Constrained variational calculus for higher order classical field theories

Cédric M Campos, Manuel de León and David Martín de Diego

Instituto de Ciencias Matemáticas, CSIC-UAM-UC3M-UCM, Serrano 123, 28006 Madrid, Spain
E-mail: cedricmc@icmat.es, mdeleon@icmat.es and david.martin@icmat.es

Received 19 May 2010, in final form 10 September 2010
Published 22 October 2010
Online at stacks.iop.org/JPhysA/43/455206

Abstract
We develop an intrinsic geometrical setting for higher order constrained field theories. As a main tool we use an appropriate generalization of the classical Skinner–Rusk formalism. Some examples of applications are studied, in particular to the geometrical description of optimal control theory for partial differential equations.

PACS numbers: 03.50.−z, 02.40.Yy, 11.10.Ef
Mathematics Subject Classification: Primary, 70S05, Secondary, 70H50, 53C80, 55R10

1. Introduction

Classical field theories can be intrinsically described in terms of the geometry of a fiber bundle \( \pi : E \to M \) and its associated higher order jet bundles, \( J^k \pi \) (\( J^1 \pi \) for the case of first-order field theories). The jet bundles give a geometrical description for higher order partial derivatives of the fiber coordinates of \( E \) with respect to those of \( M \).

As in almost every physical theory, the dynamics of a classical field theory are completely determined by a Lagrangian function, that is, a function on the corresponding jet bundle, \( L : J^1 \pi \to \mathbb{R} \). For a theory of order \( k \), the solutions are sections \( \phi \) of the fiber bundle \( \pi : E \to M \) such that they extremize the functional

\[
S(\phi) = \int_R L(j^k \phi) \eta,
\]

where \( \eta \) is a fixed volume form (it is assumed that \( M \) is orientable), \( R \subseteq M \) is a compact set and \( j^k \phi \) is the \( k \)-jet prolongation of \( \phi \).

The most basic result in variational calculus is the construction from the above functional of a set of partial differential equations, the Euler–Lagrange equations, which must be satisfied by any smooth extremal. More interestingly, the property of extremizing the problem does not depend on the particular chosen coordinate system (a fact noted by J L Lagrange during...
his studies of analytical mechanics); therefore, one must be able to write the Euler–Lagrange equations in an intrinsic way.

In this sense, and restricting ourselves to first-order field theories, when $L : J^1 \pi \rightarrow M$, the fundamental object is the so-called Poincaré–Cartan form $\Omega_L$ which is an $(m+1)$-form $(\dim M = m)$ univocally associated with the Lagrangian. This form is constructed using the geometry of the jet bundle and it is also related to the variational background [20]. Using this form, it is possible to write down the Euler–Lagrange equations in an intrinsic way. Indeed, $\phi$ satisfies the Euler–Lagrange equations (that is, it is a critical point of the action $S$) if and only if

$$(j^1 \phi)^*(i_V \Omega_L) = 0 \quad \text{for all tangent vector } V \in T J^1 \pi \, .$$

Moreover, this form plays an important role in the connection between symmetries and conservation laws (see [25]).

However, many of the Lagrangians which appear in field theories are of higher order (as for instance in elasticity or gravitation); therefore, it is interesting to also find a fully geometric setting for these field theories. Besides, in first-order field theories it is possible to define a Cartan $(m+1)$-form, whereas one of the main difficulties in higher order field theories is that the uniqueness of the Cartan $(m+1)$-form is not guaranteed (contrary to the first-order case). In other words, there will be different Cartan forms which carry out the same function in order to define an intrinsic formulation of the Euler–Lagrange equations. The main reason for this problem is the commutativity of repeated partial differentiation. In the literature, we find different approaches to fix the Cartan form for higher order field theories. For instance, the approach by Arens [5] consists of injecting the higher order Lagrangian to a first-order one in an appropriate space by introducing a great number of variables and Lagrange multipliers. Following this perspective, and from a more geometrical point of view, the reader is referred to papers by Aldaya and Azcárraga [3, 4]. A different approach is adopted by García and Muñoz, who described a method of constructing global Poincaré–Cartan forms in the higher order calculus of variations in fibered spaces by means of linear connections (see [17, 18]); in particular, they show that the Cartan forms depend on the choice of two linear connections on the base $M$ and the vertical bundle $V \pi$. In 1990, Crampin and Saunders [32] proposed the use of an operator analogous to the almost tangent structure canonically defined on the tangent bundle of a given configuration manifold $M$ for the construction of global Poincaré–Cartan forms. In [6] we have found an intrinsic formalism for higher order field theories avoiding the problem of the degree of arbitrariness in the definition of Cartan forms for a given Lagrangian function. We have proposed a differential geometric setting for the equations of motion derived from a higher order Lagrangian function, strongly based on the so-called Skinner and Rusk formalism for mechanics (see also the paper by Vitagliano [35]).

In this paper, we introduce constraints which are geometrically defined as a submanifold $C$ of $J^k \pi$ and a function $L : J^k \pi \rightarrow \mathbb{R}$. In other words, we impose the constraints on the space of sections where the action is defined. We will show that the formalism introduced in our previous paper [6] is adapted to the case of constrained field theories, deriving an intrinsic framework of the constrained Euler–Lagrange equations. For the geometrical description, we start with the fibration $\pi_{W,M} : W \rightarrow M$, where $W$ is the fibered product $J^k \pi \times_{J^{k-1} \pi} \Lambda^m_2 (J^{k-1} \pi)$, and we induce a submanifold $W^C_0$ of $W$ using the constraints given by $C$. Using that $W$ has a naturally (pre-)multisymplectic form (by pulling back the canonical multisymplectic form on $\Lambda^m_2 (J^{k-1} \pi)$), we deduce an intrinsic and unique expression of a Cartan-type equation for the Euler–Lagrange equations in constrained higher order field theories. Additionally, we obtain a resultant constraint algorithm and give conditions to
ensure that the final constraint submanifold is multisymplectic. Finally, we discuss some examples to illustrate the theory.

2. Notation

Throughout the paper, we will use the following notation. It will be recalled when necessary, but for the ease of the reader, we summarize it here.

Lowercase Latin (resp. Greek) letters will usually denote indexes that range between 1 and \( m \) (resp. 1 and \( n \)). Capital Latin letters will usually denote multi-indexes whose length ranges between 0 and \( k \). In particular and if nothing else is stated, \( I \) and \( J \) will usually denote multi-indexes whose length goes from 0 to \( k-1 \) and 0 to \( k \), respectively, and \( K \) (and sometimes \( R \)) will denote multi-indexes whose length is equal to \( k \). The Einstein notation for repeated indexes and multi-indexes is understood but, for clarity, in some cases the summation for multi-indexes will be indicated.

We denote by \((E, \pi, M)\) a fiber bundle whose base space \( M \) is a smooth manifold of dimension \( m \), and whose fibers have dimension \( n \); thus, \( E \) is \((m+n)\)-dimensional. Adapted coordinate systems on \( E \) will be of the form \((x^i, u^\alpha)\), where \((x^i)\) is a local coordinate system on \( M \) and \((u^\alpha)\) denotes fiber coordinates. The local volume form associated with \((x^i)\) is \( d^m x = d x^1 \wedge \cdots \wedge d x^m \). We write \( d^{m-1} x_i \) for the contraction \( i_\alpha \partial / \partial x^\alpha d^m x \). Note that, for a fixed index \( i \in \{ 1, \ldots, m \} \), we have \( dx^i \wedge d^{m-1} x_i = d^m x \).

In section 4 and all later sections, \( M \) shall be supposed oriented and a volume form \( \eta \) will be provided. Local coordinates on \( M \) will then be considered subject to compatibility with \( \eta \), that is, such that \( d^m x = \eta \). Any \( m \)-form along a fiber projection over \( M \) will be denoted with a curly letter. The same letter in a straight font shall denote its factor with respect to either the volume form \( \eta \) or the local one \( d^m x \), e.g. \( L = L_{d^m x}, \mathcal{H} = H_{d^m x}, \) etc.

Finally, pullbacks of forms and functions by fiber projections will still be denoted by the same symbol. For instance, \( \pi^* \eta \) will still be denoted by \( \eta \) on \( E \).

3. Jet bundles

We fix a fiber bundle \( \pi : E \to M \) with \( \text{dim } M = m \) and \( \text{dim } E = m + n \), and where neither \( M \) nor \( E \) is necessarily orientable.

Let \( \phi, \psi : M \to E \) be two smooth local sections of \( \pi \) around a given point \( x \in M \). We say that \( \phi \) and \( \psi \) are \( k \)-equivalent at \( x \) (with \( k \geq 0 \)) if the sections and all their partial derivatives until order \( k \) coincide at \( x \), that is, if

\[
\phi(x) = \psi(x) \quad \text{and} \quad \left. \frac{\partial^k \phi^\alpha}{\partial x^{i_1} \cdots \partial x^{i_k}} \right|_x = \left. \frac{\partial^k \psi^\alpha}{\partial x^{i_1} \cdots \partial x^{i_k}} \right|_x
\]

for all \( 1 \leq \alpha \leq n \), \( 1 \leq i_j \leq m \), \( 1 \leq j \leq k \). Note that this is independent of the chosen coordinate system (adapted or not) and, therefore, along a fiber projection over \( M \) will be denoted with a curly letter. The same letter in a straight font shall denote its factor with respect to either the volume form \( \eta \) or the local one \( d^m x \), e.g. \( L = L_{d^m x}, \mathcal{H} = H_{d^m x}, \) etc.

Finally, pullbacks of forms and functions by fiber projections will still be denoted by the same symbol. For instance, \( \pi^* \eta \) will still be denoted by \( \eta \) on \( E \).

**Definition 3.1.** Let \( x \in M \) given a smooth local section \( \phi \in \Gamma_x \pi \); the equivalence class of \( k \)-equivalent smooth local sections (with \( k \geq 0 \)) around \( x \) that contains \( \phi \) is called the \( k \)th jet of \( \phi \) at \( x \) and is denoted by \( j^k_x \phi \). The \( k \)th jet manifold of \( \pi \), denoted by \( J^k \pi \), is the whole collection of \( k \)th jets of arbitrary local sections of \( \pi \), that is,

\[
J^k \pi = \left\{ j^k_x \phi : x \in M , \phi \in \Gamma_x \pi \right\} .
\]

Given a (local) section of \( \pi \), its \( k \)th lift is \( (j^k_x \phi)(x) = j^k_x \phi \).
Adapted coordinates \((x^i, u^a)\) on the total space \(E\) induce coordinates \((x^i, u^a_i)\) on the \(k\)-jet manifold \(J^k\pi\) given by
\[
u^a_i(j^k_x \phi) = \left. \frac{\partial |I| \phi^a}{\partial x^I} \right|_{x} ,
\]
from where we deduce that \(J^k\pi\) is actually a smooth manifold of dimension
\[
\dim J^k\pi = m + n \cdot \sum_{l=0}^{k} \left( \begin{array}{c} m - 1 + l \\ m - 1 \end{array} \right).
\]
It is readily seen that \((J^k\pi, \pi^k, M)\) is a fiber bundle, where
\[
\pi^k(j^k_x \phi) = x^i \quad \text{(in coordinates } \pi^k(x^i, u^a_i) = (x^i)).
\]
It is also clear that the \(k\)-jet manifold \(J^k\pi\) fibers over the lower order \(l\)-jet manifolds \(J^l\pi\) (see figure 1), with \(0 \leq l < k\), where by definition \(J^0\pi = E\) and the projections are given by
\[
\pi_{k,l}(j^k_x \phi) = j^l_x \phi \quad \text{(in coordinates } \pi_{k,l}(x^i, u^a_i) = (x^i, u^a_i) \text{ with } 0 \leq |I| \leq k, 0 \leq |J| \leq l).
\]

**Definition 3.2.** The \(k\)th dual jet bundle, denoted by \(J^k\pi^\dagger\), is the subbundle of \(\Lambda^m J^{k-1}\pi\) of \(\pi^k_{k-1}\)-semi-basic \(m\)-forms over \(J^{k-1}\pi\), that is,
\[
J^k\pi^\dagger = \Lambda^m J^{k-1}\pi = \{ \omega \in \Lambda^m J^{k-1}\pi : i_{v_1,i_{v_2}} \omega = 0, v_1, v_2 \in \mathcal{V}\pi_{k-1} \}, \tag{3.2}
\]
where \(\mathcal{V}\pi_{k-1}\) is the vertical bundle of \(\pi_{k-1} : J^{k-1}\pi \to M\).

