Differential Geometry and its Applications 29 (2011) 255–270

Contents lists available at ScienceDirect

Differential Geometry and its Applications

www.elsevier.com/locate/difgeo

Equivalence of variational problems of higher order

Boris Doubrov [a,*], Igor Zelenko [b]

[a] Department of Applied Mathematics and Computer Science, Belarussian State University, Nezavisimosti Ave. 4, Minsk 220030, Belarus
[b] Department of Mathematics, Texas A&M University, College Station, TX 77843-3368, USA

ARTICLE INFO

Article history:
Received 14 April 2010
Received in revised form 24 November 2010
Available online 17 December 2010
Communicated by J. Slovák

MSC:
58A30
53A55
34C14

Keywords:
Divergent equivalence of Lagrangians
Vector distributions
Variational ODEs
Curves in projective spaces
Wilczynski invariants
Legendre transform

ABSTRACT

We show that for $n \geqslant 3$ the following equivalence problems are essentially the same: the equivalence problem for Lagrangians of order n with one dependent and one independent variable considered up to a contact transformation, a multiplication by a nonzero constant, and modulo divergence; the equivalence problem for the special class of rank 2 distributions associated with underdetermined ODEs $z' = f(x, y, y', \dots, y^{(n)})$; the equivalence problem for variational ODEs of order $2n$. This leads to new results such as the fundamental system of invariants for all these problems and the explicit description of the maximally symmetric models. The central role in all three equivalence problems is played by the geometry of self-dual curves in the projective space of odd dimension up to projective transformations via the linearization procedure (along the solutions of ODE or abnormal extremals of distributions). More precisely, we show that an object from one of the three equivalence problems is maximally symmetric if and only if all curves in projective spaces obtained by the linearization procedure are rational normal curves.

1. Introduction: three equivalence problems

The main goal of this paper is to establish a tight relationship between the following three local equivalence problems in differential geometry:

(1) equivalence of variational problems of order $\geqslant 3$ with one dependent and one independent variable considered up to a contact transformation, a multiplication by a constant, and modulo divergence;
(2) equivalence of *variational ODEs* of order $\geqslant 6$ up to contact transformations. By a variational ODE (called also variational with multiplier) we mean an ODE which is contact equivalent to an Euler–Lagrange equation for some Lagrangian;
(3) equivalence of rank 2 distributions and in particular rank 2 distributions associated with underdetermined ODEs $z' = f(x, y, y', \dots, y^{(n)})$, $n \geqslant 3$.

In particular, we shall show that equivalence problems (1), (2), and the equivalence problem for the particular class of rank 2 distributions mentioned in item (3) are essentially the same. In particular, there is a one-to-one correspondence between equivalence classes of objects in all these problems.

This one-to-one correspondence (up to above equivalence relation) between Lagrangians of order $n \geqslant 2$ and their Euler–Lagrange equations was already established earlier in works of M. Fels [15] for $n = 2$ and M. Juráš [19] for $n \geqslant 3$. It is based

* Corresponding author.
E-mail addresses: doubrov@islc.org (B. Doubrov), zelenko@math.tamu.edu (I. Zelenko).

0926-2245/$ – see front matter
doi:10.1016/j.difgeo.2010.12.004

on the characterization of variational ODEs in terms of the variational bicomplex given in [3]. We extend this correspondence to underdetermined ODEs and the corresponding rank 2 vector distributions in the case $n \geqslant 3$. This allows us to apply the results of our previous works [13,14], where more general rank 2 distributions are treated, for the description of the unique maximally symmetric Lagrangian up to the considered equivalence relation (see also the discussions on various equivalence relations for variational problems at the end of Section 1.1). Note that the one-to-one correspondence between equivalence problems (1) and (2) does not hold for $n = 2$. For example, the Lagrangian $(y'')^{1/3}\,dx$ is not equivalent to the most symmetric one $(y'')^2\,dx$, but the corresponding underdetermined ODEs and rank 2 distributions are equivalent and have 14-dimensional Lie algebra G_2 as their symmetry.

The common feature of all three problems above is that they admit linearization, which reduces them in essence to the problem of equivalence of self-dual curves in odd-dimensional projective spaces up to projective transformations. The invariants of these curves in projective spaces, the Wilczynski invariants [30], produce the invariants of the original problem. The latter are called the generalized Wilczynski invariants.

In this work we exploit an alternative (a Hamiltonian) point of view on the variational problems, which comes from the Pontryagin Maximum Principle in Optimal Control. This point of view provides us with the Hamiltonian form of the Euler–Lagrange equation and allows to construct a (generalized) Legendre transform that takes extremals of the Lagrangian (or, in other words, the solutions of the corresponding Euler–Lagrange equation) to the abnormal extremals of the corresponding optimal control problem. This immediately shows that the solution space of any variational ODE carries a natural symplectic structure ω. We show that the conformal class of this symplectic structure (i.e., all 2-forms $f\omega$ for non-vanishing functions f) can be recovered only from the self-duality of all linearizations of the given ODE along its solutions. This, in its turn, can be reformulated by vanishing of the generalized Wilczynski invariants (see [30]) of odd order. It is easy to see that for $n \geqslant 2$ there is at most one (up to constant) closed 2-form in any given conformal class of non-degenerate 2-forms on a smooth manifold. This gives a 'naive' proof that any variational ODE of order $\geqslant 4$ admits at most one Lagrangian up to a constant and divergence terms.

Another question we try to answer in this paper is whether invariants of the above three equivalence problems derived from the Wilczynski invariants of self-dual projective curves, provide the complete system of fundamental invariants. It has been known from [11,12] that the answer to the similar question for arbitrary non-linear ODEs is negative. Namely, there exist non-trivial ODEs (i.e., equations, not equivalent to $y^{(n)} = 0$ via contact transformations) such that all their linearizations are trivial and, thus, all their Wilczynski invariants vanish. However, surprisingly, for variational ODEs the answer is positive: a variational ODE is trivializable if and only if all its Wilczynski invariants vanish identically.

We show that for variational ODEs the generalized Wilczynski invariants of even order (Wilczynski invariants of odd order automatically vanish due to the self-duality of the linearization) form a fundamental set of contact invariants in the following sense. Vanishing of this fundamental set of invariants implies that any other relative differential invariant of the variational equation vanishes identically and the equation is contact equivalent to the trivial one. We note that for $n = 3$ and 4 these invariants do not generate the complete differential algebra of invariants by only differentiation and algebraic operations. In particular, for $n = 3$ (or 6-th order variational ODEs $y^{(6)} = F(x, y, y', \dots, y^{(5)})$) the differential invariant $I = F_{45}$ satisfies a non-trivial cubic equation, whose coefficients are certain derivatives of the generalized Wilczynski invariant of order 4.

To summarize vanishing of generalized Wilczynski invariants of even order gives an explicit characterization of the most symmetric models in all three equivalence problems provided that $n \geqslant 3$:

(1) all variational ODEs with vanishing generalized Wilczynski invariants are contact equivalent to the trivial equation $y^{(2n)} = 0$;
(2) all Lagrangians with vanishing generalized Wilczynski invariants are equivalent to $(y^{(n)})^2\,dx$ modulo constant multiplier, contact transformations and divergence terms;
(3) all underdetermined ODEs with vanishing generalized Wilczynski invariants equivalent to $z' = (y^{(n)})^2$.

Let us we briefly outline each of equivalence problems (1)–(3).

1.1. *Equivalence of variational problems*

This paper deals with variational problems in one dependent and one independent variable of arbitrary order. A variational problem is defined by a Lagrangian $L = f(x, y, y', \dots, y^{(n)})\,dx$ or the corresponding functional $\int f(x, y, y', \dots, y^{(n)})\,dx$.

Let us recall basic definitions from the geometry of variational problems. Let $J^\infty = J^\infty(\mathbb{R}, \mathbb{R})$ be an infinite jets of smooth functions $y(x)$. We shall use the standard coordinate system $(x, y = y_0, y_1, y_2, \dots)$ on J^∞. Denote by $\theta_i = dy_i - y_{i+1}\,dx$, $i \geqslant 0$, the basis of so-called *contact forms* on J^∞. The set of all exterior forms $\Lambda(J^\infty)$ is naturally turned into the bi-graded algebra with:

$$\Lambda^p\big(J^\infty\big) = \Lambda^{0,p} \oplus \Lambda^{1,p-1}, \quad \text{where } \Lambda^{0,p} = \langle \theta_{i_1} \wedge \dots \wedge \theta_{i_p} \rangle,\ p \geqslant 0;$$
$$\Lambda^{1,p-1} = \langle \theta_{i_1} \wedge \dots \wedge \theta_{i_{p-1}} \wedge dx \rangle, \quad p \geqslant 1.$$

The exterior derivative $d : \Lambda^p(J^\infty) \to \Lambda^{p+1}(J^\infty)$ naturally splits into the sum $d = d_H + d_V$, where for $\omega \in \Lambda^p(J^\infty)$ we have $d_H(\omega) \in \Lambda^{1,p}$ and $d_V(\omega) \in \Lambda^{0,p+1}$.

We consider variational problems $\int f(x, y, y', \dots, y^{(n)})\,dx$ of arbitrary order n up to the divergence equivalence and constant multiplier. Namely, we say that two Lagrangians $L_1 = f_1(x, y, y', \dots, y^{(n)})\,dx$ and $L_2 = f_2(x, y, y', \dots, y^{(n)})\,dx$ are equivalent if there exists a contact transformation $\phi : J^1(\mathbb{R},\mathbb{R}) \to J^1(\mathbb{R},\mathbb{R})$ with the prolongation $\Phi : J^\infty(\mathbb{R},\mathbb{R}) \to J^\infty(\mathbb{R},\mathbb{R})$, such that

$$\Phi^*(L_2) = \alpha L_1 + d_H(\mu) \mod \langle \theta_i \mid i \geqslant 0\rangle \tag{1.1}$$

for some constant nonzero $\alpha \in \mathbb{R}$ and function μ on $J^\infty(\mathbb{R},\mathbb{R})$. We shall always assume that all our Lagrangians are non-degenerate, i.e., they are non-linear in the highest derivative. It follows from [19] that two Lagrangians are equivalent under the above equivalence relation if and only if their Euler–Lagrange equations are contact equivalent.

The variational equivalence problem was treated in a number of papers using both naive approach and Cartan's equivalence method [4,8,16,20–23]. See also [26] for the symmetry classification of higher order Lagrangians.

We note that usually slightly different equivalence notion of divergence equivalence is considered, where the constant α above is equal to 1 identically. The upper bound for the variational symmetry algebra in case of n-order Lagrangian was proved to be equal to $2n + 3$ for $n \geqslant 2$ in the work of Gonzalez-Lopez [16]. Note that this upper bound is achieved in case of a family of non-equivalent Lagrangians, which is different from most of the classical local equivalence problems in differential geometry. We show in this paper, that adding this constant α in the definition (1.1) of the divergence equivalence changes this patten. In this case we get a slightly higher upper bound equal to $2n + 5$ with a unique maximally symmetric Lagrangian equivalent to $L = (y^{(n)})^2\,dx$.

1.2. Equivalence of rank 2 vector distributions

By a rank 2 vector distribution on a smooth manifold M we understand a two-dimensional subbundle D of the tangent vector bundle TM. We define its (small) derived flag $\{D^i\}$ as follows:

$$D^1 = D, \qquad D^{i+1} = D^i + [D, D^i], \quad i \geqslant 1,$$

and assume that the distribution D is regular in a sense that all D^i are smooth subbundles of the tangent bundle TM. We shall also assume in the sequel that the distribution D is *completely non-holonomic*, i.e. $D^\mu = TM$ for some sufficiently large μ.

