arXiv:math.DG/0703662v1 22 Mar 2007

ON LOCAL GEOMETRY OF NONHOLONOMIC RANK 2 DISTRIBUTIONS

BORIS DOUBROV AND IGOR ZELENKO

Abstract. In 1910 E. Cartan constructed a canonical frame and found the most symmetric case for maximally nonholonomic rank 2 distributions in $\mathbb{R}^5$. We solve the analogous problem for germs of generic rank 2 distributions in $\mathbb{R}^n$ for $n > 5$. We use a completely different approach based on the symplectification of the problem. The main idea is to consider a special odd-dimensional submanifold W_D of the cotangent bundle associated with any rank 2 distribution D. It is naturally foliated by characteristic curves, which are also called the abnormal extremals of the distribution D. The dynamics of vertical fibers along characteristic curves defines certain curves of flags of isotropic and coisotropic subspaces in a linear symplectic space. Using the classical theory of curves in projective spaces, we construct the canonical frame of the distribution D on a certain $(2n-1)$-dimensional fiber bundle over W_D with the structure group of all Möbius transformations, preserving 0. The paper is the detailed exposition of the constructions and the results, announced in the short note [8].

1. Introduction

A rank l vector distribution D on an n-dimensional manifold M or an (l, n)-distribution (where $l < n$) is a subbundle of the tangent bundle TM with l-dimensional fibers. The group of germs of diffeomorphisms of M acts naturally on the set of germs of (l, n)-distributions and defines the equivalence relation there. The question is *when two germs of distributions are equivalent?* Distributions are naturally associated with Pfaffian systems and with control systems linear in the control. So the problem of equivalence of distributions can be reformulated as the problem of equivalence of the corresponding Pfaffian systems and the state-feedback equivalence of the corresponding control systems. The obvious (but very rough in the most cases) discrete invariant of a distribution D at q is so-called *the small growth vectors* at q. It is the tuple $\{\dim D^j(q)\}_{j\in\mathbb{N}}$, where D^j is the j-th power of the distribution D, i.e., $D^j = D^{j-1} + [D, D^{j-1}]$, $D^1 = D$. A simple estimation shows that at least $l(n-l)-n$ functions of n variables are required to describe generic germs of (l, n)-distribution, up to the equivalence (see [13] and [18] for precise statements). There are only three cases, where $l(n-l)-n$ is not positive: $l = 1$ (line distributions), $l = n-1$, and $(l, n) = (2, 4)$. Moreover, it is well known that in these cases generic germs of distributions are equivalent. For $l = 1$ it is just the classical theorem about the rectification of vector fields without stationary points, for $l = n-1$ all generic germs are equivalent to Darboux's model, while for $(l, n) = (2, 4)$ they are equivalent to Engel's model (see, for example, [6]). In all other cases generic (l, n)-distributions have functional invariants.

In the present paper we restrict ourselves to the case of rank 2 distributions, although our method can be applied also for distribution of rank greater than 2 (as will be described in the

2000 *Mathematics Subject Classification.* 58A30, 53A55.

Key words and phrases. Nonholonomic distributions, equivalence problem, canonical frames, abnormal extremals, Jacobi curves, curves in projective spaces.

forthcoming paper [9]). The model examples of rank 2 distributions come from so-called underdetermined ODE's of the type

$$z'(x) = F\big(x, y(x), \dots, y^{(n-3)}(x), z(x)\big), \tag{1.1}$$

for two functions $y(x)$ and $z(x)$. Setting $p_i = y^{(i)}$, $0 \le i \le n-3$, with each such equation one can associate the rank 2 distribution on $\mathbb{R}^n$ with coordinates $(x, p_0, \dots, p_r, z)$ defined as the annihilator of the following $n-2$ 1-forms:

$$\begin{aligned} &dp_i - p_{i+1}dx,\ 0 \le i \le n-4, \\ &dz - F(x, p_0, \dots, p_{n-3}, z)dx. \end{aligned} \tag{1.2}$$

For $n = 3$ and 4 all generic germs of rank 2 distribution are equivalent to the distribution, associated with the underdetermined ODE $z'(x) = y(x)$ (Darboux and Engel models respectively). The case $n = 5$ (the smallest dimension, when functional parameters appear) was treated by E. Cartan in [7] with his reduction-prolongation procedure. First, for any $(2,5)$-distribution with the small growth vector $(2,3,5)$ he constructed the canonical coframe in some 14-dimensional manifold, which implied that the group of symmetries of such distributions is at most 14-dimensional. Second, he showed that any $(2,5)$-distribution with 14-dimensional group of symmetries is locally equivalent to the distribution, associated with the underdetermined ODE $z'(x) = \big(y''(x)\big)^2$, and its group of symmetries is isomorphic to the real split form of the exceptional Lie group G_2. Historically it was the first natural appearance of this group.

After the work of Cartan the open question was **to construct the canonical frame and to find the most symmetric cases for** $(2,n)$**-distributions with** $n > 5$. The Cartan equivalence method was systematized and generalized by N. Tanaka and T. Morimoto (see [11, 10]). Their theory is heavily based on the notion of so-called *symbol algebra* of the distribution at a point, which is a special *graded nilpotent Lie algebra*, naturally associated with the distribution at a point. The symbol algebras have to be isomorphic at different points and all constructions strongly depend on the type of the symbol. Note that already in the case of $(2,6)$-distributions with maximal possible small growth vector $(2,3,5,6)$ three different symbol algebras are possible, while for $n = 9$ the set of all possible symbol algebras depends on continuous parameters, which implies in particular that generic distributions do not have a constant symbol.

In the present paper we give the answer to the question, underlined in the previous paragraph, for rank 2 distributions from some generic class. Our constructions are based on a completely different, variational approach, developed in [4] and [15]. The paper is the detailed exposition of the constructions and the results, announced in short note [8]. The starting point of this approach is to lift a distribution D to a special odd-dimensional submanifold W_D of the cotangent bundle, foliated by the characteristic curves, which are also called the abnormal extremals of the distribution D. They are Pontryagin extremals with zero Lagrange multiplier near the functional for any extremal problem with constrains, given by the distribution D. The dynamics of the lifting (to W_D) of the distribution D w.r.t. to this 1-foliation along any abnormal extremal defines certain curve of flags of isotropic and coisotropic subspaces in a linear symplectic space. So, the problem of equivalence of distributions can be essentially reduced to the differential geometry of such curves: symplectic invariants of these curves automatically produce invariants of the distribution D itself and the canonical frame bundles associated with such curves can be in many cases effectively used to construct the canonical frames of the distribution D itself defined on a certain bundle over W_D.

In the case of a nonholonomic rank 2 distribution the submanifold W_D is nothing but the annihilator of the square of D, denoted by $(D^2)^\perp$. Under additional generic assumptions the curves of flags, associated with abnormal extremals from a certain open and dense subset of

$(D^2)^\perp$, are curves of complete flags. Moreover, these curves of complete flags can be recovered by differentiation from the curves of their one-dimensional subspaces, i.e. from curves in projective spaces. Recall that the differential geometry of curves in projective spaces was developed already in 1905 by E.J. Wilczynski ([14]). In particular, these curves (and therefore the corresponding abnormal extremals of the distribution) are endowed with the canonical projective structure, i.e., we have a distinguished set of parameterizations (called projective) such that the transition function from one such parameterization to another is a Möbius transformation. Besides, for each fixed projective parameterization on such curve one can construct the canonical moving frame in the ambient linear symplectic space.

These two facts together allow us to construct the canonical frame for any $(2, n)$-distribution D, $n > 5$, from a certain generic class. This frame lives on a certain principle bundle over $(D^2)^\perp$ with the structure group $ST(2, \mathbb{R})$ of all Möbius transformations, preserving 0. The fiber of this bundle over the point $\lambda \in (D^2)^\perp$ is just the set of all projective parameterizations of the abnormal extremal passing through λ such that the point λ corresponds to 0. In particular, it implies that the group of symmetries of such distributions is at most $(2n-1)$-dimensional.

2. Abnormal extremals

Assume that $\dim D^2(q) = 3$ and $\dim D^3(q) > 3$ for any $q \in M$. Denote by $(D^l)^\perp \subset T^*M$ the annihilator of the lth power D^l, namely

$$(D^l)^\perp = \{(p, q) \in T^*M : \ p \cdot v = 0 \ \forall v \in D^l(q)\}.$$

First we distinguish a characteristic 1-foliation on the codimension 3 submanifold $(D^2)^\perp \backslash (D^3)^\perp$ of T^*M. For this let $\pi : T^*M \mapsto M$ be the canonical projection. For any $\lambda \in T^*M$, $\lambda = (p, q)$, $q \in M$, $p \in T_q^*M$, let $\mathfrak{s}(\lambda)(\cdot) = p(\pi_*\cdot)$ be the canonical Liouville form and $\sigma = d\mathfrak{s}$ be the standard symplectic structure on T^*M. Since the submanifold $(D^2)^\perp$ has odd codimension in T^*M, the kernels of the restriction $\sigma|_{(D^2)^\perp}$ of σ on $(D^2)^\perp$ are not trivial. Moreover, as we show below, for the points of $(D^2)^\perp \backslash (D^3)^\perp$ these kernels are one-dimensional. They form the *characteristic line distribution* in $(D^2)^\perp \backslash (D^3)^\perp$, which will be denoted by $\mathcal{C}$. The line distribution $\mathcal{C}$ defines a *characteristic 1-foliation* of $(D^2)^\perp \backslash (D^3)^\perp$. Actually the leaves of this foliation are so-called *regular abnormal extremals of the distribution* D.

Recall that abnormal extremals of D are by definition Pontryagin extremals with zero Lagrange multiplier near the functional for any extremal problem with constrains, given by the distribution D (and so they depend only on D and not on a functional). Regularity means that they do not pass through $(D^3)^\perp$, which is equivalent to the fact that they satisfy so-called strong generalized Legendre–Glebsch condition ([2], [17]). In the sequel for shortness we will omit the word regular and the leaves of the characteristic foliation will be called just *abnormal extremals of* D.

Let us describe the characteristic line distribution $\mathcal{C}$ in terms of a local basis (X_1, X_2) of the distribution D, $D(q) = \text{span}\{X_1(q), X_2(q)\}$. Denote by

$$X_3 = [X_1, X_2], \ X_4 = \big[X_1, [X_1, X_2]\big], \ X_5 = \big[X_2, [X_1, X_2]\big]. \tag{2.1}$$

Let us introduce the "quasi-impulses" $u_i : T^*M \mapsto \mathbb{R}$, $1 \le i \le 5$,

$$u_i(\lambda) = p \cdot X_i(q), \ \lambda = (p, q), \ q \in M, \ p \in T_q^*M \tag{2.2}$$

Then by definitions

$$(D^2)^\perp = \{\lambda \in T^*M : u_1(\lambda) = u_2(\lambda) = u_3(\lambda) = 0\}. \tag{2.3}$$

As usual, for given function $G : T^*M \mapsto \mathbb{R}$ denote by $\vec{G}$ the corresponding Hamiltonian vector field defined by the relation $i_{\vec{G}}\sigma = -d\,G$.

