On curvatures and focal points of dynamical Lagrangian distributions and their reductions by first integrals

Andrej A. Agrachev * Natalia N. Chtcherbakova † Igor Zelenko ‡

Abstract

Pairs (Hamiltonian system, Lagrangian distribution), called *dynamical Lagrangian distributions*, appear naturally in Differential Geometry, Calculus of Variations and Rational Mechanics. The basic differential invariants of a dynamical Lagrangian distribution w.r.t. the action of the group of symplectomorphisms of the ambient symplectic manifold are *the curvature operator* and *the curvature form*. These invariants can be seen as generalizations of the classical curvature tensor in Riemannian Geometry. In particular, in terms of these invariants one can localize the focal points along extremals of the corresponding variational problems. In the present paper we study the behavior of the curvature operator, the curvature form and the focal points of a dynamical Lagrangian distribution after its reduction by arbitrary first integrals in involution. The interesting phenomenon is that the curvature form of so-called monotone increasing Lagrangian dynamical distributions, which appear naturally in mechanical systems, does not decrease after reduction. It also turns out that the set of focal points to the given point w.r.t. the monotone increasing dynamical Lagrangian distribution and the corresponding set of focal points w.r.t. its reduction by one integral are alternating sets on the corresponding integral curve of the Hamiltonian system of the considered dynamical distributions. Moreover, the first focal point corresponding to the reduced Lagrangian distribution comes before any focal point related to the original dynamical distribution. We illustrate our results on the classical N-body problem.

Key words: curvature operator and form, focal points, reduction by first integrals, curves in Lagrangian Grassmannians.

1 Introduction

In the present paper smooth objects are supposed to be C^∞. The results remain valid for the class C^k with a finite and not large k but we prefer not to specify minimal possible k.

*S.I.S.S.A., Via Beirut 2-4, 34013 Trieste Italy and Steklov Mathematical Institute, ul. Gubkina 8, 117966 Moscow Russia; email: agrachev@sissa.it

†S.I.S.S.A., Via Beirut 2-4, 34013 Trieste Italy; email: chtch@sissa.it

‡S.I.S.S.A., Via Beirut 2-4, 34013 Trieste Italy; email: zelenko@sissa.it

1.1 Dynamical Lagrangian distributions. Let W be a symplectic manifold with symplectic form σ. Lagrangian distribution $\mathcal{D}$ on W is a smooth vector sub-bundle of the tangent bundle TW such that each fiber $\mathcal{D}_\lambda$ is a Lagrangian subspace of the linear symplectic space $T_\lambda W$, i.e., $\dim D_\lambda = \frac{1}{2}\dim W$ and $\sigma_\lambda(v_1, v_2) = 0$ for all $v_1, v_2 \in \mathcal{D}_\lambda$. For example, as a symplectic manifold one can take the cotangent bundle T^*M of a manifold M with standard symplectic structure and as a Lagrangian distribution one can take the distribution $\Pi(M)$ of tangent spaces to the fibers of T^*M, namely,

$$\Pi(M)_\lambda = T_\lambda\Big(T^*_{\pi(\lambda)}M\Big), \tag{1.1}$$

where $\pi : T^*M \to M$ is the canonical projection on the base manifold M.

Let $\mathcal{H}$ be a smooth function on W. Denote by $\vec{\mathcal{H}}$ the Hamiltonian vector field, corresponding to the function $\mathcal{H}$: $d\mathcal{H}(\cdot) = \sigma(\cdot, \vec{\mathcal{H}})$, and by $e^{t\vec{\mathcal{H}}}$ the Hamiltonian flow generated by $\vec{\mathcal{H}}$. The pair $(\vec{\mathcal{H}}, \mathcal{D})$ defines the one-parametric family of Lagrangian distributions $\mathcal{D}(t) = \big(e^{t\vec{\mathcal{H}}}\big)_*\mathcal{D}$. The pair $(\vec{\mathcal{H}}, \mathcal{D})$ will be called *dynamical Lagrangian distribution.* The point $\lambda_1 = e^{t_1\vec{\mathcal{H}}}\lambda_0$ is called *focal* to λ_0 w.r.t. the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the integral curve $t \mapsto e^{t\vec{\mathcal{H}}}\lambda_0$ of $\vec{\mathcal{H}}$, if

$$\big(e^{t_1\vec{\mathcal{H}}}\big)_*\mathcal{D}_{\lambda_0} \cap \mathcal{D}_{\lambda_1} \neq 0. \tag{1.2}$$

Dynamical Lagrangian distributions appear naturally in Differential Geometry, Calculus of Variations and Rational Mechanics. The model example can be described as follows:

Example 1 On a manifold M for a given smooth function $L : TM \mapsto \mathbb{R}$, which is convex on each fiber, consider the following standard problem of Calculus of Variation with fixed endpoints q_0 and q_1 and fixed time T:

$$A\Big(q(\cdot)\Big) = \int_0^T L\big(q(t), \dot{q}(t)\big)\, dt \mapsto \min \tag{1.3}$$

$$q(0) = q_0, \quad q(T) = q_1. \tag{1.4}$$

Suppose that the Legendre transform $H : T^*M \mapsto \mathbb{R}$ of the function L,

$$H(p, q) = \max_{X \in T_qM}\Big(p(X) - L(q, X)\Big), \quad q \in M, p \in T^*_qM, \tag{1.5}$$

is well defined and smooth on T^*M. We will say that the dynamical Lagrangian distributions $\big(\vec{H}, \Pi(M)\big)$ is *associated with the problem* (1.3)-(1.4)[1]. The curve $q : [0, T] \mapsto M$, satisfying (1.4), is an extremal of the problem (1.3)-(1.4) if and only if there exists an integral curve $\gamma : [0, T] \mapsto T^*M$ of $\vec{H}$ such that $q(t) = \pi\big(\gamma(t)\big)$ for all $0 \leq t \leq T$. In this case the point $\gamma(T)$ is focal to $\gamma(0)$ w.r.t. the pair $(\vec{H}, \mathcal{D})$ if and only if q_1 is conjugate to q_0 along the extremal $q(\cdot)$ in the classical variational sense for the problem (1.3)-(1.4). $\square$

[1] In the model example the Lagrangian distributions are integrable. For application of one-parametric families of non-integrable Lagrangian distributions see [10].

The group of symplectomorphisms of W acts naturally on Lagrangian distribution and Hamiltonian vector fields, therefore it acts also on dynamical Lagrangian distributions. Dynamical Lagrangian distributions have richer geometry w.r.t. this action than just Lagrangian distribution. For example, all integrable Lagrangian distributions are locally equivalent w.r.t. the action of the group of symplectomorphisms of W, while integrable dynamical Lagrangian distributions have functional moduli w.r.t. this action.

First note that for any two vector fields Y, Z tangent to the distribution $\mathcal{D}$ the number $\sigma_\lambda\big([\vec{\mathcal{H}},Y],Z\big)$ depends only on the vectors $Y(\lambda)$, $Z(\lambda)$.[2] Therefore for a given dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ the following bilinear form $Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(\cdot,\cdot)$ is defined on each $\mathcal{D}_\lambda$:

$$\forall v,w\in\mathcal{D}_\lambda:\quad Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,w)=\sigma_\lambda\big([\vec{\mathcal{H}},Y],Z\big),\ Y(\lambda)=v,Z(\lambda)=w\,. \tag{1.6}$$

Moreover, from the fact that all $\mathcal{D}_\lambda$ are Lagrangian it follows that the form $Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ is symmetric.

A dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ is called *regular*, if the quadratic forms $v\mapsto Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,v)$ are non-degenerated for any λ. A dynamical Lagrangian distributions is called *monotone (non-decreasing or non-increasing)*, if the quadratic forms $v\mapsto Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,v)$ are sign-definite (non-negative or non-positive definite) for any λ . The regular dynamical Lagrangian distributions is called *monotone increasing (decreasing)*, if the quadratic forms $v\mapsto Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,v)$ are positive (negative) definite for any λ .

Remark 1 If $W=T^*M$ and $\mathcal{D}=\Pi(M)$ are as in (1.1), then the form $v\mapsto Q_\lambda^{\left(\vec{\mathcal{H}},\Pi(M)\right)}(v,v)$ coincides with the second differential at λ of the restriction $\mathcal{H}\big|_{T^*_{\pi(\lambda)}M}$ of the Hamiltonian $\mathcal{H}$ to the fiber $T^*_{\pi(\lambda)}M$. Therefore in this case the dynamical Lagrangian distribution $\big(\vec{\mathcal{H}},\Pi(M)\big)$ is monotone increasing if and only if the restrictions of $\mathcal{H}$ on each fiber of T^*M are strongly convex. Consequently the dynamical Lagrangian distributions $\big(\vec{H},\Pi(M)\big)$ associated with the problem (1.3)-(1.4) is monotone increasing if and only if the restrictions of the function $L:TM\mapsto\mathbb{R}$ on each fiber of TM are strongly convex.□

It turns out that under some non-restrictive assumptions on the dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ (in particular, if this dynamical Lagrangian distribution is regular) one can assign to it a special linear operator $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ on each linear spaces $\mathcal{D}_\lambda$. This operator is called the *curvature operator* of $(\vec{\mathcal{H}},\mathcal{D})$ at λ and it is the basic differential invariant of dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ w.r.t. the action of the group of symplectomorphisms of W. Moreover, the following bilinear form

$$r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,w)=Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}\big(R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}v,w\big),\quad v,w\in\mathcal{D}_\lambda \tag{1.7}$$

is symmetric. The corresponding quadratic form is called *the curvature form* of the pair $(\vec{\mathcal{H}},\mathcal{D})$. Besides, the trace of the curvature operator

$$\rho_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}=\mathrm{tr}R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})} \tag{1.8}$$

[2]Here $[v_1,v_2]$ is the Lie bracket of the vector fields v_1 and v_2, $[v_1,v_2]=v_1\circ v_2-v_2\circ v_1$.

is called *the generalized Ricci curvature* of $(\vec{\mathcal{H}},\mathcal{D})$ at λ. All these invariants where introduced in [2] (see also [3] and section 2 below) and the effective method for their calculations is given in the recent work [4]. Below we present the results of these calculations on several important examples. In all these examples $W = T^*M$ for some manifold M and $\mathcal{D} = \Pi(M)$, a smooth function $L : TM \mapsto \mathbb{R}$ is given, the functional $A\big(q(\cdot)\big)$ is as in (1.3), and $H : T^*M \mapsto \mathbb{R}$ is as in (1.5).

Example 2 *(Natural mechanical system)* $M = R^n$, $W = \mathbb{R}^n \times \mathbb{R}^n$, $\sigma = \sum\limits_{i=1}^{n} dp_i \wedge dq_i$, $\mathcal{D}_{(p,q)} = (\mathbb{R}^n, 0)$, $L(q,X) = \frac{1}{2}\|X\|^2 - U(q)$ (in this case the function $A\big(q(\cdot)\big)$ is the Action functional of the natural mechanical system with potential energy $U(q)$). Then

$$\forall 1 \le i,j \le n: \quad r^{(\vec{H},\mathcal{D})}_{(p,q)}(\partial_{p_i},\partial_{p_j}) = \frac{\partial^2 U}{\partial q_i \partial q_j}(q). \tag{1.9}$$

In other words, in this case the curvature operator can be identified with the Hessian of the potential U.

Example 3 *(Riemannian manifold)* Let a Riemannian metric G is given on a manifold M by choosing an inner product $G_q(\cdot,\cdot)$ on each subspaces T_qM for any $q \in M$ smoothly w.r.t. q. Let $L(q,X) = \frac{1}{2}G_q(X,X)$. The inner product $G_q(\cdot,\cdot)$ defines the canonical isomorphism between T_q^*M and T_qM. For any $q \in M$ and $p \in T_q^*M$ we will denote by $p^\uparrow$ the image of p under this isomorphism, namely, the vector $p^\uparrow \in T_qM$, satisfying

$$p(\cdot) = G_q(p^\uparrow,\cdot) \tag{1.10}$$

(the operation $^\uparrow$ corresponds to the operation of raising of indexes in the corresponding coordinates of co-vectors and vectors). Since the fibers of T^*M are linear spaces, one can identify $\mathcal{D}_\lambda$ $(= T_\lambda T^*_{\pi(\lambda)}M)$ with $T^*_{\pi(\lambda)}M$, i.e., the operation $^\uparrow$ is defined also on each $\mathcal{D}_\lambda$ with values in $T_{\pi(\lambda)}M$. It turns out (see [2]) that

$$\forall v \in \mathcal{D}_\lambda: \quad \big(R^{(\vec{H},\mathcal{D})}_\lambda v\big)^\uparrow = R^\nabla\big(\lambda^\uparrow, v^\uparrow\big)\lambda^\uparrow \tag{1.11}$$

where R^∇ is the Riemannian curvature tensor of the metric G. The right-hand side of (1.11) appears in the classical Jacobi equation for Jacobi vector fields along the Riemannian geodesics. Also, $\frac{1}{n-1}\mathrm{tr}R^{(\vec{\mathcal{H}},\mathcal{D})}_\lambda$ is exactly the Ricci curvature calculated at $\lambda^\uparrow$. Besides, using (1.11), the Riemannian curvature tensor R^∇ can be recovered uniquely from the curvature operator $R^{(\vec{H},\mathcal{D})}_\lambda$. Therefore studying differential invariants of the appropriate integrable dynamical Lagrangian distributions, one can obtain the classical Riemannian tensor.

