A new geometrical look at Ostrogradsky's procedure

Enrico Massa,[1, a)] Stefano Vignolo,[1, b)] Roberto Cianci,[1, c)] and Sante Carloni[2, d)]

[1)] *DIME, Sez. Metodi e Modelli Matematici, Università di Genova,*
Piazzale Kennedy, Pad. D, 16129, Genova - Italy
[2)] *Centro Multidisciplinar de Astrofisica - CENTRA,*
Instituto Superior Tecnico - IST,
Universidade de Lisboa - UL, Avenida Rovisco Pais 1,
1049-001 - Portugal.[e)]

Making use of the modern techniques of non-holonomic geometry and constrained variational calculus, a revisitation of Ostrogradsky's Hamiltonian formulation of the evolution equations determined by a Lagrangian of order ≥ 2 in the derivatives of the configuration variables is presented.

I. INTRODUCTION

About twenty years after the first formulation of Hamiltonian mechanics, Ostrogradsky proposed a generalisation of Hamilton's procedure, valid for Lagrangians involving derivatives of any order of the configuration variables[1–3].

In recent years, the study of this type of Lagrangians has been reconsidered in the context of gravitational physics and, in particular, in the development of a theoretical framework for inflation and dark energy based on modifications of General Relativity (see e.g.[4–6]).

Despite this renewed interest, to the best of the authors' knowledge, a precise geometric interpretation of Ostrogradsky's construction is still missing. In an attempt to fill this gap, we propose here a reformulation of Ostrogradsky's formalism in modern geometrical terms.

Given the event space, meant as a fibre bundle $\mathcal{V}_{n+1} \overset{t}{\to} \mathbb{R}$, we regard the Nth jet bundle $j_N(\mathcal{V}_{n+1})$ as an affine subbundle of the first jet $j_1(j_{N-1}(\mathcal{V}_{n+1}))$[7,8]. In this way, any problem involving a Lagrangian depending on the derivatives of order $\leq N$ of the configuration variables is converted into an ordinary *constrained* variational problem.

The problem is then analysed, making use of a revisitation of Pontryagin's *maximum principle* recently developed in[9] (in this connection, see also[10,11] and references therein). In the case of a non–degenerate Lagrangian $L(t, q^k, \dot{q}^k, \ddot{q}^k, \dots)$, the algorithm picks out a natural concept of "phase space", identifying it with a submanifold $\mathcal{S}$ of the contact bundle over $j_N(\mathcal{V}_{n+1})$, uniquely determined by the Pontryagin Hamiltonian associated with L.

In the resulting environment, the canonical momenta and the Ostrogradsky Hamiltonian are simply the pull–back of the coordinate functions along the fibres of the contact bundle and of the Pontryagin Hamiltonian, while the Ostrogradsky equations reproduce the Hamilton–Pontryagin equations associated with the constrained variational problem.

The layout of the paper is the following: in Section II, the geometrical setup for constrained variational calculus is briefly reviewed; Section III is then devoted to the geometric reformulation of the Ostrogradsky procedure.

II. CONSTRAINED VARIATIONAL CALCULUS

In this section, we briefly review the geometrical setup for constrained variational calculus outlined in [9]. The basic environment is a $(n+1)$–dimensional *event space*, meant as a fiber bundle $t : \mathcal{V}_{n+1} \to \mathbb{R}$ and referred to local fibred coordinates $t, q^1, \ldots, q^n$.

Every section $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$, locally described as $q^i = q^i(t)$, is interpreted as an evolution of an abstract system $\mathcal{B}$ with n degrees of freedom: for instance, if $\mathcal{B}$ is a mechanical system, the manifold $\mathcal{V}_{n+1}$ is identified with the *configuration space–time* of $\mathcal{B}$, and the fibration $t : \mathcal{V}_{n+1} \to \mathbb{R}$ with the absolute time function.

The first jet bundle $j_1(\mathcal{V}_{n+1})$, referred to local jet coordinates $t, q^i, \dot{q}^i$, is called the *kinetic space*. Every section $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$ admits a corresponding lift $j_1(\gamma) : \mathbb{R} \to j_1(\mathcal{V}_{n+1})$, locally expressed as $q^i = q^i(t), \dot{q}^i = \frac{dq^i}{dt}$.

The presence of non–holonomic constraints is accounted for through the assignment of a submanifold $i : \mathcal{A} \to j_1(\mathcal{V}_{n+1})$ fibred over $\mathcal{V}_{n+1}$, as described by the commutative diagram

$$
\begin{array}{ccc}
\mathcal{A} & \xrightarrow{\ i\ } & j_1(\mathcal{V}_{n+1}) \\
{\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\
\mathcal{V}_{n+1} & =\!=\!=\!= & \mathcal{V}_{n+1}
\end{array}
\tag{1}
$$

Referring $\mathcal{A}$ to local fibered coordinates t, q^i, z^A $(A = 1, \ldots, r < n)$, the embedding $i : \mathcal{A} \to j_1(\mathcal{V}_{n+1})$ is locally represented as

$$
\dot{q}^i = \psi^i(t, q^1, \ldots, q^n, z^1, \ldots, z^r)
\tag{2}
$$

with rank $\left\| \frac{\partial \psi^i}{\partial z^A} \right\| = r$.

