Variational calculus with constraints on general algebroids

2008 J. Phys. A: Math. Theor. 41 175204

(http://iopscience.iop.org/1751-8121/41/17/175204)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **41** (2008) 175204 (25pp) doi:10.1088/1751-8113/41/17/175204

Variational calculus with constraints on general algebroids

Katarzyna Grabowska[1] **and Janusz Grabowski**[2,3]

[1] Physics Department, Division of Mathematical Methods in Physics, University of Warsaw, Hoża 69, 00-681 Warszawa, Poland
[2] Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, PO Box 21, 00-956 Warszawa, Poland

E-mail: konieczn@fuw.edu.pl and jagrab@impan.gov.pl

Received 8 January 2008, in final form 5 March 2008
Published 15 April 2008
Online at stacks.iop.org/JPhysA/41/175204

Abstract
Variational calculus on a vector bundle E equipped with a structure of a general algebroid is developed, together with the corresponding analogs of Euler–Lagrange equations. Constrained systems are introduced in the variational and geometrical settings. The constrained Euler–Lagrange equations are derived for analogs of holonomic, vakonomic and nonholonomic constraints. This general model covers the majority of first-order Lagrangian systems which are present in the literature and reduces to the standard variational calculus and the Euler–Lagrange equations in classical mechanics for $E = TM$.

PACS numbers: 45.20.Jj, 02.40.Yy, 02.40.Ma
Mathematics Subject Classification: 70H03, 70H25, 53D17, 17B66, 53D10

1. Introduction

The classical analytical mechanics is an old and well-established part of both mathematics and physics. Nevertheless, many people still look for the best mathematical tools in describing various aspects of mechanical systems. The use of Lie algebroids and Lie groupoids for describing some systems of the classical mechanics was proposed by Libermann [23] and Weinstein [45] more than ten years ago. This turned out to be a very fruitful idea and since then much work has been done (e.g. [4, 12, 18, 19, 24, 29, 32]) making use of Lie algebroids in various aspects of classical mechanics and classical field theory. The need of extending the geometrical tools of the Lagrangian formalism from just tangent bundles to Lie algebroids is justified by the fact that reductions usually move us out of the environment of the tangent bundles (think on the rigid body). It is similar to the better-known situation of passing from the symplectic to the Poisson structures in the Hamiltonian formalism.

[3] Research financed by the Polish Ministry of Science and Higher Education under grant no N201 005 31/0115.

In [12] it was observed that, following some ideas of Tulczyjew and using general algebroids instead of just Lie algebroids, one can describe a larger class of systems in a simple and elegant way, both in the Lagrangian and in the Hamiltonian formulation. Moreover, the proposed geometric picture does not require considering prolongations of Lie algebroids we start with, as was in the case of previous approaches known in the literature. A further paper [13] was devoted, in turn, to the construction of Euler–Lagrange equations in the affine setting of the so-called *special affgebroids* which is particularly suitable for time-dependent systems.

In this paper, we concentrate on variational calculus and constraints in the algebroid setting. We work with a general algebroid, defined in [17], as a double vector bundle morphism

$$\varepsilon : \mathsf{T}^*E \to \mathsf{T}E^* \tag{1.1}$$

covering the identity on E^*. Here, $\tau : E \to M$ is a vector bundle playing the role of kinematic configurations. To some extent then, our paper can be understood as a natural generalization of [33], where a variational calculus on Lie algebroids has been developed according to the original ideas of Weinstein [45], and [5, 18], where constraints on Lie algebroids have been considered. On the other hand, our approach is definitely different from the approaches known in the literature, even when the equations we obtain cover the corresponding Euler–Lagrange equations in the Lie algebroid case. This is mainly because we adapt the framework of the *Tulczyjew triple* [38, 39, 41], working simply with the morphism (1.1) rather than following the Klein's method [21] generalized to Lie algebroids, in which the bundles tangent to E and E^* are replaced by the prolongations of E with respect to the vector bundle projections $\tau: E \to M$ and $\pi: E^* \to M$. This, in our opinion, simplifies the whole formalism substantially.

To define a variational problem on an algebroid we have to specify a manifold $\mathcal{M}$ of paths whose tangent space $\mathsf{T}\mathcal{M}$ represents all possible variations and an action functional W on $\mathcal{M}$. Then we have to choose a submanifold $\mathcal{N}$ of *admissible paths* and a set (generalized distribution) $\mathcal{D} \subset \mathsf{T}\mathcal{M}_{|\mathcal{N}}$ of *admissible variations* of admissible paths. In [33] admissible variations are constructed out of homotopies of admissible paths as defined in [8]. For general algebroids we need different ways of constructing admissible variations, since we have to accept the fact that they are not tangent to the submanifold of admissible paths in general. Therefore, we construct admissible variations for an admissible path γ in E out of *vertical variations* of γ in E, i.e. out of vertical vector fields along γ. Note that the variations are defined in E (which is $\mathsf{T}M$ in the standard variational calculus), not in M. This is because the variational calculus on algebroids leads to first-order differential equations in E rather than to second-order equations in M. This is only the case of the canonical Lie algebroid $E = \mathsf{T}M$ when paths in M are in one-to-one correspondence with admissible paths in E, this time—just tangent prolongations of paths in M, and admissible variations are tangent prolongations of variations of paths in M. For a general algebroid, the admissible variations are constructed from the vertical ones by means of the double vector bundle relation $\kappa = \kappa_\varepsilon : \mathsf{T}E \longrightarrow\!\!\triangleright\, \mathsf{T}E$ which is dual to the morphism ε. Of course, for Lie algebroids our admissible paths coincide with the infinitesimal homotopies of admissible paths associated with the lifts of time-dependent sections, as they appear in [33, 8]. We prefer a more fundamental approach which uses κ_ε to produce admissible variations out of the vertical ones, instead of lifting whole sections extending paths in E and showing that the result does not depend on the extension. In the case of $E = \mathsf{T}M$, the mapping ε defining an algebroid structure is the inverse to the Tulczyjew isomorphism $\alpha_M : \mathsf{T}\mathsf{T}^*M \to \mathsf{T}^*\mathsf{T}M$. The relation κ_ε is in this case the well-known canonical flip $\kappa_M : \mathsf{T}\mathsf{T}M \to \mathsf{T}\mathsf{T}M$. Our construction is especially convenient in the case of nonholonomic constraints where variations are not tangent to the submanifold of constraints.

It is clear from our variational picture that applying constraints must result in defining a subset of $\mathcal{D}$. In the case of a general algebroid E, our classification of the constraints is based on the way in which the constrained admissible variations are constructed. According to the tradition, we call them vakonomic, nonholonomic and holonomic constraints. Starting from a subset S of E, classically understood as a geometric constraint for velocities, we have at least two natural possibilities of constructing a constraint in admissible variations: one is to consider only admissible variations which are tangent to S (*vakonomic constraint*), the other is to consider only admissible variations coming from those vertical ones which are tangent to S (*nonholonomic constraint*). Note that our approach allows us to understand nonholonomic constraint as a constrained variational problem, contrary to the commonly accepted conviction. A nonholonomic constraint is called *holonomic* if the constrained admissible variations are tangent to S (are vakonomic). Sometimes it is hard to decide without making an experiment which method should be used to describe the real behavior of the system.

For all types of constraints, we construct analogs of the Euler–Lagrange equation for systems that are subject to those three types of constraints in a variational way. Note, however, that the corresponding equations describe 'regular' solutions rather than a general solution of the variational problem. Additionally, like for non-constrained cases in [12], we derive the equations purely geometrically, without referring to the variational calculus.

The literature concerning constraints in variational calculus is so extensive that it is impossible to cite in a complete way. We decided to list among references only those papers dealing actually with Lie algebroids or being a direct inspiration for the framework we propose. Let us also make it clear that we see the meaning of the present paper not only as a generalization of formalisms of classical mechanics. Working with the case of a general algebroid forced us to propose a geometric approach which seems to be new and illustrative even when applied to very classical situations. The main observation is that an algebroid structure is a crucial geometric ingredient in constructing the dynamics of the system. It tells us not only the configurations, velocities and inner degrees of freedom, but also contains the information on how the admissible variations should be produced from a simple geometric model of variations of paths in a vector bundle—the vertical ones. This structure is encoded in a single map (1.1) respecting double vector bundle structures. The brackets and the Jacobi identity are therefore proven to play a minor role. The Jacobi identity for an algebroid bracket ensures some integrability conditions that allow us to integrate the Lie algebroid into an (at least local) Lie groupoid (see [8]), but which is irrelevant for the possibility of constructing Euler–Lagrange equations. Fixing this geometric setting for our system, it is then the Lagrangian function which produces a concrete dynamics out of these data. However, we would like to stress that regularity of the Lagrangian is completely irrelevant for our picture. The general method of constructing dynamics out of the Lagrangian works for all Lagrangians, singular or not. The difficulty with singular Lagrangians is that the dynamics we obtain is really implicit and complicated. In other words, the difficulty with singular Lagrangians lies in the difficulty in solving equations, not in the geometric construction of the equations themselves.

Finally, if the variational calculus is considered, only admissible paths come to the play. This is because we work on the bundle E of kinematic configurations and considering only admissible paths corresponds, classically, to work with paths in the manifold M of position configurations lifted canonically to the paths in $\mathsf{T}M$. The geometrical model of (infinitesimal) variations of an admissible path γ is to consider vertical vector fields along $\gamma : [t_0, t_1] \to E$. Now, the true (mechanical) admissible variations are vector fields along γ constructed from the vertical ones out of the algebroid structure κ. This is how the algebroid structure comes to the variational picture. Note that the role of the (Lie) algebroid structure in the classical setting is usually overlooked, since it is hidden behind structures of the tangent and cotangent bundles which are viewed as a natural part of the theory.