Lemma 2.1. *The characteristic line distribution $\mathcal{C}$ satisfies*

$$\mathcal{C} = \langle u_4\vec{u}_2 - u_5\vec{u}_1\rangle. \tag{2.4}$$

Proof. Take a vector field H on $(D^2)^\perp\backslash(D^3)^\perp$ such that locally $\mathcal{C}(\lambda) = \{\mathbb{R}H(\lambda)\}$. Then by definition of $\mathcal{C}$ we have $i_H\sigma|_{(D^2)^\perp} = 0$. From this and (2.3) it follows that $i_H\sigma \in \langle du_1, du_2, du_3\rangle$, which implies that

$$H \in \langle \vec{u}_1, \vec{u}_2, \vec{u}_3\rangle. \tag{2.5}$$

On the other hand, H is tangent to $(D^2)^\perp$, i.e $du_j(H) = 0$ for $1 \le j \le 3$. This and (2.5) easily implies (2.4). □

As a consequence, we see that $\pi_*(\mathcal{C}(\lambda)) \subset D(q)$ for any $\lambda \in (D^2)^\perp\backslash(D^3)^\perp$, $\pi(\lambda) = q$. Moreover, if $\dim D^3(q) = 5$, then one-dimensional subspaces $\pi_*(\mathcal{C}(\lambda))$ generate $D(q)$:

$$\operatorname{span}\Big\{\pi_*(\mathcal{C}(\lambda)) \,:\, \lambda \in (D^2)^\perp\backslash(D^3)^\perp, \pi(\lambda) = q\Big\} = D(q). \tag{2.6}$$

In particular, in this case the original distribution can be recovered from its characteristic line distribution.

3. The curves of flags associated with abnormal extremals

Now, following [15], let $\mathcal{J}$ be the pull-back of the distribution D on $(D^2)^\perp\backslash(D^3)^\perp$ by the canonical projection π:

$$\mathcal{J}(\lambda) = \{v \in T_\lambda(D^2)^\perp \,:\, \pi_*\, v \in D(\pi(\lambda))\}. \tag{3.1}$$

Note that $\dim \mathcal{J} = n-1$ and $\mathcal{C} \subset \mathcal{J}$ by (2.4) . The distribution $\mathcal{J}$ is called the *lift of distribution D to $(D^2)^\perp\backslash(D^3)^\perp$*.

In the sequel we shall work with the lift $\mathcal{J}$ instead of the original distribution D. The crucial advantage of working with $\mathcal{J}$ is that it has the distinguished line sub-distribution $\mathcal{C}$, while the original distribution D has no distinguished sub-distributions in general.

We can produce two monotonic (by inclusion) sequences of distributions (in general of nonconstant ranks) first by making iterative Lie brackets of $\mathcal{C}$ and $\mathcal{J}$ and then by taking skew symmetric complements w.r.t. the form σ of the subspaces obtained in the previous step. Namely, first define a sequence of subspaces $\mathcal{J}^{(i)}(\lambda)$, $\lambda \in (D^2)^\perp\backslash(D^3)^\perp$, by the following recursive formulas:

$$\mathcal{J}^{(i)} = \mathcal{J}^{(i-1)} + [\mathcal{C}, \mathcal{J}^{(i-1)}], \quad \mathcal{J}^{(0)} = \mathcal{J}, \tag{3.2}$$

and then set

$$\mathcal{J}_{(i)}(\lambda) = \{v \in T_\lambda\big((D^2)^\perp\big) : \sigma(v, w) = 0\ \forall w \in \mathcal{J}^{(i)}(\lambda)\}, \tag{3.3}$$

We summarize the main properties of the sequences $\{\mathcal{J}^{(i)}\}_{i\ge 0}$ and $\{\mathcal{J}_{(i)}\}_{i\ge 0}$ in the following:

Proposition 3.1.
(1) $\sigma|_{\mathcal{J}} = 0$, $\mathcal{J}_{(0)} = \mathcal{J}^{(0)}$;
(2) $\mathcal{J}^{(i-1)}(\lambda) \subseteq \mathcal{J}^{(i)}(\lambda)$, $\mathcal{J}_{(i)}(\lambda) \subseteq \mathcal{J}_{(i-1)}(\lambda)$;
(3) $\dim \mathcal{J}^{(1)}(\lambda) - \dim \mathcal{J}(\lambda) = 1$, $\dim \mathcal{J}(\lambda) - \dim \mathcal{J}_{(1)}(\lambda) = 1$;
(4) $\dim \mathcal{J}^{(i)}(\lambda) - \dim \mathcal{J}^{(i-1)}(\lambda) \le 1$, $\dim \mathcal{J}_{(i-1)}(\lambda) - \dim \mathcal{J}_{(i)}(\lambda) \le 1$ *for* $i \ge 2$;
(5) $\dim \mathcal{J}^{(i)}(\lambda) \le 2n-4$.

Proof. First note that the second relations in the Properties (1)-(4) are direct consequences of the corresponding first relations.

Now let us prove the first relation in the Property (1). By the arguments similar to the proof of Lemma 2.4, the set of points, where the form $\sigma|_{D^\perp}$ is degenerated, coincides with $(D^2)^\perp$ and for each $\lambda \in (D^2)^\perp$ the kernel of $\sigma|_{D^\perp}(\lambda)$ satisfies

$$\ker \sigma|_{D^\perp}(\lambda) = \operatorname{span}\{\vec{u}_1(\lambda), \vec{u}_2(\lambda)\}, \tag{3.4}$$

where u_i are as in (2.2) for some local basis (X_1, X_2) of the distribution D. Also denote by $\widetilde{V}(\lambda)$ the vertical subspace of $T_\lambda D^\perp$, i.e. $\widetilde{V}(\lambda) = \{v \in T_\lambda D^\perp, \pi_* v = 0\}$. Then from (3.1) and (3.4) one gets easily that

$$\mathcal{J}(\lambda) = \big(\widetilde{V}(\lambda) + \ker \sigma|_{D^\perp}(\lambda)\big) \cap T_\lambda (D^2)^\perp.$$

This immediately implies the first relation of Property (1).

The first inclusion in Property (2) follows directly from definition (3.2) of $\mathcal{J}^{(i)}$. Further, one can easily get that

$$\mathcal{J}^{(1)}(\lambda) = \{v \in T_\lambda (D^2)^\perp : \pi_* v \in D^2(\pi(\lambda))\}. \tag{3.5}$$

Since $\dim D^2 - \dim D = 1$, we obtain the first relation in Property (3). Besides, since on each step of the recursive formula (3.2) one makes Lie brackets with the rank 1 distribution $\mathcal{C}$, Property (3) immediately implies Property (4). In order to prove Property (5) note that the line distribution $\mathcal{C}$ forms the Cauchy characteristic of the corank 1 distribution on $(D^2)^\perp$, given by the Pfaffian equation $\mathfrak{s}|_{(D^2)^\perp} = 0$, where as before $\mathfrak{s}$ is the Liouville form. Since by construction

$$\mathcal{J} \subset \{\mathfrak{s}|_{(D^2)^\perp} = 0\},$$

one has

$$\mathcal{J}^{(i)} \subset \{\mathfrak{s}|_{(D^2)^\perp} = 0\} \quad \text{for all } i \in \mathbb{N}. \tag{3.6}$$

Property (5) follows from the fact that the distribution $\{\mathfrak{s}|_{(D^2)^\perp} = 0\}$ has rank $2n-4$. □

So, by Properties (1),(2), and (3) for any $\lambda \in (D^2)^\perp \backslash (D^3)^\perp$ we get the flag

$$\ldots \subseteq \mathcal{J}_{(i)}(\lambda) \subseteq \ldots \subseteq \mathcal{J}_{(1)}(\lambda) \subset \mathcal{J}(\lambda) \subset \mathcal{J}^{(1)}(\lambda) \subseteq \ldots \subseteq \mathcal{J}^{(i)}(\lambda) \subseteq \ldots \tag{3.7}$$

in $T_\lambda (D^2)^\perp$. The dynamics of these flags along any abnormal extremal defines certain curve of flags of isotropic and coisotropic subspaces in a linear symplectic space.

More precisely, let γ be a segment of abnormal extremal of D and O_γ be a neighborhood of γ in $(D^2)^\perp$ such that the factor $N = O_\gamma/(\textit{the characteristic one-foliation})$ is a well defined smooth manifold. The quotient manifold N is a symplectic manifold endowed with the symplectic structure $\bar{\sigma}$ induced by $\sigma|_{(D^2)^\perp}$. Let $\phi : O_\gamma \to N$ be the canonical projection on the factor. For each ≥ 0 we can define the following curves of subspaces in $T_\gamma N$:

$$\lambda \mapsto \phi_*\big(\mathcal{J}^{(i)}(\lambda)\big), \quad \lambda \mapsto \phi_*\big(\mathcal{J}_{(i)}(\lambda)\big), \quad \text{for all } \lambda \in \gamma. \tag{3.8}$$

These curves describe the dynamics of the corresponding subspaces of the flag (3.7) w.r.t. the characteristic 1-foliation along the abnormal extremal γ.

Note that there exists a straight line, which is common to all subspaces appearing in (3.8) for any $\lambda \in \gamma$. So, it is more convenient to get rid of it by a factorization. Indeed, let e be the Euler field on T^*M, i.e., the infinitesimal generator of homotheties on the fibers of T^*M. Since a transformation of T^*M, which is a homothety on each fiber with the same homothety coefficient, sends abnormal extremals to abnormal extremals, we see that the vector $\bar{e} = \phi_* e(\lambda)$ is the same for any $\lambda \in \gamma$ and lies in any subspace appearing in (3.8). Let

$$J^{(i)}(\lambda) = \phi_*\big(\mathcal{J}^{(i)}(\lambda)\big)/\{\mathbb{R}\bar{e}\}, \quad J_{(i)}(\lambda) = \phi_*\big(\mathcal{J}_{(i)}(\lambda)\big)/\{\mathbb{R}\bar{e}\}. \tag{3.9}$$

For simplicity we also set $J(\lambda) = J^{(0)}(\lambda)$ $(= J_{(0)}(\lambda))$. It is clear that all subspaces appearing in (3.9) belong to the space

$$W = \{v \in T_\gamma N : \bar{\sigma}(v, \bar{e}) = 0\}/\{\mathbb{R}\bar{e}\}. \tag{3.10}$$

The space W is endowed with the natural symplectic structure induced by $\bar{\sigma}$, which for simplicity will be denoted also by $\bar{\sigma}$. Also $\dim W = 2(n-3)$. Rewriting Properties (1)-(4) of Proposition 3.1 and relation (3.3) in terms of subspaces $J^{(i)}(\lambda)$ and $J_{(i)}(\lambda)$, we get

Proposition 3.2.