A section $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$ is called *admissible* if and only if there exists a section $\hat{\gamma} : \mathbb{R} \to \mathcal{A}$ satisfying $j_1(\gamma) = i \cdot \hat{\gamma}$. A section $\hat{\gamma} : \mathbb{R} \to \mathcal{A}$ is similarly called admissible if and only if $i \cdot \hat{\gamma} = j_1(\pi \cdot \hat{\gamma})$. In coordinates, if $\hat{\gamma}$ is represented as $q^i = q^i(t), z^A = z^A(t)$, the admissibility condition is summarized into the system of first order ODE's

$$
\frac{dq^i}{dt} = \psi^i(t, q^1(t), \ldots, q^n(t), z^1(t), \ldots, z^r(t))
\tag{3}
$$

The geometry of the submanifold $\mathcal{A}$ has been extensively studied in the context of non–holonomic mechanics (see, among others,[12,13] and references therein). For the present purposes, we recall the concept of *contact bundle* $\pi : \mathcal{C}(\mathcal{A}) \to \mathcal{A}$, meant as the vector sub-bundle of the cotangent space $T^*(\mathcal{A})$ locally spanned by the contact 1-forms

$$
\omega^i := dq^i - \psi^i(t, q^k, z^A)\, dt
\tag{4}
$$

Adopting the notation $V(\mathcal{V}_{n+1}) \subset T(\mathcal{V}_{n+1})$ for the vertical bundle relative to the fibration $\mathcal{V}_{n+1} \xrightarrow{t} \mathbb{R}$ and $V^*(\mathcal{V}_{n+1})$ for the associated dual bundle, commonly referred to as the *phase space*, the contact bundle is canonically isomorphic to the fibred product $\mathcal{A} \times_{\mathcal{V}_{n+1}} V^*(\mathcal{V}_{n+1})$, giving rise to the fibred morphism

$$
\begin{array}{ccc}
\mathcal{C}(\mathcal{A}) & \xrightarrow{\ \pi\ } & V^*(\mathcal{V}_{n+1}) \\
{\scriptstyle \pi}\downarrow & & \downarrow{\scriptstyle \pi} \\
\mathcal{A} & \xrightarrow[\ \pi\ ]{} & \mathcal{V}_{n+1}
\end{array}
\tag{5}
$$

We refer $\mathcal{C}(\mathcal{A})$ to fibred coordinates t, q^i, z^A, p_i, defined according to the identification $\sigma = p_i(\sigma)\omega^i\big|_{\pi(\sigma)} \ \forall\, \sigma \in \mathcal{C}(\mathcal{A})$.

An important geometrical attribute of the contact bundle is its *Liouville* 1-*form* Θ, locally expressed as[15]

$$\Theta := p_i\,\omega^i = p_i\big[dq^i - \psi^i\left(t, q^k, z^A\right)\,dt\big] \tag{6}$$

The geometrical framework outlined above provides the mathematical setting for an intrinsic formulation of constrained variational calculus. To this end, we consider an action functional of the form

$$\mathcal{I}[\gamma] := \int_{\hat\gamma} L\,dt = \int_{t_0}^{t_1} L(t, q^i(t), z^A(t))\,dt \tag{7}$$

assigning to each admissible section $\gamma : \mathbb{R} \to \mathcal{V}_{n+1}$ a corresponding "cost", expressed as the integral of a *Lagrangian function* $L(t, q^i, z^A) \in \mathcal{F}(\mathcal{A})$ along the lift $\hat\gamma : \mathbb{R} \to \mathcal{A}$. The aim is studying the (local) extremals of the functional (7) with respect to admissible deformations of γ leaving the endpoints $\gamma(t_0), \gamma(t_1)$ fixed.

The result may be achieved observing that, under very general assumptions, the original problem is mathematically equivalent to a *free* variational problem on the contact bundle.

The procedure, outlined in [9], relies on the fact that, by means of the Liouville 1-form (6), every Lagrangian $L(t, q^i, z^A) \in \mathcal{F}(\mathcal{A})$ may be lifted to a 1-form ϑ_L over $\mathcal{C}(\mathcal{A})$ according to the prescription

$$\vartheta_L := L\,dt + \Theta = (L - p_i\psi^i)\,dt + p_i\,dq^i := -\mathcal{H}\,dt + p_i\,dq^i \tag{8}$$