The paper is organized as follows. In section 2 we set up the notation and recall the notion of general algebroid as a double vector bundle morphism. Then we introduce the relation κ that is used for defining admissible variations. In section 3 we discuss the Lagrange formalism without constraints on general algebroid. Then we pass in section 4 to the variational calculus. We derive the variation of the Lagrangian and Euler–Lagrange equations. The final section is devoted to constraints. Geometric constraints as subsets $S \subset E$ give rise to variational constraints which are classified in pure geometrical terms such as vakonomic, nonholonomic or holonomic. We derive constrained equations using variational motivations and give them pure geometric interpretations.

2. Lie algebroids as double vector bundle morphisms

We start with introducing some notation.

Let M be a smooth manifold and let (x^a), $a = 1, \ldots, n$, be a coordinate system in M. We denote by $\tau_M\colon \mathsf{T}M \to M$ the tangent vector bundle and by $\pi_M\colon \mathsf{T}^*M \to M$ the cotangent vector bundle. We have the induced (adapted) coordinate systems $(x^a, \dot{x}^b)$ in $\mathsf{T}M$ and (x^a, p_b) in T^*M. Let $\tau\colon E \to M$ be a vector bundle and let $\pi\colon E^* \to M$ be the dual bundle. Let $(e_1, \ldots, e_m)$ be a basis of local sections of $\tau\colon E \to M$ and let $(e^1_*, \ldots, e^m_*)$ be the dual basis of local sections of $\pi\colon E^* \to M$. We have the induced coordinate systems:

$$
\begin{aligned}
&(x^a, y^i), && y^i = \iota(e^i_*), && \text{in } E,\\
&(x^a, \xi_i), && \xi_i = \iota(e_i), && \text{in } E^*,
\end{aligned}
$$

where the linear functions $\iota(e)$ are given by the canonical pairing $\iota(e)(v_x) = \langle e(x), v_x\rangle$. Thus, we have local coordinates

$$
\begin{aligned}
&(x^a, y^i, \dot{x}^b, \dot{y}^j) && \text{in } \mathsf{T}E,\\
&(x^a, \xi_i, \dot{x}^b, \dot{\xi}_j) && \text{in } \mathsf{T}E^*,\\
&(x^a, y^i, p_b, \pi_j) && \text{in } \mathsf{T}^*E,\\
&(x^a, \xi_i, p_b, \varphi^j) && \text{in } \mathsf{T}^*E^*.
\end{aligned}
$$

It is well known (cf [22, 42]) that the cotangent bundles T^*E and T^*E^* are examples of double vector bundles:

$$
\begin{tikzcd}
\mathsf{T}^*E^* \arrow[r, "\mathsf{T}^*\pi"] \arrow[d, "\pi_{E^*}"'] & E \arrow[d, "\tau"] \\
E^* \arrow[r, "\pi"] & M,
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathsf{T}^*E \arrow[r, "\mathsf{T}^*\tau"] \arrow[d, "\tau_{E^*}"'] & E^* \arrow[d, "\pi"] \\
E \arrow[r, "\tau"] & M.
\end{tikzcd}
$$

Note that the concept of a double vector bundle goes back to Pradines [35, 36], see also [3, 22]. In particular, all arrows correspond to vector bundle structures and all pairs of vertical and horizontal arrows are vector bundle morphisms. The double vector bundles have been recently characterized [15] in a simple way as two vector bundle structures whose Euler vector fields commute. The above double vector bundles are canonically isomorphic with the isomorphism

$$
\mathcal{R}_\tau\colon \mathsf{T}^*E \longrightarrow \mathsf{T}^*E^* \tag{2.1}
$$

being simultaneously an anti-symplectomorphism (cf [9, 17, 22]). In local coordinates, $\mathcal{R}_\tau$ is given by

$$
\mathcal{R}_\tau(x^a, y^i, p_b, \pi_j) = (x^a, \pi_i, -p_b, y^j).
$$

This means that we can identify the coordinates π_j with ξ_j, coordinates φ^j with y^j and use the coordinates (x^a, y^i, p_b, ξ_j) in T^*E and the coordinates (x^a, ξ_i, p_b, y^j) in T^*E^*, in full agreement with (2.1).

For the standard concept and theory of *Lie algebroids* we refer to the survey article [26] (see also [14, 27]). It is well known that Lie algebroid structures on a vector bundle E correspond to linear Poisson tensors on E^*. A 2-contravariant tensor Π on E^* is called *linear* if the corresponding mapping $\widetilde{\Pi}\colon \mathsf{T}^*E^* \to \mathsf{T}E^*$ induced by contraction, $\widetilde{\Pi}(\nu) = i_\nu \Pi$, is a morphism of double vector bundles. One can equivalently say that the corresponding bracket of functions is closed on (fiber-wise) linear functions. The commutative diagram

$$\begin{tikzcd}
\mathsf{T}^*E^* \arrow[r, "\widetilde{\Pi}"] & \mathsf{T}E^* \, , \\
\mathsf{T}^*E \arrow[u, "\mathcal{R}_\tau"'] \arrow[ur, "\varepsilon"']
\end{tikzcd}$$

describes a one-to-one correspondence between linear 2-contravariant tensors Π on E^* and morphisms ε (covering the identity on E^*) of the following double vector bundles (cf [17, 22]):

$$\begin{tikzcd}
 & \mathsf{T}^*E \arrow[rrr, "\varepsilon"] \arrow[dr, "\pi_E"] \arrow[ddl, "\mathsf{T}^*\tau"'] & & & \mathsf{T}E^* \arrow[dr, "\mathsf{T}\pi"] \arrow[ddl, "\tau_{E^*}"'] & \\
 & & E \arrow[rrr, "\rho"] \arrow[ddl, "\tau"'] & & & \mathsf{T}M \arrow[ddl, "\tau_M"] \\
E^* \arrow[rrr, "id"] \arrow[dr, "\pi"'] & & & E^* \arrow[dr, "\pi"'] & & \\
 & M \arrow[rrr, "id"] & & & M &
\end{tikzcd} \quad . \tag{2.2}$$

In local coordinates, every such ε is of the form

$$\varepsilon(x^a, y^i, p_b, \xi_j) = \left(x^a, \xi_i, \rho^b_k(x) y^k, c^k_{ij}(x) y^i \xi_k + \sigma^a_j(x) p_a\right) \tag{2.3}$$

(summation convention is used) and corresponds to the linear tensor

$$\Pi_\varepsilon = c^k_{ij}(x)\xi_k \partial_{\xi_i} \otimes \partial_{\xi_j} + \rho^b_i(x)\partial_{\xi_i} \otimes \partial_{x^b} - \sigma^a_j(x)\partial_{x^a} \otimes \partial_{\xi_j}.$$

The morphism (2.2) of double vector bundles covering the identity on E^* has been called an *algebroid* in [17], while a *Lie algebroid* has turned out to be an algebroid for which the tensor Π_ε is a Poisson tensor. We can consider the *adjoint tensor* Π^+_ε, i.e. the 2-contravariant tensor obtained from Π_ε by transposition:

$$\Pi^+_\varepsilon = c^k_{ji}(x)\xi_k \partial_{\xi_i} \otimes \partial_{\xi_j} + \rho^b_i(x)\partial_{x^b} \otimes \partial_{\xi_i} - \sigma^a_j(x)\partial_{\xi_j} \otimes \partial_{x^a}$$

and the *opposite tensor* $-\Pi_\varepsilon$. It is clear that Π^+_ε and $-\Pi_\varepsilon$ are linear. They correspond therefore to new algebroid structures: the *adjoint algebroid structure* ε^+ and the *opposite algebroid structure* $\bar{\varepsilon}$. An algebroid we call a quasi-Lie algebroid if $\varepsilon^+ = \bar{\varepsilon}$.

The relation to the canonical definition of Lie algebroid is given by the following theorem (cf [16, 17]).

Theorem 1. *An algebroid structure* (E, ε) *can be equivalently defined as a bilinear bracket* $[\cdot, \cdot]_\varepsilon$ *on the space* $\mathsf{Sec}(E)$ *of sections of* $\tau\colon E \to M$*, together with vector bundle morphisms* $\rho, \sigma\colon E \to \mathsf{T}M$ *(left anchor and right anchor), such that*

$$[fX, gY]_\varepsilon = f \cdot \rho(X)(g)Y - g \cdot \sigma(Y)(f)X + fg[X, Y]_\varepsilon$$

for $f, g \in \mathcal{C}^\infty(M)$, $X, Y \in \mathsf{Sec}(E)$. *The bracket and anchors are related to the bracket* $\{\varphi, \psi\}_{\Pi_\varepsilon} = \langle \Pi_\varepsilon, \mathrm{d}\varphi \otimes \mathrm{d}\psi\rangle$ *in the algebra of functions on* E^* *which is associated with the 2-contravariant tensor* Π_ε *by the formulae*

$$\begin{aligned}
\iota([X, Y]_\varepsilon) &= \{\iota(X), \iota(Y)\}_{\Pi_\varepsilon},\\
\pi^*(\rho(X)(f)) &= \{\iota(X), \pi^* f\}_{\Pi_\varepsilon},\\
\pi^*(\sigma(X)(f)) &= \{\pi^* f, \iota(X)\}_{\Pi_\varepsilon}.
\end{aligned}$$

The algebroid (E, ε) *is a quasi-Lie algebroid if and only if the tensor* Π_ε *is skew-symmetric, and is a Lie algebroid if and only if the tensor* Π_ε *is a Poisson tensor.*

