}$-derivatives

We will derive the Jacobi equations and study them using the so called $\mathcal{L}$-derivatives. A $\mathcal{L}$-derivative is a rule that assigns to an admissible space of variations a Lagrangian plane in some symplectic space. As we add variations we can compare the relative positions of the corresponding $\mathcal{L}$-derivatives and deduce from that how the inertia indices and nullity of the Hessian change as we consider a bigger and bigger space of variations. As a result one can recover the classical theory of Jacobi and much more. This theory is applicable in a great variety of cases, even when there is no Jacobi equation at all. We begin by explaining the abstract setting and then we specialize the results to optimal control problems.

Assume that we have following constrained variational problem. Let $J : \mathcal{U} \to \mathbb{R}$ be a smooth functional and $F : \mathcal{U} \to M$ be a smooth map, where $\mathcal{U}$ is a Banach manifold and

$M$ is a finite-dimensional manifold. Given a point $q \in M$, we are interested in finding $\tilde{u} \in F^{-1}(q)$ that minimize $J$ among all other points $u \in F^{-1}(q)$. In the case of optimal control problems $\mathcal{U}$ is the space of admissible controls. The map $F$ is usually taken to be the end-point map, which we will introduce in the next section.

The first step is to apply the Lagrange multiplier rule that says that if $\tilde{u}$ is a minimal point then there exists a covector $\lambda \in T_q^* M$ and a number $\nu \in \{0, 1\}$, s.t.

$$\langle \lambda, dF[\tilde{u}](w) \rangle - \nu dJ[\tilde{u}](w) = 0, \qquad \forall w \in T_{\tilde{u}}\mathcal{U}. \tag{4}$$

A pair $(\tilde{u}, \lambda)$ that satisfies the equation above is called a *Lagrangian point* and $\tilde{u}$ is called a critical point of $(F, J)$. There are of course many critical points that a not minimal. So in order to find the minimal ones we have to apply higher order conditions for minimality. For example, we can look at the Hessian $\mathrm{Hess}(F, \nu J)[\tilde{u}, \lambda]$ at a Lagrangian point $(\tilde{u}, \lambda)$ that we define as

$$\mathrm{Hess}(F, \nu J)[\tilde{u}, \lambda] := \left( \nu d^2 J[\tilde{u}] - \langle \lambda, d^2 F[\tilde{u}] \rangle \right) |_{\ker dF[\tilde{u}]}. \tag{5}$$

In the normal case this expression coincides with the Hessian of $J$ restricted to the level set $F^{-1}(q)$. The index and the nullity of the Hessian are directly related to optimality of the critical point $\tilde{u}$ [9].

We are now ready to define $\mathcal{L}$-derivatives. We linearise (4) with respect to $\lambda$ and $u$, and obtain the following equation

$$\langle \xi, dF[\tilde{u}](w) \rangle + \langle \lambda, d^2 F[\tilde{u}](v, w) \rangle - \nu d^2 J[\tilde{u}](v, w) = 0.$$

Or if we define $Q(v, w) := \langle \lambda, d^2 F[\tilde{u}](v, w) \rangle + \nu d^2 J[\tilde{u}](v, w)$, we can rewrite this as

$$\langle \xi, dF[\tilde{u}](w) \rangle + Q(v, w) = 0. \tag{6}$$

A $\mathcal{L}$-*derivative* of a pair $(F, J)$ at a Lagrangian point $(\tilde{u}, \lambda)$ constructed over a finite-dimensional space of variations $V \subset T_{\tilde{u}}\mathcal{U}$ is the set

$$\mathcal{L}(F, J)[\tilde{u}, \lambda](V) = \{(\xi, dF[\tilde{u}](v)) \in T_\lambda(T^* M) : (\xi, v) \in (T_\lambda(T_q^* M), V) \text{ solve } (6) \text{ for } \forall w \in V\}.$$

This set is a Lagrangian plane [2]. The reason why we do not take directly $T_{\tilde{u}}\mathcal{U}$ instead of $V$ is that it is a linear equation defined on an infinite-dimensional space and it might be ill-posed. In this case $\mathcal{L}(F, J)[\tilde{u}, \lambda](V)$ is just isotropic. But if we have chosen the right topology for our space of variations, we are going to get exactly $\dim M$ independent solutions.

To define a $\mathcal{L}$-derivative over an infinite-dimensional space $V \subset T_{\tilde{u}}\mathcal{U}$ we must take a generalized limit or a limit of a net over all finite dimensional subspaces $U \subset V$:

$$\mathcal{L}(F, J)[\tilde{u}, \lambda](V) = \lim_{U \nearrow V} \mathcal{L}(F, J)[\tilde{u}, \lambda](U),$$

with the partial ordering given by inclusion. When $V$ is the whole space of available variations, we simply write $\mathcal{L}(F, J)[\tilde{u}, \lambda]$ for the corresponding $\mathcal{L}$-derivative.

We have the following important theorem proved in [1], that gives the existence of this limit and a way to compute it.

**Theorem 2.1.** *Let $(\lambda, \tilde{u})$ be a Lagrangian point of $(F, J)$.*

1. *If either the positive or the negative inertia index of $\mathrm{Hess}(F, \nu J)[\tilde{u}, \lambda]$ is finite, then $\mathcal{L}(F, J)[\tilde{u}, \lambda]$ exists;*

2. *$\mathcal{L}(F, J)[\tilde{u}, \lambda] = \mathcal{L}(F, J)[\tilde{u}, \lambda](V)$ for any $V$ dense in $T_{\tilde{u}}\mathcal{U}$.*

$\mathcal{L}$-derivatives contain information about the inertia indices and nullity of the Hessian (5) restricted to some space of variations. By comparing two $\mathcal{L}$-derivatives constructed over two subspace $V \subset W$, we can see how the inertia indices change as we add variations to our variations space [1].

In the next section we will write down explicit expressions for the optimal control problem we are studying, and define Jacobi curves. In the previous article [3] we gave a general algorithm for their approximation. In this article we focus on situations when an exact characterization is possible.

# 3 $\mathcal{L}$-derivatives for optimal control systems

Let us consider the optimal control problem (1). We introduce another type of variations called time variations. They can be used to obtain necessary [5] and sufficient [10] optimality conditions. We have discussed them in detail in [4] (see also [3]), so we just give a quick recap.

Let us consider a change of time

$$t(s) = \int_0^s (1 + \alpha(\theta))d\theta,$$

where we assume $\alpha(\theta) > -1$, in order to have invertibility of $t(s)$. Then in the new time we can write

$$\dot{q} = (1 + \alpha(s))f(q, u(t(s))),$$
$$\dot{t} = 1 + \alpha(s),$$

$$\int_0^{t^{-1}(T)} (1 + \alpha(s))L(q, u(t(s)))ds \to \min,$$

where we have added the time as a new variable in order not to added any additional constraints on the controls. One can then proceed to show that $\tilde{u}(t)$ is minimal if and only if the control $(\alpha(s), u(s)) = (0, \tilde{u}(t(s))$ is minimal in the new problem above. We will denote $\hat{U} = U \times (-1, \infty)$ the new space of control values and $\hat{u}$ a critical control in $\hat{U}$. By abusing slightly the notations we consider the original control parameter $u$ and time variations $\alpha$ as a new control, that we denote again by $u$.

In order to reformulate the optimal control problem above as a constrained variational problem of the previous section one introduces the *end-point map* $E_t : L^\infty([0, T], \hat{U}) \to M$. It takes an admissible control $u(t)$ and associates to it final point of the trajectory which

is a solution of the Cauchy problem (**??**) with $q(0) = q_0$. Then in the notations of the previous section $J = J_T$, $E = E_T$. Applying the Lagrange multiplier rule we find that if $(\tilde{u}(t), \tilde{q}(t))$ is an optimal control and the corresponding optimal trajectory, then there must exist a covector $\lambda(T) \in T^*_{\tilde{q}(T)}M$ and a number $\nu \in \{0, 1\}$, s.t.

$$\langle \lambda(T), dE_T[\tilde{u}](w) \rangle - \nu dJ_T[\tilde{u}](w) = 0,$$

for suitable variations $w$. Let $P_t^T$ be the flow of (**??**) under the control $\tilde{u}(t)$ from time $t$ to time $T$. From the definitions it clear that $dE_T|_{L^\infty([0,T],\hat{U})} = (P_t^T)_* dE_t$ and $dJ_T|_{L^\infty([0,T],\hat{U})} = dJ_t$, where $(P_t^T)_*$ is the differential of $P_t^T$. If we denote $\lambda(t) = (P_t^T)^*\lambda(T) \in T^*_{\tilde{q}(t)}M$ then, we have along $\tilde{q}(t)$

$$\langle \lambda(t), dE_t[\tilde{u}](w) \rangle - \nu dJ_t[\tilde{u}](w) = 0.$$

These first order conditions are equivalent to a weak version of the Pontryagin maximum principle which states that $\lambda(t)$ must satisfy a Hamiltonian system

$$\dot{\lambda} = \vec{h}_{\tilde{u}(t)}\lambda, \tag{7}$$

where

$$h(u, \lambda) = \langle \lambda, f(u, q) \rangle - \nu L(u, q),$$

and along the extremal curve an extremum condition

$$\left. \frac{\partial h(u, \lambda(t))}{\partial u} \right|_{u=\tilde{u}} = 0$$

is satisfied.

From the previous discussion it would seem natural to compute $\mathcal{L}$-derivatives of $(E_t, J_t)$ in order to study the second variation. In this case we would obtain a one-parametric family of Lagrangian planes which encode information about the corresponding Hessian. However in this case Lagrangian planes will lie in different symplectic spaces and thus we can not compare directly their relative position. To fix this problem, we simply take the flow of the Hamiltonian system (7) that we denote as $\Phi_t$ and compute $\Phi_t^*\mathcal{L}(E_t, J_t)[\tilde{u}, \lambda](V)$.

It remains to give an explicit expression of (6). We define

$$b(\tau) = \left. \frac{\partial^2}{\partial u^2} \right|_{u=\tilde{u}} h(u, \lambda(t))$$

and

$$X(\tau) = \left. \frac{\partial}{\partial u} \right|_{u=\tilde{u}} \overrightarrow{\Phi_t^*(h(u, \cdot))(\lambda(0))},$$

which is essentially the linearization of a pull-back of the PMP Hamiltonian using the flow $\Phi_t$. In [3] we have proven the following characterization of the mentioned previously $\mathcal{L}$-derivatives of optimal control problems.

10

**Proposition 3.1.** $\mathcal{L}(E_t, J_t)[\tilde{u}, \lambda]$ *consists of the vectors of the form*

$$\eta_t = \eta_0 + \int_0^t X(\tau)v(\tau)d\tau, \tag{8}$$

*where* $\eta_0 \in T_{\lambda_0}(T_{q_0}^* M)$ *and* $v \in V$ *satisfying*

$$\int_0^t \left( \sigma\left( \eta_0 + \int_0^\tau X(\theta)v(\theta)d\theta, X(\tau)w(\tau) \right) + b(\tau)(v(\tau), w(\tau)) \right) d\tau = 0, \qquad \forall w \in V. \tag{9}$$

One can use this proposition directly as a working definition. It immediately allows to prove an important property of $\mathcal{L}$-derivatives of optimal control problems.

**Lemma 3.1.** *Let* $0 < t_1 < t_2$ *and assume that* $\mathcal{L}(E_{t_2}, J_{t_2})[\tilde{u}, \lambda]$ *exists. We denote by* $V$ *some finite-dimensional subspace of* $L^\infty([t_1, t_2], \hat{U})$ *and we consider the following equation*

$$\int_{t_1}^{t_2} \left( \sigma\left( \lambda + \int_0^\tau X(\theta)v(\theta)d\theta, X(\tau)w(\tau) \right) + b(\tau)(v(\tau), w(\tau)) \right) d\tau = 0, \qquad \forall w \in V, \tag{10}$$

*where* $v \in V$ *and* $\lambda \in \mathcal{L}(E_{t_1}, J_{t_1})[\tilde{u}, \lambda]$. *Then we can characterize alternatively* $\mathcal{L}(E_{t_2}, J_{t_2})[\tilde{u}, \lambda]$ *as a generalized limit of Lagrangian subspaces*

$$\left\{ \lambda + \int_{t_1}^{t_2} X(\tau)v(\tau)d\tau \ : \ \lambda \in \mathcal{L}(t_1), v \in V \ \textit{satisfy} \ (10) \ \textit{for any} \ w \in V \right\}.$$

This lemma implies that we can construct $\mathcal{L}(E_t, J_t)[\tilde{u}, \lambda]$ by knowing already the same $\mathcal{L}$-derivative at time $\tau < t$. This is obvious when we have a Jacobi DE, since in this case $\mathcal{L}(E_t, J_t)[\tilde{u}, \lambda]$ can be defined using the flow of Jacobi equation.

To include maximum information about the Hessian one should take $L^\infty([0, T], \hat{U})$ as the space of admissible variations, but for simplicity and conceptional clarity we will take a smaller space $\mathcal{U} \subset L^\infty([0, T], \hat{U})$ defined in the following way. If on some interval $\tilde{u}(\tau)$ takes values on the boundary $\partial U$, then we only use time variations. Otherwise we use variations of the control $u$. Therefore we always use *only* one-dimensional variations. We define Jacobi curves as $\mathcal{L}_t = \mathcal{L}(E_t, J_t)[\tilde{u}, \lambda](\mathcal{U} \cap L^\infty([0, t], \hat{U}))$.

In this paper we make the following assumption on the regularity of our system and extremal controls:

**Assumption 1.** Functions $b(t)$, $X(t)$ and the extremal control $\tilde{u}(t)$ are piece-wise analytic as functions of $t$.

**Assumption 2.** We assume that the extremal control $\tilde{u}(t)$ takes values either on a vertex or on an edge of the polytope $U$.

**Assumption 3.** When extremal control $\tilde{u}(t)$ is on a vertex, we only use time variations, when $\tilde{u}(t)$ is on a an edge, we only use variations along that edge.

**Remark 1.** These assumptions are not as restrictives as it might seem. Time variations have no effect if the extremal control $\tilde{u}(\tau)$ is smooth, because in this situation any time variation $\alpha$ can be realised as a variation of the control parameter $u$ [4]. Therefore the effect of time variations is concentrated at discontinuities of the reference control $\tilde{u}(\tau)$. Under the piecewise analyticity assumption we can have discontinuities only at isolated points. The constructed Jacobi curve will be weaker then the one constructed using all possible variations. But even in this case we can obtain useful optimality conditions, and it is possible to generalize this argument to other situations by simply considering any subspace of one-dimensional two-sided variations.

In [3] by using Lemma 3.1 we gave an iterative algorithm for the construction of the Jacobi curve $\mathcal{L}_t$. The idea was to split the time interval $[0, T]$ into small pieces and on each piece to use the space of constant functions as $V$. This way one obtains an approximation of the Jacobi curve which by Theorem 2.1 converges point-wise to it, when the length of the biggest interval of the splitting goes to zero. This algorithm becomes particularly nice if we have a single control parameter, like in the considered problem. In this case $b(\tau)$ is just a function and $X(\tau)$ is a $\mathbb{R}^{2n}$-valued vector function.

**Proposition 3.2.** *Consider a single control parameter system. Given a $\mathcal{L}$-derivative $\mathcal{L}_t(V)$, where $V$ is some space of variations defined on $[0, t]$, we have $\mathcal{L}_t(V \oplus \mathbb{R}\chi_{[t,t+\varepsilon]}) = \mathcal{L}_t(V)^{\eta(t+\varepsilon)}$, where $\chi_{[t,t+\varepsilon]}$ is the characteristic function of the corresponding interval and $\eta(t + \varepsilon)$ is determined by one of the two alternatives*

*1. If*

$$\int_t^{t+\varepsilon} X(\tau)d\tau \in \mathcal{L}_t(V)$$

*then $\mathcal{L}_t(V \oplus \mathbb{R}\chi_{[t,t+\varepsilon]}) = \mathcal{L}_t(V)$ and we can take $\eta(t+\varepsilon)$ to be any vector from $\mathcal{L}_t(V)$*

*2. Else we fix any $\eta(t) \in \mathcal{L}_t(V)$ satisfying*

$$\sigma\left(\eta(t), \int_t^{t+\varepsilon} X(\tau)d\tau\right) \neq 0$$

*and take*

$$\eta(t + \varepsilon) = K\eta(t) + \frac{1}{\varepsilon}\int_t^{t+\varepsilon} X(\tau)d\tau$$

*where*

$$K = -\frac{\frac{1}{\varepsilon}\int_t^{t+\varepsilon}\left[\sigma\left(\int_t^\tau X(\theta)d\theta, X(\tau)\right) + b(\tau)\right]d\tau}{\sigma\left(\eta(t), \int_t^{t+\varepsilon} X(\tau)d\tau\right)}.$$

In [3] we have also proven the following useful lemma.

**Lemma 3.2.** *The Jacobi curve $\mathcal{L}_t$ is left continuous.*

At the end of this section we briefly summarize the key-points regarding $\mathcal{L}$-derivatives.

1. A $\mathcal{L}$-derivative is a map that assigns to a critical point of the functional and a subspace of admissible variations a Lagrangian plane. In the case of optimal control problems proposition 3.1 gives an effective definition;

2. A $\mathcal{L}$-derivative exists if the restriction of the Hessian at the considered critical point has a finite positive or negative inertia index;

3. The set of all $\mathcal{L}$-derivatives over all subspaces of variations contains all the information about nullity, positive and negative inertia indices. By comparing relative positions of the corresponding Lagrangian planes one can track how those numbers change as we add variations;

4. $\mathcal{L}$-derivatives do not change if we replace an infinite-dimensional space of variations by some dense subspace.