The function $\mathcal{H}(t, q^k, z^A, p_k) = p_i\,\psi^i(t, q^k, z^A) - L(t, q^k, z^A) \in \mathcal{F}(\mathcal{C}(\mathcal{A}))$ is known in the literature as the *Pontryagin Hamiltonian*. By means of the 1–form (8), to each section $\bar\gamma : [t_0, t_1] \to \mathcal{C}(\mathcal{A})$, expressed in coordinates as $q^i = q^i(t), z^A = z^A(t), p_i = p_i(t))$, we assign the action functional

$$\bar{\mathcal{I}}[\bar\gamma] := \int_{\bar\gamma} \vartheta_L = \int_{t_0}^{t_1} \left[L(t, q^k(t), z^A(t)) + p_i(t)\left(\frac{dq^i}{dt} - \psi^i(t, q^k(t), z^A(t))\right)\right] dt \tag{9}$$

The resulting setup is closely related to the original one, based on the functional (7) and on the constraints (3). In fact, denoting by $\nu : \mathcal{C}(\mathcal{A}) \to \mathcal{V}_{n+1}$ the composite projection $\mathcal{C}(\mathcal{A}) \to \mathcal{A} \to \mathcal{V}_{n+1}$, it turns out that every "ordinary" extremal of the original problem is the projection $\gamma = \nu \cdot \bar\gamma$ of a solution of the free variational problem based on the functional (9)[16]. More specifically, the requirement of stationarity of the action integral (9) under arbitrary deformations leaving the projections $\nu(\bar\gamma(t_0))$, $\nu(\bar\gamma(t_1))$ fixed leads to $2n+r$ equations

$$\frac{dq^i}{dt} = \psi^i(t, q^k, z^A) = \frac{\partial \mathcal{H}}{\partial p_i} \tag{10a}$$

$$\frac{dp_i}{dt} = \frac{\partial L}{\partial q^i} - p_k\frac{\partial \psi^k}{\partial q^i} = -\frac{\partial \mathcal{H}}{\partial q^i} \tag{10b}$$

$$p_i\frac{\partial \psi^i}{\partial z^A} - \frac{\partial L}{\partial z^A} = \frac{\partial \mathcal{H}}{\partial z^A} = 0 \tag{10c}$$

for the unknowns $q^i(t), z^A(t), p_i(t)$, identical to the *Pontryagin equations*[9–11] involved in the study of the constrained functional (7).

In order to work out the content of the system (10), it is convenient to start with eq. (10c). The latter identifies a subset of $\mathcal{C}(\mathcal{A})$, henceforth denoted by $\mathcal{S}$. The Hamiltonian $\mathcal{H}$ is called *regular* if and only if the condition

$$\det\left(\frac{\partial^2 \mathcal{H}}{\partial z^A \partial z^B}\right)_\sigma \neq 0 \tag{11}$$

holds for all $\sigma \in \mathcal{S}$. When this is the case, eqs. (10c) may be uniquely solved for the variables z^A, giving rise to a representation of the form

$$z^A = z^A(t, q^i, p_i) \tag{12}$$

Under the stated assumption, the subset $\mathcal{S}$ is therefore a $(2n+1)$–dimensional submanifold $i : \mathcal{S} \to \mathcal{C}(\mathcal{A})$, locally diffeomorphic to the phase space $V^*(\mathcal{V}_{n+1})$.

The pull–back $H := i^*(\mathcal{H})$ of the Pontryagin Hamiltonian, expressed in coordinates as

$$H(t, q^i, p_i) = p_h \, \psi^h(t, q^i, z^A(t, q^k, p_k)) - L(t, q^i, z^A(t, q^k, p_k)) \tag{13}$$

yields a proper Hamiltonian function on $\mathcal{S}$. Through the latter, eqs. (10a), (10b) may be written as ordinary Hamilton equations. On account of eqs. (10c) we have in fact the identifications

$$\frac{\partial H}{\partial p_i} \;=\; \frac{\partial \mathcal{H}}{\partial p_i} \;=\; \psi^i \tag{14a}$$

$$\frac{\partial H}{\partial q^i} \;=\; \frac{\partial \mathcal{H}}{\partial q^i} \;=\; p_k \frac{\partial \psi^k}{\partial q^i} - \frac{\partial L}{\partial q^i} \tag{14b}$$

allowing to cast eqs. (10a), (10b) into the form

$$\frac{dq^i}{dt} \;=\; \frac{\partial H}{\partial p_i} \tag{15a}$$

$$\frac{dp_i}{dt} \;=\; -\frac{\partial H}{\partial q^i} \tag{15b}$$

The original constrained variational problem is thus reduced to a free Hamiltonian problem in the submanifold $\mathcal{S}$, with Hamiltonian $H(t, q^i, p_i)$ identical to the pull–back $H = i^*(\mathcal{H})$.

III. THE OSTROGRADSKY PROCEDURE REVISITED

In this section, we present a revisitation of Ostrogradsky's construction of a Hamiltonian setup for non–degenerate Lagrangians depending on higher order derivatives of the configuration variables[1,3]. The idea is regarding any such Lagrangian as a function on a submanifold of a suitable kinetic space, thereby reducing the original problem to a constrained one, of the kind described in the Section II. As we shall see, this will result into a self–consistent interpretation of Ostrogradsky's formalism in modern geometrical terms.