Since the dual bundles of $\pi_E : \mathsf{T}^*E \to E$ and $\mathsf{T}\pi : \mathsf{T}E^* \to \mathsf{T}M$ are, respectively, $\tau_E : \mathsf{T}E \to E$ and $\mathsf{T}\tau : \mathsf{T}E \to \mathsf{T}M$, the dual to ε is a relation $\kappa = \kappa_\varepsilon : \mathsf{T}E \longrightarrow\!\!\!\!\triangleright\, \mathsf{T}E$. It is a uniquely defined smooth submanifold κ in $\mathsf{T}E \times \mathsf{T}E$ consisting of pairs (v, v') such that $\rho(\tau_E(v')) = \mathsf{T}\tau(v)$ and

$$\langle v, \varepsilon(v^*)\rangle_{\mathsf{T}\tau} = \langle v', v^*\rangle_{\tau_E}$$

for any $v^* \in \mathsf{T}^*_{\tau_E(v')}E$, where $\langle\cdot,\cdot\rangle_{\mathsf{T}\tau}$ is the canonical pairing between $\mathsf{T}E$ and $\mathsf{T}E^*$, and $\langle\cdot,\cdot\rangle_{\tau_E}$ is the canonical pairing between $\mathsf{T}E$ and T^*E. We will write $\kappa : v \longrightarrow\!\!\!\!\triangleright\, v'$ instead of $(v, v') \in \kappa$. This relation can be put into the following diagram of 'double vector bundle relations':

$$\begin{tikzcd}
 & \mathsf{T}E \arrow[dr, "\tau_E"] \arrow[ddl, "\mathsf{T}\tau"'] & & & \mathsf{T}E \arrow[lll, "\kappa"', -{Triangle[open]}] \arrow[dr, "\mathsf{T}\tau"] \arrow[ddl, "\tau_E"'] & \\
 & & E \arrow[rrr, "\rho"] \arrow[ddl, "\tau"'] & & & \mathsf{T}M \arrow[ddl, "\tau_M"] \\
\mathsf{T}M \arrow[dr, "\tau_M"'] & & & E \arrow[lll, "\sigma"'] \arrow[dr, "\tau"] & & \\
 & M \arrow[rrr, "id"] & & & M &
\end{tikzcd} \qquad (2.4)$$

The relation

$$\begin{tikzcd}
\mathsf{T}E \arrow[d, "\tau_E"] & \mathsf{T}E \arrow[l, "\kappa"', -{Triangle[open]}] \arrow[d, "\mathsf{T}\tau"] \\
E \arrow[r, "\rho"] & \mathsf{T}M
\end{tikzcd}$$

is a *vector bundle morphism of the second kind*, i.e. it is represented by linear maps of the fiber $\mathsf{T}E$ over $v \in \mathsf{T}M$ into the fibers T_eE for all $e \in E$ such that $\rho(e) = v$. This is also the simplest example of a morphism of Lie groupoids in the sense introduced and exploited by Zakrzewski [46]. To such relations we will refer therefore as to *Zakrzewski morphisms*. The expression of the Zakrzewski morphism (2.4), dual to ε, in local coordinates reads

$$\kappa : \left(x^a, Y^i, \rho^b_k(x) y^k, \dot{Y}^j\right) \longrightarrow\!\!\!\!\triangleright \left(x^a, y^i, \sigma^b_k(x) Y^k, \dot{Y}^j + c^j_{kl}(x) y^k Y^l\right). \qquad (2.5)$$

It is easy to see that the relation κ_ε^{-1} coincides with $\kappa_{\bar{\varepsilon}^+}$. Thus $\kappa = \kappa^{-1}$ for quasi-Lie algebroids.

A canonical example of a mapping ε in the case of $E = \mathsf{T}M$ is given by $\varepsilon = \varepsilon_M = \alpha_M^{-1}$—the inverse to the Tulczyjew isomorphism $\alpha_M : \mathsf{T}\mathsf{T}^*M \to \mathsf{T}^*\mathsf{T}M$ [38]. The dual Zakrzewski morphism is in this case the well-known 'canonical flip' $\kappa_M : \mathsf{T}\mathsf{T}M \to \mathsf{T}\mathsf{T}M$. Since α_M is an isomorphism, κ_M is a true map, in fact, an isomorphism of the corresponding two vector bundle structures as well.

A C^1-curve $\gamma : \mathbb{R} \to E$ (or a C^1-path $\gamma : [t_0, t_1] \to E$) in an algebroid E we call *admissible*, if the tangent prolongation $\mathsf{t}(\underline{\gamma})$ of its projection $\underline{\gamma} = \tau \circ \gamma$ coincides with its anchor:

$$\mathsf{t}(\underline{\gamma}) = \rho(\gamma(t)). \tag{2.6}$$

A curve (path) in the canonical Lie algebroid $\mathsf{T}M$ is admissible if and only if it is a tangent prolongation of its projection on M. If we denote $\mathsf{T}^{\mathrm{hol}}E$ the subset of $\mathsf{T}E$ consisting of *holonomic vectors*,

$$\mathsf{T}^{\mathrm{hol}}E = \{v \in \mathsf{T}E : \mathsf{T}\tau(v) = \rho(\tau_E(v))\}, \tag{2.7}$$

then admissible curves (paths) in the algebroid E can be characterized as those curves (paths) whose tangent prolongations lay in $\mathsf{T}^{\mathrm{hol}}E$. The set of holonomic vectors $\mathsf{T}^{\mathrm{hol}}E$ can be equivalently characterized as the subset in $\mathsf{T}E$ which is mapped via $\mathsf{T}\rho : \mathsf{T}E \to \mathsf{TT}M$ to classical holonomic vectors $\mathsf{T}^2M = \{u \in \mathsf{TT}M : \kappa_M(u) = u\}$, that justifies the name. In other words,

$$\mathsf{T}^{\mathrm{hol}}E = (\mathsf{T}\rho)^{-1}(\mathsf{T}^2M).$$

Note also that $\mathsf{T}^{\mathrm{hol}}E$ is canonically an affine bundle over E modeled on the vertical bundle $\mathsf{V}E \subset \mathsf{T}E$. In local coordinates, $\mathsf{T}^{\mathrm{hol}}E$ as submanifold in $\mathsf{T}E$ is characterized by the equations $\dot{x}^a = \rho^a_i(x)y^i$, so $(x^a, y^i, \dot{y}^j)$ can serve as local coordinates in $\mathsf{T}^{\mathrm{hol}}E$. It is easy to see that, for quasi-Lie algebroids, $\kappa(\mathsf{T}^{\mathrm{hol}}E) = \mathsf{T}^{\mathrm{hol}}E$.

Now let $\gamma : [t_0, t_1] \to E$ be a path and $\zeta : [t_0, t_1] \to \mathsf{V}E \subset \mathsf{T}E$ be a vertical vector field along γ, $\tau_E(\zeta(t)) = \gamma(t)$. It is well known that $\mathsf{V}E \simeq E \oplus_M E$, so vertical vectors at $e \in E$ can be canonically identified with vectors of the fiber $E_{\tau(e)}$. Thus, the vertical vector field ζ can be identified with a path ζ_E in E covering $\underline{\gamma}$. We can now consider the tangent prolongation $\mathsf{t}(\zeta_E)$ to get a vector field along ζ_E. The operation $\zeta \mapsto \mathsf{t}(\zeta_E)$ associates with any path ζ in $\mathsf{V}E$ a path $\mathsf{t}(\zeta_E)$ in $\mathsf{T}E$. For $v \in \mathsf{T}E$, in turn, the family $\kappa(v)$ defines a vector field over $\rho^{-1}(\mathsf{T}\tau(v))$. More precisely, for every $e \in \rho^{-1}(\mathsf{T}\tau(v))$ there is a unique vector $\kappa(v)_e \in \mathsf{T}_eE$ such that $\kappa(v)_e \in \kappa(v)$. We get the following.

Theorem 2. *If $\gamma : [t_0, t_1] \to E$ is an admissible path in E, then every vertical vector field $\zeta : [t_0, t_1] \to \mathsf{V}E$ along γ defines canonically a vector field $\delta_\zeta\gamma : [t_0, t_1] \to \mathsf{T}E$ along γ by*

$$\delta_\zeta\gamma(t) = \kappa(\mathsf{t}(\zeta_E)(t))_{\gamma(t)}. \tag{2.8}$$

In local coordinates, with $\gamma(t) = (x^a(t), y^i(t))$ and $\zeta(t) = (x^a(t), y^i(t), 0, f^i(t))$,

$$\delta_\zeta\gamma(t) = f^j(t)\sigma^b_j(x(t))\partial_{x^b}(\gamma(t)) + \left(\frac{\mathrm{d}f^k}{\mathrm{d}t}(t) + c^k_{ij}(x(t))y^i(t)f^j(t)\right)\partial_{y^k}(\gamma(t)). \tag{2.9}$$

In other words, in local coordinates in $\mathsf{T}E$,

$$\delta_\zeta\gamma(t) = \left(x^a(t), y^i(t), f^j(t)\sigma^b_j(x(t)), \frac{\mathrm{d}f^k}{\mathrm{d}t}(t) + c^k_{ij}(x(t))y^i(t)f^j(t)\right). \tag{2.10}$$

The vertical vector fields ζ along γ we will call *vertical variations* or *vertical virtual displacements* of γ and the vector fields $\delta_\zeta\gamma$ along γ—*admissible variations* or *admissible virtual displacements*. Note that the space $\mathsf{V}(\gamma)$ of vertical variations of γ is canonically an (infinite-dimensional) vector space.