Our goal is to specialize the definition of the Jacobi curve $\mathcal{L}_t$ by precomputing the corresponding generalized limits. This way we are going to obtain simple constructions of $\mathcal{L}_t$ in several cases. As a result we will recover some already known results and a completely new dynamical systems characterization of the Jacobi curve in the presence of a singularity that would be hard to guess without this general definition. Using this we can construct Jacobi curves of combinations of different extremals.

# 4 Jacobi DE under the strengthened generalized Legendre condition

Let us assume that the control $\tilde{u}(t)$ takes values in the interior of the set $U$. Our goal is going to be give a characterization of the Jacobi curve using an ODE, like in the classical theory. As we have discussed in the previous sections we will only use variations of the control parameter $u$ to construct the corresponding Jacobi curve.

We consider a sequence of functions $b^i(\tau)$ that we define as

$$b^i(\tau) = \begin{cases} b(\tau), & \text{if } i = 0, \\ \sigma\left(X^{(i+1)}(\tau), X^{(i)}(\tau)\right), & \text{if } i \geq 1. \end{cases}$$

For the sake of simplicity we will often to drop in the future the explicit dependence on time $\tau$ and simply write $b^i$ or $\sigma\left(X^{(i+1)}, X^{(i)}\right)$, when there is no confusion.

The *strengthened Legendre condition of order* $m$ is a series of identities of the form

$$b^m \leq \beta < 0, \qquad b^j \equiv 0, \qquad j < m$$

for some $m \in \mathbb{Z}_{\geq 0}$, where $\beta$ is just a constant. We say that an extremal curve $\tilde{q}(\tau)$ is a *singular curve of order* $m$, if along it the strengthened Legendre condition of order $m$ is

satisfied. If along a trajectory $b^i \equiv 0$ for all $i \in \mathbb{Z}_{\geq 0}$, we say that the trajectory has order infinity.

We define the *Goh subspaces* as

$$\Gamma^i(\tau) = \text{span}\{X^{(j)}(\tau) : j \leq i\}.$$

**Lemma 4.1.** *Assume that the strengthened Legendre condition of order $m$ is satisfied along an extremal trajectory $\tilde{q}(\tau)$. Then $\Gamma^{m-1}(\tau)$ is an isotropic subspace. Moreover $X^m(\tau) \in \Gamma^{m-2}(\tau)^{\angle}$*

*Proof.* The proof is a simple inductive argument. For $i = 1$ the statement is obvious since $\Gamma^1 = X$. Assume that the statement is true for $i < m - 1$. Then in particular we have

$$\sigma\left(X^{(i)}, X^{(j)}\right) = 0, \qquad \forall j < i.$$

Differentiating this identity and using the induction assumption we find that

$$\sigma\left(X^{(i+1)}, X^{(j)}\right) = 0, \qquad \forall j < i.$$

The equality for $j = i$ is obviously true, because $i < m - 1$ and the Legendre condition has order $m$.

The fact that $X^m(\tau) \in \Gamma^{m-2}(\tau)^{\angle}$ now follows from the differentiation of $\sigma(X^{(m-1)}, X^{(i)}) = 0$. $\qquad \square$

We are going to prove the following characterization of the Jacobi curve.

**Theorem 4.1.** *Let $\tilde{q}(\tau)$ be a regular or a singular extremal of order $m$. Then $\mathcal{L}_t$ for $t > 0$ is a linear span of $\Gamma^{m-1}(t)$ and the solutions of the following linear ODE*

$$\dot{\mu} = \frac{\sigma(X^{(m)}, \mu)}{b^m} X^{(m)},$$

*with boundary conditions $\mu(0) \in T_{\lambda(0)}(T_{q_0}^* M) \cap (\Gamma^{m-1}(0))^{\angle}$.*

*If the trajectory has infinite order, then we can define*

$$\Gamma(\tau) = \bigcup_{i=1}^{\infty} \Gamma^i(\tau)$$

*and $\mathcal{L}_t = (T_{\lambda(0)}(T_{q_0}^* M))^{\Gamma(0+)}$.*

*Proof.* The proof is based on the technique called the Goh transformations. The idea of that technique is that if an extremal is singular, then the differentials of maps and quadratic forms in the definition of the $\mathcal{L}$-derivative remain continuous in a much weaker topology. So we can extend this map by continuity to a bigger space, in which the original space is dense. From the Theorem 2.1 we know that this will not change the $\mathcal{L}$-derivative.

Let us assume first that the extremal trajectory is regular. If a vector

$$\eta(t) = \eta + \int_0^t X(\tau)v(\tau)d\tau, \qquad \eta \in \mathcal{L}_0$$

is in $\mathcal{L}_t$, then it satisfies

$$\int_0^t \sigma\left(\eta + \int_0^\tau X(\theta)v(\theta)d\theta, X(\tau)w(\tau)\right) + b(\tau)v(\tau)w(\tau)d\tau = 0, \qquad \forall w \in L^2[0,t] \quad (11)$$

From (11) we get that $v(\tau)$ must satisfy

$$\sigma(\eta(\tau), X(\tau)) + b(\tau)v(\tau) = 0 \iff v(\tau) = -b(\tau)^{-1}\sigma(\eta(\tau), X(\tau)), \qquad \text{a.e.} \tau \in [0,t]$$

But on the other hand from the definition of $\eta(\tau)$ we have

$$\dot{\eta}(\tau) = X(\tau)v(\tau) \qquad \Rightarrow \qquad \dot{\eta}(\tau) = -X(\tau)b(\tau)^{-1}\sigma(\eta(\tau), X(\tau))$$

which gives us the classical Jacobi equation [9].

We assume now that the extremal is singular of order $m$. It is clear that this derivation is not going to work anymore, since $b(\tau) \equiv 0$, so we modify it in the following way. We denote

$$P^m v(t) = \int_0^t \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} v(\tau_m)d\tau_m \dots d\tau_1, \qquad P^m w(t) = \int_0^t \int_0^{\tau_1} \dots \int_0^{\tau_{m-1}} w(\tau_m)d\tau_m \dots d\tau_1$$

the $m$-th primitives of $v$ and $m$. We integrate by parts $m$ times the first summand of (11)

$$\sigma\left(\eta, \int_0^t X(\tau)w(\tau)d\tau\right) = \sigma\left(\eta, \sum_{i=0}^{m-1}(-1)^i X^{(i)}(t)(P^{i+1}w(t)) + (-1)^m \int_0^t X^{(m)}(\tau)(P^m w(\tau))d\tau\right)$$

Now we integrate by parts the other summand of (11). Exchanging the order of integration, using Lemma 4.1 and the assumption on the order of our extremal curve:

$$\int_0^t \sigma\left(\int_0^\tau X(\theta)v(\theta)d\theta, X(\tau)w(\tau)\right)d\tau = \text{(integration by parts)} =$$

$$= \int_0^t \sigma\left(X(\tau)(Pv(\tau)) - \int_0^\tau \dot{X}(\theta)(Pv(\theta))d\theta, X(\tau)w(\tau)\right)d\tau = \text{(exchanging order)} =$$

$$= \int_0^t \sigma\left(-\dot{X}(\tau)(Pv(\tau)), \int_\tau^t X(\theta)w(\theta)d\theta\right)d\tau = \text{(integration by parts)} =$$

$$= \int_0^t \sigma\left(-\dot{X}(\tau)(Pv(\tau)), X(t)(Pw(t)) - X(\tau)(Pw(\tau)) - \int_\tau^t \dot{X}(\theta)(Pw(\theta))d\theta\right)d\tau = \text{(assumption)} =$$

$$= \sigma\left(-\int_0^t \dot{X}(\tau)(Pv(\tau))d\tau, X(t)(Pw(t))\right) + \int_0^t \sigma\left(\int_0^\tau \dot{X}(\theta)(Pv(\theta))d\theta, \dot{X}(\tau)(Pw(\tau))\right)d\tau$$

We continue to integrate by parts both summands and use Lemma 4.1, until the Legendre term $\sigma(X^{(m)}(t), X^{(m-1)}(t))$ will not appear explicitly. At the end we get

$$\int_0^t \sigma\left(\int_0^\tau X(\theta)v(\theta)d\theta, X(\tau)w(\tau)\right)d\tau =$$

$$=\sigma\left((-1)^m \int_0^t X^{(m)}(\tau)(P^m v(\tau))d\tau, \sum_{i=0}^{m-1}(-1)^i X^{(i)}(t)(P^{i+1}w(t))\right)+$$

$$+\int_0^t \sigma\left(X^{(m)}(\tau)(P^m v(\tau)), X^{(m-1)}(\tau)(P^m w(\tau))\right)d\tau+$$

$$+\int_0^t \sigma\left(\int_0^\tau X^{(m)}(\theta)(P^m v(\theta))d\theta, X^{(m)}(\tau)(P^m w(\tau))\right)d\tau$$

Thus (11) is transformed into

$$\sum_{i=0}^{m-1} \sigma\left(\eta + (-1)^m \int_0^t X^{(m)}(\tau)(P^m v(\tau))d\tau, (-1)^i X^{(i)}(t)(P^{i+1}w(t))\right)+$$

$$+\int_0^t \sigma\left((-1)^m\eta + \int_0^\tau X^{(m)}(\theta)(P^m v(\theta))d\theta, X^{(m)}(\tau)(P^m w(\tau)\right)d\tau+ \qquad (12)$$

$$+\int_0^t \sigma\left(X^{(m)}(\tau)(P^m v(\tau)), X^{(m-1)}(\tau)(P^m w(\tau))\right)d\tau = 0$$

We also integrate by parts the integral representation of $\eta(t)$, to get

$$\eta(t) = \eta + \sum_{i=0}^{m-1}(-1)^i X^i(t)(P^{i+1}v(t)) + (-1)^m \int_0^t X^{(m)}(\tau)(P^m v(\tau))d\tau$$

We can see that the right hand side of this expression and quadratic form in (12) are continuous in the topology $\hat{H}^{-m}[0,t]$ given by the norm

$$||v||_{-m} = \sqrt{\sum_{i=0}^{m-1}(P^i v(t))^2 + ||P^m v||_{L^2}^2}.$$

So we extend by continuity on $\hat{H}^{-m}[0,t]$. It is important to note that in the $\hat{H}^{-m}[0,t]$ the end-points $P^i v(t)$ represent separate variables. This implies immediately that $\Gamma^{m-1}(t) \subset \mathcal{L}_t$. Indeed, we can see that the right-hand side of (12) does not depend on $P^i v(t)$ at all. So if we take $\eta = 0$ and $P^m v(\tau) \equiv 0$, then (12) will be satisfied automatically. But then $\eta(t) \in \Gamma^{m-1}(t)$ and every vector of $\Gamma^{m-1}(t)$ can be realized this way.

This means that $\mathcal{L}_t$ actually consists of vectors

$$\mu(t) = \eta + (-1)^m \int_0^t X^{(m)}(\tau)(P^m v(\tau))d\tau$$

16

and vectors from $\Gamma^{m-1}(t)$.

The derivative of $\mu(\tau)$ is given by

$$\dot{\mu}(\tau) = (-1)^m X^{(m)}(\tau)(P^m v(\tau)).$$

We only need to find $P^m v(\tau)$. We do this by solving (12), which gives us a system of equations

$$\sigma\left(\mu(t), X^{(i)}(t)\right) = 0, \qquad 0 \le i \le m - 1, \tag{13}$$

$$(-1)^m \sigma\left(\mu(\tau), X^{(m)}(\tau)\right) + b^m(\tau) P^m v(\tau) = 0, \qquad \text{a.e. } \tau \in [0, t]. \tag{14}$$

Then from the last equation we recover

$$P^m v(\tau) = (-1)^{(m+1)} (b^m(\tau))^{-1} \sigma\left(\mu(\tau), X^{(m)}(\tau)\right)$$

and so

$$\dot{\mu}(\tau) = -X^{(m)}(\tau)(b^m(\tau))^{-1} \sigma\left(\mu(\tau), X^{(m)}(\tau)\right).$$

From (13) we recover boundary conditions

$$\mu(t) \in \Gamma^{m-1}(t)^{\angle}. \tag{15}$$

We can prove that this identity is true not only for the chosen time $t$, but for any $\tau \in [0, t]$. Indeed, from the explicit form of the Jacobi DE we can see that $X^{(m-1)}(\tau)$ is a particular solution. But since all solutions lie in $\mathcal{L}_\tau$ we have $\sigma(\mu(\tau), X^{(m-1)}(\tau)) = 0$ for any solution $\mu(\tau)$. Assume that the same is true for $X^{(i)}(\tau)$, $i \le m - 1$. Then for $X^{(i-1)}(\tau)$ we have

$$\frac{d}{d\tau}\sigma\left(\mu(\tau), X^{(i-1)}(\tau)\right) = -\frac{\sigma\left(X^{(m)}(\tau), \mu(\tau)\right)}{b^m(\tau)}\sigma\left(X^{(m)}(\tau), X^{(i-1)}(\tau)\right) + \sigma\left(\mu(\tau), X^{(i)}(\tau)\right) \equiv 0$$

by Lemma 4.1 and the induction assumption.

So we see that $\mu(\tau) \in \Gamma^{m-1}(\tau)^{\angle}$ is satisfied automatically if $\mu(0+) = (T_{\lambda(0)}(T_{q_0}^* M))^{\Gamma^{m-1}(0)}$. It means that by fixing the appropriate boundary conditions all $n$ independent solutions will lie in $\mathcal{L}_t$.

Let us now look at what can happen if the singularity is of an infinite order. Since we have already established that $\Gamma(\tau)$ is an isotropic subspace, its dimension is limited. This can happen only if higher derivatives of $X$ become dependent from the lower derivatives. Let us assume that the first $l$ derivatives are generically independent and the $(l + 1)$-th is not. Then $\Gamma^l(\tau)$ is a fixed subspace. Indeed, we can represent $\Gamma^l(\tau)$ as an element of $\wedge^l \mathbb{R}^{2n}$

$$\Gamma^l(\tau) = X(\tau) \wedge \dot{X}(\tau) \wedge ... \wedge X^{(l)}(\tau).$$

Then $\dot{\Gamma}^l(\tau) = \kappa(\tau)\Gamma^l(\tau)$ for some function $\kappa(\tau)$. Since $\wedge^l \mathbb{R}^{2n}$ is a linear space, the solution of this equation is simply

$$\Gamma^l(\tau) = e^{\int_s^\tau \kappa(\theta)d\theta}\Gamma^l(s).$$

So we see that $\Gamma(\tau) = \Gamma$ is constant except maybe a finite number of points, where the first $l$ derivatives of $X$ can become dependent. Therefore $\Gamma$ can be taken to be equal to $\Gamma(0+)$.

By assumption on the infinite order we have that the elements of $\mathcal{L}_t$ must satisfy

$$\int_0^t \sigma \left( \eta + \int_0^\tau X^{(l)}(\theta) P^l v(\theta), X^{(l)}(\tau) w(\tau) \right) d\tau = 0.$$

We look for a solution with $P^l v(\tau) = 0$, then the equation above is transformed to

$$\sigma \left( \eta, \int_0^t X^{(l)}(\tau) w(\tau) d\tau \right) = 0.$$

But we have seen that $X^{(l)}(\tau)$ never leaves $\Gamma$. Therefore all $\eta \in (T_{\lambda(0)}(T_{q_0}^* M)) \cap \Gamma^\angle$ satisfy the equation above as well as the boundary conditions (13). Therefore those vectors together with vectors from $\Gamma(\tau)$ give $n$ independent solutions whenever $\dim(\Gamma(\tau))$ is maximal. At the isolated points where the dimension of this space drops we simply use the left-continuity property from Lemma 3.2. $\qquad\square$

# 5 Bang-bang extremals

In the case of bang-bang extremals the extremal control is always on the boundary of $U$. In this case we use time variations to construct the Jacobi curve.

Let
$$f(q, \tilde{u}(t)) =: f_i(q), \qquad L(q, \tilde{u}(t)) =: L_i(q) \qquad t \in [t_i, t_{i+1}].$$
be the controled system and the minimized functional on an interval of constancy $[t_i, t_{i+1}]$ of $u(t)$. Since the system under consideration was autonomous we have that $X(\tau)$ is also a piece-wise constant function. We define $X_i = X(\tau)$, $\tau \in (t_i, t_{i+1})$. These $X_i$ have a particularly nice form, when we consider an optimal time problem. In this case

$$(P_0^\tau)_*^{-1} f(q, \tilde{u}\tau) = e_*^{-t_1 f_1}(e_*^{(t_1 - t_2) f_2}(...(e_*^{(t_{i+1} - t_i) f_{i-1}} f_i)...)), \qquad \tau \in (t_i, t_{i+1}]$$

Let $h_i(\lambda) = \langle \lambda, (P^\tau)_*^{-1} f(q, \tilde{u}(\tau)) \rangle$, $t \in (t_i, t_{i+1}]$. Then $X_i = \vec{h}_i$.

We can now apply proposition 3.2 to find an approximation of the Jacobi curve. We take $V_j$ to be the space of variations constant on the intervals $[t_i, t_{i+1}]$ and which are zero for $t \geq t_j$. We have $\mathcal{L}_t(\{0\}) = T_{\lambda(0)}(T_{q_0}^* M)$ and it is possible now to apply inductively proposition 3.2.