Example 4 *(Mechanical system on a Riemannian manifold)* Let G be the metric of the previous example and $L(q,X) = \frac{1}{2}G_q(X,X) - U(q)$ (in this case the function $A\big(q(\cdot)\big)$ is the Action functional of the mechanical system on the Riemannian manifold with potential $U(q)$). Let the operation $^\uparrow$ be as in (1.10). Then the curvature operator satisfies

$$\forall v \in \mathcal{D}_\lambda: \quad \big(R^{(\vec{H},\mathcal{D})}_\lambda v\big)^\uparrow = R^\nabla\big(\lambda^\uparrow, v^\uparrow\big)\lambda^\uparrow + \nabla_{v^\uparrow}(\mathrm{grad_G\,U})\big(\pi(\lambda)\big), \tag{1.12}$$

where $\mathrm{grad}_G U$ is the gradient of the function U w.r.t. the metric G, i.e., $\mathrm{grad}_G U = dU^{\dagger}$, and ∇ is the Riemannian covariant derivative. $\Box$

Remark 2 According to Remark 1, the dynamical Lagrangian distributions from Examples 2-4 are monotone increasing.$\Box$

The generalization of different kinds of Riemannian curvatures, using the notion of the curvature operator of dynamical Lagrangian distributions, leads to the generalization of several classical results of Riemannian geometry. In [2] for the given monotone increasing dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ the estimates of intervals between two consecutive focal points w.r.t. the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the integral curve $\gamma(t) = e^{t\vec{\mathcal{H}}}\lambda_0$ of $\vec{\mathcal{H}}$ were obtained in terms of the curvature form of the pair $(\vec{\mathcal{H}}, \mathcal{D})$. This result is the generalization of the classical Rauch Comparison Theorem in Riemannian geometry, which gives the lower and upper bounds of the interval between consecutive conjugate points along the Riemannian geodesics in terms of upper bound for the sectional curvatures and lower bound for the Ricci curvature respectively. In recent work [1] it was shown that the Hamiltonian flow, generated by a vector field $\vec{\mathcal{H}}$ on the compact level set of $\mathcal{H}$, is hyperbolic, if there exists a Lagrangian distribution $\mathcal{D}$ such that the dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ is monotone (increasing or decreasing) and the curvature form of so-called reduction of this dynamical distribution by Hamiltonian $\mathcal{H}$ on this level set is negative definite. This is an analog of the classical theorem about hyperbolicity of geodesic flows of negative sectional curvature on a compact Riemannian manifold.

1.2 The reduction by the first integrals. The subject of the present paper is the behavior of the curvature form and the focal points after the reduction of the dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ by the arbitrary s first integral $g_1, \ldots, g_s$ in involution of the Hamiltonian $\mathcal{H}$, i.e., s functions on W such that

$$\{\mathcal{H}, g_i\} = 0, \quad \{g_i, g_j\} = 0, \quad \forall 1 \le i, j \le s \tag{1.13}$$

(here $\{h, g\}$ is the Poisson bracket of the functions h and g, $\{h, g\} = dg(\vec{h})$). This problem appears naturally in the framework of mechanical systems and variational problems with symmetries. Let $\mathcal{G} = (g_1, \ldots, g_s)$ and

$$\mathcal{D}_\lambda^{\mathcal{G}} = \Big(\bigcap_{i=1}^{s} \ker\, d_\lambda g_i\Big) \cap \mathcal{D}_\lambda + \mathrm{span}\big(\vec{g}_1(\lambda), \ldots, \vec{g}_s(\lambda)\big). \tag{1.14}$$

Obviously, $\mathcal{D}^{\mathcal{G}}$ is a Lagrangian distribution. The pair $(\vec{\mathcal{H}}, \mathcal{D}^{\vec{\mathcal{G}}})$ is called *the reduction by the tuple $\mathcal{G}$ of s first integrals of $\mathcal{H}$ in involution* or shortly *the $\mathcal{G}$-reduction* of the dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$. The following example justifies the word "reduction" in the previous definition:

Example 5 Assume that we have one first integral g of $\mathcal{H}$ such that the Hamiltonian vector field $\vec{g}$, corresponding to the first integral g, preserves the distribution $\mathcal{D}_\lambda$, namely,

$$(e^{t\vec{g}})_* \mathcal{D} = \mathcal{D}. \tag{1.15}$$

Fixing some value c of g, one can define (at least locally) the following quotient manifold:

$$W_{g,c} = g^{-1}(c)/\mathcal{C}\,,$$

where $\mathcal{C}$ is the line foliation of the integral curves of the vector field $\vec{g}$. The symplectic form σ of W induces the symplectic form on a manifold $W_{g,c}$, making it symplectic too. Besides, if we denote by $\Phi : g^{-1}(c) \mapsto W_{g,c}$ the canonical projection on the quotient set, the vector field $\Phi_*(\vec{\mathcal{H}})$ is well defined Hamiltonian vector field on $W_{g,c}$, because by our assumptions the vector fields $\vec{\mathcal{H}}$ and $\vec{g}$ commute. Actually we have described the standard reduction of the Hamiltonian systems on the level set of the first integral, commonly used in Mechanics. In addition, by (1.15), $\Phi_*(\mathcal{D}^g)$ is well defined Lagrangian distribution on $W_{g,c}$. So, to any dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ on W one can associate the dynamical Lagrangian distribution $(\Phi_*\vec{\mathcal{H}}, \Phi_*\mathcal{D}^g)$ on the symplectic manifold $W_{g,c}$ of smaller dimension. It turns out (see subsection 2.2 below) that the curvature form of the g-reduction $(\vec{\mathcal{H}}, \mathcal{D}^g)$ at $\lambda \in g^{-1}(c)$ is equal to the pull-back by Φ of the curvature form of the dynamical Lagrangian distribution $(\Phi_*\vec{\mathcal{H}}, \Phi_*\mathcal{D}^g)$. So, instead of $(\vec{\mathcal{H}}, \mathcal{D}^g)$ one can work with $(\Phi_*\vec{\mathcal{H}}, \Phi_*\mathcal{D}^g)$ on the reduced symplectic space $W_{g,c}$. This is the essence of the reduction on the level set of the first integral.$\Box$

Remark 3 Suppose now that $W = T^*M$ for some manifold M and $\mathcal{D} = \Pi(M)$. In this case if g is a first integral of $\mathcal{H}$, which is "linear w.r.t. the impulses", i.e., there exists a vector field V on M such that

$$g(p,q) = p\big(V(q)\big), \quad q \in M, p \in T_q^*M, \tag{1.16}$$

then it satisfies (1.15). If we denote by $\mathcal{V}$ the line foliation of integral curves of V, then the reduced symplectic space $W_{g,c}$ can be identified with $T^*(M/\mathcal{V})$ and the distribution $\Pi(M)^g$ can be identified with $\Pi(M/\mathcal{V})$. So, after reduction we work with the dynamical Lagrangian distribution $\big(\Phi_*\vec{\mathcal{H}}, \Pi(M/\mathcal{V})\big)$ on the reduced phase space $T^*(M/\mathcal{V})$ instead of $\big(\vec{\mathcal{H}}, \Pi(M)\big)$. $\Box$

In view of the previous example the following analog of the notion of the focal points along the extremal w.r.t. the $\mathcal{G}$-reduction of the pair $(\vec{\mathcal{H}}, \mathcal{D})$ is natural: The point $\lambda_1 = e^{t_1\vec{\mathcal{H}}}\lambda_0$ is called *focal* to λ_0 w.r.t. the $\mathcal{G}$-reduction of the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the integral curve $t \mapsto e^{t\vec{\mathcal{H}}}\lambda_0$ of $\vec{\mathcal{H}}$, if

$$\Big(\big(e^{t_1\vec{\mathcal{H}}}\big)_*\mathcal{D}^{\mathcal{G}}_{\lambda_0} \cap \mathcal{D}^{\mathcal{G}}_{\lambda_1}\Big)/\mathrm{span}\big(\vec{g}_1(\lambda), \dots, \vec{g}_s(\lambda)\big) \neq 0. \tag{1.17}$$

In the situation, described in Example 5, the point $\lambda_1 = e^{t_1\vec{\mathcal{H}}}\lambda_0$ is focal to λ_0 w.r.t. the g-reduction of the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the curve $t \mapsto e^{t\vec{\mathcal{H}}}\lambda_0$ if and only if $\Phi(\lambda_1)$ is focal to $\Phi(\lambda_0)$ w.r.t. the pair $\big((\Phi)_*\vec{\mathcal{H}}, (\Phi)_*\mathcal{D}^g\big)$ along the curve $t \mapsto \Phi(e^{t\vec{\mathcal{H}}}\lambda_0)$ in $W_{g,c}$. We illustrate the meaning of the focal points of the reduction from the variational point of view on the following two examples. In both examples $W = T^*M$ for some manifold M and $\mathcal{D} = \Pi(M)$, a fiber-wise convex and smooth function $L : TM \mapsto \mathbb{R}$ is given and $H : T^*M \mapsto \mathbb{R}$ is as in (1.5).

Example 6 Assume that the Hamiltonian H admits a first integral g, satisfying (1.16). It is well known that g, satisfying (1.16), is the first integral of H if and only if the flow e^{tV} induces the one-parametric family of fiber-wise diffeomorphisms on TM, which preserve the function L, i.e., $L \circ (e^{tV})_* = L$.

Let $\mathcal{V}_1 : \mathbb{R} \mapsto M$ be an integral curve of V and $a(\cdot)$ be a function on $\mathcal{V}_1$ such that

$$a(\mathcal{V}_1(s)) = s, \quad s \in \mathbb{R}.$$

Fix some real c. Then for the given point q_0, and the time T consider the following variational problem

$$\int_0^T L\big(q(t), \dot{q}(t)\big)\, dt - c\, a\big(q(T)\big) \to \min, \tag{1.18}$$

$$q(0) = q_0, \quad q(T) = \mathcal{V}_1. \tag{1.19}$$

The curve $q : [0, T] \mapsto M$, satisfying (1.19), is an extremal of the problem (1.18)-(1.19) if and only if there exists an integral curve $\gamma : [0, T] \mapsto g^{-1}(c)$ of $\vec{H}$, such that $q(t) = \pi\big(\gamma(t)\big)$ for all $0 \le t \le T$. In this case the point $\gamma(0)$ is focal to $\gamma(T)$ w.r.t. the g-reduction of the pair $(\vec{H}, \mathcal{D})$ if and only if the point q_0 is focal to the point $q(T)$ along the extremal $q(\cdot)$ in the classical variational sense for the problem (1.18)-(1.19). $\square$

Example 7 Suppose that $g = \mathcal{H}$. For given real c and points q_0, q_1 consider the following variational problem with free terminal time

$$\int_0^T L\big(q(t), \dot{q}(t)\big)\, dt - cT \to \min, \quad T \text{ is free}, \tag{1.20}$$

$$q(0) = q_0, \quad q(T) = q_1. \tag{1.21}$$

The curve $q : [0, T] \mapsto M$, satisfying (1.21), is an extremal of the problem (1.20)-(1.21) if and only if there exists an integral curve $\gamma : [0, T] \mapsto H^{-1}(c)$ of $\vec{H}$, such that $q(t) = \pi\big(\gamma(t)\big)$ for all $0 \le t \le T$. In this case the point $\gamma(0)$ is focal to $\gamma(T)$ w.r.t. the H-reduction of the pair $(\vec{H}, \mathcal{D})$ if and only if the point q_0 is focal to the point q_1 along the extremal $q(\cdot)$ in the classical variational sense for the problem (1.20)-(1.21). Actually the considered case can be seen as a particular case of the previous example. For this one can pass to the extended (configuration) space $\overline{M} = M \times \mathbb{R}$ instead of M and take the following function $\overline{L} : T\overline{M} \mapsto \mathbb{R}$ instead of L:

$$\overline{L}(\bar{q}, \overline{X}) \stackrel{def}{=} L(q, \frac{X}{y})y,$$

where $\bar{q} \in \overline{M}$ such that $\bar{q} = (q, t)$, $q \in M$, $t \in \mathbb{R}$ and $\overline{X} \in T_{\bar{q}}\overline{M}$ such that $\overline{X} = (X, y)$, $\overline{X} = (X, y)$, $X \in T_q M$, $y \in T_t\mathbb{R} \cong \mathbb{R}$ (it is well known that (t, H) is the pair of conjugate variables for function $\overline{L}$, so as the field V one takes $\frac{\partial}{\partial t}$). $\square$

1.3 Description of main results. For the reduced dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D}^{\mathcal{G}})$ one can also define the curvature operator $R_\lambda^{(\vec{\mathcal{H}}, \mathcal{D}^{\mathcal{G}})}$ and the curvature form $r_\lambda^{(\vec{\mathcal{H}}, \mathcal{D}^{\mathcal{G}})}$ on each linear spaces $\mathcal{D}^{\mathcal{G}}_\lambda$. The natural problem is to find the relation between

$R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ (or $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$) and their reduced analogs $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}$ (or $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}$) on the linear space $\left(\bigcap_{i=1}^{s}\ker d_\lambda g_i\right)\cap\mathcal{D}_\lambda$ (which is the intersection of the corresponding spaces of definition $\mathcal{D}_\lambda$ and $\mathcal{D}_\lambda^{\mathcal{G}}$). We solve this problem in section 2 for regular dynamical distributions. It gives an effective and flexible method to compute and evaluate the curvature of Hamiltonian systems arising in Rational Mechanics and geometric variational problems. The interesting phenomenon is that *the curvature form of a monotone increasing Lagrangian dynamical distribution does not decrease after reduction.* More precisely, for such distribution the quadratic form

$$v\mapsto r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}(v,v)-r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v,v),\quad v\in\left(\bigcap_{i=1}^{s}\ker d_\lambda g_i\right)\cap\mathcal{D}_\lambda$$

is always *non-negative definite of rank not greater than* s, where s is the number of the first integrals in the tuple $\mathcal{G}$.