Remark. In [8, 33], analogs of the admissible variations $\delta_\zeta\gamma$ have been obtained (in Lie algebroid context, of course) from tangent lifts of time-dependent sections of E. The tangent lifts of sections have natural generalizations for general algebroids [11, 12]. We prefer, however, to define the admissible variation $\delta_\zeta\gamma$ directly by means of the vertical variation ζ and the relation κ, as being more fundamental and conceptually closer to the standard concepts of variations.

3. Lagrangian formalism for general algebroids

The double vector bundle morphism (2.2) can serve as geometric background for generalized Lagrangian formalisms.

The Lagrangian $L : E \to \mathbb{R}$ defines two smooth maps: the *Legendre mapping*: $\lambda_L : E \longrightarrow E^*$, $\lambda_L = \tau_{E^*} \circ \varepsilon \circ \mathrm{d}L$, which is covered by the *Tulczyjew differential* $\Lambda_L : E \longrightarrow \mathsf{T}E^*$, $\Lambda_L = \varepsilon \circ \mathrm{d}L$:

$$\begin{tikzcd} \mathsf{T}^*E \arrow[r, "\varepsilon"] & \mathsf{T}E^* \arrow[d, "\tau_{E^*}"] \\ E \arrow[u, "\mathrm{d}L"] \arrow[ur, dotted, "\Lambda_L"] \arrow[r, "\lambda_L"] & E^* \end{tikzcd} \qquad (3.1)$$

The Lagrangian function L defines therefore *the phase dynamics* $\Gamma = \Lambda_L(E) \subset \mathsf{T}E^*$ which can be understood as an implicit differential equation on E^*, solutions of which are 'phase trajectories' of the system $\beta : \mathbb{R} \to E^*$ and satisfy $\mathrm{t}(\beta)(t) \in \Gamma$. An analog of the Euler–Lagrange equation for curves $\gamma : \mathbb{R} \to E$ is then

$$(E_L): \quad \mathrm{t}(\lambda_L \circ \gamma) = \Lambda_L \circ \gamma.$$

Equation (E_L) simply means that $\Lambda_L \circ \gamma$ is an admissible curve in $\mathsf{T}E^*$, thus it is the tangent prolongation of $\lambda_L \circ \gamma$. In local coordinates, Γ has the parametrization by (x^a, y^k) via Λ_L in the form (cf (2.3))

$$\Lambda_L(x^a, y^i) = \left(x^a, \frac{\partial L}{\partial y^i}(x, y), \rho^b_k(x) y^k, c^k_{ij}(x) y^i \frac{\partial L}{\partial y^k}(x, y) + \sigma^a_j(x) \frac{\partial L}{\partial x^a}(x, y)\right) \qquad (3.2)$$

and equation (E_L), for $\gamma(t) = (x^a(t), y^i(t))$, reads

$$(E_L): \quad \frac{\mathrm{d}x^a}{\mathrm{d}t} = \rho^a_k(x) y^k, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^j}\right) = c^k_{ij}(x) y^i \frac{\partial L}{\partial y^k} + \sigma^a_j(x) \frac{\partial L}{\partial x^a}, \qquad (3.3)$$

in full agreement with [24, 29, 30, 45], if only one takes into account that, for Lie algebroids, $\sigma^a_j = \rho^a_j$. As one can see from (3.3), the solutions are automatically admissible curves in E, i.e. $\rho(\gamma(t)) = \mathrm{t}(\tau \circ \gamma)(t)$. As a curve in the canonical Lie algebroid $\mathsf{T}M$ is admissible if and only if it is a tangent prolongation of its projection on M, first-order differential equations for admissible curves (paths) in $\mathsf{T}M$ may be viewed as certain second-order differential equations for curves (paths) in M. This explains why, classically, the Euler–Lagrange equations are regarded as second-order equations.

Remark. The Tulczyjew differential $\Lambda_L : \mathsf{T}M \to \mathsf{T}\mathsf{T}^*M$ with a given Lagrangian function L on the canonical Lie algebroid $E = \mathsf{T}M$ is sometimes called the *time-evolution operator* K (see [1]), as the first ideas of this operator go back to the work by Kamimura [20]. This operator has been studied by several authors in many variational contexts, however, without recognition of its direct relation to a (Lie) algebroid structure. We named this map after Tulczyjew, since our understanding is based on his ideas [39].

The time-dependent version of the above picture is the following. Consider the direct product $\widetilde{E} = E \times \mathsf{T}\mathbb{R}$ of the algebroid E with the canonical (Lie) algebroid $\mathsf{T}\mathbb{R}$ equipped with the canonical coordinates $(t, \dot{t})$. The corresponding algebroid morphism is clearly the product of ε and the inverse of the Tulczyjew isomorphism α_R:

$$\widetilde{\varepsilon} = \left(\varepsilon, \alpha_{\mathbb{R}}^{-1}\right) : \mathsf{T}^*\widetilde{E} = \mathsf{T}^*E \times \mathsf{T}^*\mathsf{T}\mathbb{R} \to \mathsf{T}E^* \times \mathsf{T}\mathsf{T}^*\mathbb{R} = \mathsf{T}\widetilde{E}^*. \qquad (3.4)$$

The affine hyperbundle $\mathcal{A}_{\mathbb{R}} = \{(t, 1) \in \mathsf{T}\mathbb{R}\}$ of $\mathsf{T}\mathbb{R}$ is a *Lie affgebroid* in the terminology of [10, 11, 13]. Similarly, the affine hyperbundle $\widetilde{E}_1 = E \times \mathcal{A}_{\mathbb{R}}$ in $\widetilde{E}$ is an *affgebroid* (so

$\boldsymbol{E} = \widetilde{E}_1 \times \mathbb{R}$ understood as the product in fibers is canonically a *special affgebroid* in the terminology of [13]). The morphism (3.4) can be reduced then to

$$(\varepsilon, \pi_{\mathcal{A}_\mathbb{R}}) : \mathsf{T}^*\widetilde{E}_1 = \mathsf{T}^*E \times \mathsf{T}^*\mathcal{A}_\mathbb{R} \to \mathsf{T}E^* \times \mathcal{A}_\mathbb{R} \subset \mathsf{T}(E^* \times \mathbb{R}).$$

Identifying $\mathcal{A}_\mathbb{R}$ with $\mathbb{R}$ in an obvious way, we obtain a *morphism of double affine bundles* [13]

$$\bar{\varepsilon} = (\varepsilon, \bar{\pi}_\mathbb{R}) : \mathsf{T}^*(E \times \mathbb{R}) = \mathsf{T}^*E \times \mathsf{T}^*\mathbb{R} \to \mathsf{T}E^* \times \mathsf{T}\mathbb{R} = \mathsf{T}(E^* \times \mathbb{R}), \qquad (3.5)$$

where $\bar{\pi}_\mathbb{R} : \mathsf{T}^*\mathbb{R} \to \mathsf{T}\mathbb{R}$ is defined by $\bar{\pi}_\mathbb{R}(t, s) = (t, 1) \in \mathsf{T}\mathbb{R}$.

Here, we view $\bar{E} = E \times \mathbb{R}$ canonically as a vector bundle $\bar{\tau} : \bar{E} = E \times \mathbb{R} \to M \times \mathbb{R}$ over $M \times \mathbb{R}$ (the pull-back bundle of E with respect to the projection $M \times \mathbb{R} \to M$) and $E^* \times \mathbb{R}$ as its dual $\bar{E}^*$. The time-dependent analog of diagram (3.1) defining the Tulczyjew differential, for the time-dependent Lagrangian $L : E \times \mathbb{R} \to \mathbb{R}$ reads

$$\begin{tikzcd}
\mathsf{T}^*(E \times \mathbb{R}) \arrow[rr, "\bar{\varepsilon}"] & & \mathsf{T}(E^* \times \mathbb{R}) \arrow[d, "\tau_{(E^* \times \mathbb{R})}"] \\
E \times \mathbb{R} \arrow[u, "\mathrm{d}L"] \arrow[urr, dotted, "\bar{\Lambda}_L"] \arrow[rr, "\bar{\lambda}_L"] & & E^* \times \mathbb{R}
\end{tikzcd} \qquad (3.6)$$

Although there is a canonical identification $\mathcal{A}_\mathbb{R} \simeq \mathbb{R}$, the use of $\mathcal{A}_\mathbb{R}$ explains the definition of holonomic vectors in this case: since $\mathsf{T}^{\mathrm{hol}}(E \times \mathcal{A}_\mathbb{R}) = \mathsf{T}^{\mathrm{hol}}E \times \mathcal{A}_\mathbb{R}$, we assume $\mathsf{T}^{\mathrm{hol}}(E \times \mathbb{R}) = \mathsf{T}^{\mathrm{hol}}E \times \mathbb{R}$. This is due to the fact that the time-dependent picture is, in fact, an affgebroid picture (see [13, 19, 34, 37, 43]).

In other words, $\bar{\Lambda}_L : E \times \mathbb{R} \to \mathsf{T}E^* \times \mathsf{T}\mathbb{R} \simeq \mathsf{T}(E^* \times \mathbb{R})$ and $\bar{\lambda}_L : E \times \mathbb{R} \to E^* \times \mathbb{R}$ read

$$\bar{\Lambda}_L(e, t) = (\Lambda_{L^t}(e), (t, 1)), \qquad \bar{\lambda}_L(e, t) = (\lambda_{L^t}(e), t), \qquad (3.7)$$

where we put $L^t(e) = L(e, t)$ and canonically identified $\mathsf{T}\mathbb{R}$ with $\mathbb{R} \times \mathbb{R}$. If now γ is a curve in E, then the nonautonomous Euler–Lagrange equation reads

$$\left(E_L^{na}\right) : \quad \mathsf{t}(\bar{\lambda}_L \circ \bar{\gamma}) = \bar{\Lambda}_L \circ \bar{\gamma}, \qquad (3.8)$$

where $\bar{\gamma}(t) = (\gamma(t), t)$ is a natural extension of γ to $E \times \mathbb{R}$. The nonautonomous Euler–Lagrange equation in coordinates takes formally the same form (3.3), but now with L depending on t.