Since the new system and the Lagrangian are linear in controls, we obtain $b(\tau) \equiv 0$. Then

$$\sigma \left( \int_{t_i}^\tau X(\theta) d\theta, X(\tau) \right) + b(\tau) = (\tau - t_i) \sigma(X_i, X_i) = 0, \qquad \forall \tau \in [t_i, t_{i+1})$$

and so $K = 0$ on each step. Similarly we have

$$\eta(t_{i+1}) = \frac{1}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} X(\tau) d\tau = X_i.$$

This way we obtain a sequence of Lagrangian subspace $\mathcal{L}_t(V_j)$, defined inductively as

$$\mathcal{L}_t(V_0) = \mathcal{L}_t(\{0\}) = T_{\lambda(0)}(T_{q_0}^* M), \qquad \mathcal{L}_t(V_{i+1}) = \mathcal{L}_t(V_i)^{X_i}.$$

If we take a finer splitting of the interval the corresponding approximation to the Jacobi curve is the same as above, because $(\mathcal{L}_t(V_i)^{X_i})^{X_i} = \mathcal{L}_t(V_i)^{X_i}$.

The final algorithm of constructing the Jacobi curve goes as follows. One defines $\mathcal{L}_0 = T_{\lambda(0)}(T_{q_0}^* M)$. The Jacobi curves $\mathcal{L}_\tau$ is constant for $\tau \in (t_i, t_{i+1}]$ and after a switching it jumps to $\mathcal{L}(t_i+) = \mathcal{L}(t_i)^{X_i}$. This is the same algorithm that was obtained in [10].

We note that Lemma 3.1 allows to generalize the previous discussion to a wider range of situations. For example, using the results of the previous two sections we can treat the Fuller phenomena, at least when bang-bang arcs are followed by a non-degenerate singular arc. We know from the bang-bang algorithm that in order to construct the Jacobi curve we have to put $X(\tau)$ at each switching time inside of the $\mathcal{L}$-derivative. Left-continuity ensures that this procedure will give a convergent sequence if the index of the Hessian is finite. After the limit $\Lambda$ of the corresponding planes has been found, we can define the Jacobi curve using the flow of the Jacobi DE with the right boundary conditions. From Lemma 3.1 it follows that we have to replace in Theorem 4.1 $T_{\lambda(0)}(T_{q_0}^* M)$ with $\Lambda$.

# 6 Normal form for the Jacobi DE for the simplest singularity

In this section we consider the simplest singularity when $b(\tau) = 0$ for some moment of time $\tau$. Due to analyticity assumption such a moment of time must be isolated. We would like to construct the Jacobi curve after we pass the singularity. In the following sections we are going to prove the following result.

**Theorem 6.1.** *Let $\tilde{q}(\tau)$ be an extremal curve for which $b(\tau) = 0$ and $b(\tau + \varepsilon) = b_m \varepsilon^m + ...$ is negative for all $\varepsilon > 0$ sufficiently small. Assume that at $\tau$ the following conditions are satisfied*

1. *$\sigma(X(\tau), \dot{X}(\tau)) \neq 0$ and $4\sigma(X(\tau), \dot{X}(\tau)) + b_2 \neq 0$ if $m = 2$;*

2. *$\dim \operatorname{span}\{\ddot{X}(s), \dot{X}(s), X(s)\} = const$ for $s \in [\tau, \tau + \varepsilon]$.*

*Then if the right limit of the Jacobi curve $\mathcal{L}(s)$ at $s = \tau$ exists, it is equal to*

$$\mathcal{L}(\tau+) = \mathcal{L}_\tau^{X(\tau)}. \tag{16}$$

Before we proceed we would like to make some remarks about the statement of the theorem.

1. The assumptions we make allow us to have a simplest possible singularity that is in some sense is generic. It is possible to replace those conditions with different ones and to study even more singular cases;

2. The Jacobi curve may not be well defined, due to an infinite inertia index of the Hessian. We will give sufficient conditions for existence and non-existence using oscillation theorems for Hamiltonian systems in Section 8;

3. It is not true that after the singularity the Jacobi curve is determined only by its jump. Indeed, it must satisfy the Jacobi equation, but at the same time the right hand-side of the Jacobi equation is not even continuous. So we do not have neither existence nor uniqueness of solutions and we need more information to isolate the right solution. In Section 9 for $m = 1, 2$ we will prove that Jacobi curve can be uniquely characterized by a one-jet.

Due to analyticity singular points can not cluster. That is why without any loss of generality from now on we assume that $b(0) = 0$ and that $b(\tau) < 0$ for $\tau$ sufficiently small. To give a characterization of the Jacobi curve after a singularity of the considered type, we use once again the theory of $\mathcal{L}$-derivatives.

Even in this case we still have a Jacobi equation of the form

$$\dot{\eta} = \frac{\sigma(X, \eta)}{b} X,$$

and it governs the behaviour of the Jacobi curve away from singularity. In order to understand how to proceed at the singularity, we must recall that by definition $\mathcal{L}$-derivatives are constructed by adding more and more variations and in the limit we get pointwise convergence to the Jacobi curve. Assume that we use only variations whose support does not intersect $[0, \varepsilon]$. Then using the same argument as in the previous section we obtain a slightly different Jacobi equation of the form

$$\dot{\eta} = \begin{cases} \frac{\sigma(X, \eta)}{b} X, & \text{if } \tau \notin [0, \varepsilon], \\ 0, & \text{if } \tau \in [0, \varepsilon]. \end{cases}$$

It is equivalent to the following construction. We will denote by $\Lambda^\varepsilon(\tau)$ a solution of the induced Jacobi equation in the Lagrangian Grassmanian. We use the Jacobi flow to determine the Jacobi curve until time 0. We assume that the left limit exist and is equal to the corresponding $\mathcal{L}$-derivative $\Lambda^\varepsilon(0) = \mathcal{L}_0$. Then the flow does nothing for a while, meaning that $\Lambda^\varepsilon(\tau) = \mathcal{L}_0$ for $\tau \in [0, \varepsilon]$ and then we continue with the flow after the moment of time $\varepsilon$, where the dynamics in non-singular until the next zero of $b(\tau)$. This means that the Jacobi curve is going to be a point-wise limit of solutions of the Jacobi equation on the Lagrangian Grassmanian that satisfy $\Lambda^\varepsilon(\varepsilon) = \mathcal{L}_0$. Since outside of the singularity we have uniqueness and existence, for each $\varepsilon$ we obtain a unique curve in $L(n)$. The pointwise limit of these curves is the Jacobi curve we seek.

To realize this strategy we first simplify the Jacobi DE and reduce the dimension of the considered problem by separating singular and regular dynamics of the Jacobi equation.

Let $J$ be the complex structure associated to the symplectic form $\sigma$. Then we can rewrite the Jacobi DE as

$$\dot{\eta} = \frac{XX^T J}{b}\eta.$$

We make a time-dependent change of variables $\mu(\tau) = M^{-1}(\tau)\eta(\tau)$. We get

$$\dot{\mu} = -M^{-1}\dot{M}\mu + \frac{M^{-1}XX^T JM}{b}\mu.$$

First we look carefully at the second term. We assume that the matrix $M$ is symplectic. Then $M^{-1} = -JM^T J$ and we obtain

$$M^{-1}XX^T JM = (M^{-1}X)(M^T J^T X)^T = (M^{-1}X)(J^T JM^T J^T X)^T = (M^{-1}X)(M^{-1}X)^T J.$$

We also make a choice for the first column of $M$ by assuming

$$M^{-1}(\tau)X(\tau) = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

All this implies that we get an equation of the form

$$\dot{\mu} = -M^{-1}\dot{M}\mu + \frac{1}{b}\left( \begin{array}{c|c} 0 & \begin{matrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{matrix} \\ \hline 0 & 0 \end{array} \right)\mu. \tag{17}$$

Now we work with the first term $M^{-1}\dot{M}$. Since we choose $M$ to be symplectic, its columns form a Darboux basis. Denote the first $n$ columns by $e_i$ and the last by $f_i$. Then we can write

$$-M^{-1}\dot{M} = \begin{pmatrix} \sigma(f,\dot{e}) & \sigma(f,\dot{f}) \\ \sigma(e,\dot{e}) & \sigma(e,\dot{f}) \end{pmatrix} = \begin{pmatrix} \sigma(f,\dot{e}) & \sigma(f,\dot{f}) \\ \sigma(e,\dot{e}) & -\sigma(f,\dot{e}) \end{pmatrix},$$

where $\sigma(x,y)$ means a matrix whose elements are $\sigma(x_i, y_j)$, for two $n$-tuples of vectors $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$. The last equality follows from the fact that the basis is Darboux, i.e.

$$\sigma(e,f) = \mathrm{id}_n \Rightarrow \sigma(\dot{e}, f) + \sigma(e, \dot{f}) = 0.$$

Let us first assume that $n = 1$. Since by assumption of Theorem 6.1 $\sigma(X(0), \dot{X}(0)) \neq 0$, we can choose

$$e(\tau) = e_1(\tau) = X(\tau), \qquad f(\tau) = f_1(\tau) = \frac{\dot{X}(\tau)}{\sigma(X(\tau), \dot{X}(\tau))}, \tag{18}$$

21

The Jacobi DE reduces then to the following normal form

$$
\frac{d}{d\tau}\begin{pmatrix} \mu_1 \\ \mu_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & \dfrac{\sigma(\dot{X},\ddot{X})}{\sigma(X,\dot{X})^2}+\dfrac{1}{b} \\ \sigma(X,\dot{X}) & 0 \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_{n+1} \end{pmatrix} \tag{19}
$$

Let us now assume, that $n \geq 2$. We want to separate the singular dynamics from the regular dynamics. If we look at the singular part in (17), then we see that the only non-zero element is in the first row and $(n+1)$ column. So in the new coordinates we want at least some of the expressions for $\dot{\mu}^i$ for $i \neq 1, n+1$ to be independent of $\mu^1, \mu^{n+1}$. Moreover the assumption, that $M(\tau)$ is symplectic, is going to imply in addition that we will have two invariant symplectic subspaces: a subspace containing $\mu_1, \mu_{n+1}$ coordinates, where the singular dynamics happens and its complement where the dynamics is smooth.

So we look for $e_i$, $f_i$ such that

$$
\sigma(e_i, e_1) = 0, \qquad \sigma(f_i, e_1) = 0, \qquad \sigma(e_i, f_1) = 0, \qquad \sigma(f_i, f_1) = 0. \tag{20}
$$

From the assumption 2) in Theorem 6.1 that we have made it follows that $\dim \operatorname{span}\{\ddot{X}(\tau), \dot{X}(\tau), X(\tau)\}$ must be equal either to two or three. In the first case $\ddot{X}(\tau)$ is simply in the span of $X(\tau)$, $\dot{X}(\tau)$ for small $\tau \geq 0$. So we make the same choice (18) for $e_1, f_1$ and the rest of the columns we take to be a smooth Darboux basis for the symplectic space $\operatorname{span}\{X(\tau), \dot{X}(\tau)\}^{\angle}$. Then the conditions (20) are indeed satisfied and we obtain exactly the equation (19) for the singular part.

In the second case we can not guarantee that conditions (20) are satisfied for $i \geq 2$, since now $\ddot{X}(\tau)$ has to be accounted for. To isolate the singular dynamics we choose a Darboux basis $e_1, e_2, f_1, f_2$ as follows. The vectors $e_1, f_1$ are as before, $e_2$ is defined as

$$
e_2 = \ddot{X} - \frac{\sigma(\ddot{X}, X)}{\sigma(\dot{X}, X)}\dot{X} + \frac{\sigma(\ddot{X}, \dot{X})}{\sigma(\dot{X}, X)}X.
$$

and $f_2$ is chosen to be any vector such that we get a Darboux basis. In this case it just means that

$$
\sigma(X, f_2) = \sigma(\dot{X}, f_2) = \sigma(\ddot{X}, f_2) - 1 = 0.
$$

The rest of the columns of $M_\tau$ are chosen to be a smooth Darboux basis of the symplectic space $(\operatorname{span}\{e_1, e_2, f_1, f_2\})^{\angle}$. Again, the derivatives of $e_1$, $f_1$ are contained in $\operatorname{span}\{e_1, e_2, f_1, f_2\}$, so the dynamics splits. The singular dynamics takes place in the plane with $(\mu_1, \mu_2, \mu_{n+1}, \mu_{n+2})$ coordinates. Thus we get an invariant subsystem containing the singular part

$$
\frac{d}{dt}\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{n+1} \\ \mu_{n+2} \end{pmatrix} = \begin{pmatrix} \sigma(f_1,\dot{e}_1) & \sigma(f_1,\dot{e}_2) & \sigma(f_1,\dot{f}_1)+\frac{1}{b} & \sigma(f_1,\dot{f}_2) \\ \sigma(f_2,\dot{e}_1) & \sigma(f_2,\dot{e}_2) & \sigma(f_2,\dot{f}_1) & \sigma(f_2,\dot{f}_2) \\ \sigma(e_1,\dot{e}_1) & \sigma(e_1,\dot{e}_2) & \sigma(e_1,\dot{f}_1) & \sigma(e_1,\dot{f}_2) \\ \sigma(e_2,\dot{e}_1) & \sigma(e_2,\dot{e}_2) & \sigma(e_2,\dot{f}_1) & \sigma(e_2,\dot{f}_2) \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{n+1} \\ \mu_{n+2} \end{pmatrix}.
$$

and so we have proven our first result about the jump of the Jacobi curve $\mathcal{L}_\tau$

**Proposition 6.1.** *If $b(\tau) = 0$, then $\dim(\mathcal{L}_\tau \cap \mathcal{L}_{\tau+}) \geq n - 2$.*

Let us simplify this equation even more. Since $M^{-1}\dot{M}$ is a matrix from the symplectic Lie algebra, not all the entries above are independent. More precisely, the first and the last diagonal 2x2 minors are minus transpose of each other and the off diagonal 2x2 minors are symmetric. We can find explicitly

$$\sigma(f_1, \dot{e}_1) = \sigma(f_1, \dot{e}_2) = \sigma(f_2, \dot{e}_1) = \sigma(e_1, \dot{e}_2) = 0;$$

$$\sigma(e_1, \dot{f}_1) = \sigma(e_1, \dot{f}_2) = \sigma(e_2, \dot{f}_1) = \sigma(e_2, \dot{e}_1) = 0;$$

$$\sigma(f_2, \dot{f}_1) = \sigma(f_1, \dot{f}_2) = -\frac{1}{\sigma(X, \dot{X})};$$

$$\sigma(f_2, \dot{e}_2) = -\sigma(e_2, \dot{f}_2) = \sigma(f_2, \dddot{X}) + \frac{\sigma(\ddot{X}, X)}{\sigma(\dot{X}, X)};$$

$$\sigma(e_1, \dot{e}_1) = \sigma(X, \dot{X});$$

$$\sigma(e_2, \dot{e}_2) = \sigma(\ddot{X}, \dddot{X}) - \frac{\sigma(\dddot{X}, X)\sigma(\ddot{X}, \dot{X}) + \sigma(\ddot{X}, X)\sigma(\dddot{X}, \dot{X})}{\sigma(\dot{X}, X)};$$

$$\sigma(f_1, \dot{f}_1) = \frac{\sigma(\dot{X}, \ddot{X})}{\sigma(X, \dot{X})^2}.$$

So we get an equation of the form

$$\frac{d}{dt}\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{n+1} \\ \mu_{n+2} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \sigma(f_1, \dot{f}_1) + \frac{1}{b} & \sigma(f_2, \dot{f}_1) \\ 0 & \sigma(f_2, \dot{e}_2) & \sigma(f_2, \dot{f}_1) & \sigma(f_2, \dot{f}_2) \\ \sigma(e_1, \dot{e}_1) & 0 & 0 & 0 \\ 0 & \sigma(e_2, \dot{e}_2) & 0 & -\sigma(f_2, \dot{e}_2) \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{n+1} \\ \mu_{n+2} \end{pmatrix}. \quad (21)$$

Note that if $\sigma(f_2, \dot{f}_1) \equiv 0$ for $\tau$ small enough, we obtain a $n = 1$ normal form of the Jacobi DE as a subsystem. So without any loss of generality from now on we can assume that $n \geq 2$.