Further, in section 3 we show that the set of focal points to the given point along an integral curve w.r.t. the monotone increasing (or decreasing) dynamical Lagrangian distribution and the corresponding set of focal points w.r.t. its reduction by one integral are alternating sets on the curve and the first focal point w.r.t. the reduction comes before any focal point w.r.t. the original dynamical Lagrangian distribution. In view of Examples 6 and 7 this result looks natural: The reduction enlarge the set of admissible curves in the corresponding variational problems (instead of the problem with fixed endpoints ant terminal time one obtains the problem with variable endpoints or free terminal time). This justifies the fact that the first focal point of the reduction comes sooner. Besides, for the mentioned examples the last fact and the alternation of focal points are also a consequence of the classical Courant Minimax Principle, applied to the second variation along the reference extremal in the corresponding variational problems.

In addition, we demonstrate our results on the classical N - body problem.

2 Curvature and reduction

2.1 Curvature operator and curvature form. For the construction of the curvature operator of the dynamical Lagrangian distribution we use the theory of curves in the Lagrange Grassmannian, developed in [2] and [5]. The curve

$$t\mapsto J_\lambda(t)\stackrel{def}{=}e_*^{-t\vec{\mathcal{H}}}\left(\mathcal{D}_{e^{t\vec{\mathcal{H}}}\lambda}\right). \tag{2.1}$$

is called *the Jacobi curve of the curve* $t\mapsto e^{t\vec{\mathcal{H}}}\lambda$ *attached at the point* λ *(w.r.t. the dynamical distribution* $(\vec{\mathcal{H}},\mathcal{D})$*)*. It is the curve in the Lagrange Grassmannian $L(T_\lambda W)$ of the linear symplectic space $T_\lambda W$. Actually, the Jacobi curve is a generalization of the space of "Jacobi fields" along the extremal of variational problem of type (1.3)-(1.4). Note that if $\bar{\lambda} = e^{\bar{t}\vec{\mathcal{H}}}\lambda$ then by (2.1) we have

$$J_{\bar{\lambda}}(t) = e_*^{\bar{t}\vec{\mathcal{H}}} J_\lambda(t - \bar{t}).$$

In other words, the Jacobi curves of the same integral curve of $\vec{\mathcal{H}}$ attached at two different points of this curve are the same, up to symplectic transformation between the corresponding ambient linear symplectic spaces and the corresponding shift of the parameterizations. Therefore, any differential invariants of the Jacobi curve w.r.t. the action of the linear Symplectic group (in other words, any symplectic invariant of the curve) produces the invariant of the corresponding dynamical Lagrangian distributions w.r.t. the action of the group of symplectomorphisms of the ambient space W.

Now, following [2] and [5], we describe the construction of the curvature operator of the curve in the Lagrange Grassmannian. Let Σ be $2n$-dimensional linear space, endowed with symplectic form σ. The Lagrange Grassmannian $L(\Sigma)$ is real analytic manifold. Note that the tangent space $T_\Lambda L(\Sigma)$ to the Lagrangian Grassmannian at the point Λ can be naturally identified with the space of quadratic forms $\mathrm{Quad}(\Lambda)$ on the linear space $\Lambda \subset \Sigma$. Namely, take a curve $\Lambda(t) \in L(\Sigma)$ with $\Lambda(0) = \Lambda$. Given some vector $l \in \Lambda$, take a curve $l(\cdot)$ in W such that $l(t) \in \Lambda(t)$ for all t and $l(0) = l$. Define the quadratic form

$$l \mapsto \sigma(\frac{d}{dt}l(0), l). \tag{2.2}$$

Using the fact that the spaces $\Lambda(t)$ are Lagrangian, it is easy to see that this form depends only on $\frac{d}{dt}\Lambda(0)$. So, we have the map from $T_\Lambda L(\Sigma)$ to the space $\mathrm{Quad}(\Lambda)$. A simple counting of dimension shows that this mapping is a bijection.

Remark 4 In the sequel, depending on the context, we will look on the elements of $T_\Lambda L(\Sigma)$ not only as on the quadratic forms on $\Lambda(t)$, but also as on the corresponding symmetric bilinear forms on $\Lambda(t)$ or on the corresponding self-adjoint operator from $\Lambda(t)$ to $\Lambda(t)^*$ $\Box$

The curve $\Lambda(\cdot)$ in $L(\Sigma)$ is called *regular, monotone, monotone increasing (decreasing)*, if its velocity $\dot{\Lambda}(t)$ at any point t is respectively a non-degenerated, sign-definite, positive (negative) definite quadratic form on the space $\Lambda(t)$.

Proposition 1 *A dynamical distribution $(\vec{\mathcal{H}}, \mathcal{D})$ is regular, monotone, monotone increasing (decreasing) if and only if all Jacobi curves w.r.t. this distribution are respectively regular, monotone, monotone increasing (decreasing) curves in the corresponding Lagrange Grassmannians.*

Proof. Recall that for any two vector fields $\vec{\mathcal{H}}$ and ℓ in M one has

$$\frac{d}{dt}\Big((e^{-t\vec{\mathcal{H}}})_*\ell\Big) = (e^{-t\vec{\mathcal{H}}})_*[\vec{\mathcal{H}}, \ell]. \tag{2.3}$$

Let $Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ be as in (1.6). Applying this fact to the Jacobi curve $J_\lambda(t)$ and using (1.6), (2.1), and (2.2) one obtains easily that

$$Q_\lambda^{(\vec{\mathcal{H}},\mathcal{D})} = \dot{J}_\lambda(0), \tag{2.4}$$

which implies the statement of the proposition. □

Fix some $\Lambda \in L(\Sigma)$. Define the linear mapping $B_\Lambda : \Sigma \mapsto \Lambda^*$ in the following way: for given $w \in \Sigma$ one has

$$B_\Lambda(w)(v) = \sigma(w, v), \quad \forall v \in \Lambda. \tag{2.5}$$

Denote by $\Lambda^\pitchfork$ the set of all Lagrangian subspaces of Σ transversal to Λ, i.e. $\Lambda^\pitchfork = \{\Gamma \in G_m(W) : \Gamma \cap \Lambda = 0\}$. Then for any subspace $\Gamma \in \Lambda^\pitchfork$ the restriction $B_\Lambda\big|_\Gamma : \Gamma \mapsto \Lambda^*$ is an isomorphism.

Remark 5 In other words, any $\Gamma \in \Lambda^\pitchfork$ can be canonically identified with the dual space Λ^*.

Let $I_\Gamma = \left(B_\Lambda\big|_\Gamma\right)^{-1}$. Note that by construction I_Γ is linear mapping from Λ^* to Γ and

$$I_\Gamma l - I_\Delta l \in \Lambda. \tag{2.6}$$

The crucial observation is that the set $\Lambda^\pitchfork$ can be considered as an affine space over the linear space $\mathrm{Quad}(\Lambda^*)$ of all quadratic forms on the space Λ^*. Indeed, one can define the operation of subtraction on $\Lambda^\pitchfork$ with values in $\mathrm{Quad}(\Lambda^*)$ in the following way:

$$(\Gamma - \Delta)(l) = \sigma(I_\Gamma l, I_\Delta l). \tag{2.7}$$

It is not difficult to show that $\Lambda^\pitchfork$ endowed with this operation of subtraction satisfies the axioms of affine space. For example, let us prove that

$$(\Gamma - \Delta) + (\Delta - \Pi) = (\Gamma - \Pi) \tag{2.8}$$

Indeed, using skew-symmetry of σ and relation (2.6), one has the following series of identities for any $l \in \Lambda^*$

$$\sigma(I_\Gamma l, I_\Delta l) + \sigma(I_\Delta l, I_\Pi l) = \sigma(I_\Gamma l - I_\Pi l, I_\Delta l) = \sigma(I_\Gamma l - I_\Pi l, I_\Delta l - I_\Pi L) +$$

$$\sigma(I_\Gamma l - I_\Pi l, I_\Pi l) = \sigma(I_\Gamma l - I_\Pi l, I_\Pi l) = \sigma(I_\Gamma l, I_\Pi l),$$

which implies (2.8).[3]

Consider now some curve $\Lambda(\cdot)$ in $L(\Sigma)$. Fix some parameter τ. Assume that $\Lambda(t) \in \Lambda(\tau)^\pitchfork$ for all t from a punctured neighborhood of τ. Then we obtain the curve $t \mapsto \Lambda(t) \in \Lambda(\tau)^\pitchfork$ in the affine space $\Lambda(\tau)^\pitchfork$. Denote by $\Lambda_\tau(t)$ the identical embedding of $\Lambda(t)$ in the affine space $\Lambda(\tau)^\pitchfork$. Fixing an "origin" Δ in $\Lambda(\tau)^\pitchfork$ we obtain a vector function $t \mapsto \Lambda_\tau(t) - \Delta$ with values in $\mathrm{Quad}\,(\Lambda^*)$. The curve $\Lambda(\cdot)$ is called *ample* at the point τ if the function $\Lambda_\tau(t) - \Delta$ has the pole at $t = \tau$ (obviously, this definition does not depend on the choice of the "origin" Δ in $\Lambda(\tau)^\pitchfork$). In particular, if $\Lambda(\cdot)$ is a regular curve in $L(\Sigma)$, then one can show without difficulties that the function $t \mapsto \Lambda_\tau(t) - \Delta$ has a simple pole at $t = \tau$ for any $\Delta \in \Lambda(\tau)^\pitchfork$. Therefore any regular curve in $L(\Sigma)$ is ample at any point.

[3] For slightly different description of the affine structure on $\Lambda^\pitchfork$ see [2],[5], and also [4], where a similar construction is given for the Grassmannian $G_n(\mathbb{R}^{2n})$ of half-dimensional subspaces of $\mathbb{R}^{2n}$.

Suppose that the curve $\Lambda(\cdot)$ is ample at some point τ. Using only the axioms of affine space, one can prove easily that there exist a unique subspace $\Lambda^\circ(\tau) \in \Lambda^\pitchfork$ such that the free term in the expansion of the function $t \mapsto \Lambda_\tau(t) - \Lambda^\circ(\tau)$ to the Laurent series at τ is equal to zero. The subspace $\Lambda^\circ(\tau)$ is called the *derivative subspace of the curve* $\Lambda(\cdot)$ *at the point* τ. If the curve $\Lambda(\cdot)$ is ample at any point, one can consider the curve $\tau \mapsto \Lambda^\circ(\tau)$ of the derivative subspaces. This curve is called *derivative curve* of the curve $\Lambda(\cdot)$.

Now assume that the derivative curve $\Lambda^\circ(t)$ is smooth at a point τ. In particular, the derivative curve of a regular curve in the Lagrange Grassmannian is smooth at any point (see, for example, [2] and the coordinate representation below). In general, the derivative curve of the ample curve is smooth at generic points (the points of so-called constant weight, see [5]). As was mentioned already in Remark 4, one can look on $\dot\Lambda(\tau)$ and $\dot\Lambda^\circ(\tau)$ as on the corresponding self-adjoint linear mappings:

$$\dot\Lambda(\tau) : \Lambda(\tau) \mapsto \Lambda(\tau)^*, \quad \dot\Lambda^\circ(\tau) : \Lambda^\circ(\tau) \mapsto \big(\Lambda^\circ(\tau)\big)^* \tag{2.9}$$

Besides, by construction $\Lambda^\circ(\tau) \in \Lambda(\tau)^\pitchfork$. Therefore by Remark 5 the following spaces can be canonically identified:

$$\Lambda(\tau)^* \cong \Lambda^\circ(\tau), \quad \big(\Lambda^\circ(\tau)\big)^* \cong \Lambda(\tau) \tag{2.10}$$

After these identifications, the composition $\dot\Lambda^\circ(\tau) \circ \dot\Lambda(\tau)$ is well-defined linear operator on $\Lambda(t)$.

Definition 1 *The linear operator*

$$R_\Lambda(\tau) = -\dot\Lambda^\circ(\tau) \circ \dot\Lambda(\tau)\,. \tag{2.11}$$

on $\Lambda(\tau)$ *is called the curvature operator of the curve* $\Lambda(\cdot)$ *at a point* τ. *The quadratic form* $r_\Lambda(\tau)$ *on* $\Lambda(\tau)$*, defined by*

$$r_\Lambda(\tau)(v) = \big(\dot\Lambda(\tau) \circ R_\Lambda(\tau)v\big)(v), \quad v \in \Lambda(\tau) \tag{2.12}$$

is called the curvature form of the curve $\Lambda(\cdot)$ *at a point* τ.