**Remark 3.3.** The original definition of the \(k\)th dual jet bundle is as the extended dual affine bundle of the iterated jet bundle \((J^1\pi_{k-1}, (\pi_{k-1})_{1,0}, J^{k-1}\pi)\). Observe that \(J^k\pi\) is naturally embedded as an affine subbundle of \(J^1\pi_{k-1}\).

Locally, the elements of \(J^k\pi^\dagger\) are of the form
\[
p \, d^n x + p_{a_i}^{li} d u^a_i \wedge d^{m-1} x_i ,
\]
where \(0 \leq |I| \leq k - 1\). Thus, adapted coordinates \((x^i, u^a)\) on \(E\) induce coordinates \((x^i, u^a_i, p, p_{a_i}^{li})\) on \(J^k\pi^\dagger\). Note that \((x^i, u^a_i)\) are adapted coordinates on \(J^{k-1}\pi\) and that \((p, p_{a_i}^{li})\) are fiber coordinates.

While the \(k\)th jet bundle \(J^k\pi\) projects over the lower order jet bundles (figure 1), the \(k\)th dual jet bundle is ‘embedded’ into the \((k + 1)\)th dual jet bundle by means of the pullback of the affine projection \(\pi_{k+1,k}\) (figure 2).

The dual jet bundle \(J^k\pi^\dagger\) has a canonical multisymplectic structure (see [7, 8, 10]) and its elements are naturally paired with those of \(J^k\pi\). We recall that a form \(\Omega\) is multisymplectic.
if it is closed and if its contraction with a single tangent vector is injective, that is, \( i_V \Omega = 0 \) if and only if \( V = 0 \).

**Definition 3.4.** The tautological m-form on \( J^k \pi^\dagger \) is the form given by

\[
\Theta_\omega(V_1, \ldots, V_m) = (\tau_{j_1 \ldots j_m}^* \omega)(V_1, \ldots, V_m), \quad \omega \in J^k \pi^\dagger, \quad V_1, \ldots, V_m \in T_0 J^k \pi^\dagger,
\]

where \( \tau_{j_1 \ldots j_m} \) is the natural projection from \( TJ^k \pi^\dagger \) to \( J^k \pi^\dagger \). The canonical multisymplectic \((m + 1)\)-form on \( J^k \pi^\dagger \) is

\[
\Omega = -d\Theta.
\]

**Definition 3.5.** The natural pairing between \( J^k \pi \) and its dual \( J^k \pi^\dagger \) is the fibered map \( \Phi : J^k \pi \times_{\mu_{j-1}} J^k \pi^\dagger \to \Lambda^m M \) given by

\[
\Phi(j_k^x \phi, \omega) = (j_{k-1}^x)^* \omega,
\]

where \( \phi_f = f \circ \phi \circ \tilde{f}^{-1} \) (see figure 3).

Let \( (x^i, u^a_i, u^a_K) \) and \( (x^i, u^a_i, p, p^{li}_a) \), where \( |I| = 0, \ldots, k - 1 \) and \( |K| = k \), denote adapted coordinates on \( J^k \pi \) and \( J^k \pi^\dagger \), respectively. Then, the tautological form and the canonical one are locally written as

\[
\Theta = p d^m x + p^{li}_a du^a_i \wedge d^{m-1} x_i \quad \text{and} \quad \Omega = -dp \wedge d^m x - dp^{li}_a \wedge d^{m-1} x_i,
\]

and the fibered pairing between the elements of \( J^k \pi \) and \( J^k \pi^\dagger \) is locally written as

\[
\Phi(x^i, u^a_i, u^a_K, p, p^{li}_a) = (p + p^{li}_a u^{a_{K+1}}) d^m x.
\]

With the aim of clarifying the latter sections and to make this work self-content, we continue presenting some classical definitions besides some basic properties, which may be found in [23, 30, 31].

**Definition 3.6.** Let \( f : E \to F \) be a morphism between two fiber bundles \((E, \pi, M)\) and \((F, \rho, N)\), such that the induced function on the base, \( \tilde{f} : M \to N \), is a diffeomorphism. The kth prolongation of \( f \) is the map \( j^k f : J^k \pi \to J^k \rho \) given by

\[
(j^k f)(j^k_\phi) := j^k_{\tilde{f}^{-1}(\phi)} f, \quad \forall j^k_\phi \in J^k \pi,
\]

where \( \phi_f := f \circ \phi \circ \tilde{f}^{-1} \) (see figure 3).

Note that the kth prolongation \( j^k f \) of a morphism \( f \) is both a morphism between \((J^k \pi, \pi_{j_k}, E)\) and \((J^k \rho, \rho_{k}, F)\) and between \((J^k \pi, \pi_{j_k}, M)\) and \((J^k \rho, \rho_{k}, N)\). In each case, the induced functions between the base spaces are \( f \) and \( \tilde{f} \), respectively.

---

Figure 2. Chain of dual jets.
Figure 3. The kth prolongation of a morphism.

**Definition 3.7.** The (coordinate) total derivatives are the vector fields along $\pi_{k,k-1}$ locally given by

$$\frac{d}{dx^i} = \frac{\partial}{\partial x^i} + u_{i+1}^a \frac{\partial}{\partial u_a^i} + u_a^i \frac{\partial}{\partial u_a^{i+1}} + \cdots \quad (3.9)$$

for coordinates $(x^i, u_a^i)$ on $J^k\pi$.

Let $f \in C^\infty(J^k\pi)$, and then $\frac{df}{dx} \in C^\infty(J^{k+1}\pi)$ on the corresponding local chart. In general and as the coordinate total derivatives commute, $\frac{d^{(J)}f}{dx^J} \in C^\infty(J^{k+1}\pi)$ for any multi-index $J \in \mathbb{N}^m$. Observe also that total derivatives are closely related to partial derivatives by the relation

$$\frac{d^{(J)} f}{dx^J} \circ j^k\phi = \frac{\partial^{(J)} (f \circ j^k\phi)}{\partial x^J} \quad (3.10)$$

for any section $\phi \in \Gamma\pi$.

**Definition 3.8.** A contact 1-form on $J^k\pi$ is a one-form $\theta \in \Lambda^1 J^k\pi$ which is pulled back to the zero form on $M$ by the kth lift of any section $\phi$ of $\pi$, that is,

$$(j^k\phi)^* \theta = 0, \quad \forall \phi \in \Gamma\pi. \quad (3.11)$$

The set $\mathcal{C}^k$ of all the contact forms is called the Cartan distribution (of order k).

**Proposition 3.9.** Let $(x^i, u_a^i)$ be adapted coordinates on $J^k\pi$; a basis of the Cartan distribution is given by the contact forms

$$\theta_a^0 = du_a^i - u_{a+1}^i dx^i, \quad 0 \leq |I| \leq k - 1. \quad (3.12)$$

**Definition 3.10.** Given a vector field $\xi$ on $E$, its kth lift (or kth jet) is the unique vector field $\xi^{(k)}$ on $J^k\pi$ that is projectable to $\xi$ by $\pi_{k,0}$ and preserves the Cartan distribution with respect to the Lie derivative.
Proposition 3.11. Let $\xi$ be a vector field on $E$. If $\xi$ has the local expression
\[ \xi = \xi^{i} \frac{\partial}{\partial x^{i}} + \xi^{a} \frac{\partial}{\partial u^{a}} \]
in adapted coordinates $(x^{i}, u^{a})$ on $E$, then its $k$th lift $\xi^{(k)}$ has the form
\[ \xi^{(k)} = \xi^{i} \frac{\partial}{\partial x^{i}} + \xi^{a} J^{a} \frac{\partial}{\partial u^{a}} \]
for the induced coordinates $(x^{i}, u^{a} J^{a})$ on $J^{k} \pi$, where
\[ \xi^{a} J^{a} = \xi^{a} \quad \text{and} \quad \xi^{a} J^{a+1} = \frac{d\xi^{a}}{dx^{i}} - u^{a} J^{a+1} \frac{d\xi^{j}}{dx^{j}}. \]
In particular, if $\xi$ is vertical with respect to $\pi$, then $\xi^{a} J^{a} = \frac{d}{dJ} \xi^{a} / dx^{i}$.

Proposition 3.12. Let $\psi_{\varepsilon}$ be the flow of a given $\pi$-projectable vector field $\xi$ over $E$. Then, the flow of $\xi^{(k)}$ is the $k$th lift of $\psi_{\varepsilon}$, $J^{k} \psi_{\varepsilon}$.

4. Variational calculus

The dynamics in classical field theory is specified giving a Lagrangian density: a Lagrangian density is a mapping $L : J^{k} \pi \rightarrow \Lambda^{m} M$. Fixed a volume form $\eta$ on $M$, there is a smooth function $L : J^{k} \pi \rightarrow \mathbb{R}$ such that $L = L \eta$.

Definition 4.1. Given a Lagrangian density $L : J^{k} \pi \rightarrow \Lambda^{m} M$, the associated integral action is the map $S : \Gamma_{(1)} \pi \times \mathcal{K} \rightarrow \mathbb{R}$ given by
\[ S(\phi, R) = \int_{R} (J^{k} \phi)^{a} L, \]
where $\mathcal{K}$ is the collection of smooth compact regions of $M$.

Definition 4.2. Let $\phi$ be a section of $\pi$. A (vertical) variation of $\phi$ is a curve $\varepsilon \mapsto \phi_{\varepsilon} \in \Gamma \pi$ (for some interval $I \subset \mathbb{R}$) such that $\phi_{\varepsilon} = \phi \circ \hat{\phi}_{\varepsilon}^{-1}$, where $\phi_{\varepsilon}$ is the flow of a (vertical) $\pi$-projectable vector field $\xi$ on $E$.

When $\xi$ is vertical, then its flow $\phi_{\varepsilon}$ is an automorphism of fiber bundles over the identity for each $\varepsilon \in I$, i.e. $\hat{\phi}_{\varepsilon} = \text{Id}_{M}$.

Definition 4.3. We say that $\phi \in \Gamma \pi$ is a critical or stationary point of the Lagrangian action $S$ if and only if
\[ \frac{d}{d\varepsilon} |_{\varepsilon=0} S(\phi_{\varepsilon}, R) = \frac{d}{d\varepsilon} \left|_{\varepsilon=0} \left( \int_{R} (J^{k} \phi_{\varepsilon})^{a} L \right) \right| = 0 \]
for any vertical variation $\phi_{\varepsilon}$ of $\phi$ whose associated vertical field vanishes outside of $\pi^{-1}(R)$ and where $R_{\varepsilon} = \hat{\psi}_{\varepsilon}(R)$.

Lemma 4.4. Let $\phi_{\varepsilon} = \phi \circ \hat{\phi}_{\varepsilon}^{-1}$ be a variation of section $\phi \in \Gamma \pi$. If $\xi$ denotes the infinitesimal generator of $\phi_{\varepsilon}$, then
\[ \frac{d}{d\varepsilon} |_{\varepsilon=0} [(J^{k} \phi_{\varepsilon} \circ \phi)^{a} \omega] = (J^{k} \phi)^{a} \rho(\xi \ast \omega). \]
for any differential form $\omega \in \Omega(J^{k} \pi)$. 

Proof. From proposition 3.12, we have that $J$ is the infinitesimal generator of $j^k\varphi_x$. We then obtain a direct computation

$$(j^k\phi)^*_x(\xi,\omega) = (j^k\phi)^*_x\left(\frac{d}{d\varepsilon}[(j^k\varphi_x)^*\omega]\bigg|_{\varepsilon=0}\right) = \frac{d}{d\varepsilon}[(j^k\varphi_x \circ j^k\phi)^*\omega]\bigg|_{\varepsilon=0}.$$  

The following lemma will show to be useful in the variational derivation of the higher order Euler–Lagrange equations.