We can simplify the last equation even more by taking

$$Q(\tau) = \begin{pmatrix} 1 & 0 \\ 0 & \exp\left(\int_0^\tau \sigma(\dot{f}_2(s), e_2(s))ds\right) \end{pmatrix}.$$

Note that $Q(0) = \mathrm{id}_n$. We introduce new variables

$$\begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} Q^{-1} & 0 \\ 0 & Q \end{pmatrix}\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_{n+1} \\ \mu_{n+2} \end{pmatrix}.$$

If we write $p = (p_1, p_2)$, $q = (q_1, q_2)$, we obtain a normal form

$$\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & \frac{B(\tau)}{\tau^m} \\ C(\tau) & 0 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} \tag{22}$$

where $m$ is the first non zero coefficient of the Taylor expansion of $b(\tau)$,

$$C(\tau) = Q\begin{pmatrix} \sigma(e_1, \dot{e}_1) & 0 \\ 0 & \sigma(e_2, \dot{e}_2) \end{pmatrix}Q = \begin{pmatrix} c_{11}(\tau) & 0 \\ 0 & c_{22}(\tau) \end{pmatrix}$$

and

$$\frac{B(\tau)}{\tau^m} = \begin{pmatrix} 1/b & 0 \\ 0 & 0 \end{pmatrix} + Q^{-1}\begin{pmatrix} \sigma(f_1, \dot{f}_1) & \sigma(f_2, \dot{f}_1) \\ \sigma(f_2, \dot{f}_1) & \sigma(f_2, \dot{f}_2) \end{pmatrix}Q^{-1} = \begin{pmatrix} \frac{1}{b(\tau)} + b_{11}(\tau) & b_{12}(\tau) \\ b_{12}(\tau) & b_{22}(\tau) \end{pmatrix}$$

Note that (22) is still a Hamiltonian system and a normal form for the singular part of the Jacobi DE. Moreover we can choose our frame so that $B(\tau)$ is positive for small $\tau > 0$. Indeed, this is going to be true if the trace and the diagonal entries are non-negative. Since $b(\tau) = b_m\tau^m + O(\tau^{m+1}) < 0$ for $\tau > 0$ small and the frame $\{e_i, f_i\}$ was chosen to be analytic, we obviously have that the trace and the upper diagonal element are negative for $\tau > 0$ sufficiently small. We claim that $f_2$ can be chosen in such a way that also the lower diagonal term is negative as well for small $\tau > 0$. Indeed, the only freedom that we have is to replace $f_2(\tau)$ with $f_2(\tau) + a(\tau)e_2(\tau)$ for some analytic function $a(\tau)$. Then we have

$$\sigma(f_2 + ae_2, \dot{f}_2 + \dot{a}e_2 + a\dot{e}_2) = \sigma(f_2, \dot{f}_2) - \dot{a} + a^2\sigma(e_2, \dot{e}_2),$$

where we have used that $\sigma(e_2, f_2) = 1$ and $\sigma(\dot{e}_2, f_2) + \sigma(e_2, \dot{f}_2) = 0$. Then it is clear that we can simply choose $a(t) = \sin kt$, with $k$ sufficiently large and the explicit form of $Q$ implies that we can assume without any loss of generality that $B(\tau)$ is negative for $\tau$ small.

Before we start proving Theorem 6.1, it is very helpful to understand the idea of the proof using some simple heuristics in the case $n = 1$. Later we will make all the steps rigorous. In the next section we will see under which conditions $\mathcal{L}$-derivatives exist and how to characterize the Jacobi curve as a solution to the singular Jacobi equation with certain boundary conditions when $n = 1$.

# 7 An heuristic argument for $n = 1$

Assume that $n = 1$ and let $b_m\tau^m + \ldots$ be the right series of $b(\tau)$ at $\tau = 0$, with $b_m < 0$. We would like to determine whether or not $\mathcal{L}_{0+}$ exists at all. To do this we rewrite (19) as a second order ODE of the form

$$\ddot{\mu}_1 + a_1(\tau)\dot{\mu}_1 + a_0(\tau)\mu_1 = 0. \tag{23}$$

We say that this equation is oscillating on a given interval, if any solution $\mu_1(\tau)$ has an infinite number of zeroes on that interval. Equivalently the classical Sturm theory of

second order ODEs implies that this equation is oscillating whenever any solution of (19) makes an infinite number of turns around the origin in the $(\mu_1, \mu_{n+1})$ plane. Recall that for $n = 1$ the Lagrange Grassmanian is nothing but a projective line $\mathbb{P}^1$ and that the Jacobi curve is just the line $\mathbb{R}(\mu_1, \mu_{n+1})$. Therefore Jacobi curves of oscillating equations have an infinite Maslov index.

For second order ODEs there exist various oscillation and non-oscillation criteria, but among them there is a particularly simple one called the Kneser criteria [27]. It states that a second order ODE of the form

$$\ddot{x} + a(s)x = 0 \tag{24}$$

is oscillating on $[p, \infty)$, for any $p > 0$ if

$$\lim_{s \to \infty} s^2 a(s) > \frac{1}{4}$$

and it is non-oscillating if

$$\lim_{s \to \infty} s^2 a(s) < \frac{1}{4}.$$

If the limit is exactly $1/4$, Kneser criteria gives us no information and we have to use a different criteria. In order to put the equation (23) into form (24), we simply make a change of the time variable $s = 1/\tau$ and then make a change of the dependent variable

$$x(s) = \mu(s) \exp\left(\int_p^s \frac{2\theta - a_1(\theta)}{2\theta^2} d\theta\right).$$

Then we obtain exactly equation (24).

After applying the Kneser criteria, we find that for any sufficiently small interval $[0, \varepsilon]$

1. equation (23) is oscillating if $m = 2$ and $4\sigma(X(0), \dot{X}(0)) + b_2 > 0$ or if $m > 2$ and $\sigma(X(0), \dot{X}(0)) > 0$;

2. equation (23) is non-oscillating if $1 \leq m < 2$ or $m = 2$ and $4\sigma(X(0), \dot{X}(0)) + b_2 < 0$, or if $m > 2$ and $\sigma(X(0), \dot{X}(0)) >< 0$.

Thus we conclude that in the first case the Jacobi curve has no right limit and therefore it also has an infinite Maslov index.

Since we just want to give an idea of how the proof works, we assume that the Jacobi DE (19) is of the simplest form

$$\frac{d}{d\tau}\begin{pmatrix} \mu_1 \\ \mu_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\tau^2} \\ C & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_{n+1} \end{pmatrix}, \tag{25}$$

where $C$ is constant. We then make a time-dependent change of variables

$$\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \tau^{1/2} & 0 \\ 0 & \tau^{-1/2} \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_{n+1} \end{pmatrix}$$

and obtain

$$\frac{d}{d\tau}\begin{pmatrix} p \\ q \end{pmatrix} = \frac{1}{\tau}\begin{pmatrix} \frac{1}{2} & 1 \\ C & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} = -\tau^{-1}JH\begin{pmatrix} p \\ q \end{pmatrix} \tag{26}$$

Note that this change of variables can not change whether or not the Maslov index is finite. It is clear that $\tau^{-1}$ multiplier just scales the speed along solutions, but does not change the trajectories. We could have actually got rid off it using a change of time variable. This means that if we drop $\tau^{-1}$ in the equation above, the overall phase-portrait does not change. It will be completely determined by the structure of the matrix $-JH$.

A description of various phase portraits on the Lagrangian Grassmanian was given in [24]. We will use the results from that article to work out the general case. For example, for $n = 1$ we can only have fixed points or periodic trajectories, which depend on the eigenvalues and eigenvectors of the matrix $-JH$. Its eigenvalues are

$$\lambda_{1,2} = \pm\frac{\sqrt{1 + 4C}}{2}.$$

If $1 + 4C < 0$, then we have only a single closed trajectory and no equilibrium points. Thus the trajectory rotates on $\mathbb{P}^1$, and because of the $\tau^{-1}$ multiplier in (26) the curve rotates faster and faster as we get closer to $\tau = 0$ and therefore we get an infinite Maslov index. If $1 + 4C > 0$, then we have two equilibrium points: a stable and a non-stable one, that are given by two lines spanned by the eigenvectors of $-JH$. Thus all solutions except the equilibrium ones tend to the unstable equilibrium as $\tau \to 0$ and to the stable one as $\tau \to \infty$. In this case the Maslov index is finite. Note that in our example $C = \sigma(X(0), \dot{X}(0))/b_2$ and thus we recover the classical Kneser criteria. From here we can also see very well why the case $1 + 4C = 0$ is excluded. It is not stable under small perturbations and corresponds to a resonant situation when the two equilibrium points merge.

Having a small dimensional situation allows us to actually draw the extend-phase portrait. We introduce new variables $U$ and $V$ defined as $\mu_{n+1} = U\mu_1$, $p = Vq$. It is clear from the definitions that $V = \tau U^{-1}$. Since we work in a coordinate chart of the Grassmanian we can assume that $U, V \neq 0$. We differentiate expressions in the definition to obtain a couple of related Riccati equations

$$\dot{U} = C - \frac{U^2}{\tau^2},$$

$$\dot{V} = \frac{1 + V - CV^2}{\tau}.$$

In the picture 1 the non-oscillating extended phase portrait before and after the change of variables is depicted. We can see clearly, that the extended phase portrait is separated by two separatrix into three regions. After the blow up the two separatrix solutions have different initial values correspond to the equilibrium solutions. The stable solution after a blow-up can be described using an initial value problem, where as the unstable one can not. There is an infinite number of solutions that start from the unstable equilibrium.
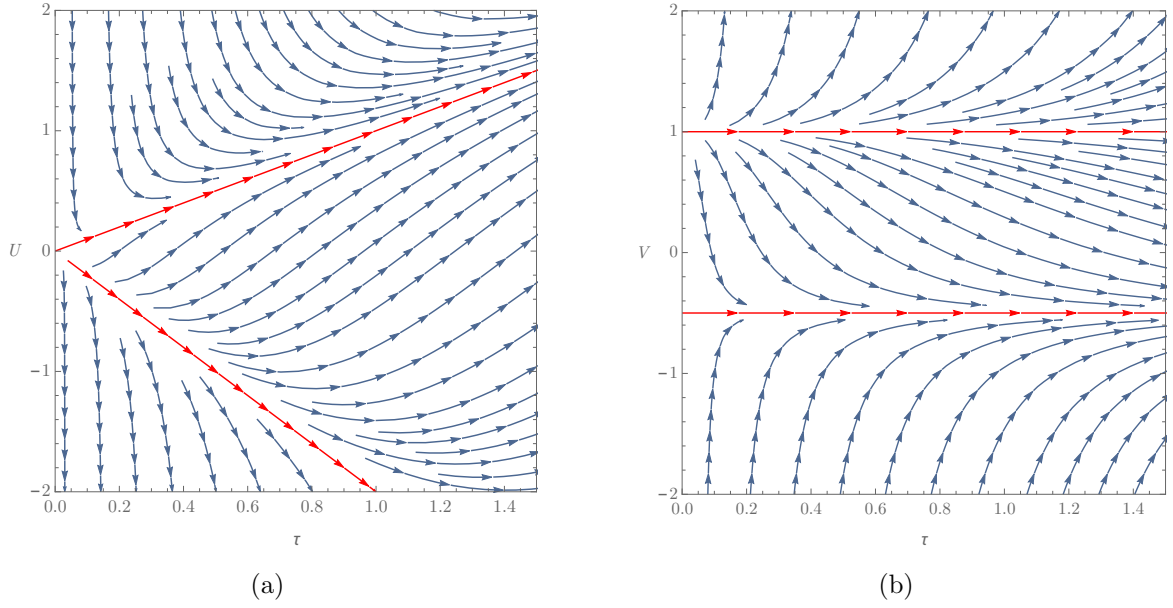
Figure 1: Phase portrait of a non-oscillating system: (a) before blow-up; (b) after blow-up.

We claim that the stable separatrix is the Jacobi curve for $\tau > 0$. To see this we do as discussed in the previous subsection. Assume that $\mathcal{L}_0$ is given by $U_0 \neq 0$. Then the Jacobi curve is the limit of solutions of the Riccati equations with $U(\varepsilon) = U_0$. Similarly after a blow up it corresponds to a limit of solutions with boundary conditions $V(\varepsilon) = \varepsilon U_0^{-1}$. On the picture 7 we can see this convergence numerically in the original phase portrait.
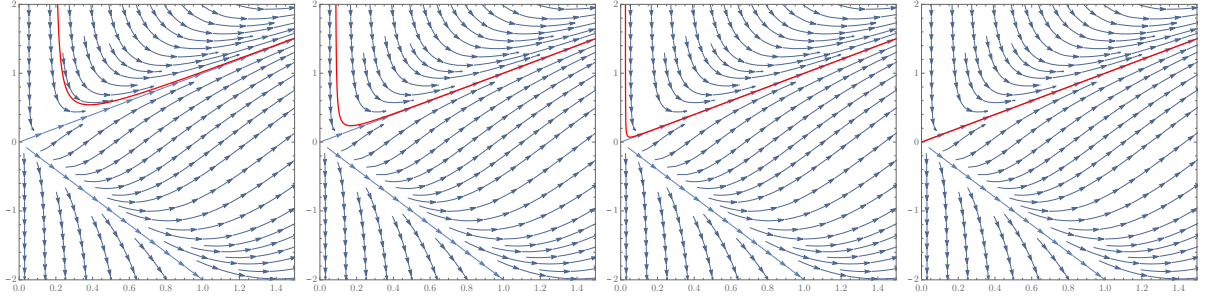


Figure 2: Point convergence of $\Lambda(\tau)$ to the Jacobi curve for $C = 2$.

In the next sections our goal is to make all the ideas from this section rigorous. Our proves are mostly perturbative and we will first prove them for the constant matrix case and then expand it to the general case. Following the outline of this section we first prove an analogue the Kneser criteria and identify the non-oscillating cases. Such a criteria is a necessary optimality condition on its own. Then using the general theory of ODEs with singular regular points we are going to characterize the jump of the Jacobi curve. Finally using the results from article [22] we are going to characterize the first derivative of the Jacobi curve for $m = 1, 2$ and prove that conditions on the first derivative guarantee uniqueness of the solution of the extended Jacobi equation that characterizes $\mathcal{L}_t$.

# 8  Kneser oscillation criteria for a simple class of Hamiltonian systems

In the case $n = 1$, the Kneser criteria gives sufficient conditions under which a second order ODE is oscillating or non-oscillating. This result is just a consequence of the Sturm comparison theorem and an explicit solution of an Euler-type linear equation [27].

Kneser criteria can be derived as a consequence of more general integral criteria and the modern theory of oscillation of ODE systems tends to generalize those. As a result, we were not able to find in the literature a similar simple point-criteria. Thus we would like to slightly generalize the Kneser criteria to a special class of Hamiltonian systems that include system (22). Our main tool will be a straightforward consequence of Theorem 1 in [11, 7], that can be seen as a generalization of the Strum-comparison theorem

**Theorem 8.1.** *Let $I$ be an open interval and $A_i(\tau), B_i(\tau), C_i(\tau)$, $i = 1, 2$ quadratic matrices whose elements are differentiable on $I$. Assume that $B_i(t)$ and $C_i(t)$ are symmetric. We consider the corresponding Hamiltonians*

$$H_i = \begin{pmatrix} C_i(\tau) & -A_i^T(\tau) \\ -A_i(\tau) & -B_i(\tau) \end{pmatrix}$$

*If the Hamiltonians $H_i(\tau)$ satisfy*

$$H_2(\tau) \geq H_1(\tau)$$

*then for any two trajectories $\Lambda_i(t)$ whose endpoints are transversal to $\Lambda$, we have the following inequality*

$$\mathrm{Mi}_\Lambda \, \Lambda_1(t) - n \leq \mathrm{Mi}_\Lambda \, \Lambda_2(t).$$

We are going to use a simple direct corollary of that result

**Corollary 8.1.** *Let $I$ be an open interval and $A_i(\tau), B_i(\tau), C_i(\tau)$, $i = 1, 2$ quadratic matrices whose elements are differentiable on any compact subset $[a, b] \subset I$. Assume that $B_i(t)$ and $C_i(t)$ are symmetric and denote by $H_i(\tau)$ the corresponding Hamiltonians, s.t. $H_2(\tau) \geq H_1(\tau)$. Then we have the following implications:*

1. *If $\exists \Lambda \in L(n)$, s.t. $H_2(\tau)|_\Lambda \leq 0$ for all $\tau \in I$ and the second system is oscillating, then the first system is oscillating as well;*

2. *If $\exists \Lambda \in L(n)$, s.t. $H_1(\tau)|_\Lambda \geq 0$ for all $\tau \in I$ and the first system is oscillating, then the second system is oscillating as well;*

3. *If $\exists \Lambda \in L(n)$, s.t. $H_i(\tau)|_\Lambda \leq 0$ for all $\tau \in I$ and the first system is non-oscillating, then the second system is non-oscillating as well;*

4. *If $\exists \Lambda \in L(n)$, s.t. $H_i(\tau)|_\Lambda \geq 0$ for all $\tau \in I$ and the second system is non-oscillating, then the first system is oscillating as well;*

*Proof.* The proof is just a corollary of Theorem 8.1 and Theorem 1.2. For example, in the first case Theorem 1.2 implies that

$$\text{Mi}_\Lambda \, \Lambda_2(\tau) \le 0.$$

The assumption that the corresponding system is oscillating means that the Maslov index of any solution is infinite. Then from the comparison Theorem 8.1 we obtain

$$\text{Mi}_\Lambda \, \Lambda_1(\tau) \le \text{Mi}_\Lambda \, \Lambda_2(\tau) + n = -\infty.$$

The remaining implications are proven in the same way. $\qquad\square$

The goal of this section is to prove the following result

**Theorem 8.2.** *Consider a Hamiltonian system*

$$\frac{d}{d\tau} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} A(\tau) & \frac{B(\tau)}{\tau^m} \\ C(\tau) & -A^T(\tau) \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix},$$

*s.t. the following assumptions are satisfied*

1. *$B(\tau)$ is a semi-definite smooth symmetric $n \times n$-matrix, s.t. $B(\tau)$ is sign-definite for $\tau > 0$;*

2. *$C(\tau)$ is a smooth symmetric $n \times n$-matrix;*

3. *$A(\tau)$ an arbitrary smooth $n \times n$-matrix;*

*Then the following statements are true:*

1. *Let $m = 2$. If all the eigenvalues of the matrix $C(0)B(0)$ are strictly greater than $-1/4$, then the system is non-oscillating on $(0, \varepsilon)$. If at least one eigenvalue is smaller than $-1/4$, then the system is oscillating on the same interval;*

2. *If $0 \le m < 2$, then the system is not oscillating on $(0, \varepsilon)$;*

3. *Let $m > 2$. If $C(0)$ is sign definite on the eigenspace that is transversal to the kernel of $B(0)$ with the same sign as $B(0)$, then the system is non-oscillating on $(0, \varepsilon)$. If $C(0)$ is not semi-definite on this subspace with the same sign as $B(0)$, then the system is oscillating on the same interval.*

We are pretty sure that this theorem can be derived as a consequence of some existing oscillation criteria for Hamiltonian systems, but we prefer to give here a simple geometric proof of the result using Theorem 8.1, that seems to be new.