Example 1. There are many examples based on Lie algebroids, see for instance [7, 18, 24, 29, 33].

(a) For instance, for the canonical Lie algebroid and the corresponding morphism—the inverse of the Tulczyjew isomorphism [38]

$$\varepsilon = \alpha_M^{-1} : \mathsf{T}^*\mathsf{T}M \to \mathsf{T}\mathsf{T}^*M,$$

with $y^a = \dot{x}^a$, we get the traditional Euler–Lagrange equations

$$\frac{\mathrm{d}x^a}{\mathrm{d}t} = \dot{x}^a, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^a}\right) = \frac{\partial L}{\partial x^a}.$$

(b) For a Lie algebroid which is just a Lie algebra with structure constants c^k_{ij} with respect to a chosen basis, we get the Euler–Poincaré equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^j}\right) = c^k_{ij} y^i \frac{\partial L}{\partial y^k}.$$

(c) True Lie algebroid examples are usually obtained as reductions of standard Lagrangian systems on tangent bundles, like the reduction of the rigid body to a dynamics on $so(3, \mathbb{R})$.

Another example of this kind is a homogeneous sphere of radius $r > 0$, mass m and inertia k^2 about any axis, moving on a horizontal table without friction (thus, is the table rotating or not makes no difference). In an obvious way, the system lives in fact on the Lie algebroid $\tau : \mathsf{T}\mathbb{R}^2 \times so(3,\mathbb{R}) \to \mathbb{R}^2$ with the product Lie algebroid structure. In standard coordinates, the algebroid morphism

$$\varepsilon : \mathsf{T}^*(\mathsf{T}\mathbb{R}^2 \times so(3,\mathbb{R})) \to \mathsf{T}(\mathsf{T}^*\mathbb{R}^2 \times so(3,\mathbb{R})^*)$$

reads

$$\begin{aligned}\varepsilon(x, y, \dot{x}, \dot{y}, \omega, p_x, p_y, p_{\dot{x}}, p_{\dot{y}}, p_\omega) = \big(&x, y, p_{\dot{x}}, p_{\dot{y}}, p_\omega, \dot{x}, \dot{y}, p_x, p_y, \omega_3 p_{\omega_2}\\ &- \omega_2 p_{\omega_3}, \omega_1 p_{\omega_3} - \omega_3 p_{\omega_1}, \omega_2 p_{\omega_1} - \omega_1 p_{\omega_2}\big).\end{aligned} \tag{3.9}$$

The pure kinetic Lagrangian

$$L = \tfrac{1}{2}m\big(\dot{x}^2 + \dot{y}^2 + k^2\big(\omega_1^2 + \omega_2^2 + \omega_3^2\big)\big)$$

induces the 'free' dynamics

$$\frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}) = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{y}) = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(mk^2\omega) = 0.$$

Later we will add nonholonomic constraints to this picture.

The above examples are associated with Lie algebroids, but some 'nonholonomic constraints' on Lie algebroids may lead to Lagrangian systems on quasi-Lie algebroids. This is related to quasi-Poisson brackets associated with nonholonomic constraints [28, 44].

Example 2 (Algebroid of linear constraints)**.** Consider an algebroid structure on a vector bundle E equipped with a Riemannian metric $\langle\cdot,\cdot\rangle_E$ and a vector subbundle C of E. Let $P : E \to C$ be the orthogonal projection. We can choose a local basis of orthonormal sections $(e_i) = (e_\alpha, e_A)$ of E such that (e_α) is a basis of local sections of C. According to the *d'Alembert principle* $\delta L(\mathrm{t}(\gamma)(t)) \in C^0$, where $C^0 \subset E^*$ is the annihilator of C, which in our case (cf (3.3)) takes the form

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^i}\right) - c^k_{\alpha i}(x)y^\alpha\frac{\partial L}{\partial y^k} - \sigma^a_i(x)\frac{\partial L}{\partial x^a}\right)e^*_i = \mu_A(x)e^*_A$$

for certain functions μ_A, the constrained dynamics is locally written as

$$y^A = 0, \qquad \frac{\mathrm{d}x^a}{\mathrm{d}t} = \rho^a_\alpha(x)y^\alpha, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^\beta}\right) - c^k_{\alpha\beta}(x)y^\alpha\frac{\partial L}{\partial y^k} - \sigma^a_\beta(x)\frac{\partial L}{\partial x^a} = 0. \tag{3.10}$$

If we deal with a Lagrangian of 'mechanical type'

$$L = \tfrac{1}{2}(y^i)^2 - V(x),$$

then $\frac{\partial L}{\partial y^A} = y^A = 0$ and equations (3.10) reduce to

$$y^A = 0, \qquad \frac{\mathrm{d}x^a}{\mathrm{d}t} = \rho^a_\alpha(x)y^\alpha, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^\beta}\right) - c^\gamma_{\alpha\beta}(x)y^\alpha\frac{\partial L}{\partial y^\gamma} - \sigma^a_\beta(x)\frac{\partial L}{\partial x^a} = 0$$

that can be viewed as the Euler–Lagrange equations of the algebroid associated with the orthogonal projection of the tensor Π_ε onto C^* according to the orthogonal decomposition $E^* = C^0 \oplus C^*$. Of course, even when E is a Lie algebroid, if C is not a Lie subalgebroid, the projected tensor is not a Poisson tensor and we deal with mechanics on a general algebroid, in fact a quasi-Lie algebroid in this case, since the projected Poisson tensor remains skew-symmetric.

4. Variational calculus

For a general algebroid structure ε on the vector bundle $\tau : E \to M$ and a smooth Lagrangian function $L : E \to \mathbb{R}$, we will define a version of a variational calculus as follows. Our (infinite-dimensional) manifold $\mathcal{M}$ will be the space of all C^1-paths $\gamma : [t_0, t_1] \to E$ in E. Of course, like in the standard variational calculus, by curves through the path $\gamma \in \mathcal{M}$ we mean C^1-maps

$$h : [t_0, t_1] \times \mathbb{R} \ni (t, s) \mapsto h(t, s) \in E$$

such that $h(t, 0) = \gamma(t)$. Thus, the tangent space $\mathsf{T}_\gamma \mathcal{M}$—the space of all possible variations of γ—is represented by $\frac{\partial h}{\partial s}(t, 0)$, i.e. by continuous paths $\delta\gamma : [t_0, t_1] \to \mathsf{T}E$ covering γ—vector fields along γ. The admissible paths form a subset $\mathcal{N}$ which is a submanifold in $\mathcal{M}$ in a natural sense, since a path γ is admissible if and only if $\mathsf{t}(\gamma) \subset \mathsf{T}^{\mathrm{hol}} E$. As easily seen (see also [33]), a vector field $\delta\gamma : [t_0, t_1] \to E$ along an admissible path γ belongs to $\mathsf{T}_\gamma \mathcal{N}$ if and only if $\kappa_E \circ \mathsf{t}(\delta\gamma)$ is tangent to $\mathsf{T}^{\mathrm{hol}} E$, i.e.

$$\kappa_E(\mathsf{t}(\delta\gamma)(t)) \in \mathsf{T}\mathsf{T}^{\mathrm{hol}} E \subset \mathsf{T}\mathsf{T}E, \tag{4.1}$$

where $\kappa_E : \mathsf{T}\mathsf{T}E \to \mathsf{T}\mathsf{T}E$ is the canonical flip.

Note that we use here 'infinite-dimensional manifold' structures in a very intuitive sense. However, we could have put rigorously a Banach manifold structure on $\mathcal{M}, \mathcal{N}$, etc, similarly as has been done in [33]. On the other hand, because the implicit function theorem will not be used, a less formal language is completely satisfactory for our purposes, so we will skip technical complications associated with the Banach manifold setting.

The Lagrangian L defines a differentiable function (action functional) $W_L : \mathcal{M} \to \mathbb{R}$ by

$$W_L(\gamma) = \int_{t_0}^{t_1} L(\gamma(t))\,\mathrm{d}t. \tag{4.2}$$

Completely classically, the differential of the action $\mathrm{d}W_L(\gamma)$, paired with the tangent vector $\delta\gamma$, gives

$$\langle \delta\gamma, \mathrm{d}W_L(\gamma)\rangle = \int_{t_0}^{t_1} \langle \delta\gamma(t), \mathrm{d}L(\gamma(t))\rangle\,\mathrm{d}t. \tag{4.3}$$

Now, we will make use of the algebroid structure on E and will reduce the differential $\mathrm{d}W_L$ to a distribution $\mathcal{D}$ over the submanifold $\mathcal{N}$ in $\mathcal{M}$ consisting of admissible paths. For an admissible path $\gamma : [t_0, t_1] \to E$, the space $\mathcal{D}(\gamma) \subset \mathsf{T}_\gamma \mathcal{M}$ of this distribution is exactly the space of admissible variations (virtual displacements) $\delta_\zeta \gamma$ as they were defined in (2.8), i.e.

$$\mathcal{D}(\gamma) = \{\delta_\zeta \gamma : \zeta \in \mathsf{V}(\gamma)\}. \tag{4.4}$$

In this sense, the space $\mathsf{V}(\gamma)$ of vertical variations, which is geometrically well-understood as the space of sections of the vertical bundle $\mathsf{V}E$ along γ, is a model space for the space $\mathcal{D}(\gamma)$ of admissible (mechanical) variations which does not have so nice geometrical description in general. The reader can easily check that in the case of the canonical Lie algebroid $E = \mathsf{T}M$ the admissible variations we have just introduced coincide with variations of tangent prolongations of paths in M (with not fixed end-points yet), as they are understood in classical mechanics. The geometrical meaning of these variations is usually not understood being hidden behind the 'obvious' Lie algebroid structure on $\mathsf{T}M$.