**Lemma 4.5** (Higher order integration by parts). Let $R \subset M$ be a smooth compact region and let $f, g : R \longrightarrow \mathbb{R}$ be two smooth functions. Given any multi-index $J \in \mathbb{N}^m$, we have

$$\int_R \frac{\partial^{[J]} f}{\partial x^J} g \, d^m x = (-1)^{|J|} \int_R f \frac{\partial^{[J]} g}{\partial x^J} \, d^m x + \sum_{I_f+I_g+1=J} \lambda(I_f, I_g, J) \int_{\partial R} \frac{\partial^{[I_f]} f}{\partial x^{I_f}} \frac{\partial^{[I_g]} g}{\partial x^{I_g}} \, d^{m-1} x_i,$$

where $\lambda$ is given by the expression

$$\lambda(I_f, I_g, J) := (-1)^{|I_f|} \frac{|I_f|! \cdot |I_g|!}{|J|!} \frac{J!}{I_f! \cdot I_g!}.$$  

Proof. In this proof, we will use the shorthand notation $f_J = \frac{\partial^{[J]} f}{\partial x^J}$ so as to save space. We proceed by induction on the length $l$ of the multi-index $J$. The case $l = |J| = 0$ is a trivial identity and the case $l = |J| = 1$ is the well-known formula of integration by parts

$$\int_R f_J g \, d^m x = \int_{\partial R} f \frac{\partial^{[1]} g}{\partial x} \, d^{m-1} x_i - \int_R f g \, d^m x.$$  

Thus, let us suppose that the result is true for any multi-index $J \in \mathbb{N}^m$ up to length $l > 1$, in order to show that it is also true for any multi-index $K \in \mathbb{N}^m$ of length $l + 1$. Let $J$ and $1 \leq j \leq m$ be such that $J + 1 = K$. Then we have

$$\int_R f_{J+1} g \, d^m x = \int_R f_{J+1} g \, d^m x + \int_{\partial R} f_J g \, d^{m-1} x_j$$

$$= (-1)^{|J|} \int_R f_{J+1} g \, d^m x + \int_{\partial R} f_J g \, d^{m-1} x_j$$

$$- \sum_{I_f+I_g+1=J} \lambda(I_f, I_g, i) \int_{\partial R} f_{I_f g_{I_g+1}} \, d^{m-1} x_i,$$

where we have used the first-order integration formula in the first place, to then apply the induction hypothesis. We now multiply each member by $(J(j) + 1)/(l + 1)$ and sum over $J + 1 = K$. Using the multi-index identity (B.2), we have

$$\int_R f_{J+1} g \, d^m x = (-1)^{|J|} \int_R f_{J+1} g \, d^m x + \sum_{J+1 = K} \frac{J(j) + 1}{l + 1} \int_{\partial R} f_J g \, d^{m-1} x_j$$

$$- \sum_{J+1 = K} \frac{J(j) + 1}{l + 1} \sum_{I_f+I_g+1=J} \lambda(I_f, I_g, i) \int_{\partial R} f_{I_f g_{I_g+1}} \, d^{m-1} x_i.$$  

It only remains to properly rearrange the last two terms to express them in the stated form. Clearly,

$$\sum_{J+1 = K} \frac{J(j) + 1}{l + 1} \int_{\partial R} f_J g \, d^{m-1} x_j$$

$$= \sum_{I_f+I_g+1=K} (-1)^{|I_f|} \frac{|I_f|! \cdot |I_g|!}{|K|!} \frac{K!}{I_f! \cdot I_g!} \int_{\partial R} \frac{\partial^{[I_f]} f}{\partial x^{I_f}} \frac{\partial^{[I_g]} g}{\partial x^{I_g}} \, d^{m-1} x_i.$$
The last term is a little bit more tricky:

\[
\sum_{J \in \mathbb{K}} \frac{J(j) + 1}{l + 1} \sum_{J_{l+1} \in J} (-1)^{J_{l+1}} \frac{|I_l|! \cdot |I_g|!}{|J|!} \cdot \frac{J(l)}{I_f} \cdot \frac{1}{I_l} \int_{\partial R} f_{l} S_{l} d^{m-1} x_{l}
\]

\[
= \sum_{J_{l+1} \in J} (-1)^{J_{l+1}} \frac{|I_l|! \cdot |I_g|!}{|J|!} \cdot \frac{1}{I_l} \int_{\partial R} f_{l} S_{l} d^{m-1} x_{l}
\]

\[
= \sum_{J_{l+1} \in J} (-1)^{J_{l+1}} \sum_{J_{l+1} \in J} \frac{I_l(j)}{I_l} \cdot \frac{|I_l|! \cdot |I_g|!}{|J|!} \cdot \frac{1}{I_l} \int_{\partial R} f_{l} S_{l} d^{m-1} x_{l}
\]

\[
= \sum_{J_{l+1} \in J} (-1)^{J_{l+1}} \frac{|I_l|! \cdot |I_g|!}{|J|!} \cdot \frac{1}{I_l} \int_{\partial R} f_{l} S_{l} d^{m-1} x_{l}
\]

where we have used the identity (B.2) again. The result is now clear. □

**Theorem 4.6** (The higher order Euler–Lagrange equations). *Given a fiber section \( \phi \in \Gamma \pi_{c} \), let us consider an infinitesimal variation \( \phi_{e} \) of it such that the support \( R \) of the associated vertical vector field \( \xi \) is contained in a coordinate chart \( (x) \). We then have that the variation of the Lagrangian action \( \mathcal{S} \) at \( \phi \) is given by

\[
\frac{d}{d \varepsilon} \mathcal{S}(\phi_{e}, R_{i}) \bigg|_{\varepsilon = 0} = \sum_{|J| = 0}^{k} (-1)^{|J|} \int_{R} (j^{2k} \phi)^{\ast} \left( x^{-d/4} \frac{dL}{dx} \right) d^{m} x
\]

\[
+ \sum_{J_{l} + J_{i} \in J} \lambda(I_{l}, I_{l}, J) \int_{\partial R} (j^{2k} \phi)^{\ast} \left( x^{-d/4} \frac{dL}{dx} \right) d^{m-1} x_{l}
\]

(4.6)

Moreover, \( \phi \) is a critical point of the Lagrangian action \( \mathcal{S} \) if and only if it satisfies the higher order Euler–Lagrange equations

\[
(j^{2k} \phi)^{\ast} \left( \sum_{|J| = 0}^{k} (-1)^{|J|} \frac{dL}{dx} \right) = 0
\]

(4.7)

on the interior of \( M \), plus the boundary conditions

\[
\frac{d}{dx} \frac{dL}{dx} = 0, \quad 0 \leq |J| < |J| \leq k,
\]

(4.8)

on the boundary \( \partial M \) of \( M \).

**Proof.** Let us denote by \( \xi \) the vertical field associated with the variation \( \phi_{e} \). Using proposition 4.4, Cartan’s formula \( \xi = d \circ i + i \circ d \) and proposition 3.11, we have

\[
\frac{d}{d \varepsilon} \mathcal{S}(\phi_{e}, R) \bigg|_{\varepsilon = 0} = \int_{R} (j^{2k} \phi)^{\ast} \left( \xi^{m} \mathcal{L} \right)
\]

\[
= \int_{R} (j^{2k} \phi)^{\ast} i_{\xi} \mathcal{L} + \int_{R} (j^{2k} \phi)^{\ast} i_{\xi} \mathcal{L} \, d\mathcal{L}
\]

\[
= \int_{\partial R} (j^{2k} \phi)^{\ast} i_{\xi} \mathcal{L} + \int_{\partial R} (j^{2k} \phi)^{\ast} (\xi^{k}(L) d^{m} x - d\mathcal{L} \wedge i_{\xi} d^{m} x)
\]

\[
= \int_{R} (j^{2k} \phi)^{\ast} \left( \sum_{|J| = 0}^{k} \frac{dL}{dx} \right) d^{m} x
\]
If we now apply the higher order integration by parts (4.4) and take into account equation (3.10), we obtain

\[
\frac{d}{d\varepsilon} S(\phi_\varepsilon, R) \bigg|_{\varepsilon=0} = k \sum_{|J|=0} (-1)^{|J|} \int_\mathbb{R} (j^{2k} \phi)^* \left( \xi_0 \frac{d^{[J]} \partial L}{dx^J \partial u^*_J} \right) d^m x 
\]

\[
+ \sum_{I_\varepsilon + I_L + I_\varepsilon = J} \lambda(I_\varepsilon, I_L, J) \int_{\partial R} (j^{2k} \phi)^* \left( \xi_0 \frac{d^{[J]} \partial L}{dx^J \partial u^*_J} \right) d^{m-1} x_i 
\]

which is the first statement of our theorem.

If we now suppose that $R$ is contained in the interior of $M$, as $\xi$ is null outside of $R$, so it is $\xi^{[J]}$ outside of $R$ and, by smoothness, on its boundary $\partial R$. Thus, if $\phi$ is a critical point of $S$, we then must have

\[
\frac{d}{d\varepsilon} S(\phi_\varepsilon, R) \bigg|_{\varepsilon=0} = k \sum_{|J|=0} (-1)^{|J|} \int_\mathbb{R} (j^{2k} \phi)^* \left( \xi_0 \frac{d^{[J]} \partial L}{dx^J \partial u^*_J} \right) d^m x = 0
\]

for any vertical field $\xi$ whose compact support is contained in $\pi^{-1}(R)$. We thus infer that $\phi$ shall satisfy the higher order Euler–Lagrange equations (4.7) on the interior of $M$.

Finally, if $R$ has a common boundary with $M$, we then have

\[
\frac{d}{d\varepsilon} S(\phi_\varepsilon, R) \bigg|_{\varepsilon=0} = k \sum_{|J|=0} \sum_{I_\varepsilon + I_L + I_\varepsilon = J} \lambda(I_\varepsilon, I_L, J) \int_{\partial R \cap \partial M} (j^{2k} \phi)^* \left( \xi_0 \frac{d^{[J]} \partial L}{dx^J \partial u^*_J} \right) d^{m-1} x_i = 0
\]

(4.9)

As this is true for any vertical field $\xi$ whose compact support is contained in $\pi^{-1}(R)$, we deduce the boundary conditions (4.8).

**Remark 4.7.** If in definition 4.3 of a critical point, one further requires the variations to be null at the boundary $\partial M$ of $M$, in the sense that the associated vector field $\xi$ be zero over $\pi^{-1}(\partial M)$, then one would no longer have the boundary condition (4.8). In such a case, the space of solutions would be broader and one could impose boundary conditions to them in order to obtain a particular one.

We also note that the theorem remains true if we choose vector variations, not necessarily vertical, although the proof would be much more cumbersome.

### 5. The Skinner–Rusk formalism

What follows in this section has already been published in [6]. Here we give a summary of the most relevant facts that are proven there, introducing some new insights and points of view. The generalization of the Skinner–Rusk formalism to higher order classical field theories will take place in the fibered product:

\[
W = J^k \pi \times J^{k-1} \pi J^1 \pi.
\]

(5.1)

The first-order case is covered in [12, 24] or [11] for time-dependent mechanics; see also [33, 34] for the original treatment by Skinner and Rusk. The projection on the $i$th factor will be denoted by $pr_i$ (with $i = 1, 2$) and the projection on the fiber bundle over $J^{k-1} \pi$ will be $\pi_W J^{k-1} \pi = \pi_{k-1} \circ pr_1$ (see figure 4). On $W$, adapted coordinates are of the form $(x^i, u^i_0, u^i_k, p, p^i_0)$, where $|I| = 0, \ldots, k-1$ and $|K| = k$. Note that $(x^i, u^i_0)$ are
coordinates on $J^{k-1}\alpha$ and that $(u^\alpha_K)$ and $(p, p_I^I)$ are fiber coordinates on $J^k\alpha \rightarrow J^{k-1}\alpha$ and $J^{k-1}\alpha \rightarrow J^{k-1}\alpha$, respectively.