We say that two such distributions D and D' on manifolds M and M' are (locally) equivalent if there exists a (local) diffeomorphism $\phi : M \to M'$, such that $\phi_*(D) = D'$.

Equivalence problem for non-holonomic distributions is an old problem, which goes back to the end of 19th century and was studied by various mathematicians including Lie, Goursat, Darboux, Engel, Elie Cartan and others. Except for several cases such as rank 2 distributions on 3- and 4-dimensional manifolds, generic rank 2 distributions have functional, and, thus, non-trivial differential invariants. In his classical paper [5] Elie Cartan associates a $(2,5)$-distribution to a system of partial differential equations of second order and constructs a canonical coframe for non-degenerate distributions of this type. This is a first example of an explicit solution for the equivalence problem of vector distributions with non-trivial functional invariants. Remarkably, the most symmetric $(2,5)$-distributions form one equivalence class and have an exceptional Lie algebra G_2 as their symmetry algebra.

The obvious (but very rough in the most cases) discrete invariant of a distribution D at a point q is the so-called *small growth vector* (s.g.v.) at q. This vector is defined as $\{\dim D^j(q)\}_{j\in\mathbb{N}}$. Furthermore, at each point $q \in M$, where $\dim D^j$ is locally constant for any j, we can consider the graded space $\mathfrak{m}_q = \sum_{i\geqslant 1} D^{j+1}(q)/D^j(q)$. It can be naturally equipped with a structure of a graded nilpotent Lie algebra and it is called a *symbol* of the distribution D at a point q. The notion of symbol is extensively used in works of N. Tanaka and his school [24,28,29,31] who systematized and generalized the Cartan equivalence method.

However, all constructions of Tanaka theory strongly depend on the algebraic structure of the symbol and they were carried out under the very restrictive assumption that symbol algebras are isomorphic at different points. An alternative approach for studying rank 2 distributions was presented by the authors in [13,14]. It is based on the ideas of the geometric control theory and uses a symplectification of the problem by lifting the distribution D to the cotangent bundle. This provides an effective way to construct a canonical coframe and, thus, solve equivalence problem for rank 2 distributions of so-called *maximal class* (this notion is defined in Section 4.4).

Rank 2 distributions of a special type are naturally associated with Lagrangians. Namely, to a variational problem with a Lagrangian $f(x, y, y', \dots, y^{(n)})\,dx$ one can assign the following (affine) control system:

$$\begin{aligned} &\dot{x}(t) = 1, \qquad \dot{y}_i(t) = y_{i+1}(t), \quad 0 \leqslant i \leqslant n-1, \\ &\dot{y}_n(t) = u(t), \qquad \dot{z}(t) = f\big(x(t), y_0(t), y_1(t), \dots, y_n(t)\big), \end{aligned} \tag{1.2}$$

on $J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ with coordinates $(x, y_0, \dots, y_n, z)$, where $u(\cdot)$ is a control function belonging say to the space L_∞. To any point $q_0 \in J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ and a control function $u(\cdot)$ consider the trajectory of the system (1.2) started at q_0. Such trajectory is called an *admissible trajectory of control system* (1.2) and its velocity at q_0 is called *an admissible velocity of control system* (1.2) *at* q_0. Taking the linear span of all admissible velocities of (1.2) we get the rank 2 distribution on $J^n(\mathbb{R},\mathbb{R})\times\mathbb{R}$ generated by the following two vector fields:

$$X_1=\frac{\partial}{\partial x}+y_1\frac{\partial}{\partial y_0}+\cdots+y_n\frac{\partial}{\partial y_{n-1}}+f\big(x,y,y',\dots,y^{(n)}\big)\frac{\partial}{\partial z},\qquad X_2=\frac{\partial}{\partial y_n}. \tag{1.3}$$

We say that this rank 2 distribution is *associated with the Lagrangian* $f(x,y,y',\dots,y^{(n)})\,dx$ *or with the underdetermined differential equation*

$$z'=f\big(x,y,y',\dots,y^{(n)}\big) \tag{1.4}$$

on two unknown functions $y(x)$ and $z(x)$. Such underdetermined ODEs and the related geometric structures have been extensively studied in [2,9,17,18,25].

It is easy to see that if $\frac{\partial^2 f}{\partial (y^{(n)})^2}\neq 0$ (i.e. the Lagrangian satisfies the Legendre condition), then $\dim D^2=3$, $\dim D^i=i+2$ for $i=3,\dots,n+1$. The case, when f is linear with respect to $y^{(n)}$ is special, since in this case we can reduce the corresponding underdetermined equation to the equation of the same type, but of lower order. Indeed, suppose that $f=f_0+f_1y^{(n)}$, where f_0 and f_1 depend only on x and the derivatives of the function $y(x)$ up to order $n-1$. Then the substitution $z=\tilde z+g(x,y,y',\dots,y^{(n-1)})$ with the function g satisfying $\frac{\partial g}{\partial y^{(n-1)}}=f_1$ reduces Eq. (1.4) to the underdetermined ODE on functions $y(x),\tilde z(x)$, where the right-hand side does not depend on $y^{(n)}$.

In the following we shall call the Lagrangian $f(x,y,y',\dots,y^{(n)})\,dx$ and the corresponding equation (1.4) *non-degenerate* if $\frac{\partial^2 f}{\partial (y^{(n)})^2}\neq 0$. In the sequel we shall always assume that all Lagrangians and the underdetermined ODEs are non-degenerate. Note that all corresponding rank 2 distributions are of maximal class (see Proposition 4.2 below).

Finally we cite the main result of [13,14] that will be needed in the sequel:

Theorem 1.1. *For any* $(2,n+3)$*-distribution,* $n>2$*, of maximal class there exists a canonical frame on a* $(2n+5)$*-dimensional bundle over* M*. The group of symmetries of such distribution is at most* $(2n+5)$*-dimensional. Any* $(2,n+3)$*-distribution of maximal class with* $(2n+5)$*-dimensional group of symmetries is locally equivalent to the distribution, associated with the Lagrangian* $(y^{(n)}(x))^2$*. The algebra of infinitesimal symmetries of this distribution is isomorphic to a semidirect sum of* $\mathfrak{gl}(2,\mathbb{R})$ *and* $(2n+1)$*-dimensional Heisenberg algebra* $\mathfrak{n}_{2n+1}$.

1.3. Equivalence of ordinary differential equations

We shall also consider the equivalence problem of scalar ordinary differential equations of the form

$$y^{(N+1)}=F\big(x,y,y',\dots,y^{(N)}\big). \tag{1.5}$$

Each such equation can be considered as a hypersurface $\mathcal{E}$ in the jet space $J^{N+1}(\mathbb{R},\mathbb{R})$. We shall always assume that our equations are solved with respect to the highest derivative, so that the restriction of the natural projection $\pi_{N+1,N}:J^{N+1}(\mathbb{R},\mathbb{R})\to J^N(\mathbb{R},\mathbb{R})$ to the hypersurface $\mathcal{E}$ is a diffeomorphism.

Two such equations $\mathcal{E}$ and $\mathcal{E}'$ are said to be *contact equivalent*, if there exists a contact transformation $\Phi: J^1(\mathbb{R},\mathbb{R})\to J^1(\mathbb{R},\mathbb{R})$ with a prolongation Φ^{N+1} to $J^{N+1}(\mathbb{R},\mathbb{R})$ such that $\Phi^{N+1}(\mathcal{E})=\mathcal{E}'$.

The equivalence problem of ordinary differential equations under contact and point transformations is yet another classical subject going back to the works of Lie, Tresse, Elie Cartan [6], Chern [7], M. Fels [15] and others. The existence of the canonical Cartan connection associated with any system of ODEs was proved in [10] based on the Tanaka theory of geometric structures on filtered manifolds [24,28,29].

Explicit formulas for the so-called fundamental contact invariants of a single ODE of arbitrary order were computed by B. Doubrov [11]. In particular, a part of these invariants comes from the linearization of the given ODE. In fact, they coincide with classical Wilczynski invariants of linear differential equations formally applied to the linearization of a non-linear ODE. See [12] for more details.

In this paper we are mainly interested in a special class of ordinary differential equations, consisting of equation which are contact equivalent to Euler–Lagrange equations of variational problems. Such equations are usually called variational (with multiplier) and have been studied in many papers [3,15,19].

From the general result of Anderson and Thompson [3, Theorem 2.6] it is known that a scalar ordinary differential equation of order $2n$ is variational, if and only if there exists a closed 2-form:

$$\omega=\sum_{i=0}^{n-1}\sum_{j=i+1}^{2n-i-1}A_{i,j}\theta_i\wedge\theta_j, \tag{1.6}$$

where $A_{n-1,n}\neq 0$ and $\theta_i=dy_i-y_{i+1}\,dx$ are contact forms on the jet space $J^{2n}(\mathbb{R},\mathbb{R})$ restricted to the equation $\mathcal{E}=\{y_{2n}=F(x,y_0,\dots,y_{2n-1})\}$.

1.4. Equivalence of curves in projective spaces

Surprisingly, the central role in all three equivalence problems is played by the geometry of self-dual curves in the projective space of odd dimension up to projective transformations.

Let $\gamma \subset P^N$ be an arbitrary curve in the projective space. We shall always assume that γ is strongly regular, i.e. its flag of osculating spaces does not have any singularities. In particular, γ itself does not lie in any proper linear subspace of the projective space. We shall not assume any distinguished parameter on γ, though there is always a distinguished family of so-called projective parameters on γ.

Let t be an arbitrary parameter on γ and let $e_0(t)$ be such a curve in $\mathbb{R}^{N+1}$ that $\gamma(t) = \mathbb{R}e_0(t)$. Define $e_i(t) = e_0^{(i)}(t)$. Then *i-th osculating space* $\gamma^{(i)}(t)$ of the curve γ at a point $\gamma(t)$ is defined as:

$$\gamma^{(i)}(t) = \langle e_0(t), \dots, e_i(t)\rangle, \quad i = 0, \dots, N. \tag{1.7}$$

It is easy to see that it does not depend on the choice of the parameter t and the curve $e_0(t)$. The $(N-1)$-st osculating spaces $\gamma^{(N-1)}$ define the curve in the dual projective space $P^{N,*}$, which is called a dual curve and is denoted by γ^*. We shall call a curve γ *self-dual*, if there exists a projective mapping $P^N \to P^{N,*}$ that maps γ to γ^* so that any point $x \in \gamma$ is mapped to the point in $P^{N,*}$ annihilating the $(N-1)$-st osculating space $\gamma^{(N-1)}$ to γ at x. We summarize the properties of self-dual curves in the following proposition. For the proofs we refer to the classical book of Wilczynski [30]:

Proposition 1.1.