1. *The subspace $J(\lambda)$ is Lagrangian subspace of W and the subspace $J_{(i)}(\lambda)$ is the skew-symmetric complement of $J^{(i)}(\lambda)$ in W for any $\lambda \in \gamma$;*
2. $J^{(i-1)}(\lambda) \subseteq J^{(i)}(\lambda)$, $J_{(i)}(\lambda) \subseteq J_{(i-1)}(\lambda)$*;*
3. $\dim J^{(1)}(\lambda) - \dim J(\lambda) = 1$, $\dim J(\lambda) - \dim J_{(1)}(\lambda) = 1$*;*
4. $\dim J^{(i)}(\lambda) - \dim J^{(i-1)}(\lambda) \le 1$, $\dim J_{(i-1)}(\lambda) - \dim J_{(i)}(\lambda) \le 1$ *for* $i \ge 2$*;*

Note also that by Property (1) of Proposition 3.1 for all $i \in \mathbb{N}$ the subspaces $J^{(i)}(\lambda)$ are coisotropic and the subspaces $J_{(i)}(\lambda)$ are isotropic in $T_\gamma N$. The curve

$$\lambda \mapsto \Big\{\ldots \subseteq J_{(i)}(\lambda) \subseteq \ldots \subseteq J_{(1)}(\lambda) \subset J(\lambda) \subset J^{(1)}(\lambda) \subseteq \ldots \subseteq J^{(i)}(\lambda) \subseteq \ldots\Big\}, \quad \lambda \in \gamma, \tag{3.11}$$

of flags of isotropic and coisotropic subspaces in a linear symplectic space W will be called the *curve of flags associated with the segment γ of an abnormal extremal.*

Clearly, any symplectic invariant of such curve automatically produces an invariant of the distribution D itself. Moreover, it turns out that under certain generic assumptions one can construct the canonical frames of the distribution D from the study of differential geometry of such curves.

Remark 3.1. As a matter of fact the whole curve of flags (3.11) can be recovered from the curve $\lambda \mapsto J(\lambda)$ of Lagrangian subspaces of W. It is clear by item 1 of Proposition 3.2 that it is sufficient to show how to recover the subspaces $J^{(i)}(\lambda)$. This can be done by making an appropriate differentiation (in a similar manner as all subspaces $\mathcal{J}^{(i)}(\lambda)$ and $\mathcal{J}_{(i)}(\lambda)$ are obtained from $\mathcal{J}(\lambda)$). More precisely, let $\Gamma(J)$ be the set of all smooth mappings ℓ from γ to the ambient symplectic space W (see (3.10)) such that $\ell(\lambda) \in J(\lambda)$ for all $\lambda \in \gamma$. In other words, $\Gamma(J)$ is the space of all smooth sections of the vector bundle over γ having the subspace $J(\lambda)$ as the fiber over a point $\lambda \in \gamma$. If $\varphi : \gamma \mapsto \mathbb{R}$ is a parameterization of γ, $\varphi(\lambda) = 0$ and $\psi = \varphi^{-1}$, then

$$J^{(i)}(\lambda) = \operatorname{span}\Big\{\frac{d^j}{dt^j}\ell\big(\psi(t)\big)|_{t=0} : \ell \in \Gamma(J),\, 0 \le j \le i\Big\}. \tag{3.12}$$

The curve $\lambda \mapsto J(\lambda)$ is called *Jacobi curve of the abnormal extremal* γ. The reason to call this curve Jacobi curve is that it can be considered as the generalization of spaces of "Jacobi fields" along Riemannian geodesics: in terms of this curve one can describe some optimality properties (so-called rigidity) of the corresponding abnormal extremal ([1] or [17]).□

Remark 3.2. Actually, one can describe the subspaces $J_{(i)}(\lambda)$, where $i \ge 1$, without using the symplectic structure on W. For this, by analogy with above, let $\Gamma(J_{(i)})$, $i \ge 0$, be the set of all smooth mappings ℓ from γ to the ambient symplectic space W such that $\ell(\lambda) \in J_{(i)}(\lambda)$ for all $\lambda \in \gamma$. Let $\varphi : \gamma \mapsto \mathbb{R}$ be a parameterization of γ, $\varphi(\lambda) = 0$ and $\psi = \varphi^{-1}$. Then it is easy to show that for any $i \geq 1$

$$J_{(i)}(\lambda) = \left\{ v \in J_{(i-1)}(\lambda) : \begin{array}{l} \exists\, \ell \in \Gamma(J_{(i-1)}) \text{ with } \ell(\lambda) = v \\ \text{such that } \frac{d}{dt}\ell\big(\psi(t)\big)|_{t=0} \in J_{(i-1)}(\lambda) \end{array} \right\} \tag{3.13}$$

The last formula allows to construct $J_{(i)}$ recursively, starting from $J_{(0)} = J$. Finally it is easy to show that identity (3.13) remains true if one replaces the quantor $\exists$ by $\forall$. $\square$

4. The class of a rank 2 distribution

Let us describe precisely the generic assumptions on a germ of rank 2 distribution necessary for constructing the canonical frames for them.

First for any point $q \in M$ denote by $(D^l)^\perp(q) = (D^l)^\perp \cap T_q^*M$ the fiber of $(D^l)^\perp$. Let us define the following integer-valued function on $(D^2)^\perp \backslash (D^3)^\perp$:

$$\nu(\lambda) = \min\{i \in \mathbb{N} : \mathcal{J}^{(i+1)}(\lambda) = \mathcal{J}^{(i)}(\lambda)\},$$
$$m(q) = \max\{\nu(\lambda) : \lambda \in (D^2)^\perp(q) \backslash (D^3)^\perp(q)\}.$$

From Properties (3), (4), (5) of Proposition 3.1, and the fact that $\dim \mathcal{J} = n-1$ it follows that $1 \leq \nu(\lambda) \leq n-3$. Furthermore, from (2.4) and the definition (3.2) of $\mathcal{J}_{(i)}$ it follows that the set $\{\lambda \in (D^2)^\perp(q) : \nu(\lambda) = m(q)\}$ is nonempty open in the Zariski topology of the fiber $(D^2)^\perp(q)$. Besides, it is easy to show that the integer-valued functions $\nu(\cdot)$ and $m(\cdot)$ are lower semicontinuous. Hence they are locally constant on the open and dense subset of $(D^2)^\perp \backslash (D^3)^\perp$ and M correspondingly and attain their maximum values on the open sets there. Moreover,

Proposition 4.1. *Germs of $(2,n)$-distributions of the maximal class $n-3$ are generic.*

This Proposition was proved in [15, Proposition 3.4]. Let us outline the proof. The distribution has maximal class at a point q_0 if and only if its jet of sufficiently high order belongs to the Zariski open set of the jet space of this order. Therefore in order to prove Proposition 4.1 it is sufficient to give just one example of a germ of $(2,n)$-distributions of the maximal class $n-3$. As such example one can take the germ at 0 of the distribution, associated with underdetermined ODE (1.1), where $F = \frac{1}{2}p_{n-3}^2$.

In the present paper we treat the germs of $(2,n)$ distributions of the maximal class $n-3$. In the cases $n=5$ and $n=6$ any rank 2 distribution has maximal class if and only if it has maximal possible small growth vector, namely, $(2,3,5)$ in the case $n=5$ and $(2,3,5,6)$ in the case $n=6$. It can be obtained by direct computations (see Propositions 3.5 and 3.6 of [15] respectively). Starting with $n=7$, distributions with different small growth vectors may have the maximal class.

Regarding completely nonholonomic rank 2 distributions of non-maximal class, it is easy to describe all distributions of minimal class 1. From [15, Remark 3.4] it follows that a rank 2 distribution D has the smallest possible class 1 at a point q if and only if $\dim D^3(q) = 4$. Moreover, it is easy to see that in this case D is either the Goursat distribution or, by the factorization of the ambient manifold by the characteristics of D^2 (or series of such factorizations), one can get a new distribution $\widetilde{D}$, satisfying $\dim \widetilde{D}^3 = 5$ (or, equivalently, having the class greater than 1). See the last section of the paper for more details.

On the other hand, ***we have no examples of completely nonholonomic rank* 2 *distributions of constant class* $2 \leq m \leq n-4$ *and our conjecture is that there are no such***

distributions. We succeeded to prove this conjecture for $m = 2, 3, 4$ by direct computation, which means, in particular, that any counter-example to our conjecture, if it exists, should live on at least a 9-dimensional manifold.

Remark 4.1. By above, for $(2, n)$-distributions of maximal class with $n > 4$ it is necessary that $\dim D^3 = 5$. Hence, each such distribution satisfies relation (2.6). □

From now on D is a $(2, n)$-distribution of maximal class $m = n - 3$. Let us study curves of flags associated with its abnormal extremals in more detail. Let

$$\mathcal{R}_D = \{\lambda \in (D^2)^\perp \backslash (D^3)^\perp : \nu(\lambda) = n - 3\}, \quad \mathcal{R}_D(q) = \mathcal{R}_D \cap T_q^*M. \tag{4.1}$$

As was already mentioned the set $\mathcal{R}_D$ is open dense subset of $(D^2)^\perp \backslash (D^3)^\perp$ and the set $\mathcal{R}_D(q)$ is a nonempty open set in Zariski topology on the linear space $(D^2)^\perp(q)$. The following Proposition follows easily from Proposition 3.1, formula (3.12), and Remark 3.2

Proposition 4.2. *Let γ be a segment of abnormal extremal such that $\gamma \subset \mathcal{R}_D$. Then the associated curve of flags is the curve of complete flags in W, i.e., it has the form*

$$\begin{aligned} &\lambda \mapsto \{0 = J_{(n-3)}(\lambda) \subset J_{(n-4)}(\lambda) \subset \ldots \subset J_{(1)}(\lambda) \subset J(\lambda) \subset \\ &J^{(1)}(\lambda) \subset \ldots \subset J^{(n-4)}(\lambda) \subset J^{(n-3)}(\lambda) = W\}, \quad \lambda \in \gamma, \end{aligned} \tag{4.2}$$

where $\dim J_{(i)}(\lambda) = n - 3 - i$ *and* $\dim J^{(i)}(\lambda) = n - 3 + i$.

Moreover, if $\lambda \mapsto \ell(\lambda)$ is a smooth curve of vectors such that $J_{(n-4)}(\lambda) = \mathbb{R}\ell(\lambda)$, $\varphi\colon \gamma \mapsto \mathbb{R}$ is a parameterization of γ, and $\psi = \varphi^{-1}$, then for $i = 0, \ldots n - 3$ and any $t \in \varphi(\gamma)$

$$\begin{aligned} J_{(i)}\big(\psi(t)\big) &= \operatorname{span}\big\{\frac{d^j}{dt^j}\ell\big(\psi(t)\big) : 0 \le j \le n - 4 - i\big\} \\ J^{(i)}\big(\psi(t)\big) &= \operatorname{span}\big\{\frac{d^j}{dt^j}\ell\big(\psi(t)\big) : 0 \le j \le n - 4 + i\big\}. \end{aligned} \tag{4.3}$$

In other words, the curve of flags, associated with the abnormal extremal $\gamma \subset \mathcal{R}_D$ can be recovered by differentiation from the curve of their one-dimensional subspaces $\lambda \mapsto J_{(n-4)}(\lambda)$, i.e., from the curve in the projective space $\mathbb{P}W$ of the $2m$-dimensional symplectic space W $(m = n-3)$. Moreover, the curve $\lambda \mapsto J_{(n-4)}(\lambda)$ is not arbitrary curve in $\mathbb{P}W$ but a curve, which can be completed by the appropriate number of differentiation to the curve of Lagrangian subspaces of γ.

The differential geometry of curves in projective spaces is the classical subject, essentially completed already in 1905 by E.J. Wilczynski ([14]). In particular, it is well known that these curves are endowed with the canonical projective structure, i.e., we have a distinguished set of parameterizations (called projective) such that the transition function from one such parameterization to another is a Möbius transformation. Let us demonstrate how to construct it for the curve $\lambda \mapsto J_{(n-4)}(\lambda)$, $\lambda \in \gamma$.