As it can be seen from the statement the matrix $A$ plays no essential role. Similarly to a change of variables in the Section 7, we make a time-dependent change of variables

$$\begin{pmatrix} p \\ q \end{pmatrix} \mapsto \begin{pmatrix} \Phi & 0 \\ 0 & (\Phi^{-1})^T \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}$$

29

where $\Phi(t)$ satisfies

$$\begin{cases} \dot{\Phi} = A\Phi, \\ \Phi(0) = \mathrm{id}_n\,. \end{cases}$$

Then $\Phi(t)$ is the fundamental matrix of the corresponding linear equation and it is smooth. Therefore our change of variables is a non-degenerate symplectic change of variables and it does not change oscillatory properties of the Hamiltonian systems.

Our Hamiltonian system now takes the form

$$\frac{d}{d\tau}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & \frac{\Phi^{-1}B(\Phi^{-1})^T}{\tau^m} \\ \Phi^T C\Phi & 0 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix},$$

We can now simply redefine matrices $B$ and $C$. This implies that without any loss of generality, we can assume that $A(\tau) \equiv 0$.

In order to apply the comparison Theorem 8.1, we need a model example, which oscillating properties we understand very well. Such model is given in the next lemma

**Lemma 8.1.** *Consider a Hamiltonian system of the form*

$$\frac{d}{d\tau}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & \frac{B}{\tau^2} \\ C & 0 \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}, \tag{27}$$

*where $B$ and $C$ are constant symmetric matrices. This Hamiltonian system is oscillating on an interval $(0,\varepsilon)$ if and only if there exists at least one real eigenvalue $\lambda$ of the matrix $BC$, s.t. $\lambda < -1/4$.*

This result is a consequence of the following theorem proven in [8].

**Theorem 8.3.** *A linear autonomous Hamiltonian system*

$$\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = -JH\begin{pmatrix} p \\ q \end{pmatrix},$$

*is oscillating on an unbounded interval if and only if the matrix $-JH$ has a purely imaginary eigenvalue.*

*Proof of the Lemma 8.1.* We do another symplectic transformation of the form

$$\begin{pmatrix} p \\ q \end{pmatrix} \mapsto \begin{pmatrix} \tau^{-1/2} & 0 \\ 0 & \tau^{1/2} \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix}$$

The transformation is smooth for $\tau > 0$ and therefore oscillating property is preserved. Our Hamiltonian system then becomes

$$\tau\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\,\mathrm{id}_n & B \\ C & -\frac{1}{2}\,\mathrm{id}_n \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} = -JH\begin{pmatrix} p \\ q \end{pmatrix}$$

Let us perform a change of time variable

$$s = \ln \tau.$$

Then we obtain a linear autonomous Hamiltonian system of the form

$$\frac{d}{ds} \begin{pmatrix} p \\ q \end{pmatrix} = -JH \begin{pmatrix} p \\ q \end{pmatrix}$$

Note that the change of time that we have made, maps the bounded interval $(0, \varepsilon)$ to an unbounded one. So by the Theorem 8.3 it just remains to compute the eigenvalues of the matrix $-JH$, i.e. to solve

$$\det(-JH - \lambda \operatorname{id}_{2n}) = 0.$$

In this case the diagonal blocks are multipliers of the identity and hence commute with all the other blocks. Under this assumption it is easy to show that

$$\det(-JH - \lambda \operatorname{id}_{2n}) = \det\left(\left(\lambda^2 - \frac{1}{4}\right) - BC\right).$$

If a matrix $-JH$ has a pair of purely complex eigenvalues $\lambda = \pm ib$, we obtain that the matrix $BC$ has an eigenvalue $-b^2 - 1/4 < -1/4$. It is obvious that the converse holds as well. So the result follows from Theorem 8.3. $\qquad \square$

Finally we need the following fact proven in (theorems in Horn's Matrix Analysis):

**Theorem 8.4.** *Let $B, C$ be two constant symmetric matrices, s.t. one of them is semidefinite. Then the spectrum of $BC$ is real.*

*Proof of the Theorem 8.2.* We assume that $B(\tau) < 0$ for sufficiently small $\tau > 0$. The case $B(\tau) > 0$ is proven in a similar way. In this case the corresponding Hamiltonian $H$ is positive semidefinite on the horizontal plane $\Sigma$ (the $q$-plane).

1) Let us start with the case $m = 2$. We know by the previous theorem that all the eigenvalues of $B(0)C(0)$ are real, and let us assume first that the minimum one is strictly less then $-1/4$. We define

$$B_1(\tau) = B(0) + \varepsilon \operatorname{id}_n, \qquad\qquad B_2(\tau) = B_\tau,$$
$$C_1(\tau) = C(0) - \varepsilon \operatorname{id}_n, \qquad\qquad C_2(\tau) = C(\tau).$$

By assumption of the theorem and Lemma 8.1 the Hamiltonian system (27) with $B = B(0)$ and $C = C(0)$ is oscillating. This implies that a system of the form with $B = B(0) + \varepsilon \operatorname{id}_n$ and $C = C(0) - \varepsilon \operatorname{id}_n$ must be oscillating as well. Indeed, eigenvalues of $BC$ are solutions of the characteristic equation whose coefficients depend continuously on the coefficients of matrices $B, C$. Therefore a small perturbation of matrices produces a small change in the eigenvalues of $BC$. But we have chosen such a perturbation in a way, that according to Theorem 8.4, the spectrum of $BC$ remains real. So all the eigenvalues shift on the real axis, and if we choose $\varepsilon > 0$ small enough, the minimum eigenvalue of the

perturbed matrix will stay strictly smaller then $-1/4$ and the corresponding Hamiltonian system (27) stays oscillating by Lemma 8.1.

So we can use the implication 2) of Corollary 8.1. By smoothness assumption, indeed, for sufficiently small times $H_1(\tau) \leq H_2(\tau)$.

2) The non-oscillating case for $m = 2$ is proven using exactly the same argument and matrices

$$
\begin{aligned}
B_1(\tau) &= B(\tau), & B_2(\tau) &= B(0) - \varepsilon \operatorname{id}_n, \\
C_1(\tau) &= C(\tau) & C_2(\tau) &= C(0) + \varepsilon \operatorname{id}_n .
\end{aligned}
$$

3) The case $0 \leq m < 2$ is now just a consequence of what we have proven so far. Indeed, we can consider the Hamiltonian as having a singularity with $m = 2$ and with a new matrix $\hat{B}(\tau) = \tau^{2-m} B(\tau)$. Then $\hat{B}(0)C(0) = 0$ and all the eigenvalues are zero. Hence the system is not oscillating.

4) Let us now assume that $m > 2$. We first we apply a symplectic transform

$$
\begin{pmatrix} R & 0 \\ 0 & R^T \end{pmatrix},
$$

where $R \in \operatorname{SO}(n)$. Then in the new coordinates $B(\tau)$ and $C(\tau)$ will be replaced by the same matrices conjugated with $R$. Let us choose $R$, s.t. in the new coordinates $B(0)$ is diagonalized.

We take

$$
\begin{aligned}
B_1(\tau) &= -k\tau^{m-2} \begin{pmatrix} \operatorname{id}_l & 0 \\ 0 & 0 \end{pmatrix} & B_2(\tau) &= B(\tau), \\
C_1(\tau) &= C(\tau), & C_2(\tau) &= C(\tau),
\end{aligned}
$$

where $k$ is some constant. If $C(0)$ is not negative semi-definite on the eigenspace of $B(0)$ that corresponds to the non-zero eigenvalue, then by taking $k$ large enough, we find that the system one is oscillating by Lemma 8.1. We also have that $B_1(\tau) - B_2(\tau)$ is non-negative for sufficiently small $\tau > 0$. Thus we can use the implication 2) of Corollary 8.1 to deduce that the second system is going to be oscillating as well.

If $C(0)$ is negative definite on the eigenspace of $B(0)$, we repeat the proof with exactly the same $B_1$, $C_1$. In this case we know from what we have proven already, that the system one is not oscillating. Then the result follows from the implication 3) from Corollary 8.1 with $\Lambda = \Pi$ being the vertical plane (the $p$-plane).

$\square$

Let us apply the theorem to our case. We are not oscillating for $0 \leq m < 2$ and for $m = 2$ whenever the eigenvalues of the

$$
B(0)C(0) = \begin{pmatrix} \frac{1}{b_2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sigma(e_1(0), \dot{e}_1(0)) & 0 \\ 0 & \sigma(e_2(0), \dot{e}_2(0)) \end{pmatrix} = \begin{pmatrix} \sigma(X(0), \dot{X}(0))/b_2 & 0 \\ 0 & 0 \end{pmatrix}
$$

are bigger then $-1/4$, i.e. whenever the only non-zero element above is bigger then $-1/4$. If $m > 2$ then we must have $\sigma(X(0), \dot{X}(0)) < 0$.

So from now on, we assume that our Hamiltonian system is non-oscillating, ensuring that the right limit $\mathcal{L}_{0+}$ exists.

# 9    Computing the jump when $n \geq 2$

We are now ready to make the first step and compute the jump of the Jacobi curve. We will give three similar but separate proofs for $m = 1$, $m = 2$ and $m \geq 3$. For $m = 1$ we will do this using the general theory of linear ODEs with regular singular points. For $m = 2$ and $m \geq 3$ the strategy of the proof is going to be very similar to the proof of the Kneser theorem in the previous section. Namely we first look at some model examples and then we use the comparison theory of Riccati equations, to obtain the result in the most general case.

## 9.1    Jump for $m = 1$

Let $\lambda = (p, q)$. We rewrite the system (22) in the following form

$$\dot{\lambda} = \left( \frac{H_{-1}}{\tau} + H(\tau) \right) \lambda, \tag{28}$$

where $H(\tau)$ is an analytic matrix function and as can be easily seen

$$H_{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{b_1} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This matrix up to a reordering of coordinates is in its Jordan normal form, and all of its eigenvalues are zero. Therefore by a well known theorem [15], the fundamental matrix $\Phi(\tau)$ of the system (28) can be written as

$$\Phi(\tau) = P(\tau) \tau^{H_{-1}},$$

where $P(\tau)$ is an analytic matrix function with $P(0) = \mathrm{id}_{2n}$. A power series expansion can be obtained by plugging this solution into (28) and expanding all the analytic functions into their Taylor series. It is easy to check that

$$\tau^{H_{-1}} = \begin{pmatrix} 1 & 0 & \frac{\ln \tau}{b_1} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Let $\mathcal{L}_0$ be the $\mathcal{L}$-derivative at moment of time $\tau = 0$. The flow of the Hamiltonian system $\Phi(\tau)$ induces a flow on the Lagrangian Grassmanian $L(2)$ that we denote using

33

the same symbol. As we have discussed previously the Jacobi curve $\mathcal{L}_\tau$ is going to be a pointwise limit of the solutions of the Jacobi DE on $L(2)$ with boundary conditions $\Lambda(\varepsilon) = \mathcal{L}_0$. Since we know explicitly the flow, we can write the solution of this boundary problem as

$$\mathcal{L}_\tau = \lim_{\varepsilon \to 0+} \Phi(\tau)\Phi^{-1}(\varepsilon)\mathcal{L}_0. \tag{29}$$

We note that $\Phi(\tau)$ is smooth and invertible for $\tau > 0$. Therefore we can exchange the limit with $\Phi(\tau)$, and we just need to compute the limit of $\Phi^{-1}(\varepsilon)\mathcal{L}_0$. To do this we use a concrete representation of Lagrangian planes as span of a couple of vectors like in Section 1

$$\Phi^{-1}(\varepsilon)\mathcal{L}_0 = \begin{bmatrix} \lambda_1(\varepsilon) & \lambda_2(\varepsilon) \end{bmatrix}.$$

Let us find the limits of $\lambda_i(\varepsilon)$ as $\varepsilon \to 0$. We have

$$(\varepsilon^{H_{-1}})^{-1} = \begin{pmatrix} 1 & 0 & -\frac{\ln \varepsilon}{b_1} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Assume that $X(0) \in \mathcal{L}_0$. Then as we have seen in the Example 1.2 we can assume

$$\mathcal{L}_0 = \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix} \tag{30}$$

Then since $P(0) = \mathrm{id}_{2n}$, we obtain

$$\lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon)\mathcal{L}_0 = \lim_{\varepsilon \to 0+} (\varepsilon^{H_{-1}})^{-1}P^{-1}(\varepsilon) \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix},$$

because $(\varepsilon^{H_{-1}})^{-1}$ acts as the identity on $\mathcal{L}_0$. For the same reason

$$\lim_{\tau \to 0+} \Lambda(\tau) = \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix}$$

and so $\mathcal{L}_{0+} = \mathcal{L}_0$ and the Jacobi curve is actually continuous.

If $X(0) \notin \mathcal{L}_0$, then again from Example 1.2 we know that we can take

$$\mathcal{L}_0 = \begin{bmatrix} \lambda_1(\varepsilon) & \lambda_2(\varepsilon) \end{bmatrix} = \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \\ 1 & 0 \\ w_1 & w_2 \end{bmatrix} \tag{31}$$

Similarly to the previous case we find that

$$\lim_{\varepsilon \to 0+} \lambda_2(\varepsilon) = \begin{pmatrix} x_2 \\ y_2 \\ 0 \\ z_2 \end{pmatrix}.$$

Let us see what happens to the limit of the first vector. We have

$$\lim_{\varepsilon \to 0+} \lambda_1(\varepsilon) = \lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon) \begin{pmatrix} x_1 \\ y_1 \\ 1 \\ w_1 \end{pmatrix} = \lim_{\varepsilon \to 0+} \begin{pmatrix} x_1 - \dfrac{\ln \varepsilon}{b_1} \\ y_1 \\ 1 \\ w_1 \end{pmatrix}$$

which is equal to infinity. As we have said before a representation of a Lagrangian plane as a span of two vectors is not unique. We can scale them as we want as we take the limit. So we take

$$\lim_{\varepsilon \to 0+} -\frac{b_1}{\ln \varepsilon} \lambda_1(\varepsilon) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

So

$$\lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon) \mathcal{L}_0 = \begin{bmatrix} 1 & x_2 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix}.$$

Then as before we find that

$$\lim_{\tau \to 0+} \Lambda(\tau) = \begin{bmatrix} 1 & x_2 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & y_2 \\ 0 & 0 \\ 0 & w_2 \end{bmatrix}.$$

So summarizing everything we have done in a more invariant manner the jump can be computed as follows. Given $\mathcal{L}_0$, the new $\mathcal{L}$-derivative $\mathcal{L}_{0+}$ is going to be a direct sum of $\mathcal{L}_0 \cap X(0)^{\measuredangle}$ and $X(0)$. But this is by definition $\mathcal{L}_0^{X(0)}$. The goal of the following subsections is to prove the same for $m \geq 2$.

## 9.2 Model examples for $m = 2$

For $m \geq 2$ we proceed in a different way. One can reduce by a change of variables the system (22) to a system with a regular singular point at $\tau = 0$. Thus all the methods from the theory of linear systems of ODE's can be used. But these techniques work well under some non-resonance conditions. In our case we can use techniques from Hamiltonian

dynamics to arrive at the results even in the presence of resonances. First we prove the result for some model problems similarly as we have done in the case $m = 1$. Then we apply Riccati comparison theorems, to prove the general result.