Suppose that for all Jacobi curves w.r.t. the dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ the curvature operator is defined. *The curvature operator* $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ *of the dynamical Lagrangian distribution* $(\vec{\mathcal{H}}, \mathcal{D})$ *at a point* λ is by definition the curvature operator of the Jacobi curve $J_\lambda(t)$ at $t = 0$, namely,

$$R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})} = R_{J_\lambda}(0). \tag{2.13}$$

By construction, it is the linear operator on $\mathcal{D}_\lambda$. *The curvature form* $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ *of the dynamical Lagrangian distribution* $(\vec{\mathcal{H}}, \mathcal{D})$ *at a point* λ is by definition the curvature form of the Jacobi curve $J_\lambda(t)$ at $t = 0$ (see also (1.7)).

Now for a regular curve $\Lambda(\cdot)$ in the Lagrange Grassmannian $L(\Sigma)$ let us give a coordinate representation of the derivative curve and the curvature operator. One can choose a basis in Σ such that

$$\Sigma \cong \mathbb{R}^n \times \mathbb{R}^n = \{(x, y) : x, y \in \mathbb{R}^n\}, \tag{2.14}$$

$$\sigma((x_1, y_1), (x_2, y_2)) = \langle x_1, y_2\rangle - \langle x_2, y_1\rangle, \tag{2.15}$$

where $\langle \cdot, \cdot\rangle$ is the standard inner product in $\mathbb{R}^n$ (such basis is called *symplectic* or *Darboux basis*). Denote by e_i the ith vector of the standard basis of $\mathbb{R}^n$.

Assume also that $\Lambda(\tau) \cap \{(0, y) : y \in \mathbb{R}^n\} = 0$. Then for any t sufficiently closed to τ there exits the symmetric $n \times n$ matrix S_t such that $\Lambda(t) = \{(x, S_t x) : x \in \mathbb{R}^n\}$. The matrix curve $t \mapsto S_t$ is *the coordinate representation of the curve* $\Lambda(\cdot)$ *(w.r.t. the chosen symplectic basis in* Σ*).*

Remark 6 *Note that from (2.15) and the fact that the subspaces* $\Lambda(t)$ *are Lagrangian it follows that the matrices* S_t *are symmetric.*

The curve $\Lambda(\cdot)$ is regular if and only if the matrices $\dot S_t$ are non-degenerated. The expression of the derivative curve and the curvature operator of the regular curve $\Lambda(\cdot)$ in terms of S_t is given by the following

Proposition 2 *The derivative curve* $\Lambda^\circ(\tau)$ *of the regular curve* $\Lambda(\tau)$ *in* $L(\Sigma)$ *satisfies*

$$\Lambda^\circ(\tau) = \{(-\frac{1}{2}\dot S_\tau^{-1}\ddot S_\tau \dot S_\tau^{-1} y, y - \frac{1}{2} S_\tau \dot S_\tau^{-1}\ddot S_\tau \dot S_\tau^{-1} y), y \in \mathbb{R}^n\}. \tag{2.16}$$

In the basis $\{(e_i, S_\tau e_i)\}_{i=1}^n$ *of* $\Lambda(\tau)$ *the curvature operator* $R_\Lambda(\tau)$ *is represented by the following matrix*

$$\mathbb{S}(S_t) = \frac{1}{2}\dot S_\tau^{-1} S_\tau^{(3)} - \frac{3}{4}(\dot S_\tau^{-1}\ddot S_\tau)^2. \tag{2.17}$$

The curvature form $r_\Lambda(\tau)$ *has the following matrix w.r.t. the same basis*

$$-\dot S_\tau \mathbb{S}(S_t) = -\frac{1}{2}\dot S_\tau^{(3)} + \frac{3}{4}\ddot S_\tau \dot S_\tau^{-1}\ddot S_\tau. \tag{2.18}$$

For the proof of (2.16) and of the matrix representation (2.17) for the curvature operator see, for example, [4]. The matrix representation (2.18) of the curvature form follows directly from (2.17) and (2.12).

Remark 7 If S_t is a scalar function (i.e., $n = 1$), then $\mathbb{S}(S_t)$ is just the classical Schwarzian derivative or Schwarzian of S_t. It is well known that for scalar functions the Schwarzian satisfies the following remarkable identity:

$$\mathbb{S}\left(\frac{a\varphi(t) + b}{c\varphi(t) + d}\right) = \mathbb{S}\big(\varphi(t)\big) \tag{2.19}$$

for any constant a, b, c, and d, $ad - bc \neq 0$. Note that by choosing another symplectic basis in Σ, we obtain a new coordinate representation $t \mapsto \widetilde S_t$ of the curve $\Lambda(\cdot)$ which is a matrix Möbious transformation of S_t,

$$\widetilde S_t = (C + DS_t)(A + BS_t)^{-1} \tag{2.20}$$

for some $n \times n$ matrix A, B, C, and D. It turns out that the matrix Schwarzian (2.17) is invariant w.r.t. matrix Möbious transformations (2.20) by analogy with identity (2.19) (the only difference is that instead of identity we obtain similarity of corresponding matrices). This is another explanation for invariant meaning of the expression (2.17), given by Proposition 2.

The coordinate representations (2.17) and (2.18) are crucial in the proof of the main theorem of this section (see Theorem 1 below).

2.2 Curvature operator and curvature form of reduction. Now fix some s vectors $l_1, \dots l_s$ in Σ such that

$$\forall 1 \le i, j \le s: \quad \sigma(l_i, l_j) = 0. \tag{2.21}$$

Denote by $\ell = (l_1, \dots l_s)$ and $\operatorname{span}\ell = \operatorname{span}(l_1, \dots l_s)$ For any $\Lambda \in L(\Sigma)$ let

$$\Lambda^\ell = \Lambda \cap (\operatorname{span}\ell)^\angle + \operatorname{span}\ell, \quad \overline{\Lambda^\ell} = \Lambda^\ell/\operatorname{span}\ell, \tag{2.22}$$

where $(\operatorname{span}\ell)^\angle \overset{def}{=} \{v \in \Sigma : \forall 1 \le i \le s \quad \sigma(v, l_i) = 0\}$ is the skew-orthogonal complement of the isotropic subspace $\operatorname{span}\ell$. Actually, $\overline{\Lambda^\ell}$ is a Lagrangian subspace of the symplectic space $(\operatorname{span}\ell)^\angle/\operatorname{span}\ell$ (with symplectic form induced by σ).

Let, as before, $\Lambda(\cdot)$ be an ample curve in $L(\Sigma)$. The curve $\Lambda(\cdot)^\ell$ is called *the reduction by the s-tuple* ℓ, satisfying (2.21), or shortly *the* ℓ*-reduction* of the curve $\Lambda(\cdot)$. Note that by (2.22) $\operatorname{span}\ell \subset \Lambda(t)^\ell$ for any t. Therefore the curve $\Lambda(\cdot)^\ell$ is not ample and the constructions of the previous subsection cannot be applied to it directly. Instead, suppose that the curve $\overline{\Lambda(\cdot)^\ell}$ is ample curve in the Lagrange Grassmannian $L((\operatorname{span}\ell)^\angle/\operatorname{span}\ell)$. Then the curvature operator $R_{\overline{\Lambda^\ell}}(t)$ of this curve is well-defined linear operator on the space $\overline{\Lambda(t)^\ell}$ (at least for a generic point t). Let $\phi : \Sigma \mapsto \Sigma/\operatorname{span}\ell$ be the canonical projection on the factor-space.

Definition 2 *The curvature operator $R_{\Lambda^\ell}(\tau)$ of the ℓ-reduction $\Lambda(\cdot)^\ell$ at a point τ is the linear operator on $\Lambda(\tau)^\ell$, satisfying*

$$R_{\Lambda^\ell}(\tau)(v) = \left(\phi|_{\Lambda(\tau)\cap(\operatorname{span}\ell)^\angle}\right)^{-1} \circ R_{\overline{\Lambda^\ell}}(\tau) \circ \phi(v), \quad v \in \Lambda(\tau)^\ell. \tag{2.23}$$

The curvature form $r_{\Lambda^\ell}(\tau)$ of the ℓ-reduction $\Lambda(\cdot)^\ell$ at a point τ is the quadratic form on $\Lambda(\tau)^\ell$, satisfying

$$r_{\Lambda^\ell}(\tau)(v) = \frac{d}{d\tau}\big(\Lambda(\tau)^\ell\big)\big(R_{\Lambda^\ell}(\tau)v, v\big). \tag{2.24}$$

All these constructions are directly related to the reduction of dynamical distributions by a tuple $\mathcal{G} = (g_1, \dots, g_s)$ of s involutive first integrals, defined in Introduction. Indeed, the Jacobi curves attached at some point λ w.r.t. the $\mathcal{G}$-reduction $(\vec{\mathcal{H}}, \mathcal{D}^\mathcal{G})$ of a dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ are exactly $\big(\vec{g}_1(\lambda), \dots, \vec{g}_s(\lambda)\big)$-reductions of the Jacobi curves attached at λ w.r.t. $(\vec{\mathcal{H}}, \mathcal{D})$ itself. *The curvature operator* $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^\mathcal{G})}$ *and the curvature form* $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^\mathcal{G})}$ *at λ of the $\mathcal{G}$-reduction* $(\vec{\mathcal{H}}, \mathcal{D}^\mathcal{G})$ are by definition the curvature operator and the curvature form of the Jacobi curves attached at λ w.r.t. $(\vec{\mathcal{H}}, \mathcal{D}^\mathcal{G})$.

To justify these definitions suppose that we are in situation of Example 5, i.e. $\mathcal{H}$ admits one first integral g, satisfying (1.15). Let a symplectic manifold $W_{g,c}$ and a mapping $\Phi : g^{-1}(c) \mapsto W_{g,c}$ be as in this example. Then directly from the definition it follows that

$$\forall v \in \mathcal{D}^g : \quad r_\lambda^{(\vec{h},\mathcal{D}^g)}(v) = r_{\Phi(\lambda)}^{(\Phi_*\vec{h},\Phi_*\mathcal{D}^g)}(\Phi_* v). \tag{2.25}$$

In other words, the curvature form of the g-reduction $(\vec{\mathcal{H}}, \mathcal{D}^g)$ at $\lambda \in g^{-1}(c)$ is equal to the pull-back by Φ of the curvature form of the dynamical Lagrangian distribution $(\Phi_*\vec{\mathcal{H}}, \Phi_*\mathcal{D}^g)$, associated to the original dynamical Lagrangian distribution $(\vec{h}, \mathcal{D})$ on the reduced symplectic space $W_{g,c}$.

The natural question is what is the relation between the curvature forms and operators of the dynamical Lagrangian distribution and its reduction on the common space of their definition. Before answering this question in the general situation, let us consider the following simple example:

Example 8 *(Kepler's problem)* Consider a natural mechanical system on $M = \mathbb{R}^2$ with the potential energy $U = -r^{-1}$, where r is the distance between a moving point in a plane and some fixed point. This system describes the motion of the center of masses of two gravitationally interacting bodies in the plane of their motion (see [7]). Let $q = (r, \varphi)$ be the polar coordinates in $\mathbb{R}^2$. Then the Hamiltonian function of the problem takes the form

$$h = \frac{p_r^2}{2} + \frac{p_\varphi^2}{2r^2} - \frac{1}{r}, \tag{2.26}$$

where p_r and p_φ are the canonical impulses conjugated to r and φ. If $\lambda = (p, q)$, where $q \in M$, $p \in T_q^*M$, then $p_r(\lambda) = p\big(\partial_r(q)\big) = d_q r$, $p_\varphi(\lambda) = r^2 p\big(\partial_\varphi(q)\big) = r^2 d_q\varphi$. Observe that $g = p_\varphi$ is nothing but the angular momentum and from (2.26) we immediately see that it is a first integral of the system. Let us compare the curvature forms $r_\lambda^{(\vec{h},\Pi(M))}$ and $r_\lambda^{(\vec{h},\Pi(M)^g)}$ on the common space $\Pi(M) \cap \ker d_\lambda g = \mathbb{R}\partial_{p_r}$ of their definition.