As before, let $\Gamma(J_{(n-4)}(\lambda))$ be the space of all smooth mappings ℓ from γ to the ambient symplectic space W such that $\ell(\lambda) \in J_{(n-4)}(\lambda)$ for all $\lambda \in \gamma$. The elements of $\Gamma(J_{(n-4)}(\lambda))$ will be called *sections of the curve* $\lambda \mapsto J_{(n-4)}(\lambda)$. Take some parameterization $\varphi\colon \gamma \mapsto \mathbb{R}$ of γ and let $\psi = \varphi^{-1}$. By Proposition 4.2 for any section ℓ one has relation

$$\operatorname{span}\big\{\frac{d^j}{dt^j}\ell\big(\psi(t)\big) \mid 0 \le j \le 2m - 1\big\} = W. \tag{4.4}$$

It is well known that there exists the unique, up to the multiplication on a nonzero constant, section E_φ such that

$$\frac{d^{2m}}{dt^{2m}}E_\varphi\big(\psi(t)\big) = \sum_{i=0}^{2m-2} B_i^\varphi(t)\frac{d^i}{dt^i}E_\varphi\big(\psi(t)\big), \tag{4.5}$$

i.e. the coefficient of the term $\frac{d^{2m-1}}{dt^{2m-1}}E_\varphi\big(\psi(t)\big)$ in the linear decomposition of $\frac{d^{2m}}{dt^{2m}}E_\varphi\big(\psi(t)\big)$ w.r.t. the basis $\big\{\frac{d^i}{dt^i}\ell\big(\psi(t)\big) : 0 \le i \le 2m-1\big\}$ vanishes.

Further, let φ_1 be another parameter and $\upsilon = \varphi \circ \varphi_1^{-1}$. Then it is not hard to show that the coefficients and B_{2m-2}^φ and $B_{2m-2}^{\varphi_1}$ in the decomposition (4.5), corresponding to parameterizations φ and φ_1, are related as follows:

$$\widetilde{B}_{2m-2}^{\varphi_1}(\tau) = \upsilon'(\tau)^2 B_{2m-2}^\varphi(\upsilon(\tau)) - \frac{m(4m^2-1)}{3}\mathbb{S}(\upsilon)(\tau), \tag{4.6}$$

where $\mathbb{S}(\upsilon)$ is a Schwarzian derivative of υ, $\mathbb{S}(\upsilon) = \frac{d}{dt}\left(\frac{\upsilon''}{2\,\upsilon'}\right) - \left(\frac{\upsilon''}{2\,\upsilon'}\right)^2$.

From the last formula and the fact that $\mathbb{S}\upsilon \equiv 0$ if and only if the function υ is Möbius it follows that *the set of all parameterizations φ of γ such that*

$$B_{2m-2}^\varphi \equiv 0 \tag{4.7}$$

defines the canonical projective structure on γ. Such parameterizations are called the *projective parameterizations of the abnormal extremal γ.*

Remark 4.2. Another description of the canonical projective structure on an abnormal extremal γ can be obtained by working with the Jacobi curve $\lambda \mapsto J(\lambda)$, using the notion of the cross-ratio of four point in Lagrangian Grassmannian ([4],[15]). This approach allows to construct the canonical projective structures in much more general situations.

Further, since in our case the m-dimensional subspaces $J\big(\psi(t)\big)$ are Lagrangian, it is easy to show that the condition (4.5) for the section $E_\varphi(t)$ is equivalent to the following one

$$\bar\sigma\left(\frac{d^m}{dt^m}E_\varphi\big(\psi(t)\big), \frac{d^{m-1}}{dt^{m-1}}E_\varphi\big(\psi(t)\big)\right) \equiv C, \quad C \in \mathbb{R}\backslash\{0\}. \tag{4.8}$$

Therefore in our case we can “kill” the freedom of the multiplication on a nonzero constant in the definition of section E_φ by setting

$$\left|\bar\sigma\left(\frac{d^m}{dt^m}E_\varphi\big(\psi(t)\big), \frac{d^{m-1}}{dt^{m-1}}E_\varphi\big(\psi(t)\big)\right)\right| \equiv 1. \tag{4.9}$$

(see also [16]). There are exactly two sections of the curve $\lambda \mapsto J_{(n-4)}(\lambda)$ with the parameterization φ, satisfying (4.9), and they are obtained one from another by the multiplication on -1. These sections are called *the canonical sections of the curve $\lambda \to J_{(n-4)}(\lambda)$ w.r.t. the parameterization φ* and both of them will be denoted in the sequel by E_φ. From these sections one can obtain the moving frame $\big\{\frac{d^i}{dt^i}E_\varphi\big(\psi(t)\big) : 0 \le i \le 2m-1\big\}$ on W, which is defined again up to a multiplication by -1 and it will be called *the canonical moving frame of the curve $\lambda \mapsto J_{(n-4)}(\lambda)$ w.r.t. the parameterization φ.*

Finally, it can be shown easily that in the case when m-dimensional subspaces $J\big(\psi(t)\big)$ are Lagrangian, from (4.7) it follows that

$$B_{2m-3}^\varphi \equiv 0 \tag{4.10}$$

5. The canonical frame

Now we are ready to describe the manifold, on which the canonical frame for $(2,n)$-distribution of maximal class, $n>5$, can be constructed. Given $\lambda\in\mathcal{R}_D$ denote by $\mathfrak{P}_\lambda$ the set of all projective parameterizations $\varphi:\gamma\mapsto\mathbb{R}$ on the characteristic curve γ , passing through λ, such that $\varphi(\lambda)=0$. Denote

$$\Sigma_D=\{(\lambda,\varphi):\lambda\in\mathcal{R}_D,\varphi\in\mathfrak{P}_\lambda\}.$$

Actually, Σ_D is a principal bundle over $\mathcal{R}_D$ with the structural group of all Möbius transformations, preserving 0 and $\dim\Sigma_D=2n-1$.

Theorem 1. *For any $(2,n)$-distribution, $n>5$, of maximal class there exist two canonical frames on the corresponding $(2n-1)$-dimensional manifold Σ_D, obtained one from another by a reflection. The group of symmetries of such distributions is at most $(2n-1)$-dimensional.*

Proof. Define the following two fiber-preserving flows on Σ_D:

$$F_{1,s}(\lambda,\varphi)=(\lambda,e^{2s}\varphi),\quad F_{2,s}(\lambda,\varphi)=\left(\lambda,\frac{\varphi}{-s\varphi+1}\right),\quad \lambda\in\mathcal{R}_D,\varphi\in\mathfrak{P}_\lambda. \tag{5.1}$$

Further, let δ_s be the flow of homotheties on the fibers of T^*M:

$$\delta_s(p,q)=(e^sp,q),\quad q\in M,\ p\in T_q^*M \tag{5.2}$$

(actually the Euler field e generates this flow). The following flow

$$F_{0,s}(\lambda,\varphi)=\big(\delta_{2s}(\lambda),\varphi\circ\delta_{2s}^{-1}\big) \tag{5.3}$$

is well-defined on Σ_D (here we use that δ_s preserves the characteristic 1-foliation). For any $0\le i\le 2$ let g_i be the vector field on Σ_D, generating the flow $F_{i,s}$. Note that g_1 and g_2 are just fundamental vector fields on Σ_D defined by the structure of the principle fiber bundle on Σ_D.

Besides, the characteristic 1-foliation on $(D^2)^\perp$ can be lifted to the *parameterized* 1-foliation on Σ_D, which gives one more canonical vector field on Σ_D. Indeed, let $u=(\lambda,\varphi)\in\Sigma_D$ and γ be the characteristic curve, passing through λ (so, φ maps γ to $\mathbb{R}$). Then the mapping

$$\Upsilon_u(t)=\big(\varphi^{-1}(t),\varphi(\cdot)-t\big)$$

defines the parameterized curve on Σ_D, the lift of γ to Σ_D, and $\Upsilon_u(0)=u$. The additional canonical vector field h on Σ_D is defined by

$$h(u)=\frac{d}{dt}\Upsilon_u(t)|_{t=0}. \tag{5.4}$$

It can be shown easily that

$$[g_1,g_2]=2g_2,\ [g_1,h]=-2h,\ [g_2,h]=g_1,\ [g_0,h]=0,\ [g_0,g_i]=0. \tag{5.5}$$

Therefore the linear span (over $\mathbb{R}$) of the vector fields g_0, g_1, g_2, and h is endowed with a structure of the Lie algebra isomorphic to $\mathfrak{gl}(2,\mathbb{R})$.

Now we will construct one more canonical, up to the sign, vector field on Σ_D. First from (3.5) it is easy to get

$$\mathcal{J}_{(1)}(\lambda)=T_\lambda\Big((D^2)^\perp\big(\pi(\lambda)\big)\Big)\oplus\mathcal{C} \tag{5.6}$$

Here $T_\lambda\Big((D^2)^\perp\big(\pi(\lambda)\big)$ is the tangent space to the fiber $(D^2)^\perp\big(\pi(\lambda)\big)$ at the point λ and it is actually equal to $\{v\in T_\lambda(D^2)^\perp,\pi_*v=0\}$, the vertical subspace of $T_\lambda(D^2)^\perp$. Let

$$V_i(\lambda)=\mathcal{J}_{(i)}(\lambda)\cap T_\lambda\Big((D^2)^\perp\big(\pi(\lambda)\big)\Big). \tag{5.7}$$

Since $\mathcal{J}_{(i+1)} \subseteq \mathcal{J}_{(i)}$, identity (5.6) yields

$$\mathcal{J}_{(i)} = V_i \oplus \mathcal{C} \quad \forall i \geq 1. \tag{5.8}$$

Note also that from Remark 3.2 and formula (5.8) it follows that

$$V_i(\lambda) = \left\{ v \in V_{i-1}(\lambda) : \begin{array}{l} \exists \text{ a vector field } \mathcal{V} \in V_{i-1} \text{ with} \\ \mathcal{V}(\lambda) = v \text{ such that } \big[\mathcal{C}, \mathcal{V}\big](\lambda) \in \mathcal{J}_{(i-1)}(\lambda) \end{array} \right\}. \tag{5.9}$$

Furthermore, identity (5.9) remains true if one replaces the quantor $\exists$ by $\forall$. In particular, from the last identity and Proposition 4.2 it follows that for any $\lambda \in \mathcal{R}_D$ the distributions $\mathcal{J}_{(i)}$ and V_i satisfy

$$\mathcal{J}_{(i)}(\lambda) = [\mathcal{C}, \mathcal{J}_{(i+1)}], \quad V_i(\lambda) = [\mathcal{C}, V_{i+1}], \quad 0 \leq i < n-4 \tag{5.10}$$

Take vector fields $E \in V_{n-4}$ and $H \in \mathcal{C}$ without stationary points and suppose that E is not collinear to the Euler field e. Then by our construction of subspaces $J_{(i)}$ and $J^{(i)}$, Proposition 4.2 and relation (5.8) it follows that on $\mathcal{R}_D$

$$\begin{aligned} &V_{n-4} = \langle e, E \rangle \quad \mathcal{J}_{(i)} = \langle H, e, E, \{(\mathrm{ad} H)^j E\}_{j=1}^{n-4-i} \rangle \quad 0 \leq i \leq n-4, \\ &\mathcal{J}^{(i)} = \langle H, e, E, \{(\mathrm{ad} H)^j E\}_{j=1}^{n-4+i} \rangle \ 0 \leq i \leq n-3. \end{aligned} \tag{5.11}$$

Now let γ be the abnormal extremal, passing through $\lambda \in \mathcal{R}_D(\lambda)$. As before, let also ϕ be the canonical projection from a sufficiently small neighborhood O_γ of γ to the factor $O_\gamma/(\textit{the characteristic one-foliation})$, $\bar{e} = \phi_* e(\lambda)$, and $E_\varphi(\lambda)$ be one of the two canonical sections of the curve $\lambda \to J_{(n-4)}(\lambda)$ w.r.t. the parameterization φ. Then, obviously, there exists a unique affine line $\mathrm{Aff}_\varphi(\lambda)$ in the plane $V_{n-4}(\lambda)$ such that

$$\phi_*\big(\mathrm{Aff}_\varphi(\lambda)\big)/\{\mathbb{R}e\} = E_\varphi(\lambda). \tag{5.12}$$

Clearly, the affine line $\mathrm{Aff}_\varphi(\lambda)$ is parallel to the vector $e(\lambda)$, but does not pass through the origin of the linear space $V_{n-4}(\lambda)$.