For $m = 2$ we choose the following Hamiltonian systems as our models

$$\frac{d}{dt}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} 0 & 0 & \frac{b_{11}}{\tau^2} & 0 \\ 0 & 0 & 0 & b_{22} \\ c_{11} & 0 & 0 & 0 \\ 0 & c_{22} & 0 & 0 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}, \tag{32}$$

where $b_{ii}, c_{ii}$ are constants, $b_{11} \neq 0$ (or else there is no singularity) and $c_{11} \neq 0$ (because we have $\sigma(X(0), \dot{X}(0)) \neq 0$). First of all we notice that this system splits into two invariant sub-systems

$$\frac{d}{dt}\begin{pmatrix} p_1 \\ q_1 \end{pmatrix} = \begin{pmatrix} 0 & \frac{b_{11}}{\tau^2} \\ c_{11} & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ q_1 \end{pmatrix}, \tag{33}$$

$$\frac{d}{dt}\begin{pmatrix} p_2 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 & b_{22} \\ c_{22} & 0 \end{pmatrix} \begin{pmatrix} p_2 \\ q_2 \end{pmatrix}. \tag{34}$$

We denote by $\Phi_i(t)$ the corresponding fundamental matrices. Without any loss of generality, we can assume that $\Phi_2(0) = \mathrm{id}_2$. In order to find $\Phi_1$, we do a symplectic change of variables

$$\begin{pmatrix} \tilde{p}_1 \\ \tilde{q}_1 \end{pmatrix} = \begin{pmatrix} \tau^{1/2} & 0 \\ 0 & \tau^{-1/2} \end{pmatrix} \begin{pmatrix} p_1 \\ q_1 \end{pmatrix}.$$

Then the first system is transformed to

$$\frac{d}{dt}\begin{pmatrix} \tilde{p}_1 \\ \tilde{q}_1 \end{pmatrix} = \frac{1}{\tau} \begin{pmatrix} 1/2 & b_{11} \\ c_{11} & -1/2 \end{pmatrix} \begin{pmatrix} \tilde{p}_1 \\ \tilde{q}_1 \end{pmatrix} = \frac{Y}{\tau} \begin{pmatrix} \tilde{p}_1 \\ \tilde{q}_1 \end{pmatrix}$$

which is a simple linear system with a regular singular point. Therefore the fundamental solution $\Phi_1$ is the following matrix function

$$\Phi_1(\tau) = \begin{pmatrix} \tau^{-1/2} & 0 \\ 0 & \tau^{1/2} \end{pmatrix} \tau^Y$$

or in a more detailed form

$$\Phi_1(\tau) = \begin{pmatrix} \dfrac{\tau^{\frac{-1-\Delta}{2}}(-1 + \Delta + (1+\Delta)\tau^\Delta)}{2\Delta} & \dfrac{b_{11}\tau^{\frac{-1-\Delta}{2}}(-1 + \tau^\Delta)}{\Delta} \\ \dfrac{c_{11}\tau^{\frac{1-\Delta}{2}}(-1 + \tau^\Delta)}{\Delta} & \dfrac{\tau^{\frac{1-\Delta}{2}}(1 + \Delta + (-1+\Delta)\tau^\Delta)}{2\Delta} \end{pmatrix},$$

where $\Delta = \sqrt{1 + 4b_{11}c_{11}}$. Under the non-oscillation assumption we have $\Delta > 0$. It is easy to check that $\det \Phi_1(\tau) = \det \Phi_2(\tau) \equiv 1$, so the inverse matrix of $\Phi(\tau)$ can be computed easily.

The Jacobi curve is given by (29) and as in the previous subsection the fundamental matrix $\Phi(\tau)$ is smooth for $\tau > 0$, so we can exchange it with the limit. So first of all we need to find the limit

$$\lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon)\mathcal{L}_0.$$

As in the previous case we are going to separate two situations: when $X(0) \in \mathcal{L}_0$ and when $X(0) \notin \mathcal{L}_0$.

As before if $X(0) \in \mathcal{L}_0$ we can assume that $\mathcal{L}_0$ is given by (30). Then the first and the second vector lie in its own invariant subspace. For example since $\Phi_2(0) = \mathrm{id}_2$, we immediately get that

$$\lim_{\tau \to 0+} \Phi(\tau) \lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon) \begin{pmatrix} 0 \\ y_2 \\ 0 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ y_2 \\ 0 \\ w_2 \end{pmatrix}.$$

Let us see what happens to the first vector. We have

$$\Phi^{-1}(\varepsilon) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \dfrac{\varepsilon^{\frac{1-\Delta}{2}}(1 + \Delta + (-1 + \Delta)\varepsilon^\Delta)}{2\Delta} \\ 0 \\ -\dfrac{c_{11}\varepsilon^{\frac{1-\Delta}{2}}(-1 + \varepsilon^\Delta)}{\Delta} \\ 0 \end{pmatrix}.$$

Then we find that

$$\lim_{\varepsilon \to 0+} \left[ \Phi^{-1}(\varepsilon) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right] = \lim_{\varepsilon \to 0+} \left[ \varepsilon^{-\frac{1-\Delta}{2}} \Phi^{-1}(\varepsilon) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right] = \begin{bmatrix} \frac{1+\Delta}{2\Delta} \\ 0 \\ \frac{c_{11}}{\Delta} \\ 0 \end{bmatrix} \in \lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon)\mathcal{L}_0$$

and

$$\lim_{\tau \to 0+} \left[ \tau^{\frac{1+\Delta}{2}} \Phi(\tau) \begin{pmatrix} \frac{1+\Delta}{2\Delta} \\ 0 \\ \frac{c_{11}}{\Delta} \\ 0 \end{pmatrix} \right] = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{35}$$

which means that the limit is up to a constant the vector $X(0)$. Thus in this case $\mathcal{L}_{0+} = \mathcal{L}_0$ and the Jacobi curve is actually continuous as expected.

We now look at the situation when $X(0) \notin \mathcal{L}_0$. Then $\mathcal{L}_0$ can be assumed to be of the form (31). We consider the vectors $\lambda_i(\varepsilon)$ and their projection onto the first invariant subspace

$$n_1(\varepsilon) = \Phi_1^{-1}(\varepsilon) \begin{pmatrix} x_1 \\ 1 \end{pmatrix}, \qquad n_2(\varepsilon) = \Phi_1^{-1}(\varepsilon) \begin{pmatrix} x_2 \\ 0 \end{pmatrix}. \tag{36}$$

We have

$$n_1(\varepsilon) = \varepsilon^{-\frac{1+\Delta}{2}} \begin{pmatrix} \dfrac{-2b_{11}(-1+\varepsilon^\Delta) + \varepsilon(1+\Delta+(-1+\Delta)\varepsilon^\Delta)x_1}{2\Delta} \\[2mm] \dfrac{-1+\Delta+2c_{11}\varepsilon x_1 + \varepsilon^\Delta(1-2c_{11}\varepsilon x_1 + \Delta)}{2\Delta} \end{pmatrix}$$

So it is clear that

$$\lim_{\varepsilon \to 0+} \varepsilon^{\frac{1+\Delta}{2}} \lambda_1(\varepsilon) = \begin{pmatrix} \dfrac{b_{11}}{\Delta} \\[1mm] 0 \\[1mm] \dfrac{-1+\Delta}{2\Delta} \\[1mm] 0 \end{pmatrix}.$$

Then the formula (35) proves that $X(0) \in \mathcal{L}_{0+}$. Now we need to find an independent from $X(0)$ limit vector that would lie in $\mathcal{L}_{0+}$.

Writing down $n_2(\varepsilon)$ we get

$$n_2(\varepsilon) = \varepsilon^{\frac{1-\Delta}{2}} \begin{pmatrix} \dfrac{(1+\Delta+\varepsilon^\Delta(-1+\Delta))x_2}{2\Delta} \\[2mm] \dfrac{c_{11}(1-\varepsilon^\Delta)x_2}{\Delta} \end{pmatrix}.$$

If $0 < \Delta < 1$ or $x_2 = 0$, then it is clear that

$$\lim_{\varepsilon \to 0+} \lambda_2(\varepsilon) = \begin{pmatrix} 0 \\ y_2 \\ 0 \\ w_2 \end{pmatrix},$$

where we have used that $\Phi_2(0) = \mathrm{id}_2$. For the same reason the very same vector is going to lie in $\mathcal{L}_{0+}$ and the result follows.

If $\Delta = 1$, then either $b_{11} = 0$ or $c_{11} = 0$. Since we have excluded these possibilities it only remains to see what happens, when $\Delta > 1$ and $x_2 \neq 0$.

We can see that the expressions for $\lambda_i(\varepsilon)$ are just sums of power series of $\varepsilon$. Therefore it is convenient to introduce the following notation

$$a(\varepsilon) = b(\varepsilon) \mod \varepsilon^{>0}$$

which means that $a(\varepsilon)$ and $b(\varepsilon)$ agree modulo terms of positive degree of $\varepsilon$. Then $\mathcal{L}_{0+} = \mathcal{L}_0^{X(0)}$ follows from the following lemma

**Lemma 9.1.** *Let $\Delta > 1$ and $x_2 \neq 0$. Then there exist constants $c_0, c_1, ..., c_l$, s.t.*

$$\lambda_2(\varepsilon) - x_2\varepsilon\lambda_1(\varepsilon) \sum_{i=0}^{l} c_i\varepsilon^i = \begin{pmatrix} k_1\varepsilon^{\frac{2l+1-\Delta}{2}} \\ y_2 \\ k_2\varepsilon^{\frac{2l+1-\Delta}{2}} \\ w_2 \end{pmatrix} \mod \varepsilon^{>0},$$

*where $k_i$ are some constants.*

Indeed, the vector on the left hand side is a linear span of $\lambda_1(\varepsilon)$ and $\lambda_2(\varepsilon)$. We can choose $l$ sufficiently big so that $2l + 1 - \Delta > 0$. Then in the limit we obtain a vector $\begin{pmatrix} 0 & y_2 & 0 & w_2 \end{pmatrix}^T$, which lies in the second invariant subspace where there is no singularity at all.

*Proof of the lemma.* We denote by $\lambda(\varepsilon)$ the vector on the left. It is easy to see why the second and fourth components of $\lambda(\varepsilon)$ have this form. It follows from the fact that $\Phi_2(\varepsilon)$ is an analytic matrix function with $\Phi_2(0) = \mathrm{id}_2$.

So it is enough to look on the projection of $\lambda(\varepsilon)$ to the singular invariant subspace. We can write

$$n_1(\varepsilon) = \frac{\varepsilon^{-\frac{1+\Delta}{2}}}{\Delta} \begin{pmatrix} b_{11} \\ \frac{-1+\Delta}{2} \end{pmatrix} + \frac{\varepsilon^{\frac{1-\Delta}{2}} x_1}{\Delta} \begin{pmatrix} \frac{1+\Delta}{2} \\ c_{11} \end{pmatrix} \quad \mathrm{mod}\ \varepsilon^{>0},$$

$$n_2(\varepsilon) = \frac{\varepsilon^{\frac{1-\Delta}{2}} x_2}{\Delta} \begin{pmatrix} \frac{1+\Delta}{2} \\ c_{11} \end{pmatrix} \quad \mathrm{mod}\ \varepsilon^{>0}.$$

Let us denote

$$\alpha(\varepsilon) = \frac{\varepsilon^{\frac{1-\Delta}{2}}}{\Delta} \begin{pmatrix} b_{11} \\ \frac{-1+\Delta}{2} \end{pmatrix}$$

From here we see that

$$x_2 \varepsilon n_1(\varepsilon) = x_2 \alpha(\varepsilon) + x_1 \varepsilon n_2(\varepsilon) \quad \mathrm{mod}\ \varepsilon^{>0}.$$

We then find an expression for the projection of $\lambda(\varepsilon)$:

$$n_2(\varepsilon) - x_2 \varepsilon n_1(\varepsilon) \sum_{i=1}^{l} c_i \varepsilon^i =$$

$$= n_2(\varepsilon) - x_2 \alpha(\varepsilon) c_0 + \sum_{i=0}^{l-1} (x_2 \alpha(\varepsilon) c_i - x_1 n_2(\varepsilon) c_{i-1}) - \varepsilon^l x_1 n_2(\varepsilon) c_{l-1} \quad \mathrm{mod}\ \varepsilon^{>0}.$$

So it is enough to choose $c_i$ to be s.t. they solve

$$n_2(\varepsilon) - x_2 \alpha(\varepsilon) c_0 = 0 \quad \mathrm{mod}\ \varepsilon^{>0},$$

$$x_2 \alpha(\varepsilon) c_{i+1} - x_1 n_2(\varepsilon) c_i = 0 \quad \mathrm{mod}\ \varepsilon^{>0}.$$

The first equality is satisfied, if

$$c_0 = \frac{1+\Delta}{2 b_{11}}, \qquad \text{(recall that } \Delta = \sqrt{1 + 4 b_{11} c_{11}}).$$

But then we can obtain an expression for $n_2(\varepsilon)$ from the first equation and plug it into the second one. We get

$$\alpha(\varepsilon) x_2 (c_{i+1} - x_1 c_0 c_i) = 0.$$

So we simply choose recursively $c_{i+1} = x_1 c_i c_0$. $\qquad\square$

## 9.3   Model examples for $m > 2$

We now consider the same model as (32) but with singularity of order $m > 2$. Recall that $B_{ii}, C_{ii}$ are constants and $b_{11}, c_{11}$ are non zero. It is convenient to define $m = 2 + \beta$. Again we have two invariant subsystems and equation (33) has the form

$$\dot{p}_1 = \frac{b_{11}}{\tau^{2+\beta}} q_1,$$
$$\dot{q}_1 = c_{11} p_1.$$

We differentiate the second equation to obtain

$$\ddot{q}_1 - \frac{b_{11} c_{11}}{\tau^{2+\beta}} q_1 = 0.$$

If we introduce a new independent variable

$$y(\tau) = \frac{q_1(\tau)}{\sqrt{\tau}}$$

and a new dependent variable

$$s(\tau) = \frac{2\sqrt{b_{11} c_{11}}}{\beta} \tau^{\frac{-\beta}{2}},$$

we obtain a modified Bessel equation

$$s^2 \frac{d^2 y}{ds^2} + s \frac{dy}{ds} - \left( s^2 + \frac{1}{\beta^2} \right) y = 0.$$

Two independent solutions of this equation are given by two modified Bessel functions $I_{\beta^{-1}}(s)$, $K_{\beta^{-1}}(s)$ [20]. Therefore the fundamental matrix $\Phi_1(\tau)$ is given by

$$\Phi_1(\tau) = \begin{pmatrix} \frac{1}{c_{11}} \frac{d}{d\tau} \sqrt{\tau} I_{\beta^{-1}}(s(\tau)) & \frac{1}{c_{11}} \frac{d}{d\tau} \sqrt{\tau} K_{\beta^{-1}}(s(\tau)) \\ \sqrt{\tau} I_{\beta^{-1}}(s(\tau)) & \sqrt{\tau} K_{\beta^{-1}}(s(\tau)) \end{pmatrix}$$

We can simplify considerably the first row using the following formulas for the derivatives of modified Bessel functions [20]

$$I_a'(x) = \frac{a}{x} I_a(x) + I_{a+1}(x),$$
$$K_a'(x) = \frac{a}{x} K_a(x) - K_{a+1}(x).$$

After some simplifications we find that

$$\Phi_1(\tau) = \begin{pmatrix} -\sqrt{\frac{b_{11}}{c_{11}}} \tau^{-\frac{1+\beta}{2}} I_{\beta^{-1}+1}(s(\tau)) & \sqrt{\frac{b_{11}}{c_{11}}} \tau^{-\frac{1+\beta}{2}} K_{\beta^{-1}+1}(s(\tau)) \\ \sqrt{\tau} I_{\beta^{-1}}(s(\tau)) & \sqrt{\tau} K_{\beta^{-1}}(s(\tau)) \end{pmatrix}.$$

Since $\beta > 0$, as $\tau \to 0+$ we get $s(\tau) \to +\infty$. Therefore we need an asymptotic expansion of modified Bessel functions as the argument goes to $+\infty$. They are given by

$$I_a(x) \sim \sqrt{\frac{1}{2\pi x}} e^x, \qquad x \to +\infty,$$

$$K_a(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x}, \qquad x \to +\infty.$$

In particular we see that the limit does not depend on the parameter, and therefore for any real $a, b$ we have

$$I_a(x) \to +\infty, \qquad\qquad \frac{I_a(x)}{I_b(x)} \to 1,$$

$$K_a(x) \to 0, \qquad\qquad \frac{K_a(x)}{K_b(x)} \to 1,$$

as $x \to +\infty$.

The matrix $\Phi_1(\tau)$ is invertible and smooth for $\tau > 0$. From the explicit form of the equation it follows that determinant of $\Phi_1(\tau)$ is constant. Using the asymptotics above we can the find that it is actually equal to $-\beta/(2c_{11})$. The very same asymptotics and an argument similar to the one for $m = 1, 2$ implies that if $X(0) \in \mathcal{L}_0$, then $\mathcal{L}_{0+} = \mathcal{L}_0$. So we assume that $X(0) \notin \mathcal{L}_0$ and consequently that $\mathcal{L}_0$ is given by (31).

If $x_2 = 0$, then it is clear that

$$\lim_{\tau \to 0+} \Phi(\tau) \lim_{\varepsilon \to 0+} \Phi^{-1}(\varepsilon) \begin{pmatrix} 0 \\ y_2 \\ 0 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ y_2 \\ 0 \\ w_2 \end{pmatrix}.$$

So it remains to find a single independent vector in $\mathcal{L}_{0+}$ in this case. Let us slightly abuse the notation and denote

$$\begin{pmatrix} x_1(\varepsilon) \\ y_1(\varepsilon) \\ z_1(\varepsilon) \\ w_1(\varepsilon) \end{pmatrix} = \Phi^{-1}(\varepsilon) \begin{pmatrix} x_1 \\ y_1 \\ 1 \\ w_1 \end{pmatrix}.$$

Using an explicit expression for the fundamental matrix, we find that $x_1(\varepsilon)$ is a linear combination of the modified Bessel $K$-functions and $z_1(\varepsilon)$ is a linear combination of $I$-functions. Due to the exponential behaviour of $I_a(x)$ and $K_a(x)$, we find that

$$\lim_{\varepsilon \to 0+} \frac{1}{z_1(\varepsilon)} \Phi^{-1}(\varepsilon) \begin{pmatrix} x_1 \\ y_1 \\ 1 \\ w_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

For the same reason

$$\lim_{\tau \to 0+} \frac{\sqrt{c_{11}} \tau^{\frac{1+\beta}{2}}}{\sqrt{b_{11}} K_{\beta^{-1}+1}(s(\tau))} \Phi(\tau) \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{37}$$

and we obtain that $\mathcal{L}_{0+} = \mathcal{L}_0^{X(0)}$.