Assume that $\mathcal{L} : J^k\alpha \rightarrow \Lambda^m M$ is a Lagrangian density. Together with the pairing $\Phi$ (equation (3.5)), we use this Lagrangian $\mathcal{L}$ to define the dynamical map $\mathcal{H}$ on $W$ (corresponding to the Hamiltonian density):

$$\mathcal{H} = \Phi - \mathcal{L} \circ pr_1. \quad (5.2)$$

Consider the canonical multisymplectic $(m + 1)$-form $\Omega$ on $J^k\alpha$ (equation (3.4)), whose pullback to $W$ shall also be denoted by $\Omega$. We define on $W$ the premultisymplectic $(m + 1)$-form

$$\Omega_N = \Omega + d\mathcal{H}. \quad (5.3)$$

In adapted coordinates

$$\mathcal{H} = (p^I J^I u^a_I + p - L(x, u^a_I, u^\alpha_J)) \, dx$$

$$\Omega_N = -d p^I J^I \wedge du^a_I \wedge d^{m-1} x_i + \left( p^I J^I du^a_I + u^a_I \, dp^I J^I - \frac{\partial L}{\partial u^a_J} \, du^a_J \right) \wedge dx, \quad (5.5)$$

where $|I| = 0, \ldots, k - 1$ and $|J| = 0, \ldots, k$.

We search for a $\pi_{W,M}$-transverse and locally decomposable $m$-multivector field $X$ on $W$ (see appendix A) that is the solution of the dynamical equation

$$i_X \Omega_N = 0. \quad (5.6)$$

**Remark 5.1.** The approach that we followed in [6] was to solve the equation

$$i_h \Omega_N = (m - 1) \Omega_N, \quad (5.7)$$

where $h : TW \rightarrow TW$ is the horizontal projector associated with a horizontal distribution in $\pi_{W,M} : W \rightarrow M$ (i.e. a connection). In fact, equations (5.6) and (5.7) are equivalent. Two solutions $X$ and $h$ of the corresponding equation are related in such a way that $X$ generates the horizontal distribution to which $h$ is associated. If a solution $h$ of (5.7) is locally written in the following form (compare with the expression (5.11) below):

$$h = dx^j \otimes \left( \frac{\partial}{\partial x^j} + A^a_j \frac{\partial}{\partial u^a_j} + B^a_j \frac{\partial}{\partial p^a_j} + C^a_j \frac{\partial}{\partial p} \right), \quad (5.8)$$

then its components will be governed by the relations (5.12)–(5.15), (5.18) and (5.19).
Remark 5.2. Another formulation of the dynamical equation (5.6) is
\[ (-1)^m i_X \Omega = -dH, \] (5.9)
which is the regular expression for classical mechanics (i.e. whenever \( M \) has dimension \( m = 1 \)). The only difference is that, in this latter one, the tangency condition to \( \mathcal{H} = 0 \) is already included: \( \Omega \) possesses explicitly \( dp \) in its local expression, while \( \Omega_{\mathcal{H}} \) does not. So (5.6) may not determine any condition on the coefficients \( C_j \) of \( X_j \), which is equivalent to establishing whether \( X \) is tangent to \( \mathcal{H} = 0 \) or not.

It is shown that solutions \( X \) of (5.6) do not exist on the whole \( W \), cf [6, 12, 24]. Because of that, we need to restrict the equation to the space where solutions do exist, that is, to
\[ W_1 = \{ w \in W : \exists X \in \Lambda^m_\pi W \text{ such that } \Lambda^m_\pi \pi W(X) \neq 0 \text{ and } i_X \Omega_{\mathcal{H}}(w) = 0 \}. \] (5.10)
Thus, a solution \( X \) of (5.6), rather than being a multivector field on \( W \), will be a multivector field along \( W_1 \).

Remark 5.3. It is important to point out that the sole purpose of the base volume form \( \eta \) is to ensure the existence of non-vanishing \( m \)-multivector fields on \( M \), which implies the existence of \( \pi W_1 \)-transverse \( m \)-multivector fields on \( W \).

An arbitrary \( \pi W_1 \)-transverse multivector \( X \in \Lambda^m_\pi T_\nu W \) may be written in the form
\[ X = f \Lambda_{j=1}^m X_j = f \Lambda_{j=1}^m \left( \frac{\partial}{\partial x_j} + A_{j}^j \frac{\partial}{\partial u_j^i} + B_{j}^l \frac{\partial}{\partial p_{l}^i} + C_{j} \frac{\partial}{\partial p}\right). \] (5.11)
If we furthermore impose the condition \( \eta(X) = 1 \) \( (f = 1) \) and compute the local expression of the equation (5.6), we then obtain
\[ A_{j}^i = u_{i+1}^j, \] with \( |I| = 0, \ldots, k - 1, \ i = 1, \ldots, \) \] (5.12)
\[ 0 = \frac{\partial L}{\partial u^a} - B_{a}^j; \] \] (5.13)
\[ \sum_{l+1=j} p_{l}^j = \frac{\partial L}{\partial u_j^a} - B_{a}^j \] with \( |J| = 1, \ldots, k - 1; \) \] (5.14)
\[ \sum_{l+1=k} p_{l}^j = \frac{\partial L}{\partial u_k^a} \] with \( |K| = k. \) \] (5.15)

We shall refer to equation (5.12) as the equation of holonomy and to equations (5.13)–(5.15) as the equations of dynamics. We will further name the equations of dynamics after the order of the multi-index in them, being the bottom-, mid- and top-level ones, respectively. We could have written these in the more compact way
\[ \sum_{l+1=j} p_{l}^j = \frac{\partial L}{\partial u_j^a} - B_{a}^j \] with \( |J| = 0, \ldots, k, \) \] (5.16)
understanding that the first summation term is empty when \( |J| = 0 \), as well as it is the last one when \( |J| = k \) (there are no \( B_{a}^j \) with \( |K| = k \)).

Note that the top-level equation of dynamics (5.15) is a constraint on the point \( w \in W \) where the multivector \( X \) stands; thus, it defines the submanifold \( W_1 \) of \( W \). The remaining equations, (5.12)–(5.14), are restrictions on the coefficients of the multivector \( X \). Note also that, for the time being, the \( A \)s with greatest order multi-index and the \( C \)s remain undetermined, as well as the most part of the \( B \)s.
The existence of a $\pi_{W,M}$-transverse and locally decomposable $m$-multivector $X \in \Lambda^m T_w W$ (with $\eta(X) = 1$) is guaranteed for every point $w \in W_1$. However, we cannot assure that such multivector $X$ is ‘tangent’ to $W_1$ at $w$. Here, ‘tangent’ means that the distribution associated with $X$ is tangent to $W_1$ at $w$. Therefore, we require $X$ to be tangent to $W_1$ by imposing the tangency condition

$$X_j \left( \sum_{I+I_i=K} p_i^{ij} - \frac{\partial L}{\partial u^i_k} \right) = 0.$$  

Furthermore, we also require $X$ to be tangent to the submanifold of $W$:

$$W_0 = \{ w \in W : \mathcal{H}(w) = 0 \} = \{ w \in W : p = L - p_i^{ij} u^{ij}_t \}$$  

by imposing the second tangency condition

$$X_j \left( p + p_i^{ij} u^{ij}_t - L \right) = 0.$$  

These two previous conditions are equivalent to the following equations of tangency:

$$\sum_{I+I_i=K} B_{ij} = \frac{\partial^2 L}{\partial x^j \partial u^i_k} + \sum_{|I|+1} u^{\beta}_{i+1} \frac{\partial^2 L}{\partial u^\beta_j \partial u^i_k} + \sum_{|R|=k} A^\rho_{ij} \frac{\partial^2 L}{\partial u^\rho_k \partial u^i_k}, \quad \text{with } |K| = k;$$

(5.18)

$$C_j = \frac{\partial L}{\partial x^j} + \sum_{|I|=0}^{k-1} u^{a}_{i+1} \left( \frac{\partial L}{\partial u^a_j} - \sum_{I+I_i=K} p_i^{ij} \right) - \sum_{|I|=0}^{k-1} u^{a}_{i+1} B_{ij}.$$  

(5.19)

Therefore, the coefficients of the multivector $X$ are governed by the equations of holonomy (5.12), the bottom- and mid-level equations of dynamics (5.13) and (5.14), and the equations of tangency (5.18) and (5.19).

Looking closer to the first equation of tangency (5.18), we may observe that, if the matrix of second-order partial derivatives of $L$ with respect to the ‘velocities’ of highest order

$$\left( \frac{\partial^2 L}{\partial u^R_k \partial u^i_k} \right)_{|R|=|K|=k}$$

is non-degenerate, then the highest order $A$s are completely determined in terms of the highest order $B$s. In the following, we will say that the Lagrangian $L : J^k \pi \longrightarrow \Lambda^m M$ is regular if, for any system of adapted coordinates, the matrix (5.20) is non-degenerate.

Note also that, thanks to the lemma B.3 and the top-level equation of dynamics (5.15), the terms in the equation of tangency (5.19) that have $A$s with a multi-index of length $k$ cancel out, while the terms that have $A$s with a lower multi-index are already determined (equation of holonomy (5.12)). So, in some sense, the $C$s depend only on the $B$s.

**Theorem 5.4.** Consider the submanifold $W_2 = W_0 \cap W_1 \hookrightarrow W$ and let $\Omega_2 = i^* \Omega_\eta$. We have that $(W_2, \Omega_2)$ is multisymplectic if and only if $\mathcal{L}$ is regular.

**Theorem 5.5.** Let $\sigma \in \Gamma_{\pi_{W,M}}$ be an integral section of a solution $X$ of the dynamical equation (5.6). Then, its ‘Lagrangian part’ $\sigma_1 = pr_1 \circ \sigma$ is holonomic, $\sigma_1 = j^k \phi$ for some section $\phi \in \Gamma \pi$, which furthermore satisfies the higher order Euler–Lagrange equations (4.7).

**Theorem 5.6.** Consider the system of linear equations in $B$s (of highest order) given by the mid-level equation of dynamics (5.14) and the tangency condition (5.18). This system always has the maximal rank but it is overdetermined when $k = 1$ or $m = 1$, completely determined when $k = m = 2$ and undetermined otherwise.
6. Variational calculus with constraints

We consider a constraint submanifold \( i : C \hookrightarrow J^l \pi \) of codimension \( l \), which is locally annihilated by \( l \) functionally independent constraint functions \( \Psi^\mu \), where \( 1 \leq \mu \leq l \). The constraint submanifold \( C \) is supposed to fiber over the whole of \( M \).

Remark 6.1. As our ultimate goal is to find holonomic jet sections that belong to \( \mathcal{C} \), one could look for a submanifold \( C' \) of \( C \) consisting of the image of such sections. The submanifold \( C' \) is given by the constraint functions \( C \) plus their consequences up to order \( k \), that is, \( \Psi^\mu \), \( \frac{\partial \Psi^\mu}{\partial \tilde{x}^r} \), \( \frac{\partial^2 \Psi^\mu}{\partial \tilde{x}^r \partial \tilde{x}^s} \), etc. Geometrically, \( C' \) is obtained as the output of the following recursive process:

\[
C^{(s,r)} := \begin{cases} 
C, & s = 0, \quad r = k; \\
\pi_{k,0}(C^{(s-1,k)}), & s > 0, \quad r = 0; \\
J^1(C^{(s,r-1)} \cap \pi_{k,1}(C^{(s-1,k)})), & s > 0, \quad 0 < r \leq k;
\end{cases}
\tag{6.1}
\]

which stops when, for some step \( s \geq 0 \), \( C^{(s+1,k)} = C^{(g,k)} \). This algorithm is a generalization to jet bundles of the method given in \([29]\) by Mendella et al to extract the integral part of a differential equation in a tangent bundle. The reader is also referred to the alternative approach by Gasqui [19].

For instance, if one considers the null divergence restriction \( u_x + v_y = 0 \) in the second-order jet manifold of \( pr_1 : \mathbb{R}^3 \times \mathbb{R}^2 \rightarrow \mathbb{R}^3 \), then the resulting manifold \( C^{(2,2)} = C^{(1,2)} \) is given by the restrictions \( u_x + v_y = 0, u_{xt} + v_{yt} = 0, u_{xx} + v_{xy} = 0 \) and \( u_{xy} + v_{yy} = 0 \) (see example 8.2).