(1) *If the mapping $P^N \to P^{N,*}$ sending γ to γ^* exists, then it is unique, up to a constant nonzero factor. It defines, a unique, up to a constant nonzero factor, non-degenerate bilinear form β on the vector space $\mathbb{R}^{N+1}$. Moreover, this form is necessarily skew-symmetric if N is odd and symmetric, if N is even.*

(2) *In case of odd N the curve γ is self-dual if and only if there exists a non-degenerate skew-symmetric (i.e. symplectic) form on $\mathbb{R}^{N+1}$ such that all osculating spaces $\gamma^{(N+1)/2}(t)$ are Lagrangian with respect to this form.*

The invariants of projective curves (up to projective transformations) were also described by Wilczynski [30]. The algebra of all invariants admits a basis of so-called fundamental invariants $W_3, \dots, W_{N+1}$ of order $3, \dots, N+1$ respectively. They can be constructed as follows. As above, let $e_0(t)$ be a curve in $\mathbb{R}^{N+1}$ such that $\gamma = \mathbb{R}e_0(t)$. If the curve γ is strongly regular, then the vectors $e_i(t) = e_0^{(i)}(t)$, $i = 0, \dots, N$, form a so-called moving frame along γ. Then the next derivative $e'_N(t) = e_0^{(N+1)}(t)$ can be uniquely expressed as a linear combination of vectors in this frame. In other words, we have a well-defined linear homogeneous differential equation on $e_0(t)$:

$$e_0^{(N+1)} = p_N(t)e_0^{(N)} + \dots + p_0(t)e_0. \tag{1.8}$$

Since $e_0(t)$ is defined up to a scale, we can always fix this scaling factor by the condition $p_N(t) = 0$. It is easy to see that this defines $e_0(t)$ uniquely up to a contact nonzero scale. Next, by reparametrizing the curve $e_t(0)$, i.e., by changing t to $\bar{t} = \lambda(t)$ we can also achieve $p_N(t) = p_{N-1}(t) = 0$. This fixes a parameter t up to projective reparametrizations $\bar{t} = (at+b)/(ct+d)$. Wilczynski proves that taking linear combinations of the derivatives of the remaining coefficients $p_i(t)$, $i = 0, \dots, N-2$, we can form $(N-2)$ (relative) invariants of the curve γ under the group of projective transformations:

$$W_k = \sum_{j=1}^{k-2} (-1)^{j+1} \frac{(2k-j-1)!(N-k+j)!}{(k-j)!(j-1)!} p_{N-k+j}^{(j-1)}, \quad k = 3, \dots, N+1.$$

Following Wilczynski, we shall say that an invariant W_i has order i, $i = 3, \dots, N+1$. Wilczynski proves that any other projective invariant of γ can be expressed as a function of invariants W_i and their derivatives. He also shows that in case of odd N the curve γ is self-dual if and only if all invariants of *odd* order vanish identically. Note also that all Wilczynski invariants of a curve $\gamma \subset P^N$ vanish if and only if γ is a rational normal curve, i.e. it can be represented as $t \mapsto [1 : t : \dots : t^N]$ in some homogeneous coordinates of P^N.

Self-dual curves γ in odd-dimensional projective spaces appear naturally in the above equivalence problems via the linearization procedure (see Section 4 for more detail). The linearization of ODE along a solution assigns a curve in projective space to the solution via identification of linear equations with curves in projective space. If the ODE is variational, then the corresponding curves in projective space are self-dual. In the case of rank 2 distributions it is not immediately clear what is the analog of solutions and what is the linearization procedure. This becomes clear if one considers distributions as the constraints for a variational problem and use the Pontryagin Maximum Principle: the analogs of solutions of ODE's are so-called abnormal extremals of the distribution and the linearization of the flow of abnormal extremal leads to the notion of Jacobi curves introduced in [13,14,34], which essentially are (or generated by) self-dual curves in a projective space. In particular, the invariants of these curves define the invariants of the original equivalence problems. For example, as shown in [35], the fundamental invariant W_4 of self-dual curves in $\mathbb{RP}^3$ can be identified with the so-called fundamental tensor of rank 2 vector distributions in 5-dimensional spaces discovered by E. Cartan [5].

2. Variational problems and rank 2 vector distributions

The aim of this section is to establish the correspondence between variational problems of order $n \geqslant 3$ and special rank 2 vector distributions associated with the underdetermined ordinary differential equations (1.4) of order n.

Lemma 2.1. *Let D be the rank 2 distribution associated with to the non-degenerate underdetermined $z' = f(x, y, y', \dots, y^{(n)})$ of order $n \geqslant 3$. Then the space of all infinitesimal symmetries of D lying in D^3 is one-dimensional (over $\mathbb{R}$) and is generated by the vector field $Z = \frac{\partial}{\partial z}$.*

Proof. Let us show that $\operatorname{sym}(D) \cap D^3$ is one-dimensional (over $\mathbb{R}$) and is generated by the vector field $\frac{\partial}{\partial z}$. Indeed, it is easy to check that the space D^3 is generated by the vector fields:

$$X_1 = \frac{\partial}{\partial x} + y_1 \frac{\partial}{\partial y_0} + \dots + y_n \frac{\partial}{\partial y_{n-1}} + f \frac{\partial}{\partial z},$$
$$X_2 = \frac{\partial}{\partial y_n}, \qquad X_3 = \frac{\partial}{\partial y_{n-1}}, \qquad X_4 = \frac{\partial}{\partial y_{n-2}}, \qquad X_5 = \frac{\partial}{\partial z}.$$

Let $Y = \sum_{i=1}^{5} a_i X_i$ be an arbitrary vector field lying in D^3. Then we have

$$[X_1, Y] = \left[X_1, \sum_{i=2}^{5} a_i X_i \right] \mod D$$
$$= \big(X_1(a_3) - a_2\big) \frac{\partial}{\partial y_{n-1}} + \big(X_1(a_4) - a_3\big) \frac{\partial}{\partial y_{n-2}} - a_4 \frac{\partial}{\partial y_{n-3}} + X_1(a_5) \frac{\partial}{\partial z} \mod D.$$

Thus, the condition $Y \in \operatorname{sym}(D)$ implies that $a_2 = a_3 = a_4 = 0$ and $X_1(a_5) = 0$. Further, we have

$$[X_2, Y] = X_2(a_5) \frac{\partial}{\partial z} - a_1 \frac{\partial}{\partial y_{n-2}} \mod D.$$

Again, the condition $Y \in \operatorname{sym}(D)$ implies that $a_1 = 0$ and $X_2(a_5) = 0$. In particular, we see that the function a_5 is a first integral of the distribution D. But since D is completely non-holonomic, a_5 should be a constant. This completes the proof of the lemma. □

Corollary 2.1. *Let $z' = f_i(x, y, y', \dots, y^{(n)})$, $i = 1, 2$, be two non-degenerate underdetermined differential equations of order $n \geqslant 3$, and let D_i, $i = 1, 2$, be the corresponding rank 2 vector distributions on $\mathbb{R}^{n+3}$. Suppose that distributions D_1 and D_2 are locally equivalent. Then the equivalence mapping ϕ maps vector field $\frac{\partial}{\partial z}$ to $c\frac{\partial}{\partial z}$ for some constant $c \in \mathbb{R}^*$.*

Let us identify the space $\mathbb{R}^{n+3}$ with the direct product of $J^n(\mathbb{R}, \mathbb{R})$ with the coordinates $(x, y_0, \dots, y_n)$ and $\mathbb{R}$ with the coordinate z and consider any equivalence mapping ϕ as a mapping from $J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R}$ to itself. Then Lemma 2.1 implies that any such mapping ϕ has the form

$$\phi : J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R} \to J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R}, \qquad (p, z) \mapsto \big(\psi(p), \alpha z + \mu(p)\big), \tag{2.1}$$

where $\psi : J^n(\mathbb{R}, \mathbb{R}) \to J^n(\mathbb{R}, \mathbb{R})$, $\alpha \in \mathbb{R}^*$ and μ is a smooth function on $J^n(\mathbb{R}, \mathbb{R})$.

Lemma 2.2. *Mapping ψ is a contact transformation and the function μ does not depend on y_n, i.e., it is a pull-back of the function on $J^{n-1}(\mathbb{R}, \mathbb{R})$.*

Proof. Since $\frac{\partial}{\partial z}$ is a symmetry of both distributions D_1 and D_2, we can consider the direct images of these distributions with respect to the natural projection $J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R} \to J^n(\mathbb{R}, \mathbb{R})$. It is easy to see that in both cases these images coincide with the contact distribution on $J^n(\mathbb{R}, \mathbb{R})$. This proves that ψ is a contact transformation.

The second statement of the lemma on the function μ follows immediately from the fact that both D_1 and D_2 contain the vector field $\frac{\partial}{\partial y_n}$. □

Note that Lemma 2.2 can be considered as a particular case of [2], where symmetries of more general classes of underdetermined ODEs are treated.

Theorem 2.1. *Suppose that $n \geqslant 3$. Two vector distributions D_1 and D_2 associated with non-degenerate Lagrangians $L_i = f_i(x, y, y', \dots, y^{(n)})\,dx$, $i = 1, 2$, for $n \geqslant 3$ are equivalent if and only if the Lagrangians L_1 and L_2 are equivalent.*

Proof. It is easy to see that the transformations (2.1) with mappings ψ and μ satisfying Lemma 2.2 induce the same equivalence relation on distributions D_1 and D_2 as the equivalence relation on Lagrangians L_1 and L_2 given by Eq. (1.1). □

As a direct consequence of Theorems 1.1 and 2.1 we get:

Theorem 2.2. *The dimension of the group of variational symmetries of a non-degenerate Lagrangian of order $n \geqslant 3$ does not exceed $2n+5$. Any non-degenerate Lagrangian with $(2n+5)$-dimensional group of variational symmetries is equivalent to the Lagrangian $(y^{(n)})^2$. The algebra of infinitesimal symmetries of the latter Lagrangian is isomorphic to a semidirect sum of $\mathfrak{gl}(2,\mathbb{R})$ and $(2n+1)$-dimensional Heisenberg algebra $\mathfrak{n}_{2n+1}$.*

From the proof of Lemma 2.2 it follows that rank 2 distribution associated with some non-degenerate Lagrangian $f(x,y,y',\dots,y^{(n)})\,dx$ can be described in the following coordinate free way:

Proposition 2.1. *A rank 2 distribution D is associated with a non-degenerate Lagrangian $f(x,y,\dots,y^{(n)})\,dx$ in a neighborhood of a generic point if and only if*

(1) $\dim D^3 = 5$;
(2) *There exists an infinitesimal symmetry X of D lying in D^3 such that the factorization by the foliation of integral curves of X sends D to the Goursat distribution on the quotient manifold.*

3. Two points of view on extremals of variational problems

In this section we introduce abnormal extremals of rank 2 distributions and show how the flow of abnormal extremals of a distribution associated with a Lagrangian $L = f(x,y,y',\dots,y^{(n)})\,dx$ can be related to the flow of extremals of the corresponding variational problems. Speaking informally, this relation is the relation between the Hamiltonian and the Lagrangian approach to variational problems and it is given by a kind of Legendre transform. The material of this section is pretty standard but, as we shall see in the next sections, it is very useful for the equivalence problem for Lagrangians and to our knowledge it was never used before in this kind of problems.

3.1. Hamiltonian form of Euler–Lagrange equation

Recall that extremals of the Lagrangian L are critical points of the corresponding functional $L = \int f(x,y,y',\dots,y^{(n)})\,dx$. That is they are solutions of the Euler–Lagrange equation

$$f_{y_0} - \frac{d}{dx}(f_{y_1}) + \cdots + (-1)^n \frac{d^n}{dx^n}(f_{y_n}) = 0. \tag{3.1}$$

If one takes a little bit more general point of view (that is standard in the Optimal Control Theory), then the extremals can be also described using the notion of the end-point mappings associated with the corresponding control system (1.2). Fix a point $q_0 \in J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ and a time $T > 0$. *The endpoint map $\mathcal{F}_{q_0,T}$* is the map from the space $L_\infty[0,T]$ to $J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ sending a control function $u(t)$ to the point of the corresponding admissible trajectory of the system (1.2) at time T. Then $y(t)$ is an extremal of the Lagrangian L if and only if *the corresponding control $\bar u(t) = y^{(n+1)}(t)$ is a critical point of the endpoint map $\mathcal{F}_{q_0,T}$ for some $T > 0$* (and therefore for any $T > 0$ as long as the corresponding trajectory is defined on $[0,T]$), where $q_0 = (0, y(0), \dots, y^n(0), z_0)$ and z_0 is an arbitrary constant. Take the admissible trajectory $q(t)$ of (1.2) corresponding to the control $\bar u(t)$ and starting at q_0. Then this trajectory can be lifted to the cotangent bundle $T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ by choosing for any $t \in [0,T]$ an appropriately normalized covectors $p(t) \in T^*_{q(t)}(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ that annihilates the image of the differential $d\mathcal{F}_{q_0,t}(\bar u(\cdot))$ of the endpoint map $\mathcal{F}_{q_0,t}$ at $\bar u(\cdot)$. This lifting constitutes one of the main fundamental ideas behind the Pontryagin Maximum Principle in Optimal Control [1,27]. As a matter of fact, the curve $(p(t), q(t))$ is an abnormal extremal of the affine control system (1.2) and also of the distribution associated with the Lagrangian L. This establish in essence the relation between extremals of the Lagrangian and the abnormal extremals of the corresponding distributions.