Assume now that $x_2 \neq 0$. Then we obtain by the same argument as above

$$\lim_{\varepsilon \to 0+} \frac{2c_{11}}{\beta x_2 \sqrt{\varepsilon} I_{\beta^{-1}}(s(\varepsilon))} \Phi^{-1}(\varepsilon) \begin{pmatrix} x_2 \\ y_2 \\ 0 \\ w_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

Exploiting once more the formula (37), we once again find that $X(0) \in \mathcal{L}_{0+}$. To find an independent vector limit let us write down explicitly the vectors $n_1(\varepsilon)$ and $n_2(\varepsilon)$ defined in (36) of the previous subsection. We have

$$n_1(\varepsilon) = \begin{pmatrix} x_1(\varepsilon) \\ z_1(\varepsilon) \end{pmatrix} = -\frac{2c_{11}}{\beta} \begin{pmatrix} -\frac{\sqrt{b_{11}} \varepsilon^{-\frac{1+\beta}{2}} K_{\beta^{-1}+1}(s(\varepsilon))}{\sqrt{c_{11}}} + \sqrt{\varepsilon} x_1 K_{\beta^{-1}}(s(\varepsilon)) \\ -\frac{\sqrt{b_{11}} \varepsilon^{-\frac{1+\beta}{2}} I_{\beta^{-1}+1}(s(\varepsilon))}{\sqrt{c_{11}}} - \sqrt{\varepsilon} x_1 I_{\beta^{-1}}(s(\varepsilon)) \end{pmatrix},$$

$$n_2(\varepsilon) = -\frac{2c_{11}}{\beta} \begin{pmatrix} \sqrt{\varepsilon} x_2 K_{\beta^{-1}}(s(\varepsilon)) \\ -\sqrt{\varepsilon} x_2 I_{\beta^{-1}}(s(\varepsilon)). \end{pmatrix}$$

To find the independent limit vector above we consider

$$\lambda(\varepsilon) = \lambda_2(\varepsilon) + \frac{\sqrt{\varepsilon} x_2 I_{\beta^{-1}}(s(\varepsilon))}{z_1(\varepsilon)} \lambda_1(\varepsilon).$$

The only component of $\lambda_1(\varepsilon)$ and $\lambda_2(\varepsilon)$ escaping to infinity are the $z$-components as can be easily seen from the explicit expression of $n_i(\varepsilon)$. But the $z$-component of $\lambda(\varepsilon)$ is equal to zero. Moreover the coefficient in front of $\lambda_1(\varepsilon)$ tends to zero as $\varepsilon \to 0+$. Thus from the explicit expressions for $x_1(\varepsilon)$ and $x_2(\varepsilon)$ we obtain that

$$\lim_{\varepsilon \to 0+} \lambda(\varepsilon) = \begin{pmatrix} 0 \\ y_2 \\ 0 \\ z_2 \end{pmatrix},$$

which is a vector that does not lie in the singular invariant subspace. Thus as in the previous sections the same vector lies in $\mathcal{L}_{0+}$ which proves the result.

## 9.4  Jump for $m \geq 2$

In the previous subsections we have seen, that for the autonomous models the Jacobi curve has the right limit $\mathcal{L}_{0+} = \mathcal{L}_0^{X(0)}$. Now we are ready to prove this for a general system (22). We use the standard Riccati comparison result from [22].

**Lemma 9.2.** *Suppose that $B(\tau)$ and $C(\tau)$ are two symmetric continuous matrix functions that satisfy $B(\tau) \geq 0$ and $C(\tau) \geq 0$ for almost every $\tau$ of any closed subinterval $[a, b]$ of a given open interval $I$. Then given a symmetric matrix $S_a \geq 0$, any Cauchy solution of*

$$\dot{S} + SA + A^T S + SBS - C = 0, \tag{38}$$
$$S(a) = S_a,$$

*satisfies $S(\tau) \geq 0$ for all $\tau \in [a, b]$.*

We consider now the general system (22). Let $q = Sp$ and we write the corresponding Riccati equation like discussed in Section 1

$$\dot{S} + \frac{SB(\tau)S}{\tau^m} - C(\tau) = 0. \tag{39}$$

If the system is not oscillating, then we have existence of the Cauchy problem with the boundary data $S(t) = S$ on the interval $(0, t]$ for any fixed symmetric matrix $S$ and for $t$ small enough.

Assume that $B_1(\tau) \leq B(\tau) \leq B_2(\tau) \leq 0$ and $C_1(\tau) \geq C(\tau) \geq C_2(\tau)$ for small $\tau \in [0, t]$. We assume that $B_i$ and $C_i$ are diagonal matrices like in our models from the previous subsection satisfying the non-oscillation conditions. Then we can define $S_i^\varepsilon$ to be solutions of the Cauchy problem

$$\dot{S} + \frac{SB_i(\tau)S}{\tau^m} - C_i(\tau) = 0, \qquad S(\varepsilon) = S_{\mathcal{L}_0},$$

where $S_{\mathcal{L}_0}$ is a symmetric matrix that corresponds to $\mathcal{L}_0$ assuming of course that $\mathcal{L}_0$ is transversal to the horizontal plane $q = 0$. Let $S^\varepsilon(\tau)$ be a solution of (39) with $S(\varepsilon) = S_{\mathcal{L}_0}$.

Let us assume, for example $W^\varepsilon(\tau) = S_1^\varepsilon(\tau) - S_2^\varepsilon(\tau)$. Then we have that $W^\varepsilon(\tau)$ satisfies

$$\dot{W}^\varepsilon + W^\varepsilon \frac{B_1(\tau)}{\tau^m} S_2^\varepsilon + S_2^\varepsilon \frac{B_1(\tau)}{\tau^m} W^\varepsilon + W^\varepsilon \frac{B_1(\tau)}{\tau^m} W^\varepsilon + \frac{S_2^\varepsilon (B_1(\tau) - B_2(\tau)) S_2^\varepsilon}{\tau^m} - (C_1(\tau) - C_2(\tau)) = 0$$

with $W_\varepsilon(\varepsilon) = 0$. But then by the Lemma 9.2 we obtain that

$$W^\varepsilon(\tau) \geq 0 \iff S_1^\varepsilon(\tau) \geq S_2^\varepsilon(\tau),$$

for any $\tau \geq \varepsilon$ as long as $S_2^\varepsilon(\tau)$ is defined.

By replacing $B_1(\tau)$ with $B(\tau)$ and then $B_2(\tau)$ with $B(\tau)$, we similarly obtain that

$$S_2^\varepsilon(\tau) \leq S^\varepsilon(\tau) \leq S_1^\varepsilon(\tau),$$

for any $\tau \geq \varepsilon$ sufficiently close to $\varepsilon$ and $\varepsilon > 0$ small. By fixing $\tau$ sufficiently small and taking limits as $\varepsilon \to 0+$ we find that

$$S_1(\tau) \leq S(\tau) \leq S_2(\tau),$$

where these matrix functions are the corresponding Jacobi curves. But we have proven in the previous subsections that for our model examples we had the same right limit. Thus $S_1(0+) = S_2(0+)$ and

$$S(0+) = S_1(0+) = S_2(0+).$$

If $\mathcal{L}_0$ or $\mathcal{L}_0^{X(0)}$ are not transversal to the horizontal space $\Sigma$, then this construction clearly does not work, because either $S^\varepsilon(\varepsilon)$ or $S(0+)$ do not exist. In this case we make a change of variables of the form

$$\begin{pmatrix} p \\ q \end{pmatrix} \mapsto M \begin{pmatrix} p \\ q \end{pmatrix}, \qquad M = \begin{pmatrix} \alpha_1 & 0 & \beta_1 & 0 \\ 0 & \alpha_2 & 0 & \beta_2 \\ \gamma_1 & 0 & \delta_1 & 0 \\ 0 & \gamma_2 & 0 & \delta_2 \end{pmatrix},$$

s.t.

$$\begin{vmatrix} \alpha_i & \beta_i \\ \gamma_i & \delta_i \end{vmatrix} = 1$$

Matrix $M$ is clearly symplectic and we want to choose it in such a way that $M\mathcal{L}_0$ and $M\mathcal{L}_0^{X(0)}$ are transversal to the horizontal subspace $\Sigma$. Such matrices $M$ are actually dense in the set of all matrices of the given form. We can prove this by an explicit computation.

If $\dim(\mathcal{L}_0 \cap \Sigma) > 0$, then following along the lines of Example 1.1 we can assume that

$$\mathcal{L}_0 = \begin{bmatrix} x & 0 \\ 0 & 0 \\ z & 0 \\ 0 & 1 \end{bmatrix}$$

Then $\dim(M\mathcal{L}_0 \cap \Sigma) = 0$ is equivalent to

$$\alpha_1 x + \beta_1 z \neq 0, \qquad \beta_2 \neq 0.$$

Similarly from examples 1.1 and 1.2 we know that if $\dim(\mathcal{L}_0^{X(0)} \cap \Sigma) > 0$, then

$$\mathcal{L}_0^{X(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Then $\dim(M\mathcal{L}_0^{X(0)} \cap \Sigma) = 0$ can be achieved by already taking $\beta_2 \neq 0$.

In the new coordinates our Jacobi equation takes the form

$$\frac{d}{d\tau} \begin{pmatrix} p \\ q \end{pmatrix} = M \begin{pmatrix} 0 & \frac{B(\tau)}{\tau^m} \\ C(\tau) & 0 \end{pmatrix} M^{-1} \begin{pmatrix} p \\ q \end{pmatrix}.$$

44

An explicit computation gives us

$$M \begin{pmatrix} 0 & \frac{B(\tau)}{\tau^m} \\ C(\tau) & 0 \end{pmatrix} M^{-1} =$$

$$= \begin{pmatrix} c_{11}\beta_1\delta_1 - \left(\frac{1}{b} + b_{11}\right)\gamma_1\alpha_1 & -b_{12}\gamma_2\alpha_1 & -c_{11}\beta_1^2 + \left(\frac{1}{b} + b_{11}\right)\alpha_1^2 & b_{12}\alpha_1\alpha_2 \\ -b_{12}\gamma_1\alpha_2 & c_{22}\delta_2\beta_2 - b_{22}\gamma_2\alpha_2 & b_{12}\alpha_1\alpha_2 & -c_{22}\beta_2^2 + b_{22}\alpha_2^2 \\ c_{11}\delta_1^2 - \left(\frac{1}{b} + b_{11}\right)\gamma_1^2 & -b_{12}\gamma_1\gamma_2 & -c_{11}\delta_1\beta_1 + \left(\frac{1}{b} + b_{11}\right)\gamma_1\alpha_1 & b_{12}\gamma_1\alpha_2 \\ -b_{12}\gamma_1\gamma_2 & c_{22}\delta_2^2 - b_{22}\gamma_2^2 & b_{12}\gamma_2\alpha_1 & -c_{22}\delta_2\beta_2 + b_{22}\alpha_2\gamma_2 \end{pmatrix}$$

Recall that our original system was such that $b < 0$ and $b_{22} < 0$ for $\tau \geq 0$ small. Thus the upper of diagonal 2x2 block will be a negative matrix function for small $\tau > 0$, if we choose $\alpha_2$ big enough. For the same reason the lower of diagonal 2x2 block will be negative if we choose $\gamma_2$ big enough. Thus we can apply the comparison lemma as before with

$$B_1(\tau) = \begin{pmatrix} \frac{\alpha_1^2}{b(\tau)} + \varepsilon & 0 \\ 0 & b_{22}(0)\alpha_2^2 - c_{22}(0)\beta_2^2 + \varepsilon \end{pmatrix},$$

$$B_2(\tau) = \begin{pmatrix} \frac{\alpha_1^2}{b(\tau)} + \varepsilon & 0 \\ 0 & b_{22}(0)\alpha_2^2 - c_{22}(0)\beta_2^2 + \varepsilon \end{pmatrix},$$

$$C_1(\tau) = \begin{pmatrix} -\frac{\gamma_1^2}{b(\tau)} - \varepsilon & 0 \\ 0 & -b_{22}(0)\gamma_2^2 + c_{22}(0)\delta_2^2 - \varepsilon \end{pmatrix},$$

$$C_2(\tau) = \begin{pmatrix} -\frac{\gamma_1^2}{b(\tau)} + \varepsilon & 0 \\ 0 & -b_{22}(0)\gamma_2^2 + c_{22}(0)\delta_2^2 + \varepsilon \end{pmatrix}$$

where $\varepsilon > 0$ is sufficiently small.

## 9.5 Jacobi curve for $m = 1, 2$

As we have already discussed before, the jump alone does not determine the Jacobi curve, because with a singular Jacobi DE we lose uniqueness. So we need to characterize the right solution of the Jacobi equation. In this section we prove the following result.

**Theorem 9.1.** *If $m = 1$ or $m = 2$ then Jacobi curve after a singularity can be characterized as a boundary value problem of the extended Jacobi DE on the Lagrangian Grassmanian with conditions on the left end-point and the first left derivative.*

This will be proven in a number of steps:

1. We change coordinates so that $\mathcal{L}_0$ and $\mathcal{L}_{0+}$ lie in the same coordinate chart and $\mathcal{L}_{0+}$ is taken to be zero;

2. We write down the corresponding Riccati equation and perform a certain blow-up procedure;

3. After the blow-up we obtain a non-autonomous Riccati equation. We then proceed in determining the Jacobi curve for the autonomous part;

4. Using a deformation argument we prove that in the non-autonomous case the Jacobi curve is well-defined by the same jet.

For the first step we are going to have three different situations as well

1. $\mathcal{L}_0$ is transversal to the horizontal plane in current coordinates and in the corresponding symmetric matrix $S^0$ either $S_{11}^0 \neq 0$ or $S_{11}^0 = S_{12}^0 = 0$;

2. $\mathcal{L}_0$ is either transversal to the horizontal plane in current coordinates and in the corresponding symmetric matrix $S_{11}^0 = 0$, $S_{22}^0 \neq 0$ or $\mathcal{L}_0$ and the horizontal plane $\Sigma$ have a common line;

3. $\mathcal{L}_0$ is either transversal to the horizontal plane in current coordinates and in the corresponding symmetric matrix $S_{11}^0 = S_{22}^0 = 0$, $S_{12}^0 \neq 0$ or $\mathcal{L}_0$ is the horizontal plane $\Sigma$.

Let

$$
S_{22}^+ = \begin{cases} S_{22}^0, & S_{11}^0 = 0 \\ S_{22}^0 - \frac{(S_{12}^0)^2}{S_{11}^0}, & S_{11}^0 \neq 0 \end{cases}
$$

Then depending on the case we apply one of the three symplectic transformations

$$
M_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -S_{22}^+ & 0 & 1 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \qquad M_3 = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.
$$

After applying those transformation $\mathcal{L}_{0+}$ will become the vertical subspace. Let us check what happens to $\mathcal{L}_0$ under these transforms. We have for case 1 either

$$
M_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ S_{11}^0 & S_{12}^0 \\ S_{12}^0 & S_{22}^0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ S_{11}^0 & S_{12}^0 \\ S_{12}^0 & \frac{(S_{12}^0)^2}{S_{11}^0} \end{bmatrix} \qquad \text{or} \qquad M_1 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & S_{22}^0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.
$$

For case 2 either

$$
M_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & S_{12}^0 \\ S_{12}^0 & S_{22}^0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ -S_{12}^0 & -S_{22}^0 \\ 0 & S_{12}^0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -\frac{(S_{12}^0)^2}{S_{22}^0} & -\frac{(S_{12}^0)}{S_{22}^0} \\ -\frac{(S_{12}^0)}{S_{22}^0} & -\frac{1}{S_{22}^0} \end{bmatrix}
$$

46

or

$$M_2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ z & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ z & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ z & 0 \\ 0 & 0 \end{bmatrix}.$$

For the case 3 either

$$M_3 \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & S_{12}^0 \\ S_{12}^0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -S_{12}^0 \\ -S_{12}^0 & 0 \\ 0 & S_{12}^0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -\frac{1}{S_{12}^0} \\ -\frac{1}{S_{12}^0} & -\frac{1}{(S_{12}^0)^2} \end{bmatrix}$$

or

$$M_3 \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

And this finishes the first step.