We now look for extremals of the Lagrangian action \((4.1)\) restricted to those sections \( \phi \in \Gamma \pi \) whose \( k \)-jet takes values in \( C \) (see \([15, 28]\)). We will use the Lagrange multiplier theorem that follows.

Theorem 6.2 (Abraham, Marsden and Ratiu [1]). Let \( \mathcal{M} \) be a smooth manifold, \( f : \mathcal{M} \rightarrow \mathbb{R} \) be \( C' \), \( r \geq 1 \), \( \mathcal{F} \) a Banach space, \( g : \mathcal{M} \rightarrow \mathcal{F} \) a smooth submersion and \( N = g^{-1}(0) \). A point \( \phi \in N \) is a critical point of \( f|_N \) if and only if there exists \( \lambda \in \mathcal{F}^* \), called a Lagrange multiplier, such that \( \phi \) is a critical point of \( f - \langle \lambda, g \rangle \).

In order to apply the Lagrange multiplier theorem, we need to define constraints as the 0-level set of some function \( g \). We therefore configure the following setting: choose the smooth manifold \( \mathcal{M} \) to be the space of local sections \( \Gamma_R \pi = \{ \phi : R \subset M \rightarrow E : \pi \circ \phi = Id_M \} \), for some compact region \( R \subset M \). The Banach space \( \mathcal{F} \) is the set of smooth functions \( C^\infty(R, \mathbb{R}^l) \), provided with the \( L^2 \)-norm, for instance. The constraint function \( \Psi \) induces a constraint function on the space of local sections \( \Gamma_R \pi \) by mapping each section \( \phi \) to the evaluation of its \( k \)-lift by the constraint, that is,

\[
g : \phi \in \Gamma_R \pi \mapsto \Psi \circ j^k \phi \in C^\infty(R, \mathbb{R}^l).
\]

Note that the 0-level set \( \mathcal{N} = g^{-1}(0) \) is the set of sections whose \( k \)-lift takes values in the constraint manifold \( \mathcal{C} \) (over \( R \)).

We therefore obtain that a section \( \phi : M \rightarrow E \) is a critical point of the integral action \( \mathcal{S} \) restricted to \( C \) if and only if there exists a Lagrange multiplier \( \lambda \in (C^\infty(R, \mathbb{R}^l))^* \) such that \( \phi \) is a critical point of \( \mathcal{S} - \langle \lambda, g \rangle \). \textit{A priori}, we cannot assure that the pairing \( \langle \lambda, g(\phi) \rangle \) has an integral expression of the type \( \int_R \lambda \mu \Psi^\mu \circ j^k \phi \, dx \) for some functions \( \lambda \mu : R \rightarrow \mathbb{R} \). Henceforth, we shall suppose that that is the case.

Remark 6.3. In theorem 6.2 some regularity conditions appear that exclude the so-called abnormal solutions. In general, given a critical point \( \phi \in \mathcal{N} = g^{-1}(0) \) of \( f|_N \), the classical...
Lagrange multiplier theorem claims that there exists a nonzero element \((\lambda_0, \lambda) \in \mathbb{R} \times F^*\) such that \(\phi\) is a critical point of
\[
\lambda_0 f - (\lambda, g).
\] (6.2)
Under the submersivity condition on \(g\), that is \(\phi\) is a regular critical point, it is possible to guarantee that \(\lambda_0 \neq 0\), and dividing by \(\lambda_0\) in (6.2) we obtain the characterization of critical points given in theorem 6.2. The critical points \(\phi\) with vanishing Lagrange multiplier, that is, \(\lambda_0 = 0\) are called abnormal critical points.

In the following we will only study the regular critical points, but our developments are easily adapted for the case of abnormality (adding the Lagrange multiplier \(\lambda_0\) and studying separately both cases \(\lambda_0 = 0\) and \(\lambda_0 = 1\)).

**Proposition 6.4 (Constrained higher order Euler–Lagrange equations).** Let \(\phi \in \Gamma_1\pi\) be a critical point of the Lagrangian action \(S\) given in (4.1) restricted to those sections of \(\pi\) whose kth lift take values in the constraint submanifold \(C \subset J^k\pi\). If the associated Lagrange multiplier \(\lambda\) is regular enough, then there must exist \(l\) smooth functions \(\lambda_\mu : R \rightarrow M\) that satisfy together with \(\phi\) the constrained higher order Euler–Lagrange equations
\[
(j^{2k}\phi)^* \left( \sum_{|J|=0}^k (-1)^{|J|} \frac{d^{|J|}}{dx^{|J|}} \left( \frac{\partial L}{\partial u_\alpha} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u_\alpha} \right) \right) = 0.
\] (6.3)

### 7. Constrained mechanics in higher order field theories

As in the previous section, we begin by considering a constraint submanifold \(i : C \hookrightarrow J^k\pi\) of codimension \(l\), which is locally annihilated by \(l\) functionally independent constraint functions \(\Psi_\mu\), where \(1 \leq \mu \leq l\). The constraint submanifold \(C\) is supposed to fiber over the whole of \(M\) and it is not necessarily generated from a previous constraint submanifold by the process shown in remark 6.1. We define in the restricted velocity–momentum space \(W_0 = \{w \in W : H(w) = 0\}\) the constrained velocity–momentum space \(W^C_0 = pr^{-1}_1(C)\), which is a submanifold of \(W_0\), whose induced embedding and whose constraint functions will still be denoted \(i : W^C_0 \hookrightarrow W\) and \(\Psi_\mu\), where \(1 \leq \mu \leq l\). The first-order case \(k = 1\) is treated in [9].

The following proposition allows us to work in local coordinates on the unconstrained velocity–momentum space \(W\).

**Proposition 7.1.** Given a point \(w \in W^C_0\), let \(X \in \Lambda^n(T_wW^C_0)\) be a decomposable multivector and denote its image, \(i_*(X)\) by \(\bar{X}\). The following statements are equivalent:

1. \(i_*(X)\) is annihilated by \(\bar{X}\),
2. \(i_*(X) \in T_wW_0\),

where \(T^0_wW^C_0\) is the annihilator of \(i_*(T_wW^C_0)\) in \(T_wW\).

We therefore look for solutions of the constrained dynamical equation
\[
(-1)^n i_\bar{X} \Omega = -\lambda_\mu d\Psi_\mu - \lambda dH,
\] (7.1)
where \(\bar{X}\) is a tangent multivector field along \(W^C_0\), the \(\lambda_\mu\)'s and \(\lambda\) are Lagrange multipliers to be determined. Here, the coefficient \((-1)^n\) is used for technical purposes.

**Remark 7.2.** It should be said that the Lagrange multipliers that appear in the dynamical equation (7.1) have a different nature to the ones that appear in proposition 6.4. The former
are locally defined on $W$, while the latter are locally defined on $M$. Although they coincide on the integral sections $\sigma \in \Gamma_{\mathcal{T}W,M}$ of a solution $X$ of the dynamical equation (7.1), since its 'Lagrangian part' $\sigma_1 = pr_1 \circ \sigma$ satisfies the constrained Euler–Lagrange equation (6.3) with $\tilde{\lambda}_\mu = \lambda_\mu \circ \sigma$ (cf proposition 7.4).

Let $\bar{X} \in \Lambda^m(T_w W)$ be a decomposable $m$-vector at a given point $w \in W$, that is, $\bar{X} = \bar{X}_1 \wedge \cdots \wedge \bar{X}_m$ for $m$ tangent vectors $\bar{X}_i \in T_w W$, which have the form

$$\bar{X}_j = \frac{\partial}{\partial x^j} + A_{ij}^\alpha \frac{\partial}{\partial u^\alpha_j} + B_{ij}^\alpha \frac{\partial}{\partial p_i^\alpha} + C_j \frac{\partial}{\partial p_i^0}$$  \hspace{1cm} (7.2)

in a given adapted chart $(x^i, u^\alpha_j, p_i^\alpha, p_i^0)$. A straightforward computation gives us

$$(-1)^m i^j_X (d p_i^\alpha \wedge du_j^\alpha \wedge d^{m-1} x_i) = \left( A_{ij}^\alpha B_j^\alpha_i - A_i^\alpha B_{ji}^\alpha \right) dx^j + A_i^\alpha d p_i^\alpha - B_{ij}^\alpha du_j^\alpha$$  \hspace{1cm} (7.3)

and

$$(-1)^m i_X^j (d p \wedge d^{m-1} x_j) = d p - C_i dx^i.$$  \hspace{1cm} (7.4)

Applying the above equations to the dynamical one (7.1), we obtain the relations

\[ \text{coefficients in } d p: \quad 1 = \lambda; \]
\[ \text{coefficients in } d p_i^\alpha: \quad A_i^\alpha = \lambda A_i^\alpha; \]
\[ \text{coefficients in } du_j^\alpha: \quad B_{ij}^\alpha = \lambda B_{ij}^\alpha; \]
\[ \text{coefficients in } dx^j: \quad A_{ij}^\alpha B_j^\alpha_i - A_i^\alpha B_{ji}^\alpha = \lambda A_i^\alpha B_{ji}^\alpha; \]
\[ \text{coefficients in } dx^j: \quad \lambda \frac{\partial L}{\partial u^\alpha_j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_j} = B_{ij}^\alpha \]
\[ \sum_{I+1,k} p_i^\alpha = \lambda \frac{\partial L}{\partial u^\alpha_j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_j} = B_{ij}^\alpha \quad \text{with} \quad |I| = 1, \ldots, k - 1; \]
\[ \sum_{I+1,k} p_i^\alpha = \lambda \frac{\partial L}{\partial u^\alpha_k} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_k} \quad \text{with} \quad |K| = k; \]
\[ C_j = \lambda \frac{\partial L}{\partial x^j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial x^j} + \sum_{|J|=0}^{k-1} u_i^\alpha_{I+1} \left( \frac{\partial L}{\partial u^\alpha_j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_j} - \sum_{I+1} p_i^\alpha \right) - u_i^\alpha_{I+1} B_{ij}^\alpha. \]

Thus, a decomposable $m$-vector $\bar{X} \in \Lambda^m(T_w W)$ at a point $w \in W$ is a solution of the dynamical equation

$$(-1)^m i_X^j \Omega = -\lambda_\mu \frac{d \Psi_\mu}{d H},$$  \hspace{1cm} (7.5)

if for any adapted chart $(x^i, u^\alpha_j, p_i^\alpha, p_i^0)$, the coefficients of $\bar{X}$ and the point $w$ satisfy the equations

$$A_{ij}^\alpha = u_i^\alpha_{I+1}, \quad \text{with} \quad |I| = 0, \ldots, k - 1, \quad i = 1, \ldots; \]

$$0 = \lambda \frac{\partial L}{\partial u^\alpha_j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_j} = B_{ij}^\alpha; \]

$$\sum_{I+1} p_i^\alpha = \lambda \frac{\partial L}{\partial u^\alpha_j} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_j} = B_{ij}^\alpha \quad \text{with} \quad |J| = 1, \ldots, k - 1; \]

$$\sum_{I+1} p_i^\alpha = \lambda \frac{\partial L}{\partial u^\alpha_k} - \lambda_\mu \frac{\partial \Psi_\mu}{\partial u^\alpha_k} \quad \text{with} \quad |K| = k; \]