For the sake of simplicity, and to better clarify the basic ideas and notations, we shall first consider Lagrangians of order 2 in the derivatives. The procedure will then be extended to higher order Lagrangians.

A. Lagrangians of order 2 in the derivatives

Given a Lagrangian of the form $L(t, q^i, \dot{q}^i, \ddot{q}^i)$, the corresponding Euler–Lagrange equations read

$$\frac{\partial L}{\partial q^i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}^i} + \frac{d^2}{dt^2} \frac{\partial L}{\partial \ddot{q}^i} = 0, \qquad i = 1, \dots, n \tag{16}$$

Assuming the validity of non-degeneracy condition $\det\left\|\frac{\partial^2 L}{\partial \ddot{q}^i \partial \ddot{q}^j}\right\| \neq 0$, the Ostrogradsky procedure relies on adopting the functions

$$q^i, \quad \dot{q}^i, \quad p_i^0 := \frac{\partial L}{\partial \dot{q}^i} - \frac{d}{dt} \frac{\partial L}{\partial \ddot{q}^i}, \quad p_i^1 := \frac{\partial L}{\partial \ddot{q}^i} \tag{17}$$

as coordinates in a $(4n + 1)$–dimensional *phase space*, with p_i^0 and p_i^1 playing the role of canonical momenta respectively conjugate to the variables q^i and $\dot{q}^i$.

Under the stated non-degeneracy condition, the last set of equations (17) can be solved for the unknowns $\ddot{q}^i$, giving rise to a representation of the form

$$\ddot{q}^i = \ddot{q}^i \left(t, q^k, \dot{q}^k, p_k^1 \right) \tag{18}$$

In this way, introducing the *Ostrogradsky Hamiltonian*

$$H := p_i^0 \dot{q}^i + p_i^1 \ddot{q}^i - L, \tag{19}$$

expressed in terms of the variables $q^i, \dot{q}^i, p_i^0, p_i^1$ through eqs. (18), a straightforward calculation yields the relations

$$\frac{\partial H}{\partial p_i^0} = \dot{q}^i, \qquad \frac{\partial H}{\partial p_i^1} = \ddot{q}^i, \qquad \frac{\partial H}{\partial q^i} = -\frac{\partial L}{\partial q^i}, \qquad \frac{\partial H}{\partial \dot{q}^i} = p_i^0 - \frac{\partial L}{\partial \dot{q}^i}$$

In view of these, the content of eqs. (16), (17) may be cast into the Hamiltonian form

$$\frac{dq^i}{dt} = \dot{q}^i = \frac{\partial H}{\partial p_i^0}, \qquad\qquad\qquad \frac{d\dot{q}^i}{dt} = \ddot{q}^i = \frac{\partial H}{\partial p_i^1} \tag{20a}$$

$$\frac{dp_i^0}{dt} = \frac{d}{dt}\frac{\partial L}{\partial \dot{q}^i} - \frac{d^2}{dt^2}\frac{\partial L}{\partial \ddot{q}^i} = \frac{\partial L}{\partial q^i} = -\frac{\partial H}{\partial q^i}, \quad \frac{dp_i^1}{dt} = \frac{d}{dt}\frac{\partial L}{\partial \ddot{q}^i} = \frac{\partial L}{\partial \dot{q}^i} - p_i^0 = -\frac{\partial H}{\partial \dot{q}^i} \tag{20b}$$

To clarify the geometrical meaning of the Ostrogradsky procedure, we focus on the fiber bundle $\mathcal{V}_{n+1} \to \mathbb{R}$ and on the associated first jet bundle. For reasons that will be clear soon, we change the notation $j_1(\mathcal{V}_{n+1})$ into $\mathcal{Q}$, q^i into q_0^i, $\dot{q}^i$ into q_1^i, and regard the bundle $t : \mathcal{Q} \to \mathbb{R}$, referred to local coordinates t, q_α^i, $\alpha = 0, 1$ as our new event space.

By its very definition, the second jet bundle $j_2(\mathcal{V}_{n+1})$ is then canonically isomorphic to an affine subbundle of the first jet bundle $j_1(\mathcal{Q})$, as expressed by the commutative diagram (see e.g.[7,8], or also[14] for the time–independent case)

$$
\begin{array}{ccc}
j_2(\mathcal{V}_{n+1}) & \overset{i}{\longrightarrow} & j_1(\mathcal{Q}) \\
\pi \downarrow & & \downarrow \pi \\
\mathcal{Q} & =\!=\!=\!= & \mathcal{Q}
\end{array}
\tag{21}
$$