Further, denote by $\Pi : \Sigma_D \mapsto \mathcal{R}_D$ the canonical projection. Let ε_1 be a vector field on Σ_D such that

$$\Pi_* \varepsilon_1(u) \in \mathrm{Aff}_\varphi(\lambda) \cup \big(-\mathrm{Aff}_\varphi(\lambda)\big), \quad \forall u = (\lambda, \varphi) \in \Sigma_D. \tag{5.13}$$

Such fields ε_1 are defined modulo $\mathcal{W}_0 = \langle g_0, g_1, g_2 \rangle$ and the sign.

The main question now is *how to choose among them the canonical field, up to the sign?* For this first let us prove the following lemma, which will be also useful in the sequel:

Lemma 5.1. *The following commutative relations hold*

$$[V_i, V_i] \subseteq V_i, \quad \forall i \geq 0; \tag{5.14}$$

$$[V_i, \mathcal{J}^{(i)}] \subseteq \mathcal{J}^{(i)}, \quad \forall i \geq 0. \tag{5.15}$$

Proof. 1) The proof of (5.14) is by induction on i. For $i = 0$ the formula is trivial, because $V_0(\lambda)$ is the tangent space to the fiber $(D^2)^\perp\big(\pi(\lambda)\big)$ of $(D^2)^\perp$. Now suppose that (5.14) holds for some i and prove it for $i+1$. Take two vector fields W_1 and W_2 being tangent to V_{i+1} and prove that $[W_1, W_2]$ is tangent to V_{i+1}. According to (5.9) it is equivalent to the fact that $\big[\mathcal{C}, [W_1, W_2]\big] \subset \mathcal{J}_{(i)}$. Note that again by (5.9) we have $[\mathcal{C}, W_j] \subset \mathcal{J}_{(i)}$, $j = 1, 2$. Besides, by construction $V_{i+1} \subset V_i$. Taking into account all this, the relations (5.8), and the induction hypothesis we obtain from Jacobi identity that

$$\big[\mathcal{C}, [W_1, W_2]\big] = \big[[\mathcal{C}, W_1], W_2\big] + \big[W_1, [\mathcal{C}, W_2]\big] \subset [V_i \oplus \mathcal{C}, V_{i+1}] \subset \mathcal{J}_{(i)}.$$

So, $[W_1, W_2] \in V_{i+1}$, i.e. $[V_{i+1}, V_{i+1}] \subseteq V_{i+1}$, which completes the proof by induction of (5.14).

2) Directly from the definitions of $\mathcal{J}$, V_0, and formula (3.5) it is easy to prove both (5.15) for $i = 0$ and the relation

$$[\mathcal{J}^{(0)}, \mathcal{J}^{(0)}] \subseteq \mathcal{J}^{(1)}. \tag{5.16}$$

Now assume that (5.15) holds for some $i = l \geq 0$ and prove it for $i = l + 1$. Since $V_{l+1} \subset V_l$, then by our assumption $[V_{l+1}, \mathcal{J}^{(l)}] \subset \mathcal{J}^{(l)}$. Therefore

$$\big[\mathcal{C}, [V_{l+1}, \mathcal{J}^{(l)}]\big] \subset \mathcal{J}^{(l+1)}. \tag{5.17}$$

On the other hand, using consequently relation (3.2), the Jacobi identity, and relations (5.17), (5.10), we obtain

$$\begin{aligned}[V_{l+1}, \mathcal{J}^{(l+1)}] &= \big[V_{l+1}, [\mathcal{C}, \mathcal{J}^{(l)}]\big] \\ &= \big[\mathcal{C}, [V_{l+1}, \mathcal{J}^{(l)}]\big] + \big[[V_{l+1}, \mathcal{C}], \mathcal{J}^{(l)}\big] \subset \mathcal{J}^{(l+1)} + [\mathcal{J}_{(l)}, \mathcal{J}^{(l)}].\end{aligned} \tag{5.18}$$

If $l = 0$, then $\mathcal{J}_{(0)} = \mathcal{J}^{(0)}$ and relations (5.16), (5.18) imply (5.15) for $i = 1$. If $l \geq 1$, then by (5.8), the induction hypothesis, and (3.2)

$$[\mathcal{J}_{(l)}, \mathcal{J}^{(l)}] = [V_l \oplus \mathcal{C}, \mathcal{J}^{(l)}] = [V_l, \mathcal{J}^{(l)}] + [\mathcal{C}, \mathcal{J}^{(l)}] \subset \mathcal{J}^{(l+1)}. \tag{5.19}$$

This together with (5.18) implies that relation (5.15) holds also for $i = l + 1$. The proof of (5.15) by induction is completed. □

Fix some vector field ε_1, satisfying (5.13). For any $\lambda \in \Sigma_D$ let

$$\mathcal{L}_i(\lambda) = \begin{cases} \{v \in T_\lambda \Sigma_D : \Pi_* v \in \mathcal{J}_{(m-i)}(\lambda)\} & 1 \leq i \leq m \\ \{v \in T_\lambda \Sigma_D : \Pi_* v \in \mathcal{J}^{(i-m)}(\lambda)\} & m \leq i \leq 2m \end{cases}, \tag{5.20}$$

$$\mathcal{W}_i(\lambda) = \{v \in T_\lambda \Sigma_D : \Pi_* v \in V_{m-i}(\lambda)\}, 1 \leq i \leq m \tag{5.21}$$

Then by (5.8)

$$\mathcal{L}_i = \mathcal{W}_i + \mathbb{R}h, \quad 1 \leq i \leq m. \tag{5.22}$$

According to the splitting (5.22), the projection $\mathrm{Pr}_{i,\lambda}$ from $\mathcal{L}_i(\lambda)$ onto $\mathcal{W}_i(\lambda)$, which is parallel to h, is well defined. Define the vector fields ε_i, $2 \leq i \leq 2m$ in addition to ε_1 by the following recursive formulas:

$$\varepsilon_i(\lambda) = \begin{cases} \mathrm{Pr}_{i,\lambda}\big([h, \varepsilon_{i-1}](\lambda)\big), & 2 \leq i \leq m-1; \\ [h, \varepsilon_{i-1}], & m \leq i \leq 2m. \end{cases} \tag{5.23}$$

Also let

$$\eta = [\varepsilon_1, \varepsilon_{2m}]. \tag{5.24}$$

Lemma 5.2. *The tuple $(h, \{g_i\}_{i=0}^2, \{\varepsilon_i\}_{i=1}^{2m}, \eta)$ is a frame on Σ_D.*

Proof. As before, take vector fields $E \in V_{n-4}$ and $H \in \mathcal{C}$ without stationary points and suppose that E is not collinear to the Euler field e. From (5.11) and the fact that $\dim \mathcal{J}^{(n-3)} = 2n - 4$ it follows that in order to prove the lemma it is sufficient to prove that $[E, (\mathrm{ad}\, H)^{2n-7} E] \notin \mathcal{J}^{(n-3)}$. Besides, by (5.15) we have $[E, \mathcal{J}^{(n-4)}] \subseteq \mathcal{J}^{(n-4)}$. Therefore in order to prove the lemma it is sufficient to prove that $[E, \mathcal{J}^{(n-3)}] \nsubseteq \mathcal{J}^{(n-3)}$. Assuming the converse and taking into account that $\mathcal{J}^{(n-3)} = \{\mathfrak{s}|_{(D^2)^\perp} = 0\}$ we get that $E \in \ker \sigma|_{(D^2)^\perp}$, which implies that E is collinear to H. We have the contradiction. □

We will say that the frame $(h, \{g_i\}_{i=0}^2, \{\varepsilon_i\}_{i=1}^{2m}, \eta)$ is *associated with the vector field* ε_1. Note that by our constructions

$$\mathcal{L}_i = \langle h, \{g_i\}_{i=0}^2, \{\varepsilon_i\}_{j=1}^i\rangle, \quad 0 \le i \le 2m, \tag{5.25}$$

$$\mathcal{W}_i = \langle \{g_i\}_{i=0}^2, \{\varepsilon_i\}_{j=1}^i\rangle \quad 0 \le i \le m-1. \tag{5.26}$$

Then by Lemma 5.1 we have

$$[\varepsilon_1, \varepsilon_2] = \kappa_1\varepsilon_2 \mod \mathcal{W}_1 \tag{5.27}$$

and, in the case $n > 5$,

$$[\varepsilon_1, \varepsilon_4] = \kappa_2\varepsilon_3 + \kappa_3\varepsilon_4 \mod \mathcal{L}_2. \tag{5.28}$$

The normalization of the field ε_1 can be done by studying how the functions κ_i, $1 \le i \le 3$, are transformed when we pass from the frame, associated with ε_1 to the frame, associated with another vector field, satisfying (5.13). For this first we need

Lemma 5.3. *The following commutative relations hold:*

$$[g_1, \varepsilon_i] = (2m - 2i + 1)\varepsilon_i \mod \mathcal{L}_0, \tag{5.29}$$

$$[g_2, \varepsilon_i] = (i-1)(2m - i + 1)\varepsilon_{i-1} \mod \mathcal{L}_0, \tag{5.30}$$

$$[g_0, \varepsilon_i] = -\varepsilon_i \mod \mathcal{L}_0 \tag{5.31}$$

In the case $1 \le i \le m-1$ *the subspace* $\mathcal{L}_0 = \operatorname{span}\{h, g_0, g_1, g_2\}$ *can be replaced by* $\mathcal{W}_0 = \operatorname{span}\{g_0, g_1, g_2\}$.