Because of the Lagrange multipliers $\lambda_\mu$, we cannot describe the submanifold of $W^C$ where solutions $X$ of the constrained dynamical equation (7.5) exist, as has been done in (5.15) for
the unconstrained dynamical equation (5.6). Therefore, we need to get rid of them. Consider
the more concise expression for the equations of dynamics (7.7)–(7.9):
\[
\sum_{I+1=J} p_\alpha^{(i)} = \frac{\partial L}{\partial u_j^\alpha} - \lambda_{J\mu} \frac{\partial \Psi^\mu}{\partial u_j^\alpha} - B^{(j)}_{\alpha j}, \quad \text{with} \quad |J| = 0, \ldots, k. \tag{7.11}
\]
where, as in (5.16), the first summation term is understood to be void when \(|J| = 0\), as well as it is the last one when \(|J| = k\). We now suppose that the constraints \(\Psi^\mu\) are of the type \(u_j^\alpha = \Phi_j^\alpha(x^i, u_j^\alpha)\), where \(u_j^\alpha\) are some constrained coordinates which depend on the free coordinates \((x^i, u_j^\alpha)\) through the functions \(\Phi_j^\alpha\). Thus, the constraint has the form \(\Phi_j^\alpha(x^i, u_j^\alpha) = u_j^\alpha - \Phi_j^\alpha(x^i, u_j^\alpha) = 0\). So writing again the previous equation (7.11) for the different sets of coordinates, the ones that are free and the ones that are not, we obtain
\[
\sum_{I+1=J} p_\alpha^{(i)} = \frac{\partial L}{\partial u_j^\alpha} - \lambda_{J\alpha j} - B^{(j)}_{\alpha j} \quad \text{with} \quad |J| = 0, \ldots, k; \tag{7.12}
\]
\[
\sum_{I+1=J} p_\alpha^{(i)} = \frac{\partial L}{\partial u_j^\alpha} - \lambda_{J\alpha j} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} - B^{(j)}_{\alpha j} \quad \text{with} \quad |J| = 0, \ldots, k. \tag{7.13}
\]
Substituting \(-\lambda_{J\alpha j}\) from (7.12) into (7.13), we have that
\[
\sum_{I+1=J} p_\alpha^{(i)} + \left( \sum_{I+1=J} p_\alpha^{(i)} \right) \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} = \frac{\partial L}{\partial u_j^\alpha} + \frac{\partial L}{\partial u_j^\alpha} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} - B^{(j)}_{\alpha j} - B^{(j)}_{\alpha j} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} \quad \text{with} \quad |J| = 0, \ldots, k. \tag{7.14}
\]
Note that, when \(|J| = k\), the term \(B^{(j)}_{\alpha j}\) disappears, but \(B^{(j)}_{\alpha j} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha}\) do not necessarily. This is circumvented by supposing that if \(|J| < k\), then \(\frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} = 0\) for any \(|K| = k\). That is the case when the constraint submanifold \(C\) has no constraint of higher order, i.e. \(C = \pi_{k,k-1}(\pi_{k,k-1}(C))\) or more generally when \(C\) fibers by \(\pi_{k,k-1}\) over its image.
Taking this into account, we expand the previous equation (7.14), obtaining constrained
equations of dynamics freed of the Lagrange multipliers
\[
\sum_{I+1=J} p_\alpha^{(i)} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} = \frac{\partial L}{\partial u_j^\alpha} - B^{(j)}_{\alpha j} - B^{(j)}_{\alpha j} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha}; \tag{7.15}
\]
\[
\sum_{I+1=J} p_\alpha^{(i)} + \sum_{I+1=J} p_\alpha^{(i)} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} = \frac{\partial L}{\partial u_j^\alpha} - B^{(j)}_{\alpha j} - B^{(j)}_{\alpha j} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha} \quad \text{with} \quad |J| = 1, \ldots, k - 1; \tag{7.16}
\]
\[
\sum_{I+1=K} p_\alpha^{(i)} + \sum_{I+1=K} p_\alpha^{(i)} \frac{\partial \Phi_k^\alpha}{\partial u_k^\alpha} = \frac{\partial L}{\partial u_k^\alpha} \quad \text{with} \quad |K| = k, \tag{7.17}
\]
where \(\frac{\partial L}{\partial u_j^\alpha} = \frac{\partial L}{\partial u_j^\alpha} + \frac{\partial L}{\partial u_k^\alpha} \frac{\partial \Phi_j^\alpha}{\partial u_j^\alpha}\), being \(\mathcal{L}^C = \mathcal{L} \circ i : C \rightarrow \Lambda^m M\) the restricted Lagrangian.
We are now in a position to define the submanifold $W_2^C$ along which solutions of the constrained dynamical equation (7.5) exist:

$$W_2^C = \{ w \in W_0^C : \text{(7.17)} \} = \left\{ w \in W : \begin{array}{l}
u_j^a = \Phi_j^a(x^i, u_j^a) \\
p = L(x^i, u_j^a) - p_j^l u_i^a, \\
\sum_{l+1=K} p_j^l + \sum_{l+1=K} p_j^l \frac{\partial \Phi_k^a}{\partial u_k^a} = \frac{\partial L^C}{\partial u_k^a}
\end{array} \right\}. \tag{7.18}$$

Tangency conditions on $X$ with respect to $W_2^C$ will give us the constrained equations of tangency

$$A_{ij} = \frac{\partial \Phi_j^a}{\partial x^i} + A_{ij}^a \frac{\partial \Phi_j^a}{\partial u_j^a}, \tag{7.19}$$

$$C_j = \frac{\partial L^C}{\partial x^j} + \sum_{|\alpha|=0}^{k-1} u_{j+1}^\alpha \left( \frac{\partial L^C}{\partial u_j^\alpha} - \sum_{l+1=K} p_j^l \right) - \sum_{\alpha=0}^{k-1} u_{j+1}^\alpha \sum_{|\alpha|=0}^{k-1} p_j^l - \sum_{\alpha=0}^{k-1} u_{j+1}^\alpha B_{ij}^l, \tag{7.20}$$

$$\sum_{l+1=K} B_{ij}^l = \frac{\partial^2 L^C}{\partial x^j \partial u_k^a} - \sum_{l+1=K} p_j^l \frac{\partial^2 \Phi_k^a}{\partial x^j \partial u_k^a} + A_{ij}^a \left( \frac{\partial^2 L^C}{\partial u_j^\alpha \partial u_k^a} - \sum_{l+1=K} p_j^l \frac{\partial^2 \Phi_k^a}{\partial u_j^\alpha \partial u_k^a} \right) - \sum_{l+1=K} B_{ij}^l \frac{\partial \Phi_k^a}{\partial u_k^a}, \quad |\tilde{K}| = k. \tag{7.21}$$

**Proposition 7.3.** Let $\Omega_2^C$ be the pullback of the premultisymplectic form $\Omega_\Pi$ to $W_2^C$ by the natural inclusion $i : W_2^C \hookrightarrow W$, that is, $\Omega_2^C = i^*(\Omega_\Pi)$. Suppose that $m = \dim M > 1$, then the $(m+1)$-form $\Omega_2^C$ is multisymplectic if and only if $L$ is regular along $W_2^C$, i.e. if and only if the matrix

$$\left( \frac{\partial^2 L^C}{\partial u_k^a \partial u_k^a} - \sum_{l+1=K} p_j^l \frac{\partial^2 \Phi_k^a}{\partial u_k^a \partial u_k^a} \right)_{|k|=|K|=k} \tag{7.22}$$

is non-degenerate along $W_2^C$.

**Proof.** First of all, let us make some considerations. By definition, $\Omega_2^C$ is multisymplectic whenever $\Omega_2^C$ has a trivial kernel, that is,

- if $v \in TW_2$, $i_v \Omega_2^C = 0 \iff v = 0$.

This is equivalent to say that

- if $v \in i_*(TW_2)$, $i_v \Omega_\Pi|_{i_*(TW_2)} = 0 \iff v = 0$.

Let $v \in TW$ be a tangent vector whose coefficients in an adapted basis are given by

$$v = v^i \frac{\partial}{\partial x^i} + A_{ij}^a \frac{\partial}{\partial u_j^a} + B_{ij}^l \frac{\partial}{\partial u_i^l} + C \frac{\partial}{\partial p_j^l}.$$
Using expression (3.6), we may compute the contraction of $\Omega_{ii}$ by $v$: 

$$i_\ast \Omega_{ii} = -B^a_B \, du^a \wedge d^{m-1} x_i + A^a_B \, dp^B_a \wedge d^{m-1} x_i - \gamma^l \, dp^B_a \wedge du^a \wedge d^{m-2} x_{ij}$$

$$+ \left( A^a_B, p^B_a + B^a_B u^a_{ij} - A^a_B \partial L / \partial u^a_j \right) \wedge d^m x$$

$$- \gamma^l \left( p^B_a \, du^a_i \wedge u^a_{ij}, dp^B_a - \partial L / \partial u^a_j \right) \wedge d^{m-1} x_j.$$

(7.23)

In addition to this, let us consider a vector $v \in TW$ tangent to $W_2$, that is, $v \in i_\ast (TW_2)$; we then have that

$$d(u^a_j - \Phi^a_j)(v) = 0, \quad \left( \sum_{l+1=k} B^l_{ij} \, \partial^2 L / \partial x^l / \partial u^a_k \right) + \left( \sum_{l+1=k} p^a_l \, \partial^2 \Phi^a_j / \partial x^l / \partial u^a_k \right)$$

$$+ A^a_j \left( \partial^2 L / \partial u^a_j / \partial x^l \right) + \left( \sum_{l+1=j} p^a_l \, \partial^2 \Phi^a_k / \partial x^l / \partial u^a_j \right) - B^a_B u^a_{ij}.$$  

(7.25)

$$C = \gamma^l \left( \partial^2 L / \partial x^l / \partial u^a_j \right) + \left( \sum_{l+1=k} p^a_l \, \partial^2 \Phi^a_j / \partial x^l / \partial u^a_k \right) - B^a_B u^a_{ij}.$$  

(7.26)

It is important to note that, even though in the previous equations (7.23)–(7.26) As explicitly appear with multi-index of length $k$, for such a vector $v \in i_\ast (TW_2)$, the terms associated with these $A$s cancel out in the development of $i_\ast \Omega_{ii}$, equation (7.23) and the third tangency relation (7.26). Thus, a tangent vector $v \in i_\ast (TW_2)$ would kill $\Omega_{ii}$ if and only if its coefficients satisfy the following relations:

$$\gamma^l = 0, \quad A^a_j = 0, \quad B^a_B = 0, \quad C = 0,$$

$$A^a_B = A^a_B \left( \partial^2 L / \partial u^a_j / \partial u^a_k \right), \quad A^a_B \left( \partial^2 L / \partial u^a_j / \partial u^a_k \right) = 0.$$  

These considerations being made, the assertion is now clear.

\(\square\)

**Proposition 7.4.** Let $\sigma \in \Gamma \pi_{W,M}$ be an integral section of a solution $X$ of the constrained dynamical equation (7.5). Then, its ‘Lagrangian part’ $\sigma_1 = pr_1 \circ \sigma$ is holonomic, $\sigma_1 = j^l \Phi$ for some section $\Phi \in \Gamma \pi$, which furthermore satisfies the constrained higher order Euler–Lagrange equations (6.3).

**Proof.** If $X$ is locally expressed as in (5.11), we know that it must satisfy the equations of dynamics (7.7)–(7.9), for unknown Lagrange multipliers $\lambda_{ij}$. If we note $\lambda_{ij} = \lambda_{ij} \circ \sigma$ and $L' = L - \lambda_{ij} \Psi_{ij}$, it suffices to follow the demonstration for $L'$ of theorem 5.5 which is proved in [6].

\(\square\)
8. Example

Here, we study an incompressible fluid under control as in [2]. The corresponding equations are the Navier–Stokes one plus the divergence-free condition:

\[
\frac{\partial \mathbf{v}}{\partial t} + \nabla \mathbf{v} + \nabla \Pi = \nu \Delta \mathbf{v} + \mathbf{f} \quad (8.1)
\]

\[
\nabla \cdot \mathbf{v} = 0, \quad (8.2)
\]

where the vector field \( \mathbf{v} \) is the velocity of the fluid, \( \mathbf{f} \) is the field of exterior forces acting on the fluid, which will be our controls, and the scalar functions \( \Pi \) and \( \nu \) are the pressure and the viscosity, respectively. In particular, our interest is in the two-dimensional case on the fluid, which will be our controls, and the scalar functions \( \nu/Delta_1 \), where the vector field \( \mathbf{v} \) are the Navier–Stokes one plus the divergence-free condition:

\[
u = \frac{\nabla \cdot \mathbf{v}}{\Pi_1}
\]

Moreover, we look for such vector fields \( \mathbf{v} = (u, v) \) that are in addition optimal in the controls for the integral action

\[
\mathcal{S}(\mathbf{v}, R) = \frac{1}{2} \int_R \|\mathbf{f}\|^2 \, dt \wedge dx \wedge dy. \quad (8.6)
\]

In order to apply the development of the present jet bundle framework, all of this is restated in the following way: we set a fiber bundle \( \pi : E \longrightarrow M \) by putting \( M = \mathbb{R} \times \mathbb{R}^2 \), \( E = \mathbb{R} \times T\mathbb{R}^2 \) and \( \pi = (pr_1, pr_{\mathbb{R}^2}) \). We fix global adapted coordinates \( (t, x, y, u, v) \) on \( E \), which induce the corresponding global adapted coordinates on \( J^2\pi \) and \( J^k\pi \). Besides, we choose the volume form \( \eta \) on \( M \) to be \( dt \wedge dx \wedge dy \). Thus, the Lagrangian function \( L : J^2\pi \rightarrow \mathbb{R} \) is nothing else but

\[
L = \frac{1}{2}(F^2 + G^2),
\]

where we obtain \( F \) and \( G \) as functions on \( J^2\pi \) using equations (8.3) and (8.4).