First, according to (1.9) of Example 2, the curvature form of $(\vec{h}, \Pi(M))$ is equal the Hessian of U at q. In particular, it implies that

$$r_\lambda^{\left(\vec{h},\Pi(M)\right)}(\partial_{p_r}) = \frac{\partial^2}{\partial r^2}U(q) = -\frac{2}{r^3}. \tag{2.27}$$

Further, let $c = g(\lambda)$. Note that g satisfies the condition (1.16) of Remark 3 with $V = r^2\partial_\varphi$. Let $W_{g,c}$ and Φ be as in Example 5. Then, following Remark 3, $W_{g,c} \cong T^*\mathbb{R}^+$ and the dynamical Lagrangian distribution $(\Phi_*\vec{h}, \Phi_*\mathbb{R}\partial_{p_r})$ is equivalent (symplectomorphic) to the dynamical Lagrangian distribution associated with the natural mechanical system with configuration space $\mathbb{R}^+$ and the potential energy

$$U_a = \frac{c^2}{2r^2} - \frac{1}{r}$$

(U_a is the so - called *amended* potential energy; it comes from the following identity: $h\big|_{g^{-1}(c)} = \frac{p_r^2}{2} + U_a(r)$). Hence by (2.25)

$$r_\lambda^{\left(\vec{h},\Pi(M)^g\right)}(\partial_{p_r}) = \frac{d^2}{dr^2}U_a(r) = \frac{3c^2}{r^4} - \frac{2}{r^3} = r_\lambda^{\left(\vec{h},\Pi(M)\right)}(\partial_{p_r}) + \frac{3c^2}{r^4}. \tag{2.28}$$

Note that from (2.28) it follows that on the common space of the definition the reduced curvature form is not less than the curvature form itself. We will show later (Corollary 2) that this is a general fact.$\Box$

2.3 The change of the curvature after the reduction. Now we give the relation between the curvature forms of the regular curve $\Lambda(\cdot)$ and its ℓ-reduction $\Lambda(\cdot)^\ell$, where ,as before, $\ell = (l_1, \dots, l_s)$ is the tuple of s vectors, satisfying (2.21). First, let us introduce some notations. Let $B_{\Lambda(t)} : \Sigma \mapsto \Lambda^*$ be as in (2.5). Looking at $\dot\Lambda(t)$ as at a linear mapping from $\Lambda(t)$ to $\Lambda(t)^*$, denote by $a_i(t)$, $1 \le i \le s$ the following vectors in $\Lambda(t)$:

$$a_i(t) = \big(\dot\Lambda(t)\big)^{-1} \circ B_{\Lambda(t)}(l_i). \tag{2.29}$$

Using definition of $\dot\Lambda(t)$ and $B_{\Lambda(t)}$ one can show that $t \mapsto a_i(t)$, $a_i(t) \in \Lambda(t)$, is a unique vector function such that for any t one has

$$\sigma(\dot a_i(t), v) = \sigma(l_i, v), \quad \forall v \in \Lambda(t),$$

or, equivalently,

$$\dot a_i(t) \equiv l_i \mod \Lambda(t). \tag{2.30}$$

Finally, let $A(t)$ be the $s \times s$ matrix with the following entries

$$A_{km}(t) = \sigma\big(l_k, a_m(t)\big), \quad 1 \le k, m \le s. \tag{2.31}$$

Note that by (2.30) and the definitions of $\dot\Lambda(t)$

$$A_{km}(\tau) = \sigma\big(\dot a_k(t), a_m(t)\big) = \dot\Lambda(t)a_k(t)\big(a_m(t)\big), \tag{2.32}$$

which implies that the matrix $A(t)$ is symmetric.

Theorem 1 *Suppose that $\Lambda(\cdot)$ is a regular curve in the Lagrange Grassmannian $L(\Sigma)$ and $\ell = (l_1, \dots, l_s)$ is a tuple of s vectors in Σ such that (2.21) holds and $\det A(\tau) \neq 0$ for some point τ. Then the curvature form $r_\Lambda(\tau)$ of the curve $\Lambda(\cdot)$ and the curvature form $r_{\Lambda^\ell}(t)$ of its ℓ-reduction $\Lambda(\cdot)^l$ at the point τ satisfy the following identity for all $v \in \Lambda(\tau) \cap \big(\mathrm{span}\, \ell\big)^\angle$:*

$$r_{\Lambda^\ell}(\tau)(v) - r_\Lambda(\tau)(v) = \frac{3}{4} \sum_{k,m=1}^{s} \big(A(\tau)^{-1}\big)_{km} \sigma\big(\ddot a_k(\tau), v\big)\sigma\big(\ddot a_m(\tau), v\big), \tag{2.33}$$

where $\big(A(\tau)^{-1}\big)_{km}$ is the km-entry of the matrix $A(\tau)^{-1}$. In addition, the curvature operator $R_\Lambda(\tau)$ of the curve $\Lambda(\cdot)$ and the curvature operator $R_{\Lambda^\ell}(\tau)$ of its ℓ-reduction $\Lambda(\cdot)^l$ at the point τ satisfy on $\Lambda(\tau) \cap \big(\mathrm{span}\, \ell\big)^\angle$ the following identity:

$$R_{\Lambda^\ell}(\tau) - R_\Lambda(\tau) = \frac{3}{4} \sum_{k,m=1}^{s} (A(\tau)^{-1})_{km} B_{\Lambda(\tau)} \ddot a_m(\tau) \otimes \Big(\big(\dot\Lambda(\tau)\big)^{-1} \circ B_{\Lambda(\tau)} \ddot a_k(\tau)\Big). \tag{2.34}$$

(As usual, for a given linear functional ξ and a given vector v by $\xi \otimes v$ we denote the following rank 1 linear operator $\xi \otimes v(\cdot) = \xi(\cdot)v$.)

Proof. First let us prove identity (2.33). As before, denote by e_i the ith vector of the standard basis of $\mathbb{R}^n$. The condition $\det A(\tau) \neq 0$ it equivalent to

$$\operatorname{span}\big(a_1(\tau), \dots, a_s(\tau)\big) \cap \big(\operatorname{span} \ell\big)^{\angle} = 0. \tag{2.35}$$

Hence one can choose a basis in Σ such that if one coordinatizes Σ w.r.t. this basis, $\Sigma \cong \mathbb{R}^n \times \mathbb{R}^n$, then the symplectic form σ satisfies (2.15) and the following relations hold

$$\Lambda(\tau) \cap \big(\operatorname{span} \ell\big)^{\angle} = \operatorname{span}\big((e_1, 0), \dots, (e_{n-s}, 0)\big), \tag{2.36}$$
$$l_i = (0, e_{n-s+i}), \quad 1 \leq i \leq s, \tag{2.37}$$
$$a_i(\tau) \in \operatorname{span}\big((e_{n-s+1}, 0), \dots, (e_n, 0)\big), \quad 1 \leq i \leq s. \tag{2.38}$$

Note that by construction

$$\big(\operatorname{span} \ell\big)^{\angle} = \operatorname{span}\big((e_1, 0), \dots, (e_{n-s}, 0), (0, e_1), , \dots, (0, e_n)\big).$$

Therefore one can make the following identification:

$$(\operatorname{span} \ell)^{\angle}/\operatorname{span} \ell \cong \operatorname{span}\big((e_1, 0), \dots, (e_{n-s}, 0), (0, e_1), \dots, (0, e_{n-s})\big). \tag{2.39}$$

Since by definition $a_i(t) \in \Lambda(t)$, there exists $b_i(t) \in \mathbb{R}^n$ such that $a(t) = \big(b_i(t), S_t b_i(t)\big)$. Note that

$$\dot a_i(t) = \big(0, \dot S_t b_i(t)\big) + \big(\dot b_i(t), S_t \dot b_i(t)\big) \equiv \big(0, \dot S_t b_i(t)\big) \bmod \Lambda(t). \tag{2.40}$$

This, together with (2.30) and (2.37), implies that $l_i = \big(0, \dot S_t b_i(t)\big)$ and then

$$b_i(t) = \dot S_t^{-1} e_{n-s+i}. \tag{2.41}$$

On the other hand, from (2.37) and (2.38), using (2.15) one can obtain that

$$b_i(\tau) = \sum_{j=1}^{s} \sigma\big(a_i(\tau), l_j\big) e_{n-s+j} = -\sum_{j=1}^{s} A_{ij}(\tau) e_{n-s+j} \tag{2.42}$$

(in the last equality we used the symmetry of the matrix $A(\tau)$). So, from (2.41), (2.42) and symmetry of $\dot S_\tau$ (see Remark 6) it follows that

$$\forall 1 \leq i \leq n-s, \quad n-s+1 \leq j \leq n: \quad (\dot S_\tau^{-1})_{ij} = (\dot S_\tau^{-1})_{ji} = 0, \tag{2.43}$$
$$\forall 1 \leq i, j \leq s: \quad (\dot S_\tau^{-1})_{n-s+i, n-s+j} = -A_{ij}(\tau). \tag{2.44}$$

Further, for given $n \times n$ matrix $\mathcal{A}$ denote by $C(\mathcal{A})$ the $(n-s) \times (n-s)$ matrix, obtained from $\mathcal{A}$ by erasing the last s columns and rows. Consider the curve $\overline{\Lambda(\cdot)^{\ell}}$ in the Lagrange Grassmannian $L((\operatorname{span} \ell)^{\angle}/\operatorname{span} \ell)$ (see (2.22) for the notation). By construction, if S_t is the coordinate representation of the curve $\Lambda(t)$ w.r.t. the chosen symplectic basis, then $C(S_t)$ is a coordinate representation of the curve $\overline{\Lambda(\cdot)^{l}}$ w.r.t. the basis of $(\operatorname{span} \ell)^{\angle}/\operatorname{span} \ell$, indicated in (2.39). Hence from (2.43), (2.44), and assumption $\det A(\tau) \neq 0$ it follows that the germ at τ of the curve $\overline{\Lambda(\cdot)^\ell}$ is regular. In particular, the curvature form of the ℓ-reduction $\Lambda(\cdot)^\ell$ is well defined at τ. Using (2.18), we obtain that the quadratic form

$$r_{\Lambda^\ell}(\tau)\Big|_{\Lambda(\tau)\cap(\operatorname{span}\ell)^\angle} - r_\Lambda(\tau)\Big|_{\Lambda(\tau)\cap(\operatorname{span}\ell)^\angle},$$

has the following matrix in the basis $\big((e_1,0),\dots,(e_{n-s},0)\big)$:

$$-\frac{d}{d\tau}C(S_\tau)\mathbb{S}\big(C(S_\tau)\big)+C\big(\dot S_\tau\mathbb{S}(S_\tau)\big).$$

Using the blocked structure of the matrix $\dot S_\tau$, given by (2.43), one can obtain from (2.18) without difficulties that

$$-\frac{d}{d\tau}C(S_\tau)\mathbb{S}\big(C(S_\tau)\big)+C\big(\dot S_\tau\mathbb{S}(S_\tau)\big)=-\frac{3}{4}\Big\{\sum_{k,m=n-s+1}^{n}(\ddot S_\tau)_{ik}(\dot S_\tau^{-1})_{km}(\ddot S_\tau)_{mj}\Big\}_{i,j=1}^{n-s}$$

In order to prove (2.33), it is sufficient to prove the following

Lemma 1 *The restriction of the quadratic form*

$$v\mapsto\sum_{k,m=1}^{s}(A(\tau)^{-1})_{km}\sigma\big(\ddot a_k(\tau),v\big)\sigma\big(\ddot a_m(\tau),v\big) \tag{2.45}$$

on $\Lambda(\tau)\cap l^\angle$ has the matrix with ij-entry equal to

$$-\sum_{k,m=n-s+1}^{n}(\ddot S_\tau)_{ik}(\dot S_\tau^{-1})_{km}(\ddot S_\tau)_{mj}$$

in the basis $\big((e_1,0),\dots,(e_{n-s},0)\big)$.

Proof. First, using the symmetry of S_t and (2.44), one has the following identity:

$$\begin{aligned}\sum_{k,m=n-s+1}^{n}(\ddot S_\tau)_{ik}(\dot S_\tau^{-1})_{km}(\ddot S_\tau)_{mj} &= \sum_{k,m=n-s+1}^{n}(\ddot S_\tau\dot S_\tau^{-1})_{ik}(\dot S_\tau)_{km}(\ddot S_\tau\dot S_\tau^{-1})_{jm} = \\ &-\sum_{k,m=1}^{s}(\ddot S_\tau\dot S_\tau^{-1})_{i,n-s+k}\big(A(\tau)^{-1}\big)_{km}(\ddot S_\tau\dot S_\tau^{-1})_{j,n-s+m}\end{aligned} \tag{2.46}$$

On the other hand, from (2.40) we have:

$$\ddot a_i(t)=\big(0,\ddot S_t b_i(t)\big)+2\big(0,\dot S_t\dot b_i(t)\big)+\big(\ddot b_i(t),S_t\ddot b_i(t)\big).$$

Substitute $t=\tau$ in the last relation. Note that $\dot S_\tau$ has the same blocked structure, as $\dot S_\tau^{-1}$, which together with (2.42) implies that $\big(0,\dot S_t\dot b_i(t)\big)\in\operatorname{span}\ell$. From this and (2.41) it follows that

$$\ddot a_i(\tau)\equiv\big(0,\ddot S_\tau\dot S_\tau^{-1}e_{n-s+i}(t)\big)\mod\operatorname{span}\big(\Lambda(\tau),l_1,\dots,l_s\big). \tag{2.47}$$

Hence

$$\forall 1 \le j \le n-s, 1 \le i \le s: \quad \sigma\big(\ddot a_i(\tau), (e_j, 0)\big) = -(\ddot S_\tau \dot S_\tau^{-1})_{j,n-s+i}, \tag{2.48}$$

The last identity together with (2.46) implies the statement of the lemma and also formula (2.33). □

Finally, identity (2.34) follows directly from (2.12). The proof of the theorem is completed. □

Note that if the curve $\Lambda(\cdot)$ is monotone increasing or decreasing , then from (2.32) it follows that the condition $\det A(\tau) \neq 0$ is equivalent to the following condition

$$\Lambda(\tau) \cap \operatorname{span} \ell = 0. \tag{2.49}$$

As a direct consequence of identity (2.33) and the last fact one has the following

Corollary 1 *If the curve $\Lambda(\cdot)$ is monotone increasing and the tuple $\ell = (l_1, \dots, l_s)$ of s vectors in Σ satisfies (2.49) for some t, then on the space $\Lambda(t) \cap \big(\operatorname{span} \ell\big)^\angle$ the curvature form of the ℓ-reduction of the curve $\Lambda(\cdot)$ is not less than the curvature form of the curve $\Lambda(\cdot)$ itself. Moreover, on the space $\Lambda(t) \cap \big(\operatorname{span} \ell\big)^\angle$ the difference between the curvature form of the ℓ-reduction of the curve $\Lambda(\cdot)$ and the curvature form of the curve $\Lambda(\cdot)$ itself is non-negative definite quadratic form of rank not greater than s.*