Proof. Let $\psi : \mathbb{R} \mapsto \mathbb{R}$, $\psi_0 = 0$. Then from (4.9) it follows easily that

$$\varepsilon_{\psi\circ\varphi}(\lambda) = (\psi'(0))^{m-\frac{1}{2}}\varepsilon_\varphi(\lambda) \mod\{\mathbb{R}e(\lambda)\}. \tag{5.32}$$

Using the last identity and (5.1) one gets without difficulties that

$$(F_{1,s}^{-1})_*(\varepsilon_1) = e^{(2m-1)s}\varepsilon_1 \mod \mathcal{W}_0, \tag{5.33}$$

$$(F_{2,s}^{-1})_*(\varepsilon_1) = \varepsilon_1 \mod \mathcal{W}_0. \tag{5.34}$$

In addition, let us show that

$$(F_{0,s}^{-1})_*(\varepsilon_1) = e^{-s}\varepsilon_1 \mod \mathcal{W}_0 \tag{5.35}$$

Indeed, the homotheties δ_s, defined by (5.2), preserve both the characteristic line distribution $\mathcal{C}$ and the distribution $\mathcal{J}$. Therefore it preserves also all $\mathcal{J}^{(i)}$. Note also that

$$\delta_s^*\sigma = e^s\sigma, \tag{5.36}$$

which implies that δ_s preserve all $\mathcal{J}_{(i)}$ as well. This together with (5.3) yields that there exists a function α such that $(F_{0,s}^{-1})_*(\varepsilon_1) = \alpha\varepsilon_1 \mod \mathcal{W}_0$. The last formula together with (5.23) implies that

$$(F_{0,s}^{-1})_*(\varepsilon_i) = \alpha\varepsilon_i \mod \mathcal{L}_{i-1}, \quad 1 \le i \le 2m. \tag{5.37}$$

From the normalization (4.9) of the vector field ε_1 it follows that

$$\Pi^*\sigma(\varepsilon_{m+1}, \varepsilon_m) = 1, \tag{5.38}$$

which together with (5.37) implies that

$$(\Pi \circ F_{0,s}^{-1})^*\sigma(\varepsilon_{m+1}, \varepsilon_m) = \alpha^2. \tag{5.39}$$

On the other hand, by (5.3), $\Pi \circ F_{0,s}^{-1} = \delta_{2s}^{-1} \circ \Pi$, which together with (5.36) implies that

$$(\Pi \circ F_{0,s}^{-1})^* \sigma(\varepsilon_{m+1}, \varepsilon_m) = e^{-2s}. \tag{5.40}$$

Combining (5.39) and (5.40), we get $\alpha = e^{-s}$, which proves (5.35).

Relations (5.33),(5.34), and (5.35) imply (5.29), (5.30), and (5.31) respectively for $i = 1$. Then relations (5.29)-(5.31) for $i > 1$ can be proved by induction, using (5.5), (5.23), and the Jacobi identity. The last sentence of the lemma follows from the fact that by our constructions the vectors $\pi_* \circ \Pi_*(\varepsilon_i) = 0$ for $1 \le i \le m-1$.□

The following lemma gives the normalization of the fields ε_1:

Lemma 5.4. *Among all vector fields , satisfying (5.13), there exists a unique, up to the sign, field $\tilde\varepsilon_1$ such that all functions κ_i, $1 \le i \le 3$, are identically zero, namely, the following commutative relations hold*

$$[\tilde\varepsilon_1, \tilde\varepsilon_2] \in \mathcal{W}_1, \quad [\tilde\varepsilon_1, \tilde\varepsilon_4] \in \mathcal{L}_2. \tag{5.41}$$

Proof. Let ε_1 and $\tilde\varepsilon_1$ be two vector field, satisfying (5.13). Then there exist functions $\{\mu_i\}_{i=0}^2$ such that

$$\tilde\varepsilon_1 = \pm\varepsilon_1 + \mu_0 g_0 + \mu_1 g_1 + \mu_2 g_2.$$

Using (5.5) it is easy to show that $[h, \tilde\varepsilon_2] = \pm[h, \varepsilon_2] + 2\mu_1 h \bmod \mathcal{W}_0$, which implies that $\tilde e_2 = e_2 \bmod \mathcal{W}_0$. In the same way one can show that

$$\begin{aligned} \varepsilon_i &= \pm\varepsilon_i \bmod \mathcal{W}_0, \quad 2 \le i \le m-1; \\ \tilde\varepsilon_i &= \pm\varepsilon_i \bmod \mathcal{L}_0, \quad m \le i \le 2m. \end{aligned} \tag{5.42}$$

Suppose that the frame, associated with ε_1, satisfies relations (5.27) and (5.28), while the frame, associated with $\tilde\varepsilon_1$ satisfies

$$\begin{aligned} [\tilde\varepsilon_1, \tilde\varepsilon_2] &= \tilde\kappa_1 \tilde\varepsilon_2 \mod \mathcal{W}_1 \\ [\tilde\varepsilon_1, \tilde\varepsilon_4] &= \tilde\kappa_2 \tilde\varepsilon_3 + \tilde\kappa_3 \tilde\varepsilon_4 \bmod \mathcal{L}_2. \end{aligned}$$

Then by direct computation, using relations (5.29)-(5.31) and (5.42), one can show without difficulties that

$$\begin{cases} \tilde\kappa_1 = \kappa_1 \pm (2m-3)\mu_1 - \mu_0 \\ \tilde\kappa_2 = \kappa_2 \pm (6m-9)\mu_2 \\ \tilde\kappa_3 = \kappa_3 \pm (2m-7)\mu_1 - \mu_0 \end{cases}. \tag{5.43}$$

Obviously, system (5.43) w.r.t. μ_i, $0 \le i \le 2$, has a unique, up to the sign, solution, when $\tilde\kappa_i = 0$, $1 \le i \le 2$, which completes the proof of the lemma. □

Suppose that $\tilde\varepsilon_1$ is one of the two vector fields, found in the previous lemma. Then two frames $(h, \{g_i\}_{i=0}^2, \{\tilde\varepsilon_i\}_{i=1}^{2m}, \eta)$ and $(h, \{g_i\}_{i=0}^2, \{-\tilde\varepsilon_i\}_{i=1}^{2m}, \eta)$ are canonically defined on Σ_D. They are *the canonical frames* we were looking for. Also, this immediately implies that the groups of symmetries of $(2, n)$-distributions, $n > 5$, of maximal class is at most $(2n-1)$-dimensional. The proof of the theorem is completed. □

Remark 5.1. The normalization, implemented above for $n > 5$, does not work in the case $n = 5$. In this case $m = 2$ and the relation (5.28) is not true (actually, in this case $[\varepsilon_1, \varepsilon_4] = \eta$). Note also that the bracket $[\varepsilon_1, \varepsilon_3]$ cannot give new conditions with compare to the bracket $[\varepsilon_1, \varepsilon_2]$, because

$$[\varepsilon_1, \varepsilon_3] = \big[h, [\varepsilon_1, \varepsilon_2]\big]. \tag{5.44}$$

It is also in accordance with the result of E. Cartan [7]. In the case $n = 5$ the canonical frame does not exist on Σ_D (which is 9-dimensional), one has to prolong further to construct it. □

Remark 5.2. One can suggest normalizations of the vector field ε_1 different from one given by Lemma 5.4. For example, among all vector fields , satisfying (5.13) there exist unique, up to the sign, vector field ε_1 such that $[h, \varepsilon_1] \in \mathcal{W}_2$, $[\varepsilon_1, \varepsilon_2] \in \mathcal{W}_1$, and $[\varepsilon_1, \varepsilon_4] \in \text{span}\{\varepsilon_4, \mathcal{L}_2\}$. In the case $n > 6$ we can get one more normalization by replacing the last condition of the previous normalization by $[h, \varepsilon_2] \in \mathcal{W}_3$. The frames, associated with such ε_1, are also intrinsically defined, up to the corresponding reflection. In particular, Theorem 2 below remains true, if one uses these frames instead of the canonical frames, constructed in the proof of Theorem 1. Note also that both of these normalizations cannot be implemented in the case $n = 5$, because in this case the vector $[h, \varepsilon_1]$ is not tangent to the fiber of Σ_D, considered as the fiber bundle over M.□

Theorem 2. *Two rank 2 distributions D_1 and D_2 of maximal class are equivalent iff there exists the diffeomorphism $\mathcal{F} : \Sigma_{D_1} \mapsto \Sigma_{D_2}$, which transform one of the canonical frames of D_1 to one of the canonical frames of D_2.*

Proof. The necessity is obvious. Let us prove the sufficiency. Let $(h^k, \{g_i^k\}_{i=0}^2, \{\tilde{\varepsilon}_i^k\}_{i=1}^{2m}, \eta^k)$ are the canonical frames of the distributions D_k respectively, where $k = 1, 2$, such that $\mathcal{F}$ transforms the frame with $k = 1$ to the frame with $k = 2$. If the maps $\Pi_k : \Sigma_{D_k} \mapsto \mathcal{R}_{D_k}$ and $\pi : T^*M \mapsto M$ are the canonical projections, then the map $\mathfrak{p}_k \stackrel{def}{=} \Pi_k \circ \pi$ defines the fiber bundle Σ_{D_k} over M. By our constructions, the tangent spaces to this fibers coincide with span$\{g_0, g_1, g_2, \varepsilon_1^k, \ldots, \varepsilon_{m-1}^k\}$. This yields that the diffeomorphism $\mathcal{F}$ is fiberwise, i.e., there exists the diffeomorphism $F : M \mapsto M$ such that

$$F \circ \mathfrak{p}_1 = \mathfrak{p}_2 \circ \mathcal{F}. \tag{5.45}$$

Note that by our constructions $\Pi_* h^1$ and $\Pi_* h^2$ span the characteristic line distributions of D and $\bar{D}$ respectively. Therefore from Remark 4.1 and relations (5.45), (2.6) it follows easily that $F_* D = \bar{D}$, which means that the distributions D and $\bar{D}$ are equivalent. □

Now we will list several properties of the canonical frames. First from (5.41) and (5.44) it follows that

$$[\tilde{\varepsilon}_1, \tilde{\varepsilon}_3] \in \mathcal{L}_2. \tag{5.46}$$

Second, directly from Lemma 5.1 it follows that

$$[\tilde{\varepsilon}_1, \tilde{\varepsilon}_i] \in \mathcal{L}_i, \quad 5 \leq i \leq 2n - 6; \tag{5.47}$$

$$[\tilde{\varepsilon}_{i_1}, \tilde{\varepsilon}_{i_2}] \in \mathcal{L}_{i_2+1}, \quad 2 \leq i_1 \leq i_2 \leq 2n - 6 - i_1 \tag{5.48}$$

where the subspaces $\mathcal{L}_i$ are as in (5.25). Further, from identities (4.5), (4.7), and (4.10) it follows that

$$[h, \tilde{\varepsilon}_{2n-6}] \in \mathcal{L}_{2n-9}. \tag{5.49}$$

Finally we have

Lemma 5.5. *The following relation holds*

$$[\tilde{\varepsilon}_i, \tilde{\varepsilon}_{2n-5-i}] = (-1)^{i+1}\eta \mod \mathcal{L}_{2n-6}, \quad 1 \leq i \leq n - 3. \tag{5.50}$$

Proof. The proof is by induction on i. For $i = 1$ formula (5.50) follows from (5.24). Now assume that (5.50) holds for some $i = j$ and prove it for $i = j + 1$. By (5.48)

$$[\tilde{\varepsilon}_j, \tilde{\varepsilon}_{2n-6-j}] \in \mathcal{L}_{2n-5-j}$$

Taking Lie brackets with h from both sides of the last inclusion and using formulas (5.23) and (5.49) together with the Jacobi identity, one gets

$$[\tilde\varepsilon_{j+1},\tilde\varepsilon_{2n-5-(j+1)}]+[\tilde\varepsilon_j,\tilde\varepsilon_{2n-6-j}]\in\mathcal{L}_{2n-6}.$$

Therefore by induction hypothesis for $i=j$

$$[\tilde\varepsilon_{j+1},\tilde\varepsilon_{2n-5-(j+1)}]=-[\tilde\varepsilon_j,\tilde\varepsilon_{2n-6-j}] \bmod \mathcal{L}_{2n-6}=(-1)^{j+2}\eta \bmod \mathcal{L}_{2n-6},$$

which proves formula (5.50) also for $i=j+1$. The proof by induction is completed. □