To make the reading easier, we change slightly the coordinate notation of jet bundles to fit in this example. The coordinate ‘velocities’ associated with \( u \) and \( v \) will still be labeled \( u \) and \( v \), respectively, with symmetric subindexes (as in the original equations); the coordinate ‘momenta’ associated with \( u \) and \( v \) will now be labeled \( p \) and \( q \), respectively, with non-symmetric subindexes. Finally and as we will focus on the equations of dynamics (7.15)–(7.17), the coefficients in the local expression (5.11) of a multivector \( X \) associated to the coordinate momenta \( p \) and \( q \) will be labeled \( B \) and \( D \), respectively.

**Example 8.1 (The Euler equation)**. We will first suppose that the fluid is Eulerian, that is, it has null viscosity. In this case, the Lagrangian function \( L = (F^2 + G^2)/2 \) associated with the integral action (8.6) is of first order when the ‘Euler equations’, (8.3) and (8.4) with \( \nu = 0 \), are taken into account. In \( J^1\pi \), we consider the divergence-free constraint submanifold \( \mathcal{C} = \{ z \in J^1\pi : u_x + u_y = 0 \} \), which introduces a single Lagrange multiplier \( \lambda \).
Proceeding with the theoretical machinery, we compute the bottom-level equations of dynamics corresponding to those of (5.13)

\[\begin{align*}
0 &= u_x \cdot F + v_x \cdot G - (B_t^x + B^x + B^y) \\
0 &= u_y \cdot F + v_y \cdot G - (D_t^x + D^x + D^y)
\end{align*}\]

and the top-level equations of dynamics (there are no middle ones) corresponding to those of (5.15)

\[\begin{align*}
p' &= F & q' &= G \\
p^x &= u \cdot F - \lambda & q^x &= u \cdot G \\
p^y &= v \cdot F & q^y &= v \cdot G - \lambda.
\end{align*}\]

We can dispose of the only Lagrange multiplier \(\lambda\) by putting

\[p^x - q^x = u \cdot F - v \cdot G,\]

which defines \(W^C_2\) together with the top-level equations of dynamics with no Lagrange multiplier.

From here, we may also compute the constrained Euler–Lagrange equations (6.3) for this problem, which are

\[\begin{align*}
d_t F + u \cdot d_x F + v \cdot d_y F + v_x \cdot F - v_x \cdot G &= 0 \\
d_t G + u \cdot d_x G + v \cdot d_y G + u_y \cdot G - u_y \cdot F &= 0
\end{align*}\]

where \(d_t = \frac{4}{\partial t}\).

Finally, we note that \(L\) is not regular along \(W^C_2\) since the square matrix, which corresponds to (7.22),

\[
\begin{pmatrix}
1 & u & v & 0 & 0 \\
u & u^2 + v^2 & u \cdot v & -v & -u \cdot v \\
v & u \cdot v & v^2 & 0 & 0 \\
0 & -v & 0 & 1 & u \\
0 & -u \cdot v & 0 & u & u^2
\end{pmatrix}
\]

has obviously rank 2. Here we have used \(u_t\) as the independent (‘check’) coordinate and \(v_y\) as the dependent (‘hat’) coordinate.

**Example 8.2** (The Navier–Stokes equation). Now we tackle the full problem of the Navier–Stokes equations. In this case, the Lagrangian function \(L = (F^2 + G^2)/2\) is of second order. In \(J^2\), we consider the constraint submanifold

\[C = \{ \zeta \in J^2 : u_x + u_y = 0, u_{xx} + v_{yy} = 0, u_{x1} + v_{y1} = 0, u_{xy} + v_{yx} = 0 \},\]

which comes from the first-order constraint (8.2), free divergence, and its consequences to second order (see remark 6.1). These constraints introduce for the Lagrange multiplier \(\lambda, \lambda_t, \lambda_x, \lambda_y\), which are associated with them respectively.

We now proceed like in the previous example by computing the equations of dynamics. In the first place, we have the bottom-level ones corresponding to those of (5.13)

\[\begin{align*}
0 &= u_x \cdot F + v_x \cdot G - (B_t^x + B^x + B^y) \\
0 &= u_y \cdot F + v_y \cdot G - (D_t^x + D^x + D^y).
\end{align*}\]

Note that they are formally the same as before. In the second place, there are the mid-level equations corresponding to those of (5.14):

\[\begin{align*}
p^t &= F - (B_t^x + B^x + B^y) & q^t &= G - (D_t^x + D^x + D^y) \\
p^x &= u \cdot F - (B_t^{xx} + B^x + B^y) + \lambda & q^x &= u \cdot G - (D_t^{xx} + D^x + D^y) \\
p^y &= v \cdot F - (B_t^{yy} + B^x + B^y) & q^y &= v \cdot G - (D_t^{yy} + D^x + D^y) + \lambda.
\end{align*}\]
Note that formally they also coincide with the top-level ones of the previous example but for the coefficients that now appear in them. And in the third place, we have the top-level equations corresponding to those of (5.15):

\[
\begin{align*}
 p_t^t &= 0 \\
 q_t^t &= 0 \\
 p_x^x &= -v \cdot F - \lambda_x \\
 q_x^x &= -v \cdot G \\
 p_y^y &= -v \cdot F \\
 q_y^y &= -v \cdot G - \lambda_y \\
 p_x^x + p_t^t &= -\lambda_x \\
 q_x^x + q_t^t &= 0 \\
 p_y^y + p_t^t &= 0 \\
 q_y^y + p_t^t &= -\lambda_t \\
 p_y^y + p_x^x &= -\lambda_x.
\end{align*}
\]

We can again get rid easily of the Lagrange multipliers by putting

\[
\begin{align*}
 p_{tx} + p_{xt} &= q_{ty} + q_{yt} \\
 p_{xx} + v \cdot F &= q_{xy} + q_{yx} \\
 p_{yy} + v \cdot G &= q_{yy} + v \cdot G.
\end{align*}
\]

which defines \( W^C \) together with the top-level equations of dynamics with no Lagrange multiplier.

From here, we may also compute the constrained Euler–Lagrange equations (6.3) for this problem, which are

\[
\begin{align*}
 2\partial_{xx}^2 \lambda_t + 2\partial_{xx}^2 \lambda_x + 2\partial_{xy}^2 \lambda_y - \partial_t \lambda = & \partial_{xx}^2 v \cdot F + 2\partial_x v \cdot d_x F + v \cdot d_x^2 F \\
 + & \partial_{yy}^2 v \cdot F + 2\partial_y v \cdot d_y F + v \cdot d_y^2 F \\
 - & d_t F - u \cdot d_x F - v \cdot d_y F - v_y \cdot F + v_x \cdot G \\
 2\partial_{xy}^2 \lambda_t + 2\partial_{xy}^2 \lambda_x + \partial_y \lambda_y - \partial_t \lambda = & \partial_{xy}^2 v \cdot G + 2\partial_y v \cdot d_y G + v \cdot d_y^2 G \\
 + & \partial_{yy}^2 v \cdot G + 2\partial_y v \cdot d_y G + v \cdot d_y^2 G \\
 - & d_t G - u \cdot d_x G - v \cdot d_y G - u_x \cdot G + u_y \cdot F.
\end{align*}
\]

As before, the Lagrangian is not regular along \( W^C \), which seems to be clear if we observe that \( L \) is highly non-degenerate: it depends only on 4 of the 12 coordinates of second order. It is worthless to show that it is ‘Hessian’, even though it is interesting to say that it is null only when \( v \) is.

9. Conclusions and future work

In this paper, we have introduced an unambiguous geometric formalism for higher order field theories subjected to constraints, with applications to optimal control of partial differential equations. Our theory is based on the classical Skinner and Rusk formalism and the theory of higher order jet bundles. In the future, we will do a detailed study of how to derive a constraint algorithm; in particular, we will pay attention to the necessary conditions in order that the algorithm does not stop at the primary constraint submanifold, for instance, restricting the dual jet bundle to an appropriate submanifold (in agreement with the ideas behind [32]). The design of variational integrators for these higher order field equations will also be analyzed (see [28] for the first-order case).

It is worth mentioning now that the BRST construction for field theory has been used systematically to deal with the problem of quantizing gauge systems (cf [21]). Gauge systems are an instance of constrained field theories where the constraints arise because of the gauge invariance of the system. The BRST construction is based on the introduction of auxiliary odd and even degrees of freedom (referred to as ‘ghosts’ in physics literature) in such a way that the supermanifold structure of the extended formalism allows for a cohomological
description of the reduced spaces, i.e. of the relevant degrees of freedom from a physical point of view (see for instance [22] for a simple description of the reduction procedure in terms of supergeometry). This construction can be put in a nice geometrical setting for first-order Lagrangian mechanics (see for instance [16]) and we are looking forward to extending such formalism to higher order Lagrangian mechanics and Lagrangian field theories.

Acknowledgments

This work has been partially supported by the MICINN, Ministerio de Ciencia e Innovación (Spain), project MTM2007-62478, project ‘Ingenio Mathematica’ (i-MATH) no. CSD 2006-00032 (Consolider-Ingenio 2010). The authors want to specially thank Professor Alberto Ibort Latre for giving us some insight into the BRST formalism. The first author (CC) also acknowledges the MICINN for an FPI grant.

Appendix A. Multivectors

Let \( P \) be an \( n \)-dimensional differentiable manifold. Sections of \( \Lambda^m(TP) \) (with \( 1 \leq m \leq n \)) are called \( m \)-\textit{multivector fields} in \( P \). We will denote by \( \mathfrak{X}^m(P) \) the set of \( m \)-\textit{multivector fields} in \( P \). Given \( X \in \mathfrak{X}^m(P) \), for every \( p \in P \), there exists an open neighborhood \( U_p \subset P \) and \( X_1, \ldots, X_r \in \mathfrak{X}(U_p) \) such that

\[
X^U_p \equiv \sum_{1 \leq i_1 < \cdots < i_m \leq r} f^{i_1 \cdots i_m} X_{i_1} \wedge \cdots \wedge X_{i_m}
\]

with \( f^{i_1 \cdots i_m} \in \mathcal{C}^\infty(U_p) \) and \( m \leq r \leq n \). A multivector field \( X \in \mathfrak{X}^m(P) \) is \textit{locally decomposable} if, for every \( p \in P \), there exists an open neighborhood \( U_p \subset P \) and \( X_1, \ldots, X_m \in \mathfrak{X}(U_p) \) such that

\[
X^U_p \equiv X_1 \wedge \cdots \wedge X_m.
\]

We will denote by \( \mathfrak{X}^m_{ld}(P) \) the set of locally decomposable \( m \)-\textit{multivector fields} in \( P \).