For the second step we have to rewrite the Jacobi equation in the new coordinates. We simply have to conjugate the right-hand side of (22) by the corresponding matrix $M_i$. Then to each case corresponds its own Jacobi equation of the form

$$\frac{d}{d\tau} \begin{pmatrix} p \\ q \end{pmatrix} = M_i \begin{pmatrix} A(\tau) & B(\tau) \\ C(\tau) & -A^T(\tau) \end{pmatrix} M_i^{-1} \begin{pmatrix} p \\ q \end{pmatrix}$$

or more precisely

Case 1 : $\frac{d}{d\tau} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 & b_{12}(\tau)S_{22}^+ & \frac{1}{b(\tau)} + b_{11}(\tau) & b_{12}(\tau) \\ 0 & b_{22}(\tau)S_{22}^+ & b_{12}(\tau) & b_{22}(\tau) \\ c_{11}(\tau) & 0 & 0 & 0 \\ 0 & c_{22}(\tau) - b_{22}(\tau)(S_{22}^+)^2 & -b_{12}(\tau)S_{22}^+ & -b_{22}(\tau)S_{22}^+ \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix},$

Case 2 : $\frac{d}{d\tau} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 0 & -b_{12}(\tau) & \frac{1}{b(\tau)} + b_{11}(\tau) & 0 \\ 0 & 0 & 0 & -c_{22}(\tau) \\ c_{11}(\tau) & 0 & 0 & 0 \\ 0 & -b_{22}(\tau) & b_{12}(\tau) & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix},$

Case 3 : $\frac{d}{d\tau} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} -c_{11}(\tau) & -b_{12}(\tau) & \frac{1}{b(\tau)} + b_{11}(\tau) - c_{11}(\tau) & 0 \\ 0 & 0 & 0 & -c_{22}(\tau) \\ c_{11}(\tau) & 0 & c_{11}(\tau) & 0 \\ 0 & -b_{22}(\tau) & b_{12}(\tau) & 0 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix}.$

We then take $q = Sp$ and obtain a Riccati equation of the form (3). We do a blow-up of this equation by taking

$$S(\tau) = \tau S_1(\tau).$$

47

Then we obtain a Riccati equation for $S_1(\tau)$. Let $\tilde{S}^0$ be the symmetric matrix that corresponds to $\mathcal{L}_0$ in the new coordinates. We denote by $S_1^\varepsilon(\tau)$ the solution of this Riccati equation with

$$S_1^\varepsilon(\varepsilon) = \frac{\tilde{S}^0}{\varepsilon}.$$

Since outside the singularity the right-hand side is analytic and we have a family of solutions converging to a solution, it is clear that

$$\dot{S}(\tau) = \lim_{\varepsilon \to 0+} S_1^\varepsilon(\tau).$$

Since in the new coordinates $S(0+) = 0$, we find that the discussed previously Riccati equation has the form

$$\tau \dot{S}_1 + S_1 + S_1 \begin{pmatrix} \frac{1}{b_2} & 0 \\ 0 & 0 \end{pmatrix} S_1 - C(0) = \tau R(\tau, S) \tag{40}$$

for $m = 2$ and

$$\tau \dot{S}_1 + S_1 - C(0) = \tau R(\tau, S) \tag{41}$$

for $m = 1$. That finishes the proof of the second step.

For the third step we are going to consider just the first case. For the second and the third case the argument is repeated word by word. We assume that the right-hand side of those equations is actually zero. Then we can understand very well the whole phase portrait of this Riccati equation. Indeed, we can extend the dynamics to the whole Lagrangian Grassmanian $L(2)$ by rewriting the corresponding Hamiltonian system.

$$\tau \frac{d}{d\tau} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{b_2} & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ c_{11}(0) & 0 & -\frac{1}{2} & 0 \\ 0 & c_{22}(0) - (S_{22}^+)^2 b_{22}(0) & 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix} = H \begin{pmatrix} p_1 \\ p_2 \\ q_1 \\ q_2 \end{pmatrix}$$

A complete description of the phase portrait of such a system was given in [24]. It is clear that the equilibrium points are spanned by the eigenvectors. In our case, $H$ has eigenvalues

$$\lambda_1 = -\frac{1}{2}\sqrt{1 + \frac{4c_{11}(0)}{b_2}}, \quad \lambda_2 = -\frac{1}{2}, \quad \lambda_3 = \frac{1}{2}, \quad \lambda_4 = \frac{1}{2}\sqrt{1 + \frac{4c_{11}(0)}{b_2}}.$$

From the assumptions we have that all four eigenvalues are real and different. Let $E_i$ be

the corresponding eigenvectors. We have

$$E_1 = \begin{pmatrix} 1 - \sqrt{1 + \frac{4c_{11}(0)}{b_2}} \\ 0 \\ 2c_{11}(0) \\ 0 \end{pmatrix}, \qquad E_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$$E_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ c_{22}(0) - b_{22}(0)(S_{22}^+)^2 \end{pmatrix}, \qquad E_4 = \begin{pmatrix} 1 + \sqrt{1 + \frac{4c_{11}(0)}{b_2}} \\ 0 \\ 2c_{11}(0) \\ 0 \end{pmatrix}.$$

We define $E_{ij} = \mathrm{span}\{E_i, E_j\}$. It is easy to see that we have four equilibrium points on the Lagrangian Grassmanian: $E_{12}, E_{13}, E_{24}, E_{34}$. For each of these equilibrium points we can find the corresponding stable and unstable manifolds $W^s(E_{ij})$ and $W^u(E_{ij})$. Then if $S_1^\varepsilon(\varepsilon)$ lies in $W^s(E_{ij})$, the Jacobi curve is going to be just the equilibrium solution $\mathcal{L}_\tau = E_{ij}$. Indeed, the Lagrangian Grassmanian is compact and therefore any trajectory in the stable manifold has finite length. But every trajectory of our Riccati equation has speed that goes to infinity as $\tau \to 0+$. So as we take $\varepsilon$ smaller and smaller for a fixed time $\tau > 0$ the corresponding curve $S^\varepsilon(\tau)$ is going to get closer and closer to the equilibrium point approaching it in the limit. It remains only to describe stable manifolds of our equilibrium points.

Luckily it was already done in [24] by M. Shayman. He proved that in order to find the stable manifolds we need to form a flag $\{0\} = V_0 \subset V_1 \subset ... \subset V_4 = \mathbb{R}^4$, where

$$V_i = \bigoplus_{j=1}^i E_j,$$

and associate to each $E_{ij}$ a sequence $l(E_{ij}) = (l_1, l_2, l_3, l_4)$, where

$$l_k = \begin{cases} 1 & \text{if } k = i, j; \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$W^s(E_{ij}) = \left\{ \Lambda \in L(2) \ : \ \dim \Lambda \cap V_m = \sum_{k=1}^m l_k, \ l_k \in l(E_{ij}), m = 1, 2, 3, 4. \right\}.$$

It remains to check for which initial data $S_1^\varepsilon(\varepsilon)$ lies in which $W^s(E_{ij})$ for small $\varepsilon > 0$ and describe the corresponding $W^s(E_{ij})$.

**Lemma 9.3.** *Suppose that the right hand side of (40) is zero. Then the curves $\Lambda_1^\varepsilon(\tau)$ that correspond to $S_1^\varepsilon(\tau)$ converge pointwise to the equilibrium solution $\Lambda(\tau) \equiv E_{34}$. Or in local coordinates we get that*

$$\lim_{\varepsilon \to 0+} S_1^\varepsilon(\tau) = \begin{pmatrix} -\frac{b_2}{2}\left(1 - \sqrt{1 + \frac{4c_{11}}{b_2}}\right) & 0 \\ 0 & c_{22}(0) - b_{22}(0)(S_{22}^+)^2 \end{pmatrix} = S_1^{34}.$$

*If $R(\tau, S) = 0$ in (40), then exists a unique solution of this equation with $S_1(0+) = S_1^{34}$.*

*Proof.* By definition we find that

$$W^s(E_{34}) = \{\Lambda \in L(2) \, : \, \dim(\Lambda \cap E_{12}) = 0\} = E_{12}^{\pitchfork},$$

which is dense in $L(2)$. So we only need to prove that $\Lambda^\varepsilon(\varepsilon) \in E_{12}^{\pitchfork}$ for $\varepsilon > 0$ small. Indeed, in this case the unstable manifold $W^u(E_{34}) = \{E_{34}\}$ and so the only solution of (40) with $S_1(0+) = S_1^{34}$ can be $S_1(\tau) \equiv S_1^{34}$.

We note that if $S_{11}^0 = S_{12}^0 = 0$, then $S_1^\varepsilon(\varepsilon) = 0$. In this case for small $\varepsilon > 0$ it is clear that $\dim(\Lambda^\varepsilon(\varepsilon) \cap E_{12}) = 0$. If $S_1^\varepsilon(\varepsilon) \neq 0$ for small $\varepsilon > 0$, then $\dim(\Lambda^\varepsilon(\varepsilon) \cap E_{12}) > 0$ if and only if

$$S_{12}^+ = 0 \quad \text{and} \quad \frac{S_{11}^+}{\varepsilon} = -\frac{b_2}{2}\left(1 + \sqrt{1 + \frac{4c_{11}}{b_2}}\right),$$

but this can happen only for a single value

$$\varepsilon = -\frac{2S_{11}^+}{b_2\left(1 + \sqrt{1 + \frac{4c_{11}}{b_2}}\right)}.$$

And so for small $\varepsilon > 0$ we indeed get $\Lambda^\varepsilon(\varepsilon) \in E_{12}^{\pitchfork}$. $\square$

Case $m = 1$ is easier, since the principal part of the equation (41) is linear and has a global stable equilibrium

$$S = \begin{pmatrix} c_{11}(0) & 0 \\ 0 & c_{22}(0) - b_{22}(0)(S_{22}^+)^2 \end{pmatrix}.$$

As for $m = 2$, we have then $S_1(0+) = S$ and a unique solution to a Cauchy problem, that characterizes our Jacobi curve.

For the case 2 and 3 we have a similar result. We obtain that

$$\lim_{\varepsilon \to 0+} S_1^\varepsilon(\tau) = \begin{pmatrix} -\frac{b_2}{2}\left(1 - \sqrt{1 + \frac{4c_{11}}{b_2}}\right) & 0 \\ 0 & -b_{22}(0) \end{pmatrix} = S_1^{34}$$

for $m = 2$ and

$$\lim_{\varepsilon \to 0+} S_1^\varepsilon(\tau) = \begin{pmatrix} c_{11}(0) & 0 \\ 0 & -b_{22}(0) \end{pmatrix}$$

for $m = 1$ and that in this case indeed the Jacobi curve is fully determined by the first jet. We keep the notation $S_1^{34}$ because in the case 2 and 3 we obtain a Hamiltonian system whose matrix has exactly the same eigenvalues as the Hamiltonian matrix of case 1 and the same eigenvectors except $E_3$ that must replaced by

$$E_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ -b_{22}(0) \end{pmatrix}$$

50

It remains now to do the last step and to show the general case. Let us assume

$$S_1 = \begin{pmatrix} S_{11} & S_{12} \\ S_{12} & S_{22} \end{pmatrix}$$

and rewrite (40) or (41) as a system on $\mathbb{R} \times L(2)$. Namely we have

$$\dot{S} = Q(S) + \tau R(\tau, S),$$
$$\dot{\tau} = \tau;$$

where $Q(S)$ is the autonomous Riccati part. It is clear that $(S_1^{34}, 0)$ is an equilibrium point of this system. Moreover, by linearising the right hand side at $(S_1^{34}, 0)$ we obtain that it is a hyperbolic equilibrium point, since the linearized operator has eigenvalues

$$\left\{ -\sqrt{1 + \frac{4c_{11}}{b_2}}, -\frac{1}{2} - \frac{1}{2}\sqrt{1 + \frac{4c_{11}}{b_2}}, -1, 1 \right\},$$

the same as for the autonomous system in all three cases. So by Grobman-Hartman theorem both systems are topologically conjugate in the neighbourhood of this equilibrium point. Since both of them have a single unstable direction it means that there exists a unique trajectory of the non-autonomous system that approaches $(S_1^{34}, 0)$ as $\tau \to 0+$. We claim that this trajectory must be a lift of the Jacobi curve to the extended phase-space. This result does not follow directly from the Grobman-Hartman theorem since $S^\varepsilon(\varepsilon)$ is far from the equilibrium and a priori we have no information about the behaviour orbits close to infinity.

The result follows from an application of the variation formulae proved in [9] that can be stated as follows. Given a non-autonomous vector field $Z_s$ we denote by $F_t[Z.]$ a flow from time 0 to time $t$ of the corresponding vector field. If $X, Y$ is a pair of autonomous vector fields, the variation formulae reads as

$$F_t[X + Y] = F_t[(F_{.-t}[X])_* Y] \circ F_t[X].$$

where $F_{s-t}[X]_* Y$ is just the push-forward of $Y$ under $F_{s-t}[X]$.

In our case

$$X = \begin{pmatrix} Q(S) \\ \tau \end{pmatrix}, \qquad Y = \begin{pmatrix} \tau R(\tau, S) \\ 0 \end{pmatrix}.$$

Due to smoothness of each flow the lift of the Jacobi curve will be then given by the limit curve

$$\lim_{\varepsilon \to 0+} F_{\tau-\varepsilon}[X + Y] \begin{pmatrix} S_1^\varepsilon(\varepsilon) \\ \varepsilon \end{pmatrix} = \lim_{\varepsilon \to 0+} \left( F_{\tau-\varepsilon}(F_{.-\tau+\varepsilon}[X])_* Y] \circ F_{\tau-\varepsilon}[X] \begin{pmatrix} S_1^\varepsilon(\varepsilon) \\ \varepsilon \end{pmatrix} \right) =$$

$$= \lim_{\varepsilon \to 0+} (F_{\tau-\varepsilon}(F_{.-\tau+\varepsilon}[X])_* Y]) \circ \lim_{\varepsilon \to 0+} \left( F_{\tau-\varepsilon}[X] \begin{pmatrix} S_1^\varepsilon(\varepsilon) \\ \varepsilon \end{pmatrix} \right) = F_\tau(F_{.-\tau}[X])_* Y] \circ \lim_{\varepsilon \to 0+} \left( F_{\tau-\varepsilon}[X] \begin{pmatrix} S_1^\varepsilon(\varepsilon) \\ \varepsilon \end{pmatrix} \right)$$

But the second limit corresponds to the lift of the Jacobi curve in autonomous case. Thus if we take a limit of this expression as $\tau \to 0+$ we would obtain

$$\begin{pmatrix} S_1^{34} \\ 0 \end{pmatrix},$$

like in the autonomous case. Which proves Theorem 9.1.

# References

[1] A. Agrachev. Feedback–invariant optimal control theory and differential geometry, ii. jacobi curves for singular extremals. *J. Dynamical and Control Systems*, pages 583–604, 1998.

[2] A. Agrachev. *Nonlinear and Optimal Control Theory*, chapter Geometry of Optimal Control Problems and Hamiltonian Systems, pages 1–59. Springer, 2004.

[3] A. Agrachev and I. Beschastnyi. Jacobi fields in optimal control: Morse and maslov indices. preprint.

[4] A. Agrachev and I. Beschastnyi. Symplectic geometry of constrained optimization. *Regular and Chaotic Dynamics*, 22:750–770, 2017.

[5] A. Agrachev and R. Gamkrelidze. *Nonlinear Controllability and Optimal Control*, chapter Symplectic geometry for optimal control, pages 263–277. CRC Press Book, 1990.

[6] A. Agrachev and R. Gamkrelidze. Feedback–invariant optimal control theory and differential geometry, i. regular extremals. *J. Dynamical and Control Systems*, pages 343–389, 1997.

[7] A. Agrachev and R. Gamkrelidze. *Geometry of Feedback and Optimal Control*, chapter Symplectic methods in optimization and control, pages 19–77. CRC Press Book, 1998.

[8] A. Agrachev, L. Rizzi, and P. Silveira. On conjugate times of lq optimal control problems. *J. Dynamical and Control Systems*, 21:625–641, 2015.

[9] A. Agrachev and Yu. Sachkov. *Control theory from a geometric point of view.* Springer, 2004.

[10] A. Agrachev, G. Steffani, and P. Zezza. Strong optimality of a bang-bang trajectory. *SIAM J. on Control and Optimization*, 41:981–1014, 2002.

[11] V. Arnolv. The sturm theorems and symplectic geometry. *Funct. Anal. Appl.*, 19:251–259, 1985.

[12] B. Bonnard and M. Chyba. *Singular Trajectories and Their Role in Control Theory.* Springer, 2003.

[13] J.-B. Caillau, J. Fejoz, M. Orieux, and R. Roussarie. Singularities of min time affine control systems. preprint, hal-01718345, version 1.

[14] S. Cappell, R. Lee, and E. Miller. On the maslov index. *Comm. Pure Appl. Math.*, 47:121–186, 1994.

[15] E. Coddington and N. Levinson. *Theory of Ordinary Differential Equations*. Krieger Pub. Co., 1984.

[16] M. de Gosson. *Symplectic Geometry and Quantum Mechanics*. Birkhauser, 2000.

[17] V. Guillemin and S. Sternberg. *Geometric Asymptotics*. American Mathematical Society, 1977.

[18] M. Morse. Singular quadratic functionals. *Math. Ann.*, 201:60–76, 1973.

[19] M. Morse and W. Leighton. Singular quadratic functionals. *Trans. Amer. Math. Soc.*, 36:252–286, 1936.

[20] F. Olver and L. Maximon. *NIST Handbook of Mathematical Functions*, chapter Bessel Functions, pages 215–286. Cambridge University Press, 2010.

[21] N. Osmolovskii and H. Maurer. *Advances in Mathematical Modeling, Optimization and Optimal Control*, chapter Second-Order Optimality Conditions for Broken Extremals and Bang-Bang Controls, pages 147–201. Springer, 2016.

[22] W. Reid. *Riccati Differential Equations*. Academic Press, 1972.

[23] H. Shättler and U. Ledzewicz. *Geometric Optimal Control*. Springer, 2012.

[24] M. Shayman. Phase portrait of the matrix riccati equation. *SIAM J. Control and Optimization*, 24(1):1–65, 1986.

[25] H. Sussmann and W. Liu. Shortest paths for sub-riemannian metrics on rank-two distributions, 1995.

[26] H. J. Sussmann. *Algebraic and Geometric Methods in Nonlinear Control Theory*, chapter Envelopes, Conjugate Points, and Optimal Bang-Bang Extremals, pages 325–346. D. Reidel Publishing company, 1986.

[27] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Society, 2012.

[28] V. Wagner. The geometrical theory of the simplest n-dimensional singular problem of the calculus of variations. *Rec. Math. [Mat. Sbornik] N.S.*, 63:321–364, 1947.