6. The most symmetric case

The present section is devoted to the following

Theorem 3. *Let $n>5$. Then any $(2,n)$-distribution of maximal class with $(2n-1)$-dimensional Lie algebra of infinitesimal symmetries is locally equivalent to the distribution, associated with the underdetermined ODE $z'(x)=\big(y^{(n-3)}(x)\big)^2$. The symmetry algebra of this distribution is isomorphic to a semidirect sum of $\mathfrak{gl}(2,\mathbb{R})$ and $(2n-5)$-dimensional Heisenberg algebra $\mathfrak{n}_{2n-5}$.*

Proof. If a $(2,n)$-distribution of maximal class has a $(2n-1)$-dimensional group of symmetries, then all structural functions of its canonical frames have to be constant. Using this fact we get the following

Lemma 6.1. *In addition to (5.5), the only nonzero commutative relations of each of the canonical frames of a $(2,n)$- distribution with a $(2n-1)$-dimensional group of symmetries are*

$$\begin{aligned}&[h,\tilde\varepsilon_i]=\tilde\varepsilon_{i+1},\ [\tilde\varepsilon_i,\tilde\varepsilon_{2m-i+1}]=(-1)^{i+1}\eta,\ [g_1,\tilde\varepsilon_i]=(2m-2i+1)\tilde\varepsilon_i,\\&[g_2,\tilde\varepsilon_i]=(i-1)(2m+1-i)\tilde\varepsilon_{i-1},\ [g_0,\tilde\varepsilon_i]=-\tilde\varepsilon_i,\ [g_0,\eta]=-2\eta.\end{aligned}\tag{6.1}$$

Proof By Lemma 5.3

$$[g_0,\tilde\varepsilon_1]=-\tilde\varepsilon_1+\sum_{i=0}^{2}\alpha_i g_i.\tag{6.2}$$

Let us prove that $\alpha_i=0$ for all $0\le i\le 2$. Indeed, from (5.41) and Lemma 5.3 it follows that

$$\big[g_0,[\tilde\varepsilon_1,\tilde\varepsilon_2]\big]\in\mathcal{W}_1,\quad \big[g_0,[\tilde\varepsilon_1,\tilde\varepsilon_4]\big]\in\mathcal{L}_2\tag{6.3}$$

Using the Jacobi identity and relation (5.31) we get easily that

$$\begin{aligned}&[g_0,[\tilde\varepsilon_1,\tilde\varepsilon_2]=\big((2m-3)\alpha_1-\alpha_0\big)\tilde\varepsilon_2 \bmod \mathcal{W}_1,\\&[g_0,[\tilde\varepsilon_1,\tilde\varepsilon_4]]=\big((2m-7)\alpha_1-\alpha_0\big)\tilde\varepsilon_4+(6m-9)\alpha_2\tilde\varepsilon_3 \bmod \mathcal{L}_2.\end{aligned}$$

Comparing the last relations with (6.3), we have immediately that $\alpha_i=0$ for all $0\le i\le 2$. In other words, $[g_0,\tilde\varepsilon_1]=-\tilde\varepsilon_1$. Let us prove that

$$[g_0,\tilde\varepsilon_i]=-\tilde\varepsilon_i,\quad 1\le i\le 2m.\tag{6.4}$$

The proof is by induction. For $i=1$ relation (6.4) is true. Suppose that it is true for some $i=j$ and prove it for $i=j+1$. By (5.23)

$$[h,\tilde\varepsilon_j]=\tilde\varepsilon_{j+1}+\tau_j h,\tag{6.5}$$

where $\tau_j\equiv 0$ for $m-1\le j\le 2m-1$. Taking the Lie brackets with g_0 from both sides of the last identity and using (5.5), the induction hypothesis, and (6.5) again, one gets easily that

$$[g_0,\tilde\varepsilon_{j+1}]=-[h,\varepsilon_j]=-\tilde\varepsilon_{j+1}-\tau_j h.\tag{6.6}$$

If $m-1 \le j \le 2m-1$ we get (6.4) for $i = j+1$, because $\tau_j = 0$. If $1 \le j \le m-2$, then from Lemma 5.3 it follows that $[g_0, \tilde{\varepsilon}_{j+1}] \in \mathcal{W}_{j+1}$. Therefore (6.6) implies that $\tau_j = 0$. Hence (6.4) holds again. The proof by induction of (6.4) is completed. Note that we have proved at the same time that

$$[h, \tilde{\varepsilon}_i] = \tilde{\varepsilon}_{i+1}, \quad 1 \le i \le 2m-1 \tag{6.7}$$

Further, from (6.4) and the Jacobi identity it follows immediately that

$$[g_0, \eta] = \big[g_0, [\tilde{\varepsilon}_1, \tilde{\varepsilon}_{2m}]\big] = -2\eta. \tag{6.8}$$

The identity (6.4) allows to show that a lot of structural constants of the canonical frame vanish. First, by Lemma 5.3

$$[g_1, \tilde{\varepsilon}_1] = (2m-1)\tilde{\varepsilon}_1 + \sum_{i=0}^{2} \beta_i g_i.$$

Taking Lie brackets with g_0 from both sides of the last identity and comparing the coefficients with the help of (6.4), (5.5), and the Jacobi identity, one obtains immediately that $\beta_i = 0$ for all $0 \le i \le 2$, which together with (6.7) implies in turn that

$$[g_1, \tilde{\varepsilon}_i] = (2m-2i+1)\tilde{\varepsilon}_i, \quad 1 \le i \le 2n-6. \tag{6.9}$$

The last identity yields also that

$$[g_1, \eta] = \big[g_1, [\tilde{\varepsilon}_1, \tilde{\varepsilon}_{2m}]\big] = 0. \tag{6.10}$$

In the same way we get

$$[g_2, \tilde{\varepsilon}_i] = (i-1)(2m-i+1)\varepsilon_i, \quad 1 \le i \le 2n-6. \tag{6.11}$$

Further, suppose that

$$[\tilde{\varepsilon}_i, \tilde{\varepsilon}_j] = \sum_{k=1}^{2n-6} a_{ij}^k \tilde{\varepsilon}_k + \sum_{k=0}^{2} b_{ij}^k g_k + c_{ij} h + d_{ij}\eta.$$

Again taking Lie brackets with the field g_0 from both sides and comparing the coefficients with the help of (5.31), (5.5) and the Jacobi identity we get immediately that $a_{ij}^k = 0$, $b_{ij}^k = 0$ and $c_{ij}^k = 0$. In other words, $[\tilde{\varepsilon}_i, \tilde{\varepsilon}_j] = d_{ij}\eta$. Taking Lie brackets with g_1 from both sides of the last identity and comparing the coefficients with the help of (6.9) and the Jacobi identity one obtains easily that $d_{ij} = 0$ for $i + j \ne 2m+1$. On the other hand, by (5.50) we have $d_{ij} = (-1)^{i+1}$ for $i + j = 2m+1$. In other words,

$$[\tilde{\varepsilon}_i, \tilde{\varepsilon}_j] = \begin{cases} (-1)^{i+1}\eta & i+j = 2m+1 \\ 0 & i+j \ne 2m+1 \end{cases}. \tag{6.12}$$

This together with the Jacobi identity immediately implies that $[\tilde{\varepsilon}_i, \eta] = \big[\tilde{\varepsilon}_i, [\tilde{\varepsilon}_1, \tilde{\varepsilon}_{2m}]\big] = 0$ for $1 < i < 2m$. To prove that $[\tilde{\varepsilon}_i, \eta] = 0$ also for $i = 1$ and $i = 2m$ we take Lie brackets with g_0 and compare the coefficients with the help of (6.4),(6.8), and the Jacobi identity. Finally, to prove that $[h, \tilde{\varepsilon}_{2m}] = 0$ we take Lie brackets with g_1 and compare the coefficients with the help of (6.9) and the Jacobi identity. This completes the prove of the lemma. □

The previous lemma and Theorem 2 imply the uniqueness, up to the equivalence, of the germ of $(2, n)$-distribution of maximal class with $(2n-1)$-dimensional group of symmetries. Besides, from these relations it follows that the algebra of infinitesimal symmetries of such distribution is isomorphic to the semi-direct sum of $\mathfrak{gl}(2, \mathbb{R})$ ($\sim \operatorname{span}_{\mathbb{R}} \{g_0, g_1, g_2, h\}$) and the Heisenberg group $\mathfrak{n}_{2m+1}$ ($\sim \operatorname{span}_{\mathbb{R}}\{\tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_{2m}, \eta\}$).

To complete the proof of the theorem it remains to show that for the $(2,n)$-distribution D_0, associated with the underdetermined ODE $z'(x)=\frac{1}{2}\big(y^{(n-3)}(x)\big)^2$, the only nonzero commutative relations for its canonical frames are (5.5) and (6.1). Let $p_i=y^{(i)}$, as in Introduction. Then the distribution D_0 is given in $M=\mathbb{R}^n$ with coordinates $(x,p_0,\dots,p_{n-3},z)$ by the intersection of the annihilators of the forms

$$\begin{aligned} &dp_i-p_{i+1}dx,\ 0\le i\le n-4,\\ &dz-\frac{1}{2}p_{n-3}^2\,dx \end{aligned} \tag{6.13}$$

Set, as before, $m=n-3$. From (6.13) the following two vector fields span the distribution D_0:

$$X_1=\partial_{p_m},\quad X_2=\partial_x+\sum_{i=0}^{m-1}p_{i+1}\partial_{p_i}+\frac{1}{2}p_m^2\partial_z. \tag{6.14}$$

Let vector fields X_3, X_4, X_5 are as in (2.1). Then

$$X_3=\partial_{p_{m-1}}+p_m\partial_z,\ X_4=\partial_z,\ X_5=-\partial_{p_{m-2}}. \tag{6.15}$$

Denote

$$X_i=(\mathrm{ad}X_2)^{i-5}X_5=(-1)^i\partial_{p_{m+3-i}},\quad 6\le i\le m+3 \tag{6.16}$$

Let $u_i:T*M\mapsto\mathbb{R}$, $1\le i\le m+3$, be the corresponding quasi-impulses, defined by (2.2). Then the tuple $(x,p_0,\dots,p_m,z,u_1,\dots,u_{m+3})$ defines the coordinates on T^*M. It is clear that in this coordinates the submanifold $(D^2)^\perp$ is given by equations $(D_0^2)^\perp=\{u_1=u_2=u_3=0\}$.