Referring $j_1(\mathcal{Q})$ to jet coordinates $t, q_\alpha^i, \dot{q}_\alpha^i$, the image $i(j_2(\mathcal{V}_{n+1})) \subset j_1(\mathcal{Q})$, henceforth denoted by $\mathcal{A}$, is locally described by the equations $\dot{q}_0^i = q_1^i$. We can therefore refer $\mathcal{A}$ to local fibred coordinates t, q_α^i, z^i, and represent the imbedding $\mathcal{A} \to j_1(\mathcal{Q})$ through the equations (analogous to eqs. (2))

$$\dot{q}_\alpha^i = \psi_\alpha^i(t, q_0^i, q_1^i, z^i), \tag{22}$$

with $\psi_0^i = q_1^i$ and $\psi_1^i = z^i$. Alternatively, we may regard $\mathcal{A}$ as a fiber bundle over $\mathcal{V}_{n+1}$, related to $j_2(\mathcal{V}_{n+1})$ by the fibred isomorphism $(t, q_0^i, q_1^i, z^i) \longleftrightarrow (t, q^i, \dot{q}^i, \ddot{q}^i)$.

Summing up, we conclude that assigning a variational problem in $\mathcal{V}_{n+1}$, based on a Lagrangian $L(t, q^i, \dot{q}^i, \ddot{q}^i) \in \mathcal{F}(j_2(\mathcal{V}_{n+1}))$, is equivalent to assigning a constrained variational problem in $\mathcal{Q}$, with constraint submanifold $\mathcal{A} \to j_1(\mathcal{Q})$ described by eqs. (22) and Lagrangian $L(t, q_\alpha^i, z^i) \in \mathcal{F}(\mathcal{A})$.

In the determination of the extremals, we can now proceed along the lines indicated in Section II. To this end, we consider once again the contact bundle $\mathcal{C}(\mathcal{A})$ over $\mathcal{A}$, and denote by $\nu : \mathcal{C}(\mathcal{A}) \to \mathcal{Q}$ the composite projection $\mathcal{C}(\mathcal{A}) \to \mathcal{A} \to \mathcal{Q}$, and by $t, q_\alpha^i, z^i, p_i^\alpha$ the local coordinates on $\mathcal{C}(\mathcal{A})$ defined by the prescription

$$\sigma = p_i^\alpha(\sigma) \left(dq_\alpha^i - \psi_\alpha^i dt \right)_{|\pi(\sigma)} \qquad \forall \sigma \in \mathcal{C}(\mathcal{A}) \tag{23}$$

the summation convention being henceforth extended to all type of indices.

Starting with the Lagrangian $L(t, q_\alpha^i, z^i) \in \mathcal{F}(\mathcal{A})$, we then construct the 1-form

$$\vartheta_L = L\,dt + p_i^\alpha \left(dq_\alpha^i - \psi_\alpha^i\,dt\right) = -\mathcal{H}\,dt + p_i^\alpha\,dq_\alpha^i \in \mathcal{C}(\mathcal{A})$$

with

$$\mathcal{H}(t, q_\alpha^i, z^i, p_i^\alpha) = p_i^\alpha \psi_\alpha^i - L(t, q_\alpha^i, z^i) = p_i^0 q_1^i + p_i^1 z^i - L(t, q_\alpha^i, z^i) \tag{24}$$

denoting the Pontryagin Hamiltonian.

Eventually, we assign to each section $\bar\gamma : [t_0, t_1] \to \mathcal{C}(\mathcal{A})$ the action functional

$$\bar{\mathcal{I}}[\bar\gamma] := \int_{\bar\gamma} \vartheta_L = \int_{t_0}^{t_1} \left(-\mathcal{H}(t, q_\alpha^i, z^i, p_i^\alpha) + p_i^\alpha \frac{dq_\alpha^i}{dt} \right) dt$$

The request of stationarity of the latter under arbitrary deformations leaving the points $\nu(\bar\gamma(t_0))$, $\nu(\bar\gamma(t_1))$ fixed leads to the Pontryagin equations

$$\frac{dq_\alpha^i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i^\alpha}, \qquad \frac{dp_i^\alpha}{dt} = -\frac{\partial \mathcal{H}}{\partial q_\alpha^i} \tag{25a}$$

$$\frac{\partial \mathcal{H}}{\partial z^i} = p_i^1 - \frac{\partial L}{\partial z^i} = 0 \tag{25b}$$

for the unknowns $q_\alpha^i(t), z^i(t), p_i^\alpha(t)$. These are the precise analogue of eqs. (10) for the case in study. In particular, eqs. (25b) reproduce the content of the last set of eqs. (17) under the morphism $(t, q_0^i, q_1^i, z^i) \longleftrightarrow (t, q^i, \dot q^i, \ddot q^i)$.