More precisely, the coordinates $q = (x, y_0, \dots, y_n, z)$ in $J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$ induce the coordinate system

$$(p,q) = (\lambda, \xi_0, \dots, \xi_n, \nu; x, y_0, \dots, y_n, z) \tag{3.2}$$

in $T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ such that the covector $p \in T^*_q(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ has the form $p = \lambda\,dx + \sum_{i=0}^n \xi_i\,dy_i + \nu\,dz$. Define the following families of scalar functions (Hamiltonians) H_u on $T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$:

$$H_u(p,q) = \lambda + \sum_{i=0}^{n-1} \xi_i y_{i+1} + \xi_n u + \nu f(x, y_0, y_1, \dots, y_n). \tag{3.3}$$

According to the weak form of the Pontryagin Maximum Principle (where the maximality condition is replaced by the stationarity condition) on has the following

Proposition 3.1. *A function $y(t)$ is an extremal of the Lagrangian L if and only if for the admissible trajectory $q(t)$ of (1.2) corresponding to the control $\bar{u}(t) = y^{(n+1)}(t)$ and starting at the point $q_0 = (0, y(0), \dots, y^{(n)}(0), z_0)$, where z_0 is an arbitrary constant, there exists a curve of nonzero covectors $p(t) \in T^*_{q(t)}(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ such that*

$$\frac{\partial}{\partial u} H_u\big(p(t), q(t)\big)\big|_{u=\bar{u}(t)} = 0 \text{ a.e.} \quad \Leftrightarrow \quad \xi_n(t) \equiv 0 \quad (\text{the stationarity condition}), \tag{3.4}$$

$$H_{\bar{u}(t)}\big(p(t), q(t)\big) \equiv 0 \quad (\text{the transversality condition}), \tag{3.5}$$

$$\dot{p}(t) = -\frac{\partial}{\partial q} H_{\bar{u}(t)}\big(p(t), q\big)\big|_{q=q(t)} \quad (\text{the adjoint equation}). \tag{3.6}$$

Note that another part of the Hamiltonian system

$$\dot{q}(t) = \frac{\partial}{\partial p} H_{\bar{u}(t)}\big(p, q(t)\big)\big|_{p=p(t)} \tag{3.7}$$

is exactly the system (1.2) with $u(t) = \bar{u}(t)$ i.e. it holds automatically. So Eqs. (3.4)–(3.7) can be considered as the *Hamiltonian form of the Euler Lagrange equation*. The curve $(p(t), q(t)) \subset T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ satisfying Proposition 3.1 is called an *abnormal extremal of affine control system* (1.2). The term "abnormal" comes again from the Pontryagin Maximum Principle applied to a functional defined on the set of admissible trajectories of system (1.2): abnormal extremals are exactly the Pontryagin extremals of this problem with vanishing Lagrange multiplier near the functional [1,27]. Roughly speaking, the extremals of our original variational problem given by the Lagrangian L become abnormal extremals of the system (1.2), because we include the Lagrangian L into this system so that it appears as a part of the constraints.

Let us analyze Eq. (3.6) in coordinates (3.2). First of all, since the Hamiltonians (3.3) do not depend on z we have $\dot{\nu} = 0$, i.e. ν is constant along an abnormal extremal. If $\nu \equiv 0$, then from other equations of (3.6) and Eq. (3.5) it follows that $p(t) \equiv 0$ but $p(t)$ cannot vanish by Proposition 3.1. So the case $\nu \equiv 0$ is impossible. Now assume that $\nu \neq 0$. From the homogeneity of Eq. (3.6) with respect to p it follows that it is enough to consider the case when $\nu \equiv -1$. Then, combining the stationarity condition (3.4) with the equation from (3.6) regarding $\dot{\xi}_n$ we will get that

$$\xi_{n-1} = f_{y_n}. \tag{3.8}$$

Writing equations for others ξ_j from (3.6) we get

$$\begin{aligned} &\dot{\xi}_0 = f_{y_0}; \\ &\dot{\xi}_j = f_{y_j} - \xi_{j-1}, \quad j = 1, \dots, n-1. \end{aligned} \tag{3.9}$$

Combining (3.8) and the equation in (3.9) corresponding $j = n-1$ we get $\xi_{n-2} = f_{n-1} - \frac{d}{dx}(f_n)$. Then using the second line of (3.9) by induction with respect to j in the decreasing order, we get

$$\xi_{j-1} = \sum_{k=j}^{n} (-1)^{k-j} \frac{d^{k-j}}{dx^{k-j}}(f_{y_k}), \quad 1 \leqslant j \leqslant n-1. \tag{3.10}$$

Finally substituting (3.10) with $j = 1$ into the first line of (3.9) we get the Euler–Lagrange equation (3.1), as expected.

If X is a vector field without stationary points or a line distribution, denote by $\mathrm{Fol}(X)$ the one-dimensional foliation of integral curves of X. Consider a codimension 4 submanifold $\mathcal{H}$ of $T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ given by Eqs. (3.4), (3.5), (3.8), and $\nu = -1$. It is foliated by abnormal extremals of system (1.2) with $\nu = -1$. Besides by constructions the group of translations along z-axis, $z \mapsto z + c$, preserves this foliation. Therefore this foliation induces the one-dimensional foliation on the quotient manifold $\mathcal{H}/\mathrm{Fol}(\frac{\partial}{\partial z})$ by the foliation $\mathrm{Fol}(\frac{\partial}{\partial z})$ or, equivalently, on the manifold of the orbits of the group of these translations. The tuple $(x, y_0, \dots, y_n, \xi_0, \dots, \xi_{n-2})$ constitute a coordinate system on the manifold $\mathcal{H}/\mathrm{Fol}(\frac{\partial}{\partial z})$. On the other hand, the Euler–Lagrange equation (3.1) defines a codimension one submanifold $\mathcal{E}(L)$ of $J^{2n}(\mathbb{R},\mathbb{R})$ foliated by a one-dimensional foliation of prolongations of its solutions to $J^{2n}(\mathbb{R},\mathbb{R})$, and the tuple $(x, y_0, \dots, y_{2n-1})$ constitute a coordinate system on $\mathcal{E}(L)$. This foliation is called the *foliation of solutions* of the Euler–Lagrange equation.

By above, the map $\mathfrak{L} : \mathcal{E}(L) \mapsto \mathcal{H}/\mathrm{Fol}(\frac{\partial}{\partial z})$, defined by

$$(x, y_0, \dots, y_{2n-1}) \mapsto (x, y_0, \dots, y_n, \xi_0, \dots, \xi_{n-2}), \tag{3.11}$$

with ξ_j satisfying (3.10), sends the one-dimensional foliation on $\mathcal{E}(L)$ to the one-dimensional foliation on $\mathcal{H}/\mathrm{Fol}(\frac{\partial}{\partial z})$. In other words, this map transforms the extremals of our variational problem obtained in the Lagrangian form to the extremal obtained in the Hamiltonian form. Therefore we call it the (*generalized*) *Legendre transform*. Note that the Legendre transform depend on the choice of coordinates on $J^0(\mathbb{R},\mathbb{R}) = \mathbb{R}^2$, which induces the coordinates on $J^{2n}(\mathbb{R},\mathbb{R})$. Once we use the Legendre transform in the sequel it will mean that such choice is already done.

3.2. Abnormal extremals of rank 2 distributions

Now we are going to describe abnormal extremals for a distribution D on a manifold M. We shall use more geometric language. Let $\tilde{\pi}: T^*M \mapsto M$ be the canonical projection. For any $\lambda \in T^*M$, $\lambda = (p, q)$, $q \in M$, $p \in T_q^*M$, let $\mathfrak{s}(\lambda)(\cdot) = p(\tilde{\pi}_* \cdot)$ be the tautological Liouville 1-form and $\sigma = d\mathfrak{s}$ be the standard symplectic structure on T^*M. Denote by $(D^l)^\perp \subset T^*M$ the annihilator of the lth power D^l, namely

$$(D^l)^\perp = \{(q, p) \in T^*M\colon\ p \cdot v = 0,\ \forall v \in D^l(q)\}. \tag{3.12}$$

Finally let $\mathcal{S}_0$ be the zero section of T^*M. With this notation the Pontryagin Maximum Principle in the coordinate-free form [1] implies immediately the following description of abnormal extremals of the distribution D:

Definition 3.1. An absolutely continuous curve $\gamma \subset T^*M$ is an abnormal extremal of a distribution D if the following two conditions hold:

1. $\gamma \subset D^\perp \backslash \mathcal{S}_0$,
2. $\dot{\gamma}(t)$ belongs to $\mathrm{Ker}(\sigma|_{D^\perp})$ a.e., i.e., to the kernel of the restriction of the canonical symplectic form σ to the annihilator $D^\perp$ of D.

From now on we will consider only rank 2 distributions. From direct computations [32, Proposition 2.2] it follows that $\mathrm{Ker}(\sigma(\lambda)|_{D^\perp}) \neq 0$ if and only if $\lambda \in (D^2)^\perp$. This implies the following characterization of abnormal extremals of rank 2 distribution.

Proposition 3.2. *An absolutely continuous curve $\gamma \subset T^*M$ is abnormal extremal of a rank 2 distribution D with $\dim D^2 = 3$ if and only if the following two conditions hold*:

1. $\gamma \subset (D^2)^\perp \backslash \mathcal{S}_0$,
2. $\dot{\gamma}(t)$ *belongs to* $\mathrm{Ker}(\sigma|_{(D^2)^\perp})$ *a.e., i.e., to the kernel of the restriction of the canonical symplectic form σ to the annihilator $(D^2)^\perp$ of D^2.*

Further, if $\lambda \in (D^2)^\perp \backslash (D^3)^\perp$ then $\mathrm{Ker}(\sigma|_{(D^2)^\perp})$ is one-dimensional. These kernels form a special line distribution on $\lambda \in (D^2)^\perp \backslash (D^3)^\perp$, called the *characteristic distribution*. It will be denoted by $\widetilde{\mathcal{C}}$. The abnormal extremals of D, lying in $\lambda \in (D^2)^\perp \backslash (D^3)^\perp$, are smooth and they are exactly the integral curves of the line distribution $\mathcal{C}$ (in some literature these abnormal extremals are called regular).

Remark 3.1. For any $\lambda \in (D^2)^\perp \backslash (D^3)^\perp$ let

$$\widetilde{\mathcal{J}}(\lambda) = \{v \in T_\lambda (D^2)^\perp\colon\ \tilde{\pi}_* v \in D(\tilde{\pi}(\lambda))\}. \tag{3.13}$$

A simple count shows that $\dim \widetilde{\mathcal{J}}(\lambda) = n + 2$. Then from constructions it follows immediately that the restriction of the form $\sigma(\lambda)$ to $\mathcal{J}(\lambda)$ is identically equal to zero.