Let us consider now the differential $\mathrm{d}W_L$ being restricted to $\mathcal{D}$. Our aim is to show its special realization, very similar to the one present in the standard variational calculus of analytical mechanics. Of special interest are variations $\delta_\zeta \gamma$ coming from the set $\mathsf{V}(\gamma)_0$ of paths ζ that vanish at the end points, $\zeta(t_0) = 0, \zeta(t_1) = 0$. They form a submanifold $\mathcal{D}_0$ of

$\mathcal{D}$ and analogs of the standard Euler–Lagrange equations are obtained as equations for critical points of $(\mathrm{d}W_L)_{|\mathcal{D}_0}$, i.e. for such $\gamma \in \mathcal{N}$ that $\mathrm{d}W_L(\gamma)$ vanishes on $\mathcal{D}_0(\gamma)$. Note, however, that in contrast with what has been done in [33], being interested in the infinitesimal picture only, we do not care about global homotopies inside the manifold of admissible paths. In fact, our distribution is not tangent to $\mathcal{N}$ in general, so even 'infinitesimal homotopies' go outside $\mathcal{N}$ in the case of a general algebroid. This is due to the following observation.

Theorem 3. *The distribution $\mathcal{D}$ is tangent to the submanifold $\mathcal{N}$ of admissible paths if and only if the right and the left anchor coincide, $\rho = \sigma$, and they induce a homomorphism of brackets:*

$$\rho([X, Y]_\varepsilon) = [\rho(X), \rho(Y)]_{vf}, \tag{4.5}$$

where $[\cdot, \cdot]_{vf}$ is the bracket of vector fields. In particular, $\mathcal{D} \subset \mathsf{T}\mathcal{N}$ if (E, ε) is a Lie algebroid.

Proof. It is a matter of easy calculations to show that, according to (2.10), the vector field $\delta_\xi\gamma$ along $\gamma(t) = (x(t), y(t))$ satisfies (4.1) if and only if

$$\frac{\mathrm{d}f^j}{\mathrm{d}t}(t)\left(\sigma^b_j - \rho^b_j\right)(x(t)) + f^j(t)y^i(t)\left(\frac{\partial\sigma^b_j}{\partial x^a}\rho^a_i - \frac{\partial\rho^b_i}{\partial x^a}\sigma^a_j - c^k_{ij}\rho^b_k\right)(x(t)) = 0.$$

Since the above should be satisfied for any admissible γ and for any given $x(t) = x(t_0)$, we can take $f^j(t_0)$, $\frac{\mathrm{d}f^j}{\mathrm{d}t}(t_0)$ and $y(t_0)$ arbitrary. Hence we get $\rho = \sigma$ and

$$\frac{\partial\sigma^b_j}{\partial x^a}\rho^a_i - \frac{\partial\rho^b_i}{\partial x^a}\sigma^a_j - c^k_{ij}\rho^b_k = \frac{\partial\rho^b_j}{\partial x^a}\rho^a_i - \frac{\partial\rho^b_i}{\partial x^a}\rho^a_j - c^k_{ij}\rho^b_k = 0.$$

The latter can be rewritten in the form $\rho([e_i, e_j]_\varepsilon) = [\rho(e_i), \rho(e_j)]_{vf}$, whence (4.5). □

Remark. One develops often a variational calculus introducing homotopies as 'paths in path spaces' satisfying certain boundary conditions—this is exactly how the variational calculus on Lie algebroids has been developed in [33]. However, this approach is much more restrictive when passing to constraints. Let us only mention the existence of singular paths in the theory of linear nonholonomic constraints. In this case no real variation of a singular path is possible, so the differential calculus does not make sense any longer. On the other hand, the standard Euler–Lagrange equations are obtained as critical points of the action—so in fact only 'infinitesimal homotopies', i.e. admissible variations are used. In the Lie algebroid case, the admissible homotopies can be taken as integral curves of variations. Crainic and Fernandes have related homotopies of admissible paths to flows of the complete lifts of time-dependent sections of the Lie algebroid in their work [8] on integration of Lie algebroids. They did not mention the variational calculus, but this integration is actually finding a manifold $G(E)$ (Lie groupoid) that allows us to represent the variational calculus on the Lie algebroid E as a reduction of standard variational calculus on $\mathsf{T}G(E)$. Let us also point out that, contrary to the approaches by Crainic,Fernandes and Martínez [8, 33], we work in full generality and we do not assume at the beginning that admissible variations come from vertical variations vanishing at the end points.

Since calculating $\mathrm{d}W_L(\gamma)$ on $\mathcal{D}$ according to (4.3), we can divide our path into a finite number of smaller parts if needed, we can assume for simplicity that the path γ lies in a single coordinate chart (x^a, y^i), so we can write $\gamma(t) = (x^a(t), y^i(t))$. That our path is admissible means now that

$$\frac{\mathrm{d}x^a}{\mathrm{d}t}(t) = \rho^a_i(\underline{\gamma}(t))y^i(t). \tag{4.6}$$

For an admissible variation $\delta_\zeta\gamma$, with $\zeta(t) = f^i(t)e_i(\underline{\gamma}(t))$, we have then

$$\langle\delta_\zeta\gamma(t), \mathrm{d}L(\gamma(t))\rangle = \left[f^k(t)\cdot\sigma^a_k(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial x^a}(\gamma(t)) + \left(y^i(t)\cdot c^j_{ik}(\underline{\gamma}(t))\cdot f^k(t) + \frac{\mathrm{d}f^j}{\mathrm{d}t}(t)\right)\cdot\frac{\partial L}{\partial y^j}(\gamma(t))\right]$$
$$= f^k(t)\left(\sigma^a_j(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial x^a}(\gamma(t)) + y^i(t)\cdot c^k_{ij}(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial y^k}(\gamma(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^k}(\gamma(t))\right) + \frac{\mathrm{d}}{\mathrm{d}t}\left(f^j(t)\frac{\partial L}{\partial y^j}(\gamma(t))\right).$$

Writing $\lambda_L : E \to E^*, \lambda_L(x, y) = \frac{\partial L}{\partial y^j}(x, y)e^j_*$, for the vertical derivative (Legendre map) associated with L, and the *variation of the Lagrangian* along γ:

$$\delta L(\mathsf{t}(\gamma)(t)) = \left(\sigma^a_j(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial x^a}(\gamma(t)) + y^i(t)\cdot c^k_{ij}(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial y^k}(\gamma(t)) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^j}(\gamma(t))\right)e^j_*, \tag{4.7}$$

we can write

$$\langle\delta_\zeta\gamma(t), \mathrm{d}L(\gamma(t))\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\zeta_E(t), \lambda_L(\gamma(t))\rangle + \langle\zeta_E(t), \delta L(\mathsf{t}(\gamma)(t))\rangle. \tag{4.8}$$

According to (3.2), it is clear that $\delta L(\mathsf{t}(\gamma)(t)) = 0$ if and only if the image of the path $\mathrm{d}L(\gamma(t))$ under ε is admissible in $\mathsf{T}E^*$, i.e. if and only if γ satisfies the Euler–Lagrange equations (3.3).

In a more explicit form, the variation of the Lagrangian can be viewed as a map $\delta L : \mathsf{T}^{\mathrm{hol}}E \to E^*$ which in coordinates reads

$$\delta L(x, y, \dot{y}) = \left(\sigma^a_j(x)\frac{\partial L}{\partial x^a}(x, y) + y^i c^k_{ij}(x)\frac{\partial L}{\partial y^k}(x, y) - y^i\rho^a_i(x)\frac{\partial^2 L}{\partial x^a\partial y^j}(x, y) - \dot{y}^k\frac{\partial^2 L}{\partial y^k\partial y^j}(x, y)\right)e^j_*. \tag{4.9}$$

A geometrical description of the variation of the Lagrangian is as follows. If $v \in \mathsf{T}E$ is a holonomic vector, $v \in \mathsf{T}^{\mathrm{hol}}E$, then, as easily seen,

$$\Lambda_L \circ \tau_E, \mathsf{T}\lambda_L : \mathsf{T}E \to \mathsf{T}E^*$$

are bundle maps over $\lambda_L : E \to E^*$ and $\hat{\delta}L(v) = \Lambda_L(\tau_E(v)) - \mathsf{T}\lambda_L(v)$ is a vertical vector in $\mathsf{T}_{\tau_E(v)}E^*$. As the vertical bundle $\mathsf{V}E^* \subset \mathsf{T}E^*$ is canonically isomorphic to $E^* \oplus_M E^*$ by means of the vertical lift, we can identify $\hat{\delta}L(v)$ with a vector $\delta L(v) = (\hat{\delta}L(v))_{E^*}$ from the fiber of E^* over $\tau(\tau_E(v)) \in M$ which, expressed in coordinates, is exactly (4.9). In other words,

$$\delta L = ((\Lambda_L \circ \tau_E - \mathsf{T}\lambda_L)_{|\mathsf{T}^{\mathrm{hol}}E})_{E^*} = ((\varepsilon \circ \mathrm{d}L \circ \tau_E - \mathsf{T}(\tau_{E^*} \circ \varepsilon \circ \mathrm{d}L))_{|\mathsf{T}^{\mathrm{hol}}E})_{E^*}. \tag{4.10}$$