Denoting by $\mathcal{S}$ the subset of $\mathcal{C}(\mathcal{A})$ described by eqs. (25b) and taking eq. (24) into account, it is readily seen that the non–degeneracy condition $\det\left\|\frac{\partial^2 L}{\partial \ddot q^i \partial \ddot q^j}\right\| \neq 0$, here rephrased as $\det\left\|\frac{\partial^2 L}{\partial z^i \partial z^j}\right\| \neq 0$, automatically ensures the regularity of the Pontryagin Hamiltonian (24). We can therefore solve eqs. (25b) for the variables z^i, getting an expression of the form

$$z^i = z^i(t, q_\alpha^k, p_k^1) \tag{26}$$

formally identical to eq. (18)

Exactly as it happened in Section II, eq. (26) allows to regard $\mathcal{S}$ as a submanifold $\mathcal{S} \xrightarrow{i} \mathcal{C}(\mathcal{A})$, locally diffeomorphic to the phase space $V^*(\mathcal{Q})$. In view of eqs. (17), (25b), the pull back of the Pontryagin Hamiltonian (24) to the submanifold $\mathcal{S}$ yields the function

$$H(t, q_\alpha^i, p_i^\alpha) := p_i^0 q_1^i + p_i^1 z^i(t, q_\alpha^k, p_k^1) - L(t, q_\alpha^i, z^i(t, q_\alpha^k, p_k^1))$$

identical to the Ostrogradsky Hamiltonian (19) and satisfying the relations

$$\frac{\partial H}{\partial p_i^\alpha} = \frac{\partial \mathcal{H}}{\partial p_i^\alpha}, \qquad \frac{\partial H}{\partial q_\alpha^i} = \frac{\partial \mathcal{H}}{\partial q_\alpha^i}$$

On account of the latter, eqs. (25a) may be cast into the canonical Hamiltonian form

$$\frac{dq_0^i}{dt} = \frac{\partial H}{\partial p_i^0}, \qquad \frac{dq_1^i}{dt} = \frac{\partial H}{\partial p_i^1}$$

$$\frac{dp_i^0}{dt} = -\frac{\partial H}{\partial q_0^i}, \qquad \frac{dp_i^1}{dt} = -\frac{\partial H}{\partial q_1^i}$$

reproducing the content of the Ostrogradsky equations (20).

B. Lagrangians of order N in the derivatives

The Ostrogradsky construction is easily extended to Lagrangians depending on higher order derivatives. To this end, let $j_N(\mathcal{V}_{n+1})$ denote the Nth jet–bundle of the event space, referred to fibred coordinates $t, q^i, q_1^i, \ldots, q_N^i$. Setting $q_0^i = q^i$, the Euler–Lagrange equations associated with a Lagrangian $L(t, q^i, q_1^i, \ldots, q_N^i)$ are synthetically written as

$$\sum_{\alpha=0}^{N} (-1)^\alpha \frac{d^\alpha}{dt^\alpha} \frac{\partial L}{\partial q_\alpha^i} = 0, \qquad i = 1, \ldots, n \tag{27}$$

For each $\alpha = 0, \ldots, N-1$, let the canonical momentum p_i^α conjugate to the coordinate q_α^i be defined according to the prescription

$$p_i^\alpha := \sum_{\beta=\alpha}^{N-1} (-1)^{\beta-\alpha} \frac{d^{\beta-\alpha}}{dt^{\beta-\alpha}} \frac{\partial L}{\partial q_{\beta+1}^i} \tag{28}$$

whence, in particular

$$p_i^{N-1} := \frac{\partial L}{\partial q_N^i}(t, q^i, q_1^i, \ldots, q_N^i) \tag{28'}$$

The variables $t, q_\alpha^i, p_i^\alpha$, $\alpha = 0, \ldots, N-1$ are regarded as coordinates in a $(2nN+1)$–dimensional *phase space*.

Under the non-degeneracy assumption $\det \left\| \dfrac{\partial^2 L}{\partial q_N^i \partial q_N^j} \right\| \neq 0$, eqs. (28') may be solved for q_N^i, giving rise to expressions of the form

$$q_N^i = q_N^i(t, q_0^k, \ldots, q_{N-1}^k, p_k^{N-1}) \tag{29}$$

The Ostrogradsky Hamiltonian is then defined as

$$H(t, q_0^i, \ldots, q_{N-1}^i, p_i^0, \ldots, p_i^{N-1}) := \sum_{\alpha=0}^{N-1} p_i^\alpha q_{\alpha+1}^i - L(t, q_0^i, \ldots, q_{N-1}^i, q_N^i) \tag{30}$$

with q_N^i given by eq. (29). In this way, taking eqs. (28') into account, the Hamilton equations generated by the Hamiltonian (30) are easily recognized to be

$$\frac{dq_\alpha^i}{dt} = \frac{\partial H}{\partial p_i^\alpha} = q_{\alpha+1}^i \quad (\alpha = 0, \ldots, N-2), \tag{31a}$$

$$\frac{dq_{N-1}^i}{dt} = \frac{\partial H}{\partial p_i^{N-1}} = q_N^i\left(t, q_0^k, \ldots, q_{N-1}^k, p_k^{N-1}\right) \tag{31b}$$

$$\frac{dp_i^0}{dt} = -\frac{\partial H}{\partial q_0^i} = \frac{\partial L}{\partial q_0^i}, \qquad \frac{dp_i^\alpha}{dt} = -\frac{\partial H}{\partial q_\alpha^i} = -p_i^{\alpha-1} + \frac{\partial L}{\partial q_\alpha^i} \quad (\alpha = 2, \ldots, N) \tag{31c}$$

Conversely, a straightforward check shows that eqs. (31c), together with eqs. (28'), imply the validity of the Euler–Lagrange equations (27).