Now consider the distribution D on $M = J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R}$ associated with the Lagrangian $L = f(x, y, y', \dots, y^{(n)})\,dx$ with $\frac{\partial^2 f}{\partial y_n^2} \neq 0$. We would like to rewrite the constructions of the end of the previous subsection in more geometric form. First of all in the considered case $D^\perp$ is a codimension 2 submanifold of T^*M given by Eqs. (3.4) and (3.5), $(D^2)^\perp$ is a codimension 1 submanifold of $D^\perp$ satisfying in additional equation (3.8), and $(D^3)^\perp$ is a codimension 2 submanifold of $(D^2)^\perp$ satisfying to additional equations $\nu = 0$ and $\xi_{n-2} = 0$.

Further, given a vector field X on M denote by $H_X: T^*M \to \mathbb{R}$ the corresponding quasi-impulse

$$H_X(p, q) = p(X(q)), \quad q \in M,\ p \in T_q^*M, \tag{3.14}$$

and by $\vec{H}_X$ the corresponding Hamiltonian vector field on T^*M, i.e. the vector field satisfying $i_{\vec{H}_X}\sigma = -dH_X$. It is clear that if X is an infinitesimal symmetry of the distribution D, then the flow $e^{t\vec{H}_X}$, generated by $\vec{H}_X$, sends an abnormal extremal of D to an abnormal extremal of D. Moreover, any abnormal extremal lies on a level set of the function H_X.

In particular, let as in Lemma 2.1 Z be the infinitesimal symmetry of D lying in D^3. The submanifold $\mathcal{H}$ introduced in the previous subsection is equal to $\{H_Z = -1\} \cap (D^2)^\perp$ and the abnormal extremals of the distribution D lying on $\mathcal{H}$ coincide (as unparametrized curves) with the abnormal extremals of system (1.2) having $\nu = -1$. The distribution $\widetilde{\mathcal{C}}$ induces a rank 1 distribution $\bar{\mathcal{C}}$ on the quotient manifold $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$, where as before $\operatorname{Fol}(H_Z)$ is the foliation of integral curves of the field $\vec{H}_Z$. A Legendre transform $\mathfrak{L}: \mathcal{E}(L) \to \mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$, defined in the previous subsection, sends the one-dimensional foliation of solutions of Euler–Lagrange equations to the one-dimensional foliation of the integral curves of distribution $\bar{\mathcal{C}}$.

Remark 3.2. The Legendre transform $\mathfrak{L}$ satisfies another important property. To describe it in geometric terms let $\pi_{i,j} : J^i(\mathbb{R},\mathbb{R}) \to J^j(\mathbb{R},\mathbb{R})$, where $i > j$, and

$$\bar{\pi} : T^*\big(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}\big) \to J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}$$

denote the canonical projections. The mapping $\bar{\pi}$ induces the mapping

$$\pi_Z : T^*\big(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}\big)/\operatorname{Fol}(\vec{H}_Z) \mapsto \big(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R}\big)/\operatorname{Fol}(Z) \sim J^n(\mathbb{R},\mathbb{R})$$

in the obvious way. If the submanifolds $\mathcal{E}(L)$ and $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$ are considered as fiber bundles over $J^n(\mathbb{R},\mathbb{R})$ with the projections $\pi_{2n,n}|_{\mathcal{E}(L)}$ and $\pi_Z|_{\mathcal{H}/e^{t\vec{H}_Z}}$, respectively, then from (3.11) it follows immediately that the Legendre transform $\mathfrak{L}$ is fiberwise mapping over the identity on the base manifold $J^n(\mathbb{R},\mathbb{R})$.

Remark 3.3. The tautological Liouville 1-form $\mathfrak{s}$ and the standard symplectic structure σ on $T^*(J^n(\mathbb{R},\mathbb{R}) \times \mathbb{R})$ induce the 1-form $\bar{\mathfrak{s}}$ and the closed 2-form $\bar{\sigma} = d\bar{\mathfrak{s}}$, respectively, on $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$. By constructions, the rank 1 distribution $\bar{\mathcal{C}}$ satisfies $\bar{\mathcal{C}} = \operatorname{Ker}\bar{\sigma}$. Besides, using condition (3.5) it is easy to show that

$$\bar{\mathfrak{s}} = L + \sum_{i=1}^{n-2} \xi_i\theta_i + f_{y_n}\theta_{n-1} \tag{3.15}$$

in the coordinates $(x, y_0, \dots, y_n, \xi_0, \dots, \xi_{n-2})$ on $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$, where, as before, $\theta_i = dy_i - y_{i+1}\,dx$. Finally, for any $\lambda \in \mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$ we denote

$$\bar{\mathcal{J}}(\lambda) = \big\{v \in T_\lambda\mathcal{H}/\operatorname{Fol}(\vec{H}_Z){:}\ (\pi_Z)_*v \in D\big(\tilde{\pi}(\lambda)\big)\big\}. \tag{3.16}$$

Then from the last sentence of Remark 3.1 the restriction of the form $\bar{\sigma}(\lambda)$ to the subspace $\bar{\mathcal{J}}(\lambda)$ is identically equal to zero.

4. Linearization of variational ODEs and Jacobi curves of rank 2 distribution

Let us outline the content of this section. As was mentioned before both in the equivalence problem for variational ODEs and in the equivalence problem for rank 2 distributions self-dual curves in a projective space play a crucial role. They appear via the linearization along the "flow" of solutions in the first case and along the "flow" of abnormal extremals in the second case. Using the Legendre transform introduced in Section 3.1 we show that the self-dual curves in a projective space obtained by the linearization along a solution of the Euler–Lagrange equations of a Lagrangian and by the linearization along the corresponding abnormal extremal of the associated rank 2 distribution are actually isomorphic. This observation leads to the description of the fundamental system of invariants for rank 2 distributions associated with Lagrangians given in the next section.

4.1. General linearization procedure

Let us first clarify what do we mean by the linearization procedure in a general geometric setting. Let $\mathcal{M}$ be an arbitrary smooth manifold, let $\mathcal{G}$ and $\mathcal{V}$ be a pair of vector distributions on $\mathcal{M}$ of rank l and k, respectively, where one of them, say $\mathcal{G}$, is integrable and $\mathcal{V} \cap \mathcal{G}$ is a distribution of rank r.

Similarly to above, let $\operatorname{Fol}(\mathcal{G})$ be a foliation of $\mathcal{M}$ by maximal integral submanifolds of $\mathcal{G}$. Then we can define the *linearization of the distribution* $\mathcal{V}$ *along the foliation* $\operatorname{Fol}(\mathcal{G})$ in the following way. Locally we can assume that there exists a quotient manifold $\mathcal{M}/\operatorname{Fol}(\mathcal{G})$, whose points are leaves of $\operatorname{Fol}(\mathcal{G})$. Let Γ by any such leaf. Then we define the map ϕ of Γ into the Grassmannian $\operatorname{Gr}_{k-r}(T_\Gamma(\mathcal{M}/\operatorname{Fol}(\mathcal{G})))$ of $(k-r)$-dimensional subspaces of $(\mathcal{M}/\operatorname{Fol}(\mathcal{G}))$ or, under additional regularity assumptions, an l-dimensional submanifold of $\operatorname{Gr}_{k-r}(T_\Gamma(\mathcal{M}/\operatorname{Fol}(\mathcal{G})))$ as follows: $\phi(x) = \operatorname{pr}_*(\mathcal{V}_x)$, $x \in \Gamma$, where $\operatorname{pr} : \mathcal{M} \to \mathcal{M}/\operatorname{Fol}(\mathcal{G})$ is a natural projection. The map ϕ or its image in $\operatorname{Gr}_{k-r}(T_\Gamma(\mathcal{M}/\operatorname{Fol}(\mathcal{G})))$ is called the *linearization of the distribution* $\mathcal{V}$ *along the foliation* $\operatorname{Fol}(\mathcal{G})$ *at the leaf* Γ or the *linearization of the distribution* $\mathcal{V}$ *along the leaf* Γ *(of* $\operatorname{Fol}(\mathcal{G})$*)*. In the cases under consideration $l = 1$ so that the linearizations are curves in projective spaces.

The main idea of using the linearization procedure in the equivalence problem for the structures given by the pair of distribution $(\mathcal{V},\mathcal{G})$ on $\mathcal{M}$ with respect to the action of group of diffeomorphisms of M is that it allows to construct the invariants of such structures from invariants of submanifold in an appropriate Grassmannian with respect to the natural action of the General Linear Group on this Grassmannian.

4.2. Linearization procedure for ODEs

As in Section 1.3 an ODE of order $N+1$, resolved with respect to the highest derivative, is given by a hypersurface $\mathcal{E}$ in the jet space $J^{N+1}(\mathbb{R},\mathbb{R})$ so that the restriction of the natural projection $\pi_{N+1,N} : J^{N+1}(\mathbb{R},\mathbb{R}) \to J^N(\mathbb{R},\mathbb{R})$ to the hypersurface $\mathcal{E}$ is a diffeomorphism. Further, the Cartan distribution of $J^{N+1}(\mathbb{R},\mathbb{R})$ (i.e. the rank 2 distribution defined by contact forms $\theta_i = dy_i - y_{i+1}dx$, $i = 0,\dots,N$, in the standard coordinates $(x, y_0, \dots, y_{N+1})$ in $J^{N+1}(\mathbb{R},\mathbb{R})$) defines the line distribution $\mathcal{S}$ on $\mathcal{E}$. This distribution is obtained by the intersection of the Cartan distribution with the tangent space to $\mathcal{E}$ at every point of $\mathcal{E}$. Note that the corresponding foliation $\mathrm{Fol}(S)$ is the foliation of solutions of our ODE (more precisely, the foliation of the prolongations of the solutions to $J^N(\mathbb{R},\mathbb{R})$). If the hypersurface $\mathcal{E}$ has the form (1.5) in coordinates $(x, y_0, \dots, y_{N+1})$ on $J^{N+1}(\mathbb{R},\mathbb{R})$, then in the coordinates $(x, y_0, \dots, y_N)$ on $\mathcal{E}$:

$$S = \left\langle \frac{\partial}{\partial x} + \sum_{i=1}^{N-1} y_{i+1}\frac{\partial}{\partial y_i} + F(x, y_0, y_1, \dots, y_N)\frac{\partial}{\partial y_N}\right\rangle. \tag{4.1}$$

The distribution $\mathcal{S}$ will play the role of the distribution $\mathcal{G}$ from the previous subsection.