Using the obvious pairing between $\mathsf{V}E$ and $\mathsf{V}E^*$ based on the fact that the fibers over e and e^*, respectively, are canonically dual spaces if $\tau(e) = \pi(e^*)$, we can write (4.8) equivalently in the form

$$\langle\delta_\zeta\gamma(t), \mathrm{d}L(\gamma(t))\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\langle\zeta(t), \mathrm{d}L(\gamma(t))\rangle + \langle\zeta(t), \hat{\delta}L(\mathsf{t}(\gamma)(t))\rangle. \tag{4.11}$$

Integrating (4.8) (or (4.11)) we get

$$\begin{aligned}\langle\delta_\zeta\gamma, \mathrm{d}W_L(\gamma)\rangle &= \int_{t_0}^{t_1}\langle\delta_\zeta\gamma(t), \mathrm{d}L(\gamma(t))\rangle\,\mathrm{d}t\\ &= \langle\zeta_E(t), \lambda_L(\gamma(t))\rangle|_{t_0}^{t_1} + \int_{t_0}^{t_1}\langle\zeta_E(t), \delta L(\mathrm{t}(\gamma)(t))\rangle\,\mathrm{d}t\\ &= \zeta(L)(\gamma(t))|_{t_0}^{t_1} + \int_{t_0}^{t_1}\langle\zeta(t), \widehat{\delta}L(\mathrm{t}(\gamma)(t))\rangle\,\mathrm{d}t. \end{aligned}\tag{4.12}$$

Now, if $\zeta \in \mathsf{V}(\gamma)_0$, then $\langle\mathrm{d}W_L(\gamma), \delta_\zeta\gamma\rangle = \int_{t_0}^{t_1}\langle\zeta_E(t), \delta L(\mathrm{t}(\gamma)(t))\rangle\,\mathrm{d}t$. If $\zeta \in \mathsf{V}(\gamma)_0$, then $r(t)\zeta(t) \in \mathsf{V}(\gamma)_0$ for any function $r : [t_0, t_1] \to \mathbb{R}$, so $\mathrm{d}W_L(\gamma)$ vanishes on $\mathcal{D}_0(\gamma)$ if and only if $\delta L(\mathrm{t}(\gamma)) = 0$.

We can summarize the above observations as follows.

Theorem 4. *By means of the variational calculus for a general algebroid, one can define the velocities–momenta correspondence (Legendre map): $\lambda_L : E \to E^*$ and the variation of the Lagrangian $\delta L : \mathsf{T}^{\mathrm{hol}}E \to E^*$, such that the derivative of the action functional $\mathrm{d}W_L(\gamma)$ is represented by*

$$\langle\mathrm{d}W_L(\gamma), \delta_\zeta\gamma\rangle = \langle\zeta_E(t), \lambda_L(\gamma(t))\rangle|_{t_0}^{t_1} + \int_{t_0}^{t_1}\langle\zeta_E(t), \delta L(\mathrm{t}(\gamma)(t))\rangle\,\mathrm{d}t.$$

Moreover, formula (4.10) defines the Tulczyjew differential $\Lambda_L : E \to \mathsf{T}E^*$ associated with the Lagrangian L. An admissible path $\gamma(t) = (x(t), y(t))$ in E satisfies $\delta L(\mathrm{t}(\gamma)) = 0$ if and only if $\mathrm{d}W_L$ vanishes on $\mathcal{D}_0(\gamma)$ and if and only if γ satisfies the Euler–Lagrange equations (3.3).

For a given admissible path $\gamma : [t_0, t_1] \to E$, the values $p(t_0) = \lambda_L(\gamma(t_0))$ and $p(t_1) = \lambda_L(\gamma(t_1))$ represent the *initial* and the *final momenta*, and $\eta_\gamma(t) = \delta L(\mathrm{t}(\gamma)(t))$—the *external force* that we have to apply to make the system moving along the path γ. A standard way to obtain the dynamics in analytical mechanics is to look for critical points of the action functional with respect to admissible variations $\delta_\zeta\gamma$ which vanish at the end points. In this way, we obtain the Euler–Lagrange equations (3.3) for admissible curves in the form $\delta L(\mathrm{t}(\gamma)(t)) = 0$. In a more general setting, one can view the force-defining equation

$$\delta L(\mathrm{t}(\gamma)(t)) = \eta_\gamma(t) \tag{4.13}$$

as a differential equation for γ if the external force $\eta_\gamma(t)$ is given. In many cases, this force is defined in a path-independent way as a time-dependent field of forces $F : E \times \mathbb{R} \to E^*$, $\pi(F(e, t)) = \tau(e)$, i.e. $\eta_\gamma(t) = F(\gamma(t), t)$.

There is no real difference when we admit time-dependent Lagrangians $L : E \times \mathbb{R} \to \mathbb{R}$, so that the action reads

$$W_L(\gamma) = \int_{t_0}^{t_1} L(\gamma(t), t)\,\mathrm{d}t. \tag{4.14}$$

Formula (4.12) just takes the form

$$\langle\mathrm{d}W_L(\gamma), \delta_\zeta\gamma\rangle = \left(f^j(t)\cdot\frac{\partial L}{\partial y^j}(\gamma(t), t)\right)\bigg|_{t_0}^{t_1} + \int_{t_0}^{t_1}\langle\zeta_E(t), \delta L(\mathrm{t}(\gamma)(t), t)\rangle\,\mathrm{d}t, \tag{4.15}$$

where

$$\begin{aligned}\delta L(\mathrm{t}(\gamma)(t), t) = \bigg(&\sigma_j^a(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial x^a}(\gamma(t), t)\\ &+ y^i(t)\cdot c_{ij}^k(\underline{\gamma}(t))\cdot\frac{\partial L}{\partial y^k}(\gamma(t), t) - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^j}(\gamma(t), t)\bigg)e_*^j. \end{aligned}\tag{4.16}$$

In local coordinates,

$$\delta L(x,y,t,\dot y)=\bigg(\sigma^a_j(x)\frac{\partial L}{\partial x^a}(x,y,t)+y^i c^k_{ij}(x)\frac{\partial L}{\partial y^k}(x,y,t)$$
$$-y^i\rho^a_i(x)\frac{\partial^2 L}{\partial x^a\partial y^j}(x,y,t)-\dot y^k\frac{\partial^2 L}{\partial y^k\partial y^j}(x,y)-\frac{\partial^2 L}{\partial t\partial y^j}(x,y,t)\bigg)e^j_*. \tag{4.17}$$

The geometrical picture is based on (3.6). Now, $\delta L:\mathsf{T}^{\mathrm{hol}}E\times\mathbb{R}\to E^*$ is defined as the map whose vertical lift is

$$\hat\delta L=\mathsf{v}_\pi\circ\delta L=(\bar\Lambda_L\circ\tau_{\bar E}-\mathsf{T}\bar\lambda_L)_{|\mathsf{T}^{\mathrm{hol}}E\times\mathbb{R}}=(\bar\varepsilon\circ\mathrm{d}L\circ\tau_{\bar E}-\mathsf{T}(\tau_{\bar E^*}\circ\bar\varepsilon\circ\mathrm{d}L))_{|\mathsf{T}^{\mathrm{hol}}E\times\mathbb{R}} \tag{4.18}$$

and the standard Euler–Lagrange equation for the time-dependent Lagrangian in the presence of external forces takes the form

$$\delta L(\mathsf{t}(\gamma)(t),t)=\eta_\gamma(t).$$

Again, the equation $\delta L(\mathsf{t}(\gamma)(t),t)=0$ means that the image of the path $\mathrm{d}L(\gamma(t),t)$ in $\mathsf{T}^*(E\times\mathbb{R})$ under $\bar\varepsilon$ is an admissible path (tangent prolongation) in $\mathsf{T}(E^*\times\mathbb{R})$.

5. Constraints

In view of the just developed variational calculus on general algebroids we can introduce, in principle, two types of constraints: the *configuration constraints* which are put in the 'bundle of velocities' E, i.e. constrains for paths in $\mathcal{N}$, and the *virtual displacement constraints* put for variations, i.e. for fibers of the admissible distribution $\mathcal{D}$. As the admissible variations are also related to paths in E, the latter constraints can be defined also via constraints in E that often leads to misunderstandings. In this way, a constrained submanifold in E (classically in $E=\mathsf{T}M$) is sometimes referred to as a *nonholonomic constraint*. Note, however, that in general, speaking on a submanifold (in general—subset) of E as of a constraint does not make much sense before we decide how the constrained submanifold produces true constraints in the variational calculus. To put some order in the subject, we will start with describing our understanding of constraints in the variational calculus for general algebroid that will motivate a description of constraints in the pure geometric setting.

Definition 1. *A constraint in the variational calculus for a general algebroid is a subset $\mathcal{C}$ of the bundle $\mathcal{D}$. The corresponding (dynamical) configuration constraint is the subset $\mathcal{C}_\mathcal{N}$ obtained from $\mathcal{C}$ by the projection to $\mathcal{N}$. The constrained variational calculus is the study of the differential of the action functional* $\mathrm{d}W_L$ *restricted to $\mathcal{C}$, or $\mathcal{C}_0=\mathcal{C}\bigcap\mathcal{D}_0$.*

It seems that the true variational constraints in physics strongly depend on the actual system we work with. In theory, however, the variational constraints are often derived from geometric constraints of different types in a more or less canonical way. A *geometric constraint* will be understood as a submanifold (more generally, a subset) S in E. Of course, as we have already mentioned (see also [40]), the submanifold (subset) $S\subset E$ does not define a true variational constraint without additional specifications. There are at least two geometrically justified ways of deriving variational constraints out of S. According to the tradition (see the review article [6]), we will refer to them, respectively, as to *vakonomic* and *nonholonomic* constraints. In the vakonomic case we accept only admissible variations (virtual displacements) which are tangent to the constraint, while in the nonholonomic case we admit only vertical variations which are tangent to the constraint, i.e. which belong to $\mathsf{V}(S)=\mathsf{T}S\bigcap\mathsf{V}E$.