A deeper insight into the geometrical meaning of eqs. (28), (30) is gained denoting by $\mathcal{Q} := j_{N-1}(\mathcal{V}_{n+1})$ the $(N-1)^{th}$ jet bundle of the fibration $t : \mathcal{V}_{n+1} \to \mathbb{R}$, regarded as a fibre bundle $t : \mathcal{Q} \to \mathbb{R}$, and by $j_1(\mathcal{Q})$ the corresponding first jet bundle.

The Nth jet bundle $j_N(\mathcal{V}_{n+1})$ is then canonically isomorphic to an affine subbundle of $j_1(\mathcal{Q})$, as summarized into the commutative diagram

$$
\begin{array}{ccc}
j_N(\mathcal{V}_{n+1}) & \xrightarrow{\ i\ } & j_1(\mathcal{Q}) \\
\pi \downarrow & & \downarrow \pi \\
\mathcal{Q} & =\!=\!= & \mathcal{Q}
\end{array}
$$

Adopting t, q_α^i, $(\alpha = 0, \ldots, N-1)$ as local coordinates in $\mathcal{Q}$, and referring $j_1(\mathcal{Q})$ to jet coordinates $t, q_\alpha^i, \dot{q}_\alpha^i$, the submanifold $\mathcal{A} := i(j_N(\mathcal{V}_{n+1})) \subset j_1(\mathcal{Q})$ is locally described by the equations $\dot{q}_\alpha^i = q_{\alpha+1}^i$, $\alpha = 0, \ldots, N-2$. We can therefore refer $\mathcal{A}$ to local fibred coordinates $t, q_0^i, \ldots, q_{N-1}^i, z^i$, and represent the imbedding $\mathcal{A} \to j_1(\mathcal{Q})$ through the equations

$$\dot{q}_\alpha^i = \psi_\alpha^i(t, q_0^i, \ldots, q_{N-1}^i, z^i), \qquad \alpha = 0, \ldots, N-1 \tag{32}$$

with $\psi_\alpha^i = q_{\alpha+1}^i$, $\alpha = 0, \ldots, N-2$, and $\psi_{N-1}^i = z^i$.

Alternatively, we may regard $\mathcal{A}$ as a fiber bundle over $\mathcal{V}_{n+1}$, isomorphic to $j_N(\mathcal{V}_{n+1})$ through the fibred morphism $(t, q_0^i, \ldots, q_{N-1}^i, z^i) \longleftrightarrow (t, q_0^i, \ldots, q_{N-1}^i, q_N^i)$.

Once again, collecting all results, we conclude that assigning a variational problem in $\mathcal{V}_{n+1}$, based on a Lagrangian $L(t, q_0^i, , \ldots, q_N^i) \in \mathcal{F}(j_N(\mathcal{V}_{n+1}))$, is equivalent to assigning a constrained variational problem in $\mathcal{Q}$, with constraint submanifold $\mathcal{A} \to j_1(\mathcal{Q})$ described by eqs. (32) and Lagrangian $L(t, q_0^i, \ldots, q_{N-1}^i, z^i) \in \mathcal{F}(\mathcal{A})$.

The constrained problem in $\mathcal{Q}$ may then be lifted to a free variational problem on the contact bundle $\mathcal{C}(\mathcal{A})$, referred to fibred coordinates $t, q_\alpha^i, z^i, p_i^\alpha$, $\alpha = 0, \ldots, N-1$. The procedure, identical to the one exploited in Sect. III A, culminates in the introduction of the Pontryagin Hamiltonian

$$\mathcal{H}(t, q_\alpha^i, z^i, p_i^\alpha) = \sum_{\alpha=0}^{N-1} p_i^\alpha \psi_\alpha^i - L(t, q_\alpha^i, z^i) =$$
$$= \sum_{\alpha=0}^{N-2} p_i^\alpha q_{\alpha+1}^i + p_i^{N-1} z^i - L(t, q^i, \ldots, q_{N-1}^i, z^i) \tag{33}$$

Preserving the notation $\mathcal{C}(\mathcal{A}) \overset{\nu}{\to} \mathcal{Q}$ for the composite map $\mathcal{C}(\mathcal{A}) \to \mathcal{A} \to \mathcal{Q}$, to each section $\bar{\gamma} : [t_0, t_1] \to \mathcal{C}(\mathcal{A})$ we now assign the action functional

$$\bar{\mathcal{I}}[\bar{\gamma}] := \int_{\bar{\gamma}} -\mathcal{H}\, dt + \sum_{\alpha=0}^{N-1} p_i^\alpha \, dq_\alpha^i = \int_{t_0}^{t_1} \left(-\mathcal{H} + \sum_{\alpha=0}^{N-1} p_i^\alpha \frac{dq_\alpha^i}{dt} \right) dt$$