Let us translate the results of Theorem 1 in terms of a regular dynamical Lagrangian distribution $(\vec{\mathcal H}, \mathcal D)$ on a symplectic space W. Suppose that the Hamiltonian $\mathcal H$ admits a s-tuple $\mathcal G = (g_1, \dots, g_s)$ of involutive first integrals. Similarly to (2.5) denote by $\mathcal B_{\mathcal D_\lambda} : T_\lambda W \mapsto \mathcal D_\lambda^*$ the linear mapping such that for given $Y \in T_\lambda W$ the following identity holds

$$\mathcal B_{\mathcal D_\lambda} Y(Z) = \sigma(Y, Z), \quad \forall Z \in \mathcal D_\lambda. \tag{2.50}$$

Let us look at $\dot J_\lambda(0)$ (the velocity at $t=0$ of the Jacobi curve attached at λ) as at a linear mapping from $\mathcal D_\lambda$ to $\mathcal D_\lambda^*$. Then using regularity by analogy with (2.29) one can define the following s vector fields $\mathcal X_i$ on W:

$$\mathcal X_i(\lambda) = \big(\dot J_\lambda(0)\big)^{-1} \circ \mathcal B_{\mathcal D_\lambda}\big(\vec g_i(\lambda)\big). \tag{2.51}$$

Using relation (2.3) one can obtain by analogy with (2.30) that $\mathcal X_i$ is a unique vector field, satisfying $\mathcal X_i(\lambda) \in \mathcal D_\lambda$ and

$$[\vec{\mathcal H}, \mathcal X_i](\lambda) \equiv \vec g_i(\lambda) \mod \mathcal D_\lambda \tag{2.52}$$

for all $\lambda \in W$. Finally let $\Upsilon(\lambda)$ be the $s \times s$ matrix with the following entries

$$\Upsilon(\lambda)_{km} = \sigma_\lambda\big(\vec g_k, \mathcal X_m\big), \quad 1 \le k, m \le s. \tag{2.53}$$

From Theorem 1 and relation (2.3) one gets immediately

Theorem 2 *Suppose that $(\vec{\mathcal{H}},\mathcal{D})$ is a regular Lagrangian dynamical distribution on a symplectic space W and the Hamiltonian $\mathcal{H}$ admits a tuple $\mathcal{G}=(g_1,\ldots,g_s)$ of s involutive first integrals such that* $\det\Upsilon(\lambda)\neq 0$. *Then the curvature form $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ of the dynamical distribution $(\vec{\mathcal{H}},\mathcal{D})$ and the curvature form $r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}$ of its $\mathcal{G}$-reduction $(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})$ satisfy the following identity for all $v\in\Big(\bigcap\limits_{i=1}^{s}\ker d_\lambda g_i\Big)\cap\mathcal{D}_\lambda$*

$$r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}(v)-r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}(v)=\frac{3}{4}\sum_{k,m=1}^{s}\big(\Upsilon(\lambda)^{-1}\big)_{km}\sigma_\lambda\big([\vec{\mathcal{H}},[\vec{\mathcal{H}},\mathcal{X}_k]],v\big)\sigma_\lambda\big([\vec{\mathcal{H}},[\vec{\mathcal{H}},\mathcal{X}_m]],v\big), \tag{2.54}$$

while the curvature operator $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$ of the dynamical distribution $(\vec{\mathcal{H}},\mathcal{D})$ and the curvature operator $R_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}$ of its $\mathcal{G}$-reduction $(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})$ satisfy on $\Big(\bigcap\limits_{i=1}^{s}\ker d_\lambda g_i\Big)\cap\mathcal{D}_\lambda$ the following identity:

$$R_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}-R_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}=$$
$$\frac{3}{4}\sum_{k,m=1}^{s}\big(\Upsilon(\lambda)^{-1}\big)_{km}\mathcal{B}_{\mathcal{D}_\lambda}[\vec{\mathcal{H}},[\vec{\mathcal{H}},\mathcal{X}_m]](\lambda)\otimes\Big(\big(\dot{J}_\lambda(0)\big)^{-1}\circ\mathcal{B}_{\mathcal{D}_\lambda}[\vec{\mathcal{H}},[\vec{\mathcal{H}},\mathcal{X}_k]](\lambda)\Big). \tag{2.55}$$

Also, by analogy with Corollary 1 we have

Corollary 2 *If the dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ is monotone increasing and the Hamiltonian $\mathcal{H}$ admits a tuple $\mathcal{G}=(g_1,\ldots,g_s)$ of s involutive first integrals such that*

$$D_\lambda\cap\operatorname{span}\big(\vec{g}_1(\lambda),\ldots,\vec{g}_s(\lambda)\big)=0, \tag{2.56}$$

then on the space $\Big(\bigcap\limits_{i=1}^{s}\ker d_\lambda g_i\Big)\cap\mathcal{D}_\lambda$ the curvature form of the $\mathcal{G}$-reduction of the dynamical Lagrangian distribution $(\vec{\mathcal{H}},\mathcal{D})$ is not less than the curvature form of $(\vec{\mathcal{H}},\mathcal{D})$ itself. Moreover, on the space $\Big(\bigcap\limits_{i=1}^{s}\ker d_\lambda g_i\Big)\cap\mathcal{D}_\lambda$ the difference

$$r_\lambda^{(\vec{\mathcal{H}},\mathcal{D}^{\mathcal{G}})}-r_\lambda^{(\vec{\mathcal{H}},\mathcal{D})}$$

is non-negative definite quadratic form of rank not greater than s.

Now let us give the coordinate representation of the vector fields $\mathcal{X}_i$, $1\leq i\leq s$ from Theorem 2 in the case, when $W=T^*M$ and $\mathcal{D}=\Pi(M)$. Let $q=(q^1,\ldots,q^n)$ be local coordinates in some open subset $\mathcal{N}$ of M and $p=(p_1,\ldots,p_n)$ be induced coordinates in the fiber of $T^*\mathcal{N}$ so that the canonical symplectic form is given by $\sigma=\sum\limits_{i=1}^{n}dp_i\wedge dq^i$.

It gives the identification of $T^*\mathcal{N} \cong \mathbb{R}^n \times \mathbb{R}^n = \{(p,q), p, q \in \mathbb{R}^n)$ (so, $\mathcal{N} = 0 \times \mathbb{R}^n$). Also the tangent space $T_\lambda(T^*\mathcal{N})$ to $T^*\mathcal{N}$ at any λ is identified with $\mathbb{R}^n \times \mathbb{R}^n$. Under this identification $\vec{\mathcal{H}} = \left(-\frac{\partial \mathcal{H}}{\partial q}, \frac{\partial \mathcal{H}}{\partial p}\right)$, where for given function h on $T^*\mathcal{N}$ we denote by $\frac{\partial h}{\partial q} = \left(\frac{\partial h}{\partial q^1}, \dots, \frac{\partial h}{\partial q^n}\right)^T$ and $\frac{\partial h}{\partial p} = \left(\frac{\partial h}{\partial p_1}, \dots, \frac{\partial h}{\partial p_n}\right)^T$. Denote by $\mathcal{H}_{pp}$ the Hessian matrix of the restriction of $\mathcal{H}$ to the fibers. Then from Remark 1 and relation (2.51) we have

$$\mathcal{X}_i = (\mathcal{H}_{pp}^{-1} \frac{\partial g_i}{\partial p}, 0) \tag{2.57}$$

Now suppose for simplicity that the dynamical Lagrangian distribution is associated with a natural mechanical system (Example 2) or, more generally, with a mechanical system on a Riemannian manifolds (Example 3), which admits one or several first integrals being in involution and linear w.r.t. the impulses. One way to compute the reduced curvatures is to pass to the reduced phase space, as was described in Remark 3, and apply the method of computation of the curvatures from [4] to the corresponding dynamical Lagrangian distribution in the reduced phase space (this way was actually implemented in Example 8). But in order to apply the method of [4] we need to find a new canonical coordinates in the reduced phase space, which is not just a trivial exercise. Moreover, very often the new Hamiltonian system on the reduced phase space has more complicated form than the original one. Both these facts make the computation in this way quite tricky. Theorem 2 gives another method to compute all reduced curvatures without passing to the reduced phase space: to do this one can combine (1.9) or (1.12) with (2.54) (or (2.55)) and (2.57). This method is more effective from the computational point of view, especially if the number of the involutive first integrals is essentially less than the number of the degrees of freedom in the problem. We illustrate the effectiveness of this method on the following example:

Example 9 *(Plane N-body problem with equal masses)* Let us consider the motion of N bodies of unit mass in $\mathbb{R}^2$ endowed with the standard Cartesian coordinates so that $r_i = (q_{2i-1}, q_{2i}) \in \mathbb{R}^2$ represents the radius vector of the i-th body with respect to some inertial frame. It is described by a natural mechanical system on $M = \mathbb{R}^{2N}$ with potential energy

$$U(r_1, \dots, r_N) = -\sum_{i<j}^{N} \frac{1}{r_{ij}}, \qquad r_{ij} = \|r_i - r_j\|. \tag{2.58}$$

Then $T^*M \cong \mathbb{R}^{2N} \times \mathbb{R}^{2N} = \{(p,q),\ p, q \in \mathbb{R}^{2N}\}$, $p_1 \dots, p_{2N}$ are the canonical impulses conjugated to $q_1, \dots, q_{2N}$ ($p_i \sim \dot{q}_i$). The systems has the following first integral

$$g = \sum_{i=1}^{N} (p_{2i} q_{2i-1} - p_{2i-1} q_{2i}) \tag{2.59}$$

which is nothing but the angular momentum (in the considered planar case the angular momentum is scalar). From Example 2 we know that the generalized curvature form of the dynamical Lagrangian distribution $(\vec{H}, \Pi(M))$ is just the Hessian of the potential energy

U and the generalized Ricci curvature (see (1.8) for the definition) is the Laplacian of U, which can be calculated without difficulties:

$$\rho_\lambda^{\left(\vec{H},\Pi(M)\right)} = \Delta U = -2\sum_{i<j}^{N}\frac{1}{r_{ij}^3}, \quad \lambda=(p,q). \tag{2.60}$$

Our goal is to compute the reduced generalized Ricci curvatures $\rho_\lambda^{\left(\vec{H},\Pi(M)^g\right)}$, using the formula (2.54). In our case $s=1$. Let $\mathcal{X}$ be as in (2.51) with g instead of g_i. Note that by definition the vector $\mathcal{X}(\lambda)$ is orthogonal to the subspace $\Pi(M)_\lambda\cap\ker d_\lambda g$ w.r.t. the inner product $Q_\lambda^{\left(\vec{H},\Pi(M)\right)}(\cdot,\cdot)$. Therefore

$$\begin{aligned}\rho_\lambda^{\left(\vec{H},\Pi(M)^g\right)} = \rho_\lambda^{\left(\vec{H},\Pi(M)\right)} - \frac{r_\lambda^{\left(\vec{H},\Pi(M)\right)}(\mathcal{X})}{Q_\lambda^{\left(\vec{H},\Pi(M)\right)}(\mathcal{X},\mathcal{X})}+\\ \mathrm{tr}\left[\left(R_\lambda^{\left(\vec{H},\Pi(M)^g\right)} - R_\lambda^{\left(\vec{H},\Pi(M)\right)}\right)\Big|_{\Pi(M)_\lambda\cap\ker d_\lambda g}\right]\end{aligned} \tag{2.61}$$

Further, the Hamiltonian vector fields corresponding to the functions H and g are given by $\vec{H}=(U_q,p)$ and $\vec{g}=(Jp,-Jq)$ with $U_q=\left(\frac{\partial U}{\partial q_1},\dots,\frac{\partial U}{\partial q_n}\right)^T$ and J being the unit symplectic $2N\times 2N$ matrix:

$$J=\begin{pmatrix} 0 & 1 & 0 & \dots & 0\\ -1 & 0 & 0 & \dots & 0\\ & & \ddots & & \\ 0 & \dots & 0 & 0 & 1\\ 0 & \dots & 0 & -1 & 0\end{pmatrix}.$$

Applying formula (2.57) we find that

$$\mathcal{X}=(Jq,0). \tag{2.62}$$

Denote $\overline{\mathcal{X}}=Jq$. Also, let $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ be the standard Euclidean inner product and norm. Using again (1.9), one can obtain by direct computation that

$$r_\lambda^{(\vec{H},\Pi(M))}(\mathcal{X}) = \langle U_{qq}\overline{\mathcal{X}},\overline{\mathcal{X}}\rangle = -U. \tag{2.63}$$

Further, using (2.54) (or (2.55)) and (2.62), one can obtain without difficulties that

$$\begin{aligned}&\mathrm{tr}\left[\left(R_\lambda^{\left(\vec{H},\Pi(M)^g\right)} - R_\lambda^{\left(\vec{H},\Pi(M)\right)}\right)\Big|_{\Pi(M)_\lambda\cap\ker d_\lambda g}\right]=\\ &\frac{3}{4\sigma(\mathcal{X},\vec{g})}\left(\sum_{i=1}^{2N}\sigma([\vec{H},[\vec{H},\mathcal{X}]],\partial_{p_i})^2 - \frac{\sigma([\vec{H},[\vec{H},\mathcal{X}]],\mathcal{X})^2}{Q_\lambda^{\left(\vec{H},\Pi(M)\right)}(\mathcal{X},\mathcal{X})}\right) = \frac{3}{4}\left(\frac{\|p\|^2}{\|q\|^2}-\frac{\langle p,q\rangle^2}{\|q\|^4}\right).\end{aligned} \tag{2.64}$$

Substituting (2.60), (2.63), and (2.64) into (2.61) we obtain finally that

$$\rho_\lambda^{\left(\vec{H},\Pi(M)^g\right)}=-2\sum_{i<j}^{N}\frac{1}{r_{ij}^3}-\frac{U}{I}+\frac{3}{I^2}(2TI-\frac{1}{4}\{H,I\}^2)\,, \tag{2.65}$$

where $I=\|q\|^2$, $T=\frac{1}{2}\|p\|^2$ are the central momentum of inertia and the kinetic energy of the system of N bodies. Note that the sum of the first two terms in (2.65) is the trace of the restriction of the curvature operator $R_\lambda^{\left(\vec{H},\Pi(M)\right)}$ on the space $\Pi(M)_\lambda\cap\ker d_\lambda g$. So, by Remark 2 and Corollary 2, the last term in (2.65) has to be nonnegative. Actually this term contains the right-hand side of the famous Sundman's inequality $2TI-\frac{1}{4}\{H,I\}^2\geq 0$ and it is nothing but the generalized area of the parallelogram formed by two $2N$-dimensional vectors p and q. $\Box$

3 Focal points and Reduction

In the present section we study the relation between the set of focal points to the given point w.r.t. the monotone increasing (or decreasing) dynamical Lagrangian distribution and the set of focal points w.r.t. its reduction. As before, first we prove the corresponding result for the curves in Lagrange Grassmannians and then reformulate it in terns of the dynamical Lagrangian distributions.