Further let, as before, $\pi_{N+1,N-i} : J^{N+1}(\mathbb{R},\mathbb{R}) \mapsto J^{N-i}(\mathbb{R},\mathbb{R})$ be the canonical projection. For any $\varepsilon \in \mathcal{E}$ we can define the filtration $\{V^i_\varepsilon\}_{i=0}^N$ of $T_\varepsilon\mathcal{E}$ as follows:

$$V^i_\varepsilon = \ker d_\varepsilon \pi_{N+1,N-i} \cap T_\varepsilon\mathcal{E}. \tag{4.2}$$

Then V^i is a rank i distribution on $\mathcal{E}$. In the coordinates $(x, y_0, \dots, y_N)$ on $\mathcal{E}$ we have

$$V^i = \left\langle \frac{\partial}{\partial y_{N-i+1}}, \dots, \frac{\partial}{\partial y_N}\right\rangle. \tag{4.3}$$

Let Sol denote the quotient manifold $\mathcal{E}/\mathrm{Fol}(\mathcal{S})$, i.e. the manifold of solutions of the equation $\mathcal{E}$. Fix a point $\Gamma \in \mathrm{Sol}$. In other words, Γ is a leaf of $\mathrm{Fol}(\mathcal{S})$ or a solution of the equation $\mathcal{E}$. Consider the linearization Lin^i_Γ of the distribution V^i along Γ. It is a curve in $\mathrm{Gr}_i(T_\Gamma \mathrm{Sol})$. In particular, Lin^1_Γ is a curve in the projective space $\mathbb{P}(T_\Gamma \mathrm{Sol})$. Moreover, if Γ is considered as the leaf of $\mathrm{Fol}(\mathcal{S})$, then from (4.1) and (4.2) one gets immediately the following

Lemma 4.1. *For any $\varepsilon \in \Gamma$ the i-dimensional subspace $\mathrm{Lin}^i_\Gamma(\varepsilon)$ of $T_\Gamma \mathrm{Sol}$ is exactly the i-th osculating space of the curve Lin^1_Γ at ε (as defined in (1.7)).*

The Wilczynski invariants of the linearizations Lin^1_Γ taken for every solution Γ define the invariants of the original ODE $\mathcal{E}$ under the group of contact transformations (see [11]). We call these invariants the *generalized Wilczynski invariants of $\mathcal{E}$* and denote them also by W_i, $i = 3, \dots, N+1$.

4.3. The case of variational ODEs

Now assume that $\mathcal{E}$ is a variational ODE, $\mathcal{E} = \mathcal{E}(L)$ for some Lagrangian L. Let $\bar\sigma$ be the closed 2-form on $H/\mathrm{Fol}(\vec{H}_Z)$ introduced in Remark 3.3 and let $\mathfrak{L} : \mathcal{E}(L) \to \mathcal{H}/\mathrm{Fol}(\vec{H}_Z)$ be a Legendre transform. Then we can define the closed 2-form ω on $\mathcal{E}(L)$ as follows:

$$\omega = \mathfrak{L}^*\bar\sigma. \tag{4.4}$$

Note that from relation (4.3) and Remark 3.2 it follows that the distribution $\bar{\mathcal{J}}$ defined by (3.16) satisfies

$$\bar{\mathcal{J}} = \mathfrak{L}_*(V_n) \oplus \bar{\mathcal{C}}. \tag{4.5}$$

From this and Remark 3.3 it follows that the form ω satisfies the following two properties:

(1) $\operatorname{Ker}\omega = \mathcal{S}$, where, as before, $\mathcal{S}$ is the rank 1 distribution generating the foliation of solutions of $\mathcal{E}(L)$;
(2) The restriction of ω to the distribution V^n vanishes.

From property (1) it follows that ω induces the symplectic form $\bar\omega$ on the manifolds Sol of solutions of $\mathcal{E}(L)$. Moreover from the property (2) it follows that

(2′) For any $\Gamma \in \mathcal{E}(L)$ the linearization Lin^n_Γ of the distribution V^n along G is the curve of Lagrangian subspaces with respect to the symplectic form $\omega(\Gamma)$.

From item (2) of Proposition 1.1 we get immediately the following

Corollary 4.1. *For a variational ODE the linearizations Lin^1_Γ along any solution Γ is a self-dual curve in the corresponding projective space.*

Further, as an immediate consequence of item (1) of Proposition 1.1 and the fact that ω is closed, we get

Proposition 4.1. *For a variational ODE $\mathcal{E}(L)$ there exists a unique, up to a constant nonzero factor, closed 2-form ω on $\mathcal{E}(L)$ satisfying conditions* (1) *and* (2) *above or, equivalently, a unique, up to a constant nonzero factor, symplectic structure $\bar{\omega}$ on the manifold of solutions* Sol, *satisfying condition* (2′) *above.*

Remark 4.1. It is easy to see that this symplectic structure $\bar{\omega}$ from Proposition 4.1 is exactly the 2-form ω given by (1.6) and prescribed by Theorem 2.6 of [3]. Indeed, the condition $d\omega = 0$ implies $d_H\omega = 0$. In particular, this means that ω projects to the solution space of the equation $\mathcal{E}(L)$. Proposition gives an alternative construction for this 2-form based on the generalized Legendge transform.

Now we are ready to prove the following theorem (which is also proved in [15, Theorem 5.1] in the case $n = 2$ and in [19, Corollary 2.6] for $n \geqslant 3$):

Theorem 4.1. *Lagrangians are equivalent if and only if their Euler–Lagrange equations are contact equivalent.*

Proof. If Lagrangians L_1 and L_2 are equivalent then directly from (1.1) it follows that their Euler–Lagrange equations are contact equivalent.

In the other direction let ψ be the mapping establishing the equivalence of two Euler–Lagrange equations $\mathcal{E}_1$ and $\mathcal{E}_2$. Let ω_1 and ω_2 be the closed 2-forms from Proposition 4.1, corresponding to $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively. Then by this proposition there exists a constant $\alpha \neq 0$ such that

$$\psi^*\omega_2 = \alpha\omega_1. \tag{4.6}$$

Now assume that $\bar{\mathfrak{s}}_1$ and $\bar{\mathfrak{s}}_2$ are 1-forms from Remark 3.3 corresponding to Lagrangians L_1 and L_2, respectively. Assume that $\mathfrak{L}_i$ are the corresponding Legendre transforms and let $\rho_i = (\mathfrak{L}_i)^*\mathfrak{s}_i$, $i = 1, 2$. Then by construction $\omega_i = d\rho_i$. Hence from (4.6) there exists a function μ on $\mathcal{E}_1$ such that

$$\psi^*\rho_2 = \alpha\rho_1 + d\mu.$$

Taking into account the coordinate expressions for the forms $\bar{\mathfrak{s}}_i$ given by (3.15), we get immediately that the last relation is equivalent to (1.1), i.e. the Lagrangians L_1 and L_2 are equivalent. □

4.4. Linearization procedure for rank 2 distributions: Jacobi curves

The presentation of this subsection is rather closed to our previous works [13,14] but it is considered here in relation with the linearization procedure of ODE's from Section 4.2. Let D be an arbitrary rank 2 vector distribution on an $(n+3)$-dimensional manifold M. We shall assume that D is completely non-holonomic and that $\dim D^3 = 5$. It is more convenient to work with the projectivization of $\mathbb{P}T^*M$ rather than with T^*M. Here $\mathbb{P}T^*M$ is the fiber bundle over M with the fibers that are the projectivizations of the fibers of T^*M. The canonical projection $\Pi : T^*M \to \mathbb{P}T^*M$ sends the characteristic distribution $\widetilde{\mathcal{C}}$ on $(D^2)^\perp \backslash (D^3)^\perp$ to the line distribution $\mathcal{C}$ on $\mathbb{P}(D^2)^\perp \backslash \mathbb{P}(D^3)^\perp$, which will be also called the *characteristic distribution* of the latter manifold. The manifold $\mathbb{P}(D^2)^\perp \backslash \mathbb{P}(D^3)^\perp$ and the distribution $\mathcal{C}$ play the role of $\mathcal{M}$ and $\mathcal{G}$, respectively, from the general linearization procedure of Section 4.1.

Further note that the corank 1 distribution on $T^*M \backslash S_0$ annihilating the tautological Liouville form $\mathfrak{s}$ on T^*M induces a contact distribution on $\mathbb{P}T^*M$, which in turns induces the even-contact (quasi-contact) distribution Δ on $\mathbb{P}(D^2)^\perp \backslash \mathbb{P}(D^3)^\perp$. The characteristic line distribution $\mathcal{C}$ is exactly the Cauchy characteristic distribution of Δ, i.e. it is the maximal subdistribution of Δ such that

$$[C, \Delta] \subset \Delta. \tag{4.7}$$

Now let $\widetilde{\mathcal{J}}$ be as in (3.13). Let $\mathcal{J}(\lambda) = \Pi_*\widetilde{\mathcal{J}}$ and define a sequence of subspaces $\mathcal{J}^{(i)}(\lambda)$, $\lambda \in \mathbb{P}(D^2)^\perp \backslash \mathbb{P}(D^3)^\perp$, by the following recursive formulas:

$$\mathcal{J}^{(i)}(\lambda) := \mathcal{J}^{(i-1)}(\lambda) + \left[\mathcal{C}, \mathcal{J}^{i-1}\right](\lambda). \tag{4.8}$$

By [34, Proposition 3.1], we have $\dim \mathcal{J}^{(1)}(\lambda) - \dim \mathcal{J}(\lambda) = 1$, which implies easily that

$$\dim \mathcal{J}^{(i)}(\lambda) - \dim \mathcal{J}^{(i-1)}(\lambda) \leqslant 1, \quad \forall i \in \mathbb{N}.$$

Besides, from (4.7) it follows that $\mathcal{J}^{(i)} \subset \Delta$ for all natural i. Simple counting of dimensions implies that $\operatorname{rank} \Delta = 2n+1$ so that $\dim \mathcal{J}^{(i)}(\lambda) \leqslant 2n+1$.

Note that for any $\lambda \in \mathbb{P}(D^2)^\perp \backslash \mathbb{P}(D^3)^\perp$ the subspace $\Delta(\lambda)$ is equipped canonically, up to a constant nonzero factor, with a skew-symmetric form with the kernel equal to $\mathcal{C}(\lambda)$: for this take the restriction to $\Delta(\lambda)$ of the differential of any 1-form annihilating Δ. For a given subspace W of $\Delta(\lambda)$ denote by $W^\angle$ the skew-symmetric complement of W with respect to this form. Note that by Remark 3.1 $(\mathcal{J}^{(0)})^\angle = \mathcal{J}^{(0)}$. Then set

$$\mathcal{J}^{(-i)}(\lambda) = \left(\mathcal{J}^{(i)}(\lambda)\right)^{\angle} \quad \text{for any } i > 0. \tag{4.9}$$

The sequence of subspaces $\{\mathcal{J}^{(i)}(\lambda)\}_{i\in\mathbb{Z}}$ defines the filtration of $\Delta(\lambda)$.

Further, define the following two integer-valued functions:

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

The number $m(q)$ is called *the class of distribution D at the point q*. By above, $1 \leqslant m(q) \leqslant n$. It is easy to show that *germs of $(2, n+3)$-distributions of the maximal class n are generic* (see [34, Proposition 3.4]). The following proposition shows that all distributions associated with Lagrangians are of maximal class:

Proposition 4.2. *A rank 2 distribution associated with the non-degenerate underdetermined differential equation*

$$z' = f\left(x, y, y', \dots, y^{(n)}\right) \tag{4.10}$$

is of maximal class at any point.