Imposing stationarity of the latter under arbitrary deformations leaving the projections $\nu(\bar{\gamma}(t_0))$, $\nu(\bar{\gamma}(t_1))$ fixed leads to the Pontryagin equations

$$\frac{dq_\alpha^i}{dt} = \frac{\partial \mathcal{H}}{\partial p_i^\alpha}, \qquad \frac{dp_i^\alpha}{dt} = -\frac{\partial \mathcal{H}}{\partial q_\alpha^i} \tag{34a}$$

$$\frac{\partial \mathcal{H}}{\partial z^i} = p_i^{N-1} - \frac{\partial L}{\partial z^i} = 0 \tag{34b}$$

for the unknowns $q_\alpha^i(t), z^i(t), p_i^\alpha(t)$. Eqs. (34b) reproduce the content of eqs. (28') under the morphism $(t, q_0^i, \ldots, q_{N-1}^i, z^i) \longleftrightarrow (t, q_0^i, \ldots, q_{N-1}^i, q_N^i)$.

Denoting by $\mathcal{S}$ the subset of $\mathcal{C}(\mathcal{A})$ described by eqs. (34b), it is readily seen that the non–degeneracy condition $\left\| \frac{\partial^2 L}{\partial q_N^i \partial q_N^j} \right\| \neq 0$, here rephrased as $\det \left\| \frac{\partial^2 L}{\partial z^i \partial z^j} \right\| \neq 0$, automatically ensures the regularity of the Pontryagin Hamiltonian (33). We can therefore solve eqs. (34b) for the variables z^i, getting the expression

$$z^i = z^i(t, q_0^k, \ldots, q_{N-1}^k, p_k^{N-1}) \tag{35}$$

formally identical to eq. (29).

Eqs. (35) point out that the subset $\mathcal{S} \subset \mathcal{C}(\mathcal{A})$ is in fact a *submanifold* $i : \mathcal{S} \to \mathcal{C}(\mathcal{A})$, locally diffeomorphic to the phase space $V^*(Q)$. A straightforward check shows that the pull back $H := i^*(\mathcal{H})$ of the Pontryagin Hamiltonian (33) determines a proper Hamiltonian function on $V^*(Q)$, identical to the Ostrogradsky Hamiltonian (30), and that the Hamilton equations generated by H coincide with the Ostrogradsky equations (31).

REFERENCES

[1] M. Ostrogradsky, Mem. Ac. St. Petersbourg, **VI 4**, 385 (1850).

[2] R. P. Woodard, Lect. Notes Phys. **720**, 403 (2007).

[3] R. P. Woodard, arXiv:1506.02210v2 (2015).

[4] see e.g. S. Capozziello, S. Carloni and A. Troisi, Recent Res. Dev. Astron. Astrophys. **1**, 625 (2003); T. P. Sotiriou and V. Faraoni, Rev. Mod. Phys. **82**, 451 (2010); A. De Felice and S. Tsujikawa, Living Rev. Rel. **13**, 3 (2010); S. Capozziello, T. Harko, T. S. Koivisto, F. S. N. Lobo and G. J. Olmo, Universe **1**, no. 2, 199 (2015); S. Nojiri and S. D. Odintsov, Phys. Lett. B **631**, 1 (2005).

[5] D. Langlois and K. Noui, JCAP **1602**, no. 2, 34 (2016).

[6] T. J. Chen, M. Fasiello, E. A. Lim and A. J. Tolley, JCAP **1302**, 42 (2013).

[7] J. F. Pommaret *Systems of partial differential equations and Lie pseudogroups*, Gordon and Breach Science Publishers, New York, London (1978).

[8] D. J. Saunders *The geometry of jet bundles*, Cambridge University Press (1989).

[9] E. Massa, D. Bruno, G. Luria and E. Pagani, *Int. J. Geom. Methods Mod. Phys.*, **12**, 1550061 (2015).

[10] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, *The mathematical theory of optimal processes*, Interscience, New York (1962).

[11] M. Giaquinta and S. Hildebrandt, *Calculus of variations I, II*, Springer-Verlag, Berlin Heidelberg New York (2004).

[12] E. Massa and E. Pagani, *Ann. Inst. Henry Poicaré*, Physique théorique, **66**, 1-36 (1997).

[13] E. Massa, S. Vignolo and D. Bruno, *J. Phys. A: Math. Gen.*, **35**, 6713-6742 (2002).

[14] C. Godbillon C. *Géométrie différentielle et mécanique analytique* Hermann, Paris (1969).

[15] For simplicity, we preserve the same notation for covariant objects on $\mathcal{A}$ and for their pull–back on $\mathcal{C}(\mathcal{A})$.

[16] For a precise definition of ordinariness, see[9].