Denote by $\overline{X}_i$ the vector field on $(D_0^2)^\perp$, which is the lift of the vector field X_i (i.e., $\pi_*\overline{X}_i=X_i$) such that $du_j(\overline{X}_i)=0$ for all $1\le j\le n$. In addition to (2.1) the only nonzero commutators generated by the vector fields $\{X_i\}_{i=1}^5$ are $(\mathrm{ad}X_2)^jX_5$, $1\le m-2$. This together with (2.4) implies that the characteristic line distribution $\mathcal{C}$ satisfies

$$\mathcal{C}=\Big\{\mathbb{R}\big(u_4\overline{X}_2-u_5\overline{X}_1+u_4\sum_{i=5}^{m+2}u_{i+1}\partial_{u_i}\big)\Big\}.$$

From this it is not difficult to show that

$$\mathcal{R}_{D_0}=\{\lambda\in(D_0^2)^\perp:u_4(\lambda)\ne0\} \tag{6.17}$$

(if $u_4(\lambda)=0$, then $\mathcal{J}^{(3)}(\lambda)=\mathcal{J}^{(2)}(\lambda)$). Define the following vector field H on $\mathcal{R}_{D_0}$, which generates the characteristic line distribution $\mathcal{C}$:

$$H=\overline{X}_2-\frac{u_5}{u_4}\overline{X}_1+\sum_{i=5}^{m+2}u_{i+1}\partial_{u_i} \tag{6.18}$$

Assume that φ_λ is the parameterization of the characteristic curve γ, passing through λ, such that

$$\varphi_\lambda(e^{tH}\lambda)=t. \tag{6.19}$$

Then by direct computations one can show that the following vector

$$\varepsilon_{\varphi_\lambda}(\lambda)=|u_4(\lambda)|^{1/2}\partial_{u_{m+3}}(\lambda)$$

satisfies

$$\varepsilon_{\varphi_\lambda}(\lambda)\in\mathrm{Aff}_{\varphi_\lambda}(\lambda)\cup\big(-\mathrm{Aff}_{\varphi_\lambda}(\lambda)\big) \tag{6.20}$$

Denote by ε_H the vector field, satisfying $\varepsilon_H(\lambda) = \varepsilon_{\varphi_\lambda}(\lambda)$ for all $\lambda \in \mathcal{R}_{D_0}$. By direct calculation it is easy to show that

$$(6.21)\qquad \begin{aligned} &(\mathrm{ad}H)^i\varepsilon_H = (-1)^i|u_4|^{1/2}\partial_{u_{m+3-i}}, \quad 0 \le i \le m-2, \\ &(\mathrm{ad}H)^{m-1}\varepsilon_H = (-1)^{m-2}|u_4|^{-1/2}\overline{X}_1, \quad (\mathrm{ad}H)^m\varepsilon_H = (-1)^{m-1}|u_4|^{-1/2}\overline{X}_3, \\ &(\mathrm{ad}H)^{m+i}\varepsilon_H = (-1)^{m-1}|u_4|^{-1/2}\big(\overline{X}_{m+i} - \frac{u_{4+i}}{u_4}\overline{X}_4\big), \quad 1 \le i \le m-1, \end{aligned}$$

and finally

$$(6.22)\qquad (\mathrm{ad}H)^{2m}\varepsilon_H = 0.$$

Then by (4.7) the parameterizations φ_λ defined by (6.19) are projective[1]; namely, $\varphi_\lambda(\cdot) \in \mathfrak{P}_\lambda$. Therefore if some parameterization $\bar{\varphi}$ belongs to $\mathfrak{P}_\lambda$, then $\bar{\varphi} = \frac{a\varphi_\lambda}{b\varphi_\lambda+1}$. Let us introduce the coordinates on Σ_{D_0} in the following way:

$$\left(\lambda, \frac{a\varphi_\lambda}{b\varphi_\lambda + 1}\right) \mapsto (x, p_0, \ldots, p_m, x, z, u_4, \ldots, u_n, a, b)$$

Then from (5.1), (5.3), and (5.4 it is not difficult to show that

$$(6.23)\qquad \begin{aligned} &g_1 = 2a\partial_a, \quad g_2 = -a\partial_b, \quad g_0 = 2\sum_{i=4}^n u_i\partial_{u_i} \\ &h = \overline{H} - 2b\partial_a - \frac{b^2}{a}\partial_b, \end{aligned}$$

where the vector field $\overline{H}$ is the lift of the vector field H on Σ_{D_0} (i.e., $\Pi_*\overline{H} = H$) such that $da(H) = db(H) = 0$. From (5.32) and (6.20) it follows that $\varepsilon_{\bar{\upsilon}f} = a^{m-1/2}\varepsilon_H \bmod\{\mathbb{R}e\}$, where $\bar{\varphi} = \frac{a\varphi_\lambda}{b\varphi_\lambda+1}$ and e is the Euler field on T^*M. Therefore the vector field

$$(6.24)\qquad \varepsilon_1 = a^{m-1/2}|u_4|^{1/2}\partial_{u_{m+3}}$$

satisfies (5.13). Consider the frame $(h, \{g_i\}_{i=0}^2, \{\varepsilon_i\}_{i=1}^{2m}, \eta)$ associated with the vector field ε_1. Using commutative relations (6.21) and formulas (6.23), (6.24), it is easy to show that this frame is canonical and the only its nonzero commutative relations are (5.5) and (6.1). This completes the proof of the theorem. □

Remark 6.1. Note that for the most symmetric case the frames, indicated in Remark 5.2 coincide (up to the reflection) with the canonical frames, introduced in the proof of Theorem 1. □

7. Discussion

7.1. **Distributions of non-maximal class.** As was already mentioned before, a rank 2 distribution D has the smallest possible class 1 at a point q iff dim $D^3(q) = 4$ (see [15, Remark 3.4]). Suppose that D satisfies dim $D^3(q) = 4$ on some open set M^o. It is easy to show that the distribution D^2 has a one-dimensional characteristic distribution C. Then (locally) we can consider the quotient M' of the manifold M^o by the corresponding one-dimensional foliation together with a new rank 2 distribution D' obtained by the factorization of D^2.

[1] Actually from (6.22) it follows that the curves in a projective space associated with all abnormal extremals of the distribution D_0 are so-called normal rational curves. The interesting question is whether the distribution D_0 is a unique, up to the equivalence, $(2, n)$-distribution, having this property.

In fact, D can be uniquely reconstructed from D'. Let $P(D')$ be a submanifold in $P(TM')$ consisting of all lines lying in D'. Similarly to the canonical contact system on $P(TM')$, we can define lifts of integral curves of D' to $P(D')$ and a canonical rank 2 distribution on $P(D')$ generated by tangent vectors to these lifts. It can be proved that this contact system on $P(D')$ is locally equivalent to D.

Iterating this procedure, we end up either at a nonholonomic rank 2 distribution on a three-dimensional manifold or at a distribution $\widetilde{D}$, satisfying $\dim \widetilde{D}^3 = 5$. In the former case the original distribution D is locally equivalent to the Goursat distribution and has an infinite-dimensional symmetry algebra. In other words, the case of non-Goursat distributions of constant class 1 can be reduced to the case of distributions of class greater than 1.

This leaves the following question open: *Do there exist completely nonholonomic rank* 2 *distributions of constant class* $2 \le m \le n-4$*?* We know only that the answer is negative for $m = 2$ ($n > 5$), which means that any such example, if it exists, should live on at least 7-dimensional manifold.

7.2. **Connection with Tanaka theory.** After the symplectification procedure described above, the results of this paper can be interpreted in terms of Tanaka–Morimoto theory of structures on filtered manifolds [10, 11]. The original distribution D (even of maximal class) has, in general, a non-constant symbol, which makes this theory very difficult to apply to the filtered manifold defined by the distribution D itself. However, given rank 2 distribution D of maximal class there is a natural rank 2 distribution on the manifold $P(\mathcal{R}_D)$ obtained from $\mathcal{R}_D$ via the factorization by the trajectories of the Euler vector field (or, in other words, by the projectivization of the fibers of $\mathcal{R}_D$). It is generated by the projection of the sum $V_{n-4} \oplus \mathcal{C}$ w.r.t. this factorization. It is possible to show that this distribution has already a fixed symbol isomorphic to the Lie algebra generated by the vector fields $\{h, \tilde{\varepsilon}_1, \dots, \tilde{\varepsilon}_{2m}, \eta\}$ from the proof of the main theorem (see equation (6.1)).

Moreover, there is a natural decomposition of this distribution into the sum of two line distributions equal to the projections of V_{n-4} and $\mathcal{C}$. This decomposition can be interpreted as a G-structure on a filtered manifold in terms of Tanaka theory and is called a *pseudo-product* structure [12]. The prolongation of this structure (in terms of filtered manifolds) is of finite type and is isomorphic to the maximal symmetry algebra from the main theorem.

We shall dwell into the details of this approach in the forthcoming paper.

References

[1] A.A. Agrachev, Feedback-invariant optimal control theory - II. Jacobi Curves for Singular Extremals, J. Dynamical and Control Systems, **4**(1998), No. 4 , 583-604.

[2] A.A. Agrachev, Yu.L. Sachkov Control Theory from the Geometric Viewpoint Series: Encyclopaedia of Mathematical Sciences , Vol. 87, 412 pages.

[3] A.A. Agrachev, I. Zelenko, Principal Invariants of Jacobi Curves, Lecture Notes in Control and Information Sciences 258, Springer, 2001, 9-21.

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

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

[6] R.L. Bryant, S.S. Chern, R. B. Gardner, H.L. Goldschmidt, P. A. Griffiths, Exterior Differential Systems, Mathematical Sciences Research Institute Publications, vol. 18, Springer-Verlag.

[7] E. Cartan, *Les systemes de Pfaff a cinque variables et les equations aux derivees partielles du second ordre*, Oeuvres completes, Partie II, vol.2, Paris, Gautier-Villars, 1953, 927-1010.

[8] B. Doubrov, I. Zelenko, *A canonical frame for nonholonomic rank two distributions of maximal class*, C.R. Acad. Sci. Paris, Ser. I, Vol. 342, Issue 8 (15 April 2006), 589-594.

[9] B. Doubrov, I. Zelenko, *Geometry of rank 3 distributions*, in preparation.

[10] T. Morimoto, *Geometric structures on filtered manifolds*, Hokkaido Math. J.,**22**(1993), pp. 263-347.
[11] N. Tanaka, *On the equivalence problems associated with simple graded Lie algebras*, Hokkaido Math. J.,**6**(1979), pp 23-84.
[12] N. Tanaka, *On affine symmetric spaces and the automorphism groups of product manifolds*, Hokkaido Math. J., **14**(1985), pp 277-351.
[13] A. Vershik, V. Gershkovich, *Determination of the functional dimension of the orbit space of generic distributions*, Mat. Zametki 44, 596-603 (in Russian); English transl: Math. Notes 44, 806-810 (1988).
[14] E.J. Wilczynski, *Projective Differential Geometry of Curves and Ruled Surfaces*, Teubner, Leibzig, 1905.
[15] I. Zelenko, *Variational Approach to Differential Invariants of Rank 2 Vector Distributions*, Differential Geometry and Its Applications, Vol. 24, Issue 3 (May 2006), 235-259
[16] I. Zelenko, *Complete systems of invariants for rank 1 curves in Lagrange Grassmannians*, Differential Geom. Application, Proc. Conf. Prague, 2005, pp 365-379, Charles University, Prague
[17] I. Zelenko, *Nonregular abnormal extremals of 2-distribution: existence, second variation, and rigidity*, J. Dynamical and Control Systems, **5**(1999), No. 3, 347-383.
[18] M. Zhitomirskii, *Normal forms of germs of smooth distributions*, Mat. Zametki 49 (1991), no. 2, 36–44, 158 (in Russian); English translation in Math. Notes 49 (1991), no. 1-2, 139–144.

BELARUSSIAN STATE UNIVERSITY, NEZAVISIMOSTI AVE. 4, MINSK 220050, BELARUS; E-MAIL: DOUBROV@ISLC.ORG

S.I.S.S.A., VIA BEIRUT 2-4, 34014, TRIESTE, ITALY; E-MAIL: ZELENKO@SISSA.IT