Let $\Lambda(\cdot)$ be a curve in the Lagrange Grassmannian $L(\Sigma)$, defined on the interval $[0,T]$. The time t_1 is called *focal to the time* 0 *w.r.t. the curve* $\Lambda(\cdot)$, if $\Lambda(t_1)\cap\Lambda(0)\neq 0$. The dimension of the space $\Lambda(t_1)\cap\Lambda(0)$ is called the *multiplicity* of the focal time t_1. Denote by $\#\mathrm{foc}_0\Lambda(\cdot)\big|_I$ the number of focal times to 0 on the subset I w.r.t. $\Lambda(\cdot)$, counted with their multiplicity. If the curve $\Lambda(\cdot)$ is monotone increasing, then $\#\mathrm{foc}_0\Lambda(\cdot)\big|_I$ is finite and one can write

$$\#\mathrm{foc}_0\Lambda(\cdot)\big|_I=\sum_{t\in I}\dim\bigl(\Lambda(t)\cap\Lambda(0)\bigr). \tag{3.1}$$

Fix some tuple $\ell=(l_1\ldots,l_s)$ of s linearly independent vectors in Σ, satisfying (2.21). The time t_1 is called *focal to the time* 0 *w.r.t. the* ℓ*-reduction* $\Lambda(\cdot)^\ell$ *of the curve* $\Lambda(\cdot)$, if $\Lambda^\ell(t_1)\cap\Lambda^\ell(0)\neq\mathrm{span}\,\ell$ or, equivalently, t_1 is the focal time to 0 w.r.t. the curve $\overline{\Lambda(\cdot)^\ell}$ in the Lagrange Grassmannian $L((\mathrm{span}\,\ell)^\angle/\mathrm{span}\,\ell)$. The multiplicity of the focal time t_1 to 0 w.r.t. the ℓ-reduction $\Lambda(\cdot)^\ell$ is equal by definition to

$$\dim\left(\overline{\Lambda(t_1)^\ell}\cap\overline{\Lambda(0)^\ell}\right)=\dim\left(\Lambda(t_1)^\ell\cap\Lambda(0)^\ell\right)-s. \tag{3.2}$$

So, the number of focal times to 0 on the subset I w.r.t. the ℓ-reduction $\Lambda(\cdot)^\ell$, counted with their multiplicity, is equal to $\#\mathrm{foc}_0\overline{\Lambda(\cdot)^\ell}\big|_I$.

It is not hard to see that if $\Lambda(\cdot)$ is monotone increasing, then $\overline{\Lambda(\cdot)^\ell}$ is monotone increasing too. So, the number of focal times to 0 w.r.t. the ℓ-reduction is also finite. The question is what is the relation between the set of points, which are focal to 0 w.r.t. the curves $\Lambda(\cdot)$ and its ℓ-reduction $\Lambda(\cdot)^\ell$? To answer this question, we use the fact that if $\Lambda(\cdot)$ is monotone curve, then the number $\#\mathrm{foc}\Lambda(\cdot)\big|_{(0,T]}$ can be represented as the intersection index of this curve with a certain cooriented hypersurface in $L(\Sigma)$. The advantage of this representation is that the intersection index is a homotopic invariant.

More precisely, for given Lagrangian subspace Λ_0 denote by $\mathcal{M}_{\Lambda_0}$ the following subset of $L(\Sigma)$:

$$\mathcal{M}_{\Lambda_0} = L(\Sigma)\backslash\Lambda_0^{\pitchfork} = \{\Lambda \in L(\Sigma) :\ \Lambda \cap \Lambda_0 \neq 0\}\,.$$

Following [6], the set $\mathcal{M}_{\Lambda_0}$ is called the *train* of the Lagrangian subspace Λ_0. The set $\mathcal{M}_{\Lambda_0}$ is a hypersurface in $L(\Sigma)$ with singularities, consisting of the Lagrangian subspaces Λ such that $\dim\,(\Lambda\cap\Lambda_0) \geq 2$. The set of singular points has codimension 3 in $L(\Sigma)$. As we have already seen, the tangent space $T_\Lambda L(\Sigma)$ has a natural identification with the space of quadratic forms on Λ. If Λ is a non-singular point of the train $\mathcal{M}_{\Lambda_0}$, then vectors from $T_\Lambda L(\Sigma)$ that correspond to positive or negative definite quadratic forms are not tangent to the train. It defines the canonical coorientation of the hyper-surface $\mathcal{M}_{\Lambda_0}$ at a non-singular point Λ by taking as a positive side the side of $\mathcal{M}_{\Lambda_0}$ containing positive definite forms. The defined coorientation permits to define correctly the intersection index $\Lambda(\cdot)\cdot\mathcal{M}_{\Lambda_0}$ of an arbitrary continuous curve in the Lagrangian Grassmannian $\Lambda(\cdot)$, having endpoints outside $\mathcal{M}_{\Lambda_0}$: If $\Lambda(\cdot)$ is smooth and transversally intersecting $\mathcal{M}_{\Lambda_0}$ in non-singular points, then, as usual, every intersection point $\Lambda(\bar t)$ with $\mathcal{M}_{\Lambda_0}$ adds $+1$ or -1 into the value of the intersection index according to the direction of the vector $\dot\Lambda(\bar t)$ respectively to the positive or negative side of $\mathcal{M}_{\Lambda_0}$. Further, an arbitrary continuous curve $\Lambda(\cdot)$ with endpoints outside $\mathcal{M}_{\Lambda_0}$ can be (homotopically) perturbed to a curve which is smooth and transversally intersects $\mathcal{M}_{\Lambda_0}$ in non-singular points. Since the set of singular points of $\mathcal{M}_{\Lambda_0}$ has codimension 3 in $L(\Sigma)$, any two curves, obtained by such perturbation, can be deformed one to another by homotopy, which avoids the singularities of $\mathcal{M}_{\Lambda_0}$. Hence the intersection index of the curve, obtained by the perturbation, does not depend on the perturbation and can be taken as the intersection index of the original curve. The intersection index $\Lambda(\cdot)\cdot\mathcal{M}_{\Lambda_0}$ can be calculated using the notion of the Maslov index of the triple of Lagrangian subspaces (see [3], [6]) for details). This implies in particular that if the curve $\Lambda:[0,T]\mapsto L(\Sigma)$ is monotone increasing with endpoints outside $\mathcal{M}_{\Lambda_0}$, then

$$\Lambda(\cdot)\cdot\mathcal{M}_{\Lambda_0} = \sum_{0\leq t\leq T} \dim\big(\Lambda(t)\cap\Lambda_0\big) \tag{3.3}$$

Now we are ready to formulate the main result of this section:

Theorem 3 *Let $\Lambda:[0,T]\mapsto L(\Sigma)$ be a monotone increasing curve and $\ell=(l_1\ldots,l_s)$ be a tuple of s linearly independent vectors in Σ, satisfying (2.21) and (2.49) at 0. Then on the set $(0,T]$ the difference between the number of focal times to 0 w.r.t. the ℓ-reduction $\Lambda(\cdot)^\ell$ and the number of focal times to 0 w.r.t. the curve $\Lambda(\cdot)$ itself, counted with their multiplicity, is nonnegative and does not exceed s, namely*

$$0 \leq \#\mathrm{foc}_0\overline{\Lambda(\cdot)^\ell}\big|_{(0,T]} - \#\mathrm{foc}_0\Lambda(\cdot)\big|_{(0,T]} \leq s\,. \tag{3.4}$$

Proof. Since the curves $\Lambda(\cdot)$ and $\overline{\Lambda(\cdot)^\ell}$ are monotone increasing, for sufficiently small $\varepsilon > 0$ the set $(0,\varepsilon]$ does not contain the times focal to 0 w.r.t. both of these curves. Also, without loss of generality, one can assume that T is not focal to 0 w.r.t. both of these curves (otherwise, one can extend $\Lambda(\cdot)$ as a monotone increasing curve to a slightly bigger interval $[0, T+\tilde{\varepsilon}]$ such that $\Lambda(T+\tilde{\varepsilon}) \cap \Lambda(0) = 0$ and $\overline{\Lambda(T+\tilde{\varepsilon})^\ell} \cap \overline{\Lambda(0)^\ell} = 0$). So, by relations (3.1), (3.3), we have first that

$$\#\mathrm{foc}_0\Lambda(\cdot)\big|_{(0,T]} = \Lambda(\cdot)\big|_{[\varepsilon,T]} \cdot \mathcal{M}_{\Lambda(0)}. \tag{3.5}$$

Besides, from (2.49) it follows that

$$\dim\big(\Lambda(t)^\ell \cap \Lambda(0)^\ell\big) - s = \dim\big(\Lambda^\ell(t) \cap \Lambda(0)\big). \tag{3.6}$$

Hence, combining (3.1) and (3.3) with (3.2) and (3.6), we get

$$\#\mathrm{foc}_0\overline{\Lambda(\cdot)^\ell}\big|_{(0,T]} = \Lambda(\cdot)^\ell\big|_{[\varepsilon,T]} \cdot \mathcal{M}_{\Lambda(0)} \tag{3.7}$$

Now we prove the theorem in the case $s = 1$. In this case $\ell = l_1$. We use the invariance of the defined intersection index under homotopies, preserving the endpoints.

Let $a_1(t)$ be as in (2.29). Denote

$$F(\tau, t) = \mathrm{span}\big(\Lambda(t) \cap l_1^\angle, (1-\tau)a_1(t) + \tau l_1\big) \tag{3.8}$$

Note that all subspaces $F(\tau,t)$ are Lagrangian. Let $\Phi_\tau : [0,T] \mapsto L(\Sigma)$ and $\Gamma_t : [0,1] \mapsto L(\Sigma)$ be the curves, satisfying

$$\begin{aligned} &\Phi_\tau(\cdot) = F(\tau,\cdot), \quad 0 \le \tau \le 1; \\ &\Gamma_t(\cdot) = F(\cdot, t), \quad 0 \le t \le T. \end{aligned} \tag{3.9}$$

Then $\Phi_0(\cdot) = \Lambda(\cdot)$, $\Phi_1(\cdot) = \Lambda^{l_1}(\cdot)$, and the curves

$$\Gamma_\varepsilon(\cdot)\big|_{[0,\tau]} \cup \Phi_\tau(\cdot)\big|_{[\varepsilon,T]} \cup \Big(-\Gamma_T(\cdot)\big|_{[0,\tau]}\Big)$$

define the homotopy between $\Lambda(\cdot)\big|_{[\varepsilon,T]}$ and $\Gamma_\varepsilon(\cdot) \cup \Lambda(\cdot)^{l_1}\big|_{[\varepsilon,T]} \cup \big(-\Gamma_T(\cdot)\big)$, preserving the endpoints (here $-\gamma(\cdot)$ means the curve, obtained from a curve $\gamma(\cdot)$ by inverting the orientation). Therefore,

$$\Lambda(\cdot)\big|_{[\varepsilon,T]} \cdot \mathcal{M}_{\Lambda(0)} = \Gamma_\varepsilon(\cdot) \cdot \mathcal{M}_{\Lambda(0)} + \Lambda(\cdot)^{l_1}\big|_{[\varepsilon,T]} \cdot \mathcal{M}_{\Lambda(0)} - \Gamma_T(\cdot) \cdot \mathcal{M}_{\Lambda(0)}$$

Using (3.5) and (3.7), the last relation can be rewritten in the following form

$$\#\mathrm{foc}_0\overline{\Lambda(\cdot)^{l_1}}\big|_{(0,T]} - \#\mathrm{foc}_0\Lambda(\cdot)\big|_{(0,T]} = \Gamma_T(\cdot) \cdot \mathcal{M}_{\Lambda(0)} - \Gamma_\varepsilon(\cdot) \cdot \mathcal{M}_{\Lambda(0)}. \tag{3.10}$$

So, in order to prove the theorem in the considered case it is sufficient to prove the following two relations:

$$0 \le \Gamma_T(\cdot) \cdot \mathcal{M}_{\Lambda(0)} \le 1, \tag{3.11}$$

$$\exists \varepsilon_0 > 0 \text{ s.t. } \forall \varepsilon_0 \ge \varepsilon > 0: \quad \Gamma_\varepsilon(\cdot) \cdot \mathcal{M}_{\Lambda(0)} = 0. \tag{3.12}$$

a) Let us prove (3.11). If $l_1 \in \Lambda(T)$, then by definition $\Gamma_T(\tau) \equiv \Lambda(T)$. Since, by our assumptions, $\Lambda(T) \cap \Lambda(0) = 0$, we obviously have $\Gamma_T(\cdot) \cdot \mathcal{M}_{\Lambda(0)} = 0$.