Proof. Let X_1 and X_2 be the basis of our distribution as in (1.3). Define recursively

$$\begin{aligned} &X_{i+1} = [X_1, X_i], \quad i = 2, \dots, n+1, \\ &Y = [X_2, X_3]. \end{aligned} \tag{4.11}$$

Then $D^2 = \operatorname{span}\{X_1, X_2, X_3\}$, $D^i = \operatorname{span}\{X_1, \dots, X_{i+1}, Y\}$, $i = 1, \dots, n+1$, and $D^{n+1} = TM$ (here $M = J^n(\mathbb{R}, \mathbb{R}) \times \mathbb{R}$). Note that $Y = -\frac{\partial^2 f}{\partial y_n^2}\frac{\partial}{\partial z}$. Further, let

$$\begin{aligned} &u_i = H_{X_i}, \quad i = 1, \dots, n+1, \\ &v = H_Y \end{aligned} \tag{4.12}$$

be the corresponding quasi-impulses (see (3.14)). Let $\pi : \mathbb{P}T^*M \to M$ be the canonical projection. First, from direct computations [34, Proposition 3.1] it follows that

$$\widetilde{C} = \mathbb{R}(u_4\vec{u}_2 - v\vec{u}_1), \tag{4.13}$$

$$\mathcal{J}^{(1)}(\lambda) = \left\{w \in T_\lambda\left(\mathbb{P}D^2\right)^{\perp}\colon \pi_* w \in D^2\left(\pi(\lambda)\right)\right\}. \tag{4.14}$$

Further, let $\Pi : T^*M \to \mathbb{P}T^*M$ be the canonical projections. For any $\lambda \in \mathbb{P}(D^2)^{\perp}\backslash\mathbb{P}(D^3)^{\perp}$ take $\widetilde{\lambda} \in (D^2)^{\perp}\backslash(D^3)^{\perp}$ such that $\Pi(\widetilde{\lambda}) = \lambda$. Then by direct computations one can get from (4.8) and (4.13) that

$$\mathcal{J}^{(2)}(\lambda) = \mathcal{J}^{(1)}(\lambda) + \left\{w \in T_\lambda\left(\mathbb{P}D^2\right)^{\perp}\colon \pi_* w \in \mathbb{R}\left(u_4(\widetilde{\lambda})Y\left(\pi(\lambda)\right) - v(\widetilde{\lambda})X_4\left(\pi(\lambda)\right)\right)\right\}, \tag{4.15}$$

Note that formulas (4.13)–(4.15) are valid for any rank 2 distributions generated by some vector fields X_1 and X_2 (with the notations as in (4.11) and (4.12)) and not only for distributions, generated by (1.3). Finally, using (1.3), for any $\lambda \in \mathbb{P}(D^2)^{\perp}\backslash\mathbb{P}(D^3)^{\perp}$ and $\widetilde{\lambda} \in (D^2)^{\perp}\backslash(D^3)^{\perp}$ such that $\Pi(\widetilde{\lambda}) = \lambda$ one gets by straightforward computations that

$$\mathcal{J}^{(i+1)}(\lambda) = \mathcal{J}^{(i)}(\lambda) + \left\{w \in T_\lambda\left(\mathbb{P}D^2\right)^{\perp}\colon \pi_* w \in \mathbb{R}v(\widetilde{\lambda})X_{i+3}\left(\pi(\lambda)\right)\right\}, \quad i = 2, \dots, n-1.$$

Hence if $v(\widetilde{\lambda}) \neq 0$ then $\dim \mathcal{J}^{i+1} - \dim \mathcal{J}^i = 1$ for $i = 0, \dots, n-1$ and therefore $\nu(\lambda) = n$. So, the distribution D is of maximal class at any point. $\square$

Remark 4.2. The same result with literally the same proof holds for distributions associated with any underdetermined ordinary differential equation $z' = f(x, y, y', \dots, y^{(n)}, z)$ with $\frac{\partial^2 f}{\partial (y^{(n)})^2} \neq 0$.

From now on we assume that D is a $(2, n+3)$-distribution of maximal constant class $m = n$. Let $\mathcal{R} = \{\lambda \in \mathbb{P}(D^2)^{\perp}\backslash\mathbb{P}(D^3)^{\perp}\colon \nu(\lambda) = n\}$. Then on $\lambda \in \mathcal{R}$ the subspaces $\mathcal{J}^{(i)}$ form a distribution of rank $(i+n+1)$ for all integer i between $-n$ and n (in particular $\mathcal{J}^{(n)} = \Delta$). If we denote by Abn the quotient of manifold $\mathbb{P}(D^2)^{\perp}\backslash\mathbb{P}(D^3)^{\perp}$ by $\mathrm{Fol}(\mathcal{C})$ then the distribution Δ induces the distribution $\bar{\Delta}$ on Abn equipped with the canonical, up to a constant nonzero factor, symplectic form. Given any segment Υ of abnormal extremal (a leaf of $\mathrm{Fol}(\mathcal{C})$) consider the linearization $J^{(i)}_{\Upsilon}$ of the distribution $\mathcal{J}^{(i)}$ along Υ. It is a curve in $\mathrm{Gr}_{n+i}(\bar{\Delta}(\Upsilon))$. In particular, $J^{(1-n)}_{\Upsilon}$ is a curve in the projective space $\mathbb{P}(\bar{\Delta}(\Gamma))$. The curve $J^{(0)}_{\Upsilon}$ is called the *Jacobi curve* along the abnormal extremal Γ. By Remark 3.1 it is the curve of Lagrangian subspaces of $\bar{\Delta}$. Further, it is not hard to see [34] that for any $\lambda \in \Gamma$ the $(n+i)$-dimensional subspace $J^{(i)}_{\Upsilon}(\lambda)$ of $T_{\Upsilon}\mathrm{Abn}$ is exactly the $(i+n-1)$-st osculating space of the curve $J^{(1-n)}_{\Upsilon}$ at λ (as defined in (1.7)). So, by item (2) of Proposition 1.1, the curve $J^{(1-n)}_{\Upsilon}$ is self-dual.

Remark 4.3. It can be shown [33] that $J_\Upsilon^{(1-n)}$ is the only curve in $\mathbb{P}(\bar{\Delta}(\Upsilon))$ such that the Jacobi curve at $\lambda \in \Upsilon$ is its $(n-1)$-st osculating space at λ.

The Wilczynski invariants of the linearizations $J_\Upsilon^{(1-n)}$ taken for every abnormal Υ define the invariants of the distribution D. The latter invariants are called the *generalized Wilczynski invariants of* $\mathcal{E}$. The Jacobi curve along the abnormal extremal is called *flat* if the corresponding curve $J_\Upsilon^{(1-n)}$ is a rational normal curve in $\mathbb{P}(\bar{\Delta}(\Upsilon))$ or, equivalently, all Wilczynski invariants of $J_\Upsilon^{(1-n)}$ vanish identically. For the maximally symmetric $(2, n+3)$-distribution of maximal class which is locally equivalent to the distribution associated with the Lagrangian $(y^{(n)})^2\,dx$ (Theorem 1.1) all Jacobi curves are flat or, equivalently, all generalized Wilczynski invariants vanish. The general question is

Question. Is it true that if all generalized Wilczynski invariants of $(2, n+3)$-distribution of maximal class vanish or, equivalently, all its Jacobi curves are flat, then the distribution is locally equivalent to the distribution associated with the Lagrangian $(y^{(n)})^2\,dx$?

Since, as was shown in [35], in the case $n = 2$ the generalized Wilczynski invariant coincides with Cartan's covariant binary biquadratic form introduced in [5], then the fact that this form is the fundamental invariant of a $(2, 5)$-distribution (proved in [5] as well) gives the positive answer to our question in this case. We show in the next section that for $n \geqslant 3$ the answer to this question is positive if we restrict ourselves to distributions associated with Lagrangians (see Theorem 5.1 below). Note that this class of distributions is rather restrictive and the general question for $n \geqslant 3$ remains open.

Assume that the distribution D is associated with some Lagrangian L with $\frac{\partial^2 f}{\partial y_n^2} \neq 0$. Then from the proof of Proposition 4.2 it follows that

$$\mathcal{R} = \Pi\big(\big\{\widetilde{\lambda} \in \big(D^2\big)^\perp \backslash \big(D^3\big)^\perp \colon v(\widetilde{\lambda}) \neq 0\big\}\big).$$

It implies that the manifold $\mathcal{H}$ introduced in Section 3 can be identified with $\mathcal{R}$. Let $\phi_Z : \mathcal{H}_Z \to \mathcal{H}_Z/\operatorname{Fol}(\vec{H}_Z)$ be the canonical projection. Then, as was already mentioned in Section 3.2 the characteristic distribution $\widetilde{C}$ is reduced to the line distribution $\bar{C} = (\phi_Z)_* C$ on $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$. Moreover, the filtration $\{\mathcal{J}^{(i)}\}_{i\in\mathbb{Z}}$ on $\mathcal{H}$ induces the filtration $\{\bar{\mathcal{J}}^{(i)}\}_{i\in\mathbb{Z}}$ on $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$. Note that in this notation the distribution $\bar{\mathcal{J}}^{(0)}$ coincides with $\bar{\mathcal{J}}$ defined by (3.16). If Υ_1 and Υ_2 are two abnormal extremals on $\mathcal{H}$ such that $\phi_Z(\Upsilon_1) = \phi_Z(\Upsilon_2)$, then the curves $J_{\Upsilon_1}^{(1-n)}$ and $J_{\Upsilon_2}^{(1-n)}$ coincide up to a projective transformation. As a matter of fact they coincide, up to a projective transformation, with the linearization of the distribution $\bar{\mathcal{J}}^{(1-n)}$ along $\bar{\Upsilon} = (\phi_Z)_*\Upsilon_1$ on $\mathcal{H}/\operatorname{Fol}(\vec{H}_Z)$. From this, Lemma 4.1, Remark 4.3, and (4.5) we have the following

Proposition 4.3. *Given a solution Γ of the Euler–Lagrange equation $\mathcal{E}(L)$ and the abnormal extremal Υ such that $\phi_Z(\Upsilon) = \mathfrak{L}(\Gamma)$, where $\mathfrak{L}$ is a Legendre transform, the curves $\operatorname{Lin}^1_\Gamma$ and $J_\Upsilon^{(1-n)}$ coincide up to a projective transformation.*

5. Fundamental invariants for our equivalence problems

Now we are ready to prove the following

Theorem 5.1. *For any $n \geqslant 3$ a $(2, n+3)$-distribution D associated with a non-degenerate Lagrangian $f(x, y, y', \ldots, y^{(n)})\,dx$ (or the Lagrangian L itself) is locally equivalent to the distribution associated with the Lagrangian $(y^{(n)})^2\,dx$ if and only if any of the following two equivalent conditions are satisfied*:

(1) *all Jacobi curves of the distribution D are flat*;
(2) *all generalized Wilczynski invariants of D vanish identically.*

The proof of the theorem immediately follows from Theorem 4.1 and the following

Theorem 5.2. *Let $y^{(2n)} = F(x, y, y', \ldots, y^{(2n-1)})$ be the Euler–Lagrange equation of the non-degenerate Lagrangian $L = f(x, y, y', \ldots, y^{(n)})\,dx$. This equation is contact equivalent to the trivial equation $y^{(2n)} = 0$ if and only if all its generalized Wilczynski invariants $W_4, W_6, \ldots, W_{2n}$ vanish identically.*

Proof. In the sequel we will denote by F_i the partial derivative F_{y_i}. The higher order derivatives of F will be denoted in a similar way. According to [11], any ordinary differential equation of order $2n$, $n \geqslant 3$, is trivializable if and only if all its generalized Wilczynski invariants vanish identically and, in addition, the following conditions hold:

$$F_{55} = F_{45} = 0, \quad \text{for } n = 3; \tag{5.1}$$

$$F_{2n-1,2n-1} = F_{2n-1,2n-2} = F_{2n-2,2n-2} = 0, \quad \text{for } n \geqslant 4. \tag{5.2}$$

Since our equation is variational, its generalized Wilczynski invariants of odd degree vanish automatically. The generalized Wilczynski invariants of even degree vanish by assumption of the lemma. So, we need to prove that the above conditions also hold.

Since the equation $y^{(2n)} = F(x, y, y', \dots, y^{(2n-1)})$ is variational, then according to [3] F is a polynomial in $y^{(n+1)}, \dots, y^{(2n-1)}$ of weighted degree $\leqslant n$, where these derivatives have weights $1, \dots, n-1$ respectively. In particular, we see that the polynomials $(y^{(2n-1)})^2$, $y^{(2n-1)}y^{(2n-2)}$, $(y^{(2n-2)})^2$ have weighted degree $2n-2$, $2n-3$ and $2n-4$ respectively. Assume that $n \geqslant 5$. Then $2n-4 > n$, and these terms cannot appear in F. Thus, the condition (5.2) holds automatically for $n \geqslant 5$. So, it remains to consider only the cases $n = 3, 4$.