Let \( D \subseteq TP \) be an \( m \)-dimensional distribution. The sections of \( \Lambda^m D \) are locally decomposable \( m \)-\textit{multivector fields} in \( P \). A locally decomposable \( m \)-\textit{multivector field} \( X \in \mathfrak{X}^m_{ld}(P) \) and an \( m \)-dimensional distribution \( D \subseteq TP \) are \textit{associated} whenever \( X \) is a section of \( \Lambda^m D \). If \( X, X' \in \mathfrak{X}^m_{ld}(P) \) are non-vanishing multivector fields associated with the same distribution \( D \), then there exists a non-vanishing function \( f \in \mathcal{C}^\infty(P) \) such that \( X' = fX \). This fact defines an equivalence relation in the set of non-vanishing \( m \)-\textit{multivector fields} in \( P \), whose equivalence classes will be denoted by \( \mathcal{D}(X) \). There is a bijective correspondence between the set of \( m \)-dimensional orientable distributions \( D \) in \( TP \) and the set of the equivalence classes \( \mathcal{D}(X) \) of non-vanishing, locally decomposable \( m \)-\textit{multivector fields} \( X \) in \( P \). By abuse of notation, \( \mathcal{D}(X) \) will also denote the \( m \)-dimensional orientable distribution \( D \) in \( TP \) with whom \( X \) is associated.

An \( m \)-dimensional submanifold \( S \hookrightarrow P \) is said to be an \textit{integral manifold} of \( X \in \mathfrak{X}^m(P) \) (resp. of an \( m \)-\textit{dimensional distribution} \( D \) in \( TP \)) if \( X \) spans \( \Lambda^m TS \) (resp. if \( TS = D \)). In such a case, \( X \) (resp. \( D \)) is said to be \textit{integrable}. Integrable multivector fields shall be locally decomposable. A non-vanishing, locally decomposable \( m \)-\textit{multivector field} \( X \in \mathfrak{X}^m_{ld}(P) \) is \textit{involutive} if its associated distribution \( \mathcal{D}(X) \) is involutive, that is, if \( [\mathcal{D}(X), \mathcal{D}(X)] \subseteq \mathcal{D}(X) \). If a non-vanishing multivector field \( X \in \mathfrak{X}^m_{ld}(P) \) is involutive, so is every other in its equivalence class \( \mathcal{D}(X) \). By Frobenius’ theorem, a non-vanishing and locally decomposable multivector field is integrable if, and only if, it is involutive.
Now, let $\pi : P \rightarrow M$ be a fiber bundle with dim $M = m$. A multivector field $X \in \mathfrak{X}^m(P)$ is said to be $\pi$-transverse if $\Lambda^m \pi_*(X)$ does not vanish at any point of $M$; hence, $M$ must be orientable. If $X \in \mathfrak{X}^m(P)$ is integrable, then $X$ is $\pi$-transverse if, and only if, its integral manifolds are sections of $\pi : P \rightarrow M$. In this case, if $S$ is an integral manifold of $X$, then there exists a section $\phi \in \Gamma \pi$ such that $S = \text{Im}(\phi)$.

For more details on multivector fields and their relation with field theories, refer to [13, 14].

Appendix B. The multi-index notation

Given a function $f : \mathbb{R}^m \rightarrow \mathbb{R}$, its partial derivatives are classically denoted as

$$ f_{i_1i_2\ldots i_k} = \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}. $$

When smooth functions are considered, their crossed derivatives coincide. Thus, the order in which the derivatives are taken is not important, but the number of times with respect to each variable is.

Another notation to denote partial derivatives is defined through 'symmetric' multi-indexes (see [31]). A multi-index $I$ will be an $m$-tuple of non-negative integers. The $i$th component of $I$ is denoted as $I(i)$. Addition and subtraction of multi-indexes are defined component-wise (whenever the result is still a multi-index)

$$ (I \pm J)(i) = I(i) \pm J(i). $$

The length of $I$ is the sum $|I| = \sum I(i)$, and its factorial $I! = \prod I(i)!$. In particular, $1^i$ will be the multi-index that is zero everywhere except at the $i$th component which is equal to 1.

Keeping in mind the above definition, we shall denote the partial derivatives of a function $f : \mathbb{R}^m \rightarrow \mathbb{R}$ by

$$ f_I = \frac{\partial^{|I|} f}{\partial x_{I(1)} \partial x_{I(2)} \ldots \partial x_{I(m)}}. $$

Thus, given a multi-index $I$, $I(i)$ denotes the number of times the function is differentiated with respect to the $i$th component. The former notation should not be confused with the latter one. For instance, the third-order partial derivative $\frac{\partial^3 f}{\partial x^2 \partial x^3}$ (with $f : \mathbb{R}^4 \rightarrow \mathbb{R}$) is denoted by $f_{232}$ and $f_{0,2,1,0}^3$, respectively.

Here we present some simple but useful results on multi-indexes.

**Lemma B.1.** Let $\{a_{I,i}\}_{I,i}$ be a family of real numbers indexed by a multi-index $I \in \mathbb{N}^m$ and by an integer $i$ such that $1 \leq i \leq m$. Given an integer $k \geq 1$, we have that

$$ \sum_{|I|=k} \sum_{i=1}^m a_{I,i} = \sum_{|I|=k} \sum_{I+1,i} a_{I,i}. \tag{B.1} $$

**Lemma B.2.** Let $J \in \mathbb{N}^m$ be a non-zero multi-index. We have that

$$ \sum_{I+1,J} I(i) + 1 \frac{1}{|I| + 1} = 1. \tag{B.2} $$

**Lemma B.3.** Let $\{a_J\}_J$ be a family of real numbers indexed by a multi-index $J \in \mathbb{N}^m$. Given a positive integer $l \geq 1$, we have

$$ \sum_{|J|=l} a_J = \sum_{|J|=l} \sum_{i=1}^m I(i) + 1 \frac{1}{|I| + 1} a_{I+1}. \tag{B.3} $$
Lemma B.4. Let \( \{a_J, b^J\}_J \) be a family of real numbers indexed by a multi-index \( J \in \mathbb{N}^m \). Given an integer \( l \geq 1 \), we have

\[
\sum_{|J|=l} b^J a_J = \sum_{|I|=l-1} \sum_{i=1}^m \frac{I(i) + 1}{|I| + 1} (b^I, + Q^I_J) a_{I+1,i},
\]

where \( \{Q^I_J\}_I,J \) is a family of real numbers such that for any multi-index \( J \in \mathbb{N}^m \) (with \( |J| \geq 1 \)) we have

\[
\sum_{I+1,i=J} \frac{I(i) + 1}{|I| + 1} Q^I_J = 0.
\]

References

[1] Abraham R, Marsden J E and Ratiu T 1988 Manifolds, Tensor Analysis, and Applications 2nd edn (Applied Mathematical Sciences vol 75) (New York: Springer)

[2] Agrachev A and Sarychev A 2008 Solid controllability in fluid dynamics Instability in Models Connected with Fluid Flows I (Int. Math. Ser. (NY) vol 6) (New York: Springer) pp 1–35

[3] Aldaya V and de Azcárraga J A 1978 Variational principles on rth order jets of fibre bundles in field theory J. Phys. A: Math. Gen. 13 2545–51

[4] Campos C M, de León M, Martín de Diego D and Vankerschaver J 2009 Unambiguous formalism for higher order Lagrangian field theories J. Phys. A: Math. Theor. 42 475207

[7] Cantrijn F, Ibort A and de León M 1996 Hamiltonian structures on multisymplectic manifolds Rend. Sem. Mat. Univ. Politec. Torino 54 225–36

[8] Cantrijn F, Ibort A and de León M 1999 On the geometry of multisymplectic manifolds J. Austral. Math. Soc. Ser. A 66 303–30

[9] Cantrijn F and Vankerschaver J 2007 The Skinner–Rusk approach for vakonomic and nonholonomic field theories Differential Geometric Methods in Mechanics and Field Theory (Volume in Honour of Willy Sarlet) (Ghent: Academia Press) pp 1–14

[10] Cariñena J F, Crampin M and Ibort L A 1991 On the multisymplectic formalism for first order field theories Differ. Geom. Appl. 1 345–74

[11] Cortés J, Martínez S and Cantrijn F 2002 Skinner–Rusk approach to time-dependent mechanics Phys. Lett. A 300 250–8

[12] Echeverría-Enríquez A, López C, Marín-Solano J, Muñoz-Lecanda M C and Román-Roy N 2004 Lagrangian–Hamiltonian unified formalism for field theory J. Math. Phys. 45 360–80

[13] Echeverría-Enríquez A, Muñoz-Lecanda M C and Román-Roy N 1998 Multivector fields and connections: setting Lagrangian equations in field theories J. Math. Phys. 39 4578–603

[14] Echeverría-Enríquez A, Muñoz-Lecanda M C and Román-Roy N 1999 Multivector field formulation of Hamiltonian field theories: equations and symmetries J. Phys. A: Math. Gen. 32 8461–84

[15] Fernández Martínez A, García Pérez P L and Rodrigo C 2004 Stress–energy–momentum tensors for natural constrained variational problems J. Geom. Phys. 49 1–20

[16] Forger M and Kellendonk J 1992 Classical BRST cohomology and invariant functions on constraint manifolds: I Commun. Math. Phys. 143 235–51

[17] García Pérez P L and Muñoz Masqué J 1983 On the geometrical structure of higher order variational calculus Proc. IUTAM–ISIMM Symp. on Modern Developments in Analytical Mechanics Vol 1 (Torino) vol 117 pp 127–47

[18] García Pérez P L and Muñoz Masqué J 1991 Higher order regular variational problems Symplectic Geometry and Mathematical Physics (Aix-en-Provence, 1990) (Png. Math. vol 99) (Boston, MA: Birkhäuser) pp 136–59

[19] Gasiuk J 1985 Formal integrability of systems of partial differential equations Nonlinear Equations in Classical and Quantum Field Theory (Meadow/Paris, 1983/1984) (Lecture Notes in Phys. vol 226) (Berlin: Springer) pp 21–36

[20] Gotay M J, Marsden J E, Isenberg J A, Montgomery R, Śniatycki J and Yasskin P B 1996 Covariant Field Theory (Momentum Maps and Classical Fields vol 1) arXiv:physics/9801019v2 [math-ph]
[21] Henneaux M and Teitelboim C 1992 Quantization of Gauge Systems (Princeton, NJ: Princeton University Press)
[22] Ibort A and Martín-Solano J 1994 Geometrical reduction and Parisi-Sourlas supersymmetry Phys. Lett. B 332 83–7
[23] Kropka D and Saunders D (eds) 2008 Handbook of Global Analysis (Amsterdam: Elsevier)
[24] de León M, Marrero J C and Martín de Diego D 2003 A new geometric setting for classical field theories Classical and Quantum Integrability (Warsaw, 2003) (Banach Center Publication vol 59) (Warsaw: Polish Academy of Science) pp 189–209
[25] de León M, Martín de Diego D and Santamaría-Merino A 2004 Symmetries in classical field theory Int. J. Geom. Methods Mod. Phys. 1 651–710
[26] de León M and Rodrigues P R 1985 Generalized Classical Mechanics and Field Theory (North-Holland Mathematics Studies vol 112) (Amsterdam: North-Holland)
[27] de León M and Rodrigues P R 1989 Methods of Differential Geometry in Analytical Mechanics (North-Holland Mathematics Studies vol 158) (Amsterdam: North-Holland)
[28] Marsden J E, Pekarsky S, Shkoller S and West M 2001 Variational methods, multisymplectic geometry and continuum mechanics J. Geom. Phys. 38 253–84
[29] Mendella G, Marno G and Tulczyjew W M 1995 Integrability of implicit differential equations J. Phys. A: Math. Gen. 28 149–65
[30] Olver P J 1986 Applications of Lie Groups to Differential Equations (Graduate Texts in Mathematics vol 107) (New York: Springer)
[31] Saunders D J 1989 The Geometry of Jet Bundles (London Mathematical Society Lecture Note Series vol 142) (Cambridge: Cambridge University Press)
[32] Saunders D J and Crampin M 1990 On the Legendre map in higher-order field theories J. Phys. A: Math. Gen. 23 3169–82
[33] Skinner R and Rusk R 1983 Generalized Hamiltonian dynamics: I. Formulation on $T^*Q \oplus TQ$ J. Math. Phys. 24 2589–94
[34] Skinner R and Rusk R 1983 Generalized Hamiltonian dynamics: II. Gauge transformations J. Math. Phys. 24 2595–601
[35] Vitagliano L 2010 The Lagrangian–Hamiltonian formalism for higher order field theories J. Geom. Phys. 60 857–73