If $l_1 \notin \Lambda(T)$, then $\dim\,(\Lambda(0) + \Lambda(T) \cap (l_1)^\angle) = 2n - 1$. In particular, it implies that

$$0 \leq \dim\,\big(\Gamma_T(\tau) \cap \Lambda(0)\big) \leq 1. \tag{3.13}$$

Further, let $p : \Sigma \mapsto \Sigma/(\Lambda(0) + \Lambda(T) \cap (l_1)^\angle)$ be the canonical projection on the factor space. Then from (3.8) and (3.9), using standard arguments of Linear Algebra, it follows that $\Gamma_T(\tau) \cap \Lambda(0) \neq 0$ if and only if

$$(1-\tau)p\big(a_1(T)\big) + \tau p(l_1) = 0 \tag{3.14}$$

Since, by assumptions, $\Gamma_T(0) \cap \Lambda(0) = 0$ (recall that $\Gamma_T(0) = \Lambda(T)$), the equation (3.14) has at most one solution on the segment $[0,1]$. In other words, the curve $\Gamma_T(\cdot)$ intersects the train $\mathcal{M}_{\Lambda(0)}$ at most ones and according to (3.13) the point of intersection is non-singular.

Finally, the curve $\Gamma_T(\cdot)$ is monotone non-decreasing, i.e. its velocities $\frac{d}{d\tau}\Gamma_T(\tau)$ are non-negative definite quadratic forms for any τ. Indeed, since $\Lambda(T) \cap (l_1)^\angle$ is the common space for all $\Gamma_T(\tau)$, one has $\frac{d}{d\tau}\Gamma_T(\tau)\big|_{\Lambda(T)\cap(l_1)^\angle} \equiv 0$. On the other hand, if we denote by $c(\tau) = (1-\tau)a_1(t) + \tau l_1$, then by (2.2), one has

$$\frac{d}{d\tau}\Gamma_T(\tau)\big(c(\tau)\big) = \sigma\big(c'(\tau), c(\tau)\big) = \sigma\big(l_1 - a_1(T), (1-\tau)a_1(T) + \tau l_1\big) = \sigma(l_1, a_1) > 0$$

(the last inequality follows from (2.31), (2.32) and the assumption about monotonicity of $\Lambda(\cdot)$).

So, $\frac{d}{d\tau}\Gamma_T(\tau)$ are non-negative definite quadratic forms. Hence in the uniquely possible point of intersection of $\Gamma_T(\cdot)$ with the train $\mathcal{M}_{\Lambda(0)}$ the intersection index becomes equal to 1. This proves (3.11).

b) Let us prove (3.12). Take a Lagrangian subspace Δ such that $l_1 \in \Delta$ and $\Delta \cap \Lambda(0) = 0$. Then there exists ε_0 such that

$$\Lambda(\cdot)\big|_{[0,\varepsilon_0]} \subset \Delta^\pitchfork. \tag{3.15}$$

Similarly to the arguments in **a)**, for any $0 < \varepsilon \leq \varepsilon_0$ the curve $\Gamma_\varepsilon(\cdot)$ intersects the train $\mathcal{M}_\Delta$ once. But by construction this unique intersection occurs at $\tau = 1$. Indeed, $\Gamma_\varepsilon(1) = \Lambda(\varepsilon)^{l_1}$, hence $l_1 \in \Gamma_\varepsilon(1) \cap \Delta$. In other words,

$$\Gamma_\varepsilon(\cdot)\big|_{[0,1)} \subset \Delta^\pitchfork. \tag{3.16}$$

Further, one can choose a symplectic basis in Σ such that $\Sigma = \mathbb{R}^n \times \mathbb{R}^n$, the symplectic form σ is as in (2.15), $\Lambda(0) = 0 \times \mathbb{R}^n$, and $\Delta = \mathbb{R}^n \times 0$. By (3.15) and (3.16), there exists two one parametric families of symmetric matrices S_t, $0 \leq t \leq \varepsilon_0$ and C_τ ,$0 \leq \tau < 1$ such that $\Lambda(t) = \{(S_t p, p) : p \in \mathbb{R}^n\}$ and $\Gamma_\varepsilon(\tau) = \{(C_\tau p, p) : p \in \mathbb{R}^n\}$. Since the curve $\Lambda(\cdot)$ is monotone increasing and the curve $\Gamma_\varepsilon(\cdot)$ is monotone nondecreasing, for any $0 \leq \tau < 1$ the quadratic forms $p \mapsto \langle C_\tau p, p\rangle$ are positive definite, while $S_0 = 0$. It implies that

$$\forall 0 \leq \tau < 1: \quad \Gamma_\epsilon(\tau) \cap \Lambda(0) = 0. \tag{3.17}$$

Note also that for sufficiently small $\varepsilon > 0$

$$\Gamma_\varepsilon(1) \cap \Lambda(0) = 0. \tag{3.18}$$

Indeed, $\Gamma_\varepsilon(1) = \Lambda^{l_1}(\varepsilon)$ and a sufficiently small $\varepsilon > 0$ is not a focal time for the l_1-reduction $\Lambda^{l_1}(\cdot)$, which according to (3.6) is equivalent to the fact that $\Lambda^{l_1}(\varepsilon) \cap \Lambda(0) = 0$ and hence to (3.18). By (3.17) and (3.18), for sufficiently small $\varepsilon > 0$ the curve $\Gamma_\epsilon(\cdot)$ does not intersect the train $\mathcal{M}_{\Lambda(0)}$. The relation (3.12) is proved, which completes the proof of our theorem in the case $s = 1$.

The case of arbitrary s can be obtained immediately from the case $s = 1$ by induction, using the fact that

$$\Lambda(\cdot)^{(l_1,\dots,l_s)} = \left(\Lambda(\cdot)^{(l_1,\dots,l_{s-1})}\right)^{l_s}. \quad \Box \tag{3.19}$$

Remark 8 Note that in the case $s = 1$ from Theorem 3 it follows immediately that the sets of focal times (to 0) w.r.t monotone increasing curve and its reduction are alternating. Also, for any s the first focal time to 0 w.r.t. the reduction does exceed the first focal time w.r.t. the curve itself. $\Box$

All constructions above are directly related to the notion of focal points of a dynamical Lagrangian distributions and $(\vec{\mathcal{H}}, \mathcal{D})$ and its reduction by a tuple $\mathcal{G} = (g_1, \dots, g_s)$ of s involutive first integrals, defined in Introduction. Note that the point $\lambda_1 = e^{t_1\vec{\mathcal{H}}}\lambda_0$ is focal to λ_0 w.r.t. the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the integral curve $t \mapsto e^{t\vec{\mathcal{H}}}\lambda_0$ of $\vec{\mathcal{H}}$ if and only if the time t_1 is focal to 0 w.r.t the Jacobi curve $J_{\lambda_0}(\cdot)$ attached at the point λ_0, while $\lambda_1 = e^{t_1\vec{\mathcal{H}}}\lambda_0$ is focal to λ_0 w.r.t. the $\mathcal{G}$-reduction of the pair $(\vec{\mathcal{H}}, \mathcal{D})$ along the integral curve $t \mapsto e^{t\vec{\mathcal{H}}}\lambda_0$ of $\vec{\mathcal{H}}$ if and only if t_1 is focal to 0 w.r.t $\big(\vec{g}_1(\lambda_0), \dots, \vec{g}_s(\lambda_0)\big)$-reductions of the Jacobi curves $J_{\lambda_0}(\cdot)$ attached at λ_0. Translating Theorem 3 into the terms of dynamical Lagrangian distribution, we have immediately the following

Corollary 3 *If the dynamical Lagrangian distribution $(\vec{\mathcal{H}}, \mathcal{D})$ is monotone increasing and the Hamiltonian $\mathcal{H}$ admits a tuple $\mathcal{G} = (g_1, \dots, g_s)$ of s involutive first integrals satisfying (2.56), then along any segment of the integral curve of $\vec{\mathcal{H}}$ the difference between the number of the focal points to the starting point of the segment w.r.t. the $\mathcal{G}$-reduction of the pair $(\vec{\mathcal{H}}, \mathcal{D})$ and the number of the focal points to the starting point of the segment w.r.t. the pair $(\vec{\mathcal{H}}, \mathcal{D})$ itself, counted with their multiplicity* [4], *is nonnegative and does not exceed s.*

Example 10 *(Plane 3-body problem with equal masses: focal points of the 8-shaped orbit)* The following example illustrates the Theorem 3. In 2000, A. Chenciner and R.Montgomery proved the existence of a new periodic solution of the planar 3-body problem with equal masses - the 8-shaped orbit or just *the Eight* [9]. In the plane of the motion each body moves along the same 8-shaped orbit, symmetric w.r.t. the point of its self-intersection, coinciding with the center of mass of the bodies. The configuration space is $M = \mathbb{R}^6$. As an initial point in the phase space T^*M we take the point λ_0 such that

[4] Here we do not count the starting point of the segment as the focal point to itself.

its projection on the configuration space is a collinear configuration, i.e. one of the bodies lies in the middle of the segment, connecting the other two.

In [8] there were found the focal points to λ_0 along the Eight w.r.t. the g-reduction of the Lagrangian dynamical distribution $(\vec{H}, \Pi(M)^g)$, where H, g are as in Example 9. In particular it was shown numerically that the 8-shaped orbit contains three such focal points along its period T, and the first focal time $\tau_1 \approx 0.52T$.

Let $e^{t_i\vec{H}}\lambda_0$ be the ith focal point w.r.t. $(\vec{H}, \Pi(M))$ along the Eight, and let $e^{\tau_i\vec{H}}\lambda_0$ be the ith focal point w.r.t. its g-reduction $(\vec{H}, \Pi(M)^g)$ along the same curve. In the following table we present the result of the numerical computation of t_i and τ_i on the interval $(0, 3T]$ (all this focal points have the multiplicity 1):

i	1	2	3	4	5	6	7	8	9	10	11
$\tau_i/T \approx$	0.52	0.76	0.95	1.08	1.52	1.56	1.88	2.05	2.29	2.49	2.65
$t_i/T \approx$	0.76	0.95	1.08	1.42	1.54	1.88	2.05	2.28	2.45	2.65	

We observe that $\tau_i \le t_i \le \tau_{i+1}$, $1 \le i \le 10$, as was expected by Theorem 3. $\square$

References

[1] A.A. Agrachev, N.N. Chtcherbakova *Hamiltonian Systems of negative curvature are Hyperbolic*, to appear in Math. Doklady, preprint SISSA 39/2004/M

[2] A.A. Agrachev and R.V. Gamkrelidze *Feedback - invariant optimal control theory and differential geometry -I. Regular extremals.* Journal of Dynamical and Control Systems, vol.3, 3, pp.343-389 (1997)

[3] A.A. Agrachev and R.V. Gamkrelidze *Symplectic Methods for Optimization and Control in " Geometry of feedback and optimal control*, ed. B.Jakubczyk, V.Respondek, pp.19-77, Marcel Dekker, 1998

[4] A.A. Agrachev, R.V. Gamkrelidze *Vector Fields on n-foliated $2n$-dimensional Manifolds*, to appear in "Journal of Mathematical Sciences", preprint SISSA 25/2004/M

[5] A.A. Agrachev and I. Zelenko,*Geometry of Jacobi curves, I and II*, J. Dynamical and Control Systems, **8**, 2002, No. 1, 93-140 and No.2, 167-215

[6] V.I. Arnold *The Sturm theorems and Symplectic geometry*, Funkcional. Anal. i Prilozen, **19**, No. 4 (1985), 1-10; English translation in Functional Anal. Appl. 19 (1985),251-259.

[7] V.I. Arnold, V.V. Kozlov, A.I. Neishtadt *Mathematical aspects of Classical and Celestial Mechanics (IIed.), Springer, 1997*

[8] N.N.Chtcherbakova.*On the minimizing properties if the 8-shaped solution of the 3-body problem, SISSA preprint, SISSA 108/2003/AF*

[9] A. Chenciner and R. Montgomery *A remarkable periodic solution of the three body problem in the case of equal masses. Annals of Mathematics, 152, pp.881-901 (2000)*

[10] I. Zelenko *Variational Approach to Differential Invariants of Rank 2 Vector Distributions*, submitted to "Differential Geometry and its Applications", preprint SISSA 12/2004/M