Let $n = 3$. Then the term $y^{(5)}$ has weighted degree 2, while the function F is of weighted degree $\leqslant 3$. So, we see that $F_{55} = 0$. Let us prove that the condition $W_4 = 0$ implies also that $F_{45} = 0$. The direct computation shows that $I = F_{45} = -3f_{333}/f_{33}$. Let us denote W_4 simply by W. From [11] we have:

$$W = -\frac{5}{36}F_{5xxx} + \frac{2}{21}F_4F_{55} - \frac{5}{12}F_{3x} + \frac{1}{3}F_{4xx} + \frac{5}{18}F_5F_{5xx} + \frac{5}{36}F_3F_5 - \frac{5}{21}F_5^2F_{5x} - \frac{37}{126}F_4F_{5x} + \frac{5}{252}F_5^4 + \frac{37}{630}F_4^2 + \frac{25}{84}F_{5x}^2 + \frac{5}{18}F_2 - \frac{5}{18}F_5F_{4x},$$

where F_i denotes the partial derivative by $y^{(i)}$ and F_x denotes the total derivative. Then the direct computation shows that:

$$W_{55} = \frac{1}{35}\frac{57f_{333}^2 - 35f_{33}f_{3333}}{f_{33}^2};$$

$$W_{355} = -\frac{1}{35}\frac{35f_{33}f_{33333} - 149f_{33}f_{333}f_{3333} + 144f_{333}^2}{f_{33}^3};$$

$$W_{445} = -\frac{2}{35}\frac{35f_{33}f_{33333} - 162f_{33}f_{333}f_{3333} + 135f_{333}^2}{f_{33}^3},$$

and we have the following syzygy:

$$210W_{355} - 105W_{445} + 26IW_{55} - \frac{4}{105}I^3 = 0.$$

In particular, if $W = 0$ we get $I^3 = 0$ and hence $I = 0$. In particular, the function f is actually quadratic in the highest derivative. This proves the case $n = 3$.

For $n = 4$, we see that $F_{77} = F_{76} = 0$ due to the weighted degree argument. So, it remains to prove that $I = F_{66} = -6f_{444}/f_{44}$ vanishes if all generalized Wilczynski invariants vanish identically. Again, denote by $W = W_4$ the first non-trivial Wilczynski invariant for Euler–Lagrange equation. Then we have:

$$W = \frac{35}{528}F_5F_7 + \frac{49}{176}F_7F_{7xx} + \frac{7}{22}F_{6xx} - \frac{35}{176}F_7F_{6x} - \frac{1127}{6336}F_7^2F_{7x} + \frac{161}{3168}F_6F_7^2 + \frac{931}{3168}F_{7x}^2 + \frac{7}{66}F_4 + \frac{47}{1584}F_6^2 + \frac{1127}{101376}F_7^4 - \frac{329}{1584}F_6F_{7x} - \frac{49}{264}F_{7xxx}.$$

Again, direct computation shows that:

$$W_{75} = -\frac{7}{66}\frac{8f_{44}f_{4444} - 13f_{444}^2}{f_{44}^2};$$

$$W_{66} = -\frac{1}{198}\frac{252f_{44}f_{4444} - 437f_{444}^2}{f_{44}^2},$$

and we have the following syzygy:

$$3W_{75} - 2W_{66} + \frac{5}{648}I^2 = 0.$$

Hence, the equality $W = 0$ implies also that $I = 0$. Again, we see that vanishing of the first non-trivial generalized Wilczynski invariant implies that the function f is actually quadratic in the highest derivative. This completes the case $n = 4$ and the proof of the lemma. □

Remark 5.1. Note that for $n = 2$ Theorem 5.2 does not hold. In particular, the Lagrangian $L = (y'')^{1/3}\,dx$ has trivial Wilczynski invariants, but the associated Euler–Lagrange equation $3y''y^{(4)} - 5(y''')^2 = 0$ is not trivializable and has only 6-dimensional contact symmetry group consisting of all affine transformations on the plane (see [26, Chapter 6]). On the other hand, Theorem 5.1 is valid for $n = 2$ and not only for distributions associated with second order Lagrangians but for any rank 2 distribution in $\mathbb{R}^5$ with the small growth vector $(2,3,5)$. The rank 2 distributions D corresponding to the equations $z' = (y'')^{1/3}$ and $z' = (y'')^2$ are equivalent and have the 14-dimensional symmetry algebra, which is the maximal possible algebra for distributions under the consideration.

Remark 5.2. Direct analysis of the symmetry classification of all ordinary differential equations (see [26]) shows that all Euler–Lagrange equations of order $2n$ with the symmetry algebra of dimension at least $2n+1$ are exhausted (modulo contact transformations) by the following ones:

Equation $\mathcal{E}$	Lagrangian L	$\dim \operatorname{sym}(\mathcal{E})$
$y_{2n} = 0$	$y_n^2\,dx$	$2n+4$
$y_n^2 + \sum_{i=0}^{n-1} c_i y_{2i} = 0$	$(y_n^2 + \sum_{i=0}^{n-1} c_i y_i^2)\,dx$	$2n+2$
$9y_3^2 y_6 - 45 y_3 y_4 y_5 + 40 y_4^3 = 0$	$y_3^{1/3}\,dx$	7

Note also that we always have $\dim \operatorname{sym}(L) = \dim \operatorname{sym}(\mathcal{E}) + 1$.

Acknowledgements

We would like to thank Andrei Agrachev, Ian Anderson, Mark Fels, Eugene Ferapontov for valuable discussions on the subject of this paper.

References

[1] A. Agrachev, Yu. Sachkov, Control Theory from the Geometric Viewpoint. Control Theory and Optimization, II, Encyclopaedia of Mathematical Sciences, vol. 87, Springer-Verlag, Berlin, 2004, xiv+412 pp.
[2] I. Anderson, B. Kruglikov, Rank 2 distributions of Monge equations: symmetries, equivalences, extensions, arXiv:0910.5946v1.
[3] I. Anderson, G. Tompson, The inverse problem of the calculus of variations for ordinary differential equations, Mem. Amer. Math. Soc. 98 (1992).
[4] R. Bryant, On notions of equivalence of variational problems with one independent variable, in: Differential Geometry: The Interface Between Pure and Applied Mathematics, San Antonio, Tex., 1986, in: Contemp. Math., vol. 68, Amer. Math. Soc., Providence, RI, 1987, pp. 65–76.
[5] E. Cartan, Les systemes de Pfaff a cinque variables et les equations aux derivees partielles du second ordre, Ann. Sci. Ec. Norm. 27 (1910) 109–192; See also E. Cartan, Oeuvres Completes, Partie II, vol. 2, Gautier–Villars, Paris, 1953, pp. 927–1010.
[6] E. Cartan, Sur les variétés à connexion projective, Bull. Soc. Math. France 52 (1924) 205–241.
[7] S.-S. Chern, The geometry of the differential equation $y''' = F(x, y, y', y'')$, Sci. Rep. Nat. Tsing Hua Univ. 4 (1950) 97–111.
[8] J. Chrastina, On the equivalence of variational problems. I, J. Differential Equations 98 (1992) 76–90.
[9] J. Chrastina, The Formal Theory of Differential Equations, Masaryk University, Brno, 1998.
[10] B. Doubrov, B. Komrakov, T. Morimoto, Equivalence of holonomic differential equations, Lobachevskij J. Math. 3 (1999) 39–71.
[11] B. Doubrov, Contact trivialization of ordinary differential equations, in: Differential Geometry and Its Applications, Proc. Conf., Opava (Czech Republic), Silestian University, Opava, 2001, pp. 73–84.
[12] B. Doubrov, Generalized Wilczynski invariants for non-linear ordinary differential equations, in: Symmetries and Overdetermined Systems of Partial Differential Equations, in: IMA, vol. 144, Springer, New York, 2008, pp. 25–40.
[13] B. Doubrov, I. Zelenko, A canonical coframe of non-holonomic rank two distributions of maximal class, C. R. Acad. Sci. Paris, Ser. I 342 (8) (2006) 589–594.
[14] B. Doubrov, I. Zelenko, On local geometry of nonholonomic rank 2 distributions, J. Lond. Math. Soc. 80 (3) (2009) 545–566.
[15] M. Fels, The inverse problem of the calculus of variations for scalar fourth-order ordinary differential equations, Trans. Amer. Math. Soc. 348 (1996) 5007–5029.
[16] A. Gonzalez-Lopez, Symmetry bounds of variational problems, J. Phys. A 27 (1994) 1205–1232.
[17] E. Goursat, Lecons sur le problem de Pfaff, Hermann, Paris, 1923.
[18] D. Hilbert, Über den Begriff der Klasse von Differentialgleichungen, Math. Ann. 73 (1912) 95–108.
[19] M. Juráš, The inverse problem of the calculus of variations for sixth- and eighth-order scalar ordinary differential equations, Acta Appl. Math. 66 (2001) 25–39.
[20] N. Kamran, P. Olver, Equivalence of higher order Lagrangians. I. Formulation and reduction, J. Math. Pures Appl. 70 (1991) 369–391.
[21] L. Hsu, N. Kamran, P. Olver, Equivalence of higher-order Lagrangians. II. The Cartan form for particle Lagrangians, J. Math. Phys. 30 (1989) 902–906.
[22] N. Kamran, P. Olver, Equivalence of higher-order Lagrangians. III. New invariant differential equations, Nonlinearity 5 (1992) 601–621.
[23] A. Moor, Über äquivalente Variationsprobleme erster und zweiter Ordnung, J. Reine Angew. Math. 223 (1966) 131–137.
[24] T. Morimoto, Geometric structures on filtered manifolds, Hokkaido Math. J. 22 (1993) 263–347.
[25] P. Nurowski, Differential equations and conformal structures, J. Geom. Phys. 55 (2005) 19–49.
[26] P.J. Olver, Equivalence, Invariants, and Symmetry, Cambridge University Press, Cambridge, 1995.
[27] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mischenko, The Mathematical Theory of Optimal Processes, Wiley, New York, 1962.
[28] N. Tanaka, On differential systems, graded Lie algebras and pseudo-groups, J. Math. Kyoto Univ. 10 (1970) 1–82.
[29] N. Tanaka, On the equivalence problems associated with simple graded Lie algebras, Hokkaido Math. J. 6 (1979) 23–84.
[30] E.J. Wilczynski, Projective Differential Geometry of Curves and Ruled Surfaces, Teubner, Leipzig, 1905.
[31] K. Yamaguchi, Differential systems associated with simple graded lie algebras, Adv. Stud. Pure Math. 22 (1993) 413–494.
[32] I. Zelenko, Nonregular abnormal extremals of 2-distributions: existence, second variation, and rigidity, J. Dyn. Control Syst. 5 (1999) 347–383.
[33] I. Zelenko, Complete systems of invariants of rank 1 curves in Lagrange Grassmanians, in: Differential Geom. Application, Proc. Conf. Prague, 2005, pp. 365–379.
[34] I. Zelenko, Variational approach to differential invariants of rank 2 vector distributions, Differential Geom. Appl. 24 (2006) 235–259.
[35] I. Zelenko, Fundamental form and Cartan's tensor of $(2,5)$-distributions coincide, J. Dyn. Control Syst. 12 (2) (2006) 247–276.