Definition 2.

- *The vakonomic constraint associated with $S \subset E$ is the variational constraint $\mathcal{C}^{vk}(S)$ consisting of these admissible variations $\delta_\zeta\gamma$ which are tangent to S, i.e. $\delta_\zeta\gamma(t) \in \mathsf{T}S$. In particular, the admissible path γ lies in S.*
- *The nonholonomic constraint associated with $S \subset E$ is the variational constraint $\mathcal{C}^{nh}(S)$ consisting of admissible variations $\delta_\zeta\gamma$ associated with vertical variations ζ which are tangent to S. In other words, $\zeta(t) \in \mathsf{T}S$ (thus, $\zeta(t) \in \mathsf{V}(S)$). In particular, the admissible path γ lies in S.*
- *A geometric constraint $S \subset E$ we call holonomic, if the nonholonomic variational constraint associated with S implies the vakonomic constraint, i.e. $\mathcal{C}^{nh}(S) \subset \mathcal{C}^{vk}(S)$.*

Note that the variational constraints associated with S can be very small or even empty, e.g. when there are no admissible paths in S. To avoid pathologies like that, certain additional integrability conditions can be introduced. A natural integrability condition we will use is $\rho(S) \subset \mathsf{T}S_M$, where $S_M = \tau(S)$. It is assumed in the sequel that the geometric constraints are integrable.

Remark. We should stress here the obvious fact that $\mathsf{T}S$ is well defined in a general setting even if S is not a submanifold of E, since it makes sense to speak about smooth curves in E with values in S. Note that, just by definition, for the vakonomic constraint only the restriction of the Lagrangian function L to S plays the role in the variational problem. The latter is not the case for nonholonomic constraints, except for the holonomic case. We can say that holonomic constraints are those nonholonomic constraints for which only the restrictions of the Lagrange functions to S play the role in the corresponding variational problems. One can easily derive from the form of lift (2.10) that a linear (integrable) constraint S, i.e. a vector subbundle $S \subset E$, is holonomic if and only if the algebroid bracket $[\cdot,\cdot]_\varepsilon$ is closed on sections of S, i.e. S a subalgebroid in E.

5.1. Vakonomic constraints—variational approach

The variational problem now depends on studying the differential of the action functional on $\mathcal{C}^{vk}(S)$. A naive but instructive approach is that the corresponding constrained Euler–Lagrange equations describe admissible paths γ in S which are *critical points of W_L relative to the generalized distribution* $\mathcal{C}_0^{vk}(S) = \mathcal{C}^{vk}(S) \bigcap \mathcal{D}_0$, i.e. such that $\mathrm{d}W_L(\gamma)$ vanishes on all $\delta_\zeta\gamma \in \mathcal{C}_0^{vk}(S)(\gamma)$:

$$\langle \mathrm{d}W_L(\gamma), \delta_\zeta\gamma\rangle = \int_{t_0}^{t_1} \langle \mathrm{d}L(\gamma(t)), \delta_\zeta\gamma(t)\rangle\,\mathrm{d}t = \int_{t_0}^{t_1} \langle \zeta_E(t), \delta L(\mathrm{t}(\gamma)(t))\rangle\,\mathrm{d}t = 0 \tag{5.1}$$

for all vertical vector fields ζ along γ, with $\zeta(t_0) = 0$, $\zeta(t_1) = 0$, and such that $\delta_\zeta\gamma$ is tangent to S. Of course, it is hard to decide how large is $\mathcal{C}_0^{vk}(S)$. This is related to the difficult questions of the existence of singular or abnormal paths, etc, which cannot be solved in the whole generality. Leaving these questions aside, we will reduce ourselves to natural and geometric sufficient conditions ensuring that a given admissible path satisfies (5.1). Namely, let us observe that if Φ is a function vanishing on S, then, as $\delta_\zeta\gamma$ is tangent to S,

$$\langle \mathrm{d}\Phi(\gamma(t)), \delta_\zeta\gamma(t)\rangle = 0.$$

Thus, if

$$\delta(L - \mu_k\Phi^k)(\mathrm{t}(\gamma)(t), t) = 0, \qquad \gamma(t) \in S, \tag{5.2}$$

for certain $\mu_k : [t_0, t_1] \to \mathbb{R}$ and for certain functions Φ^k vanishing on S (e.g. defining S according to the implicit function theorem), then, according to (5.1) applied to $L := L - \mu_k \Phi^k$,

$$\begin{aligned}\langle \mathrm{d}W_L(\gamma), \delta_\zeta \gamma\rangle &= \int_{t_0}^{t_1} \langle \mathrm{d}L(\gamma(t)), \delta_\zeta\gamma(t)\rangle\, \mathrm{d}t \\ &= \int_{t_0}^{t_1} \langle (\mathrm{d}L - \mu_k(t)\, \mathrm{d}\Phi^k)(\gamma(t)), \delta_\zeta\gamma(t)\rangle\, \mathrm{d}t \\ &= \int_{t_0}^{t_1} \langle \zeta_E(t), \delta(L - \mu_k\Phi^k)(\mathrm{t}(\gamma)(t), t)\rangle\, \mathrm{d}t = 0,\end{aligned}$$

so (5.1) is satisfied. Such γ we will call a *normal solution* of the vakonomic variational problem associated with $S \subset E$. In the above procedure we can take as well a time-dependent Lagrangian L satisfying (5.2). The latter does not depend directly on how big is $\mathcal{C}_0^{vk}(S)$ and it simply means that the image of the path $\mathrm{d}\left(L - \mu_k\Phi^k\right)(\gamma(t), t)$ in $\mathsf{T}^*(E \times \mathbb{R})$ under $\bar{\varepsilon}$ is an admissible path (tangent prolongation) in $\mathsf{T}(E^* \times \mathbb{R})$. Motivated by the tradition we will regard equation (5.2) as the *vakonomically constrained Euler–Lagrange equation*. There is a clear analog of the above procedure also for time-dependent constraints. The nonautonomous vakonomic Euler–Lagrange equation takes in coordinates the form

$$\Phi^k(x, y) = 0, \qquad \frac{\mathrm{d}x^a}{\mathrm{d}t} = \rho^a_k(x) y^k, \tag{5.3}$$

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^j}(x, y, t) - c^l_{ij}(x) y^i \frac{\partial L}{\partial y^l}(x, y, t) - \sigma^a_j(x)\frac{\partial L}{\partial x^a}(x, y, t) \\ &\qquad = \dot{\mu}_k(t)\frac{\partial \Phi^k}{\partial y^j}(x, y) + \mu_k(t)\left(\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \Phi^k}{\partial y^j}(x, y)\right. \\ &\qquad\quad \left. - c^l_{ij}(x) y^i \frac{\partial \Phi^k}{\partial y^l}(x, y) - \sigma^a_j(x)\frac{\partial \Phi^k}{\partial x^a}(x, y)\right)\end{aligned} \tag{5.4}$$

and reduces to the classical one for the canonical Lie algebroid $E = \mathsf{T}M$ (see, e.g. [6]). In the above form, however, the vakonomic Euler–Lagrange equation is not easily seen to depend only on the restriction of L to S. Below we present a geometric approach clarifying this question. On the other hand, this approach seems to be more transparent not only at this point.

5.2. Vakonomic constraints—geometric approach

Of course, one can take (5.3) as the Euler–Lagrange equation for the vakonomic constraints without referring to the variational calculus. It has the advantage that we do not care about possibly complicate constrained admissible variations. There is a nice geometric interpretation of these equations. For simplicity we will reduce to the autonomous case, so that L does not depend on t. Let us first recall that with any submanifold S in E and any function $L : S \to \mathbb{R}$ one can associate canonically a Lagrangian submanifold S_L in T^*E defined by

$$S_L = \{\alpha_e \in \mathsf{T}^*_e E : e \in S \text{ and } \langle \alpha_e, v_e\rangle = \mathrm{d}L(v_e) \text{ for every } v_e \in \mathsf{T}_e S\}.$$

If $S = E$, then $S_L = \mathrm{d}L(E)$, i.e. S_L reduces to the image of $\mathrm{d}L$. We have the following.

Theorem 5. *A curve $\gamma : \mathbb{R} \to E$ satisfies the vakonomic Euler–Lagrange equations (5.2) associated with an autonomous Lagrangian $L : E \to \mathbb{R}$ if and only if it is a projection to S of a curve $\gamma^* : \mathbb{R} \to S_L$ whose image under $\varepsilon : \mathsf{T}^*E \to \mathsf{T}E^*$ is admissible (is a tangent* *prolongation of a curve in E^*). In particular, the vakonomic Euler–Lagrange equations depend on the restriction of the Lagrangian to the constraint only.*

Proof. If a curve γ satisfies (5.2), then γ is admissible and lies in S. Moreover, the curve $\gamma^*(t) = \mathrm{d}L(\gamma(t)) - \mu_k(t)\,\mathrm{d}\Phi^k(\gamma(t))$ in T^*E lies in S_L, projects on γ and is mapped through ε to an admissible curve.

Conversely, if a curve $\gamma^*(t)$ has the above properties, then there are $\mu_k(t)$ such that $\gamma^*(t) = \mathrm{d}L(\gamma(t)) - \mu_k(t)\,\mathrm{d}\Phi^k(\gamma(t))$. Since admissibility of $\varepsilon(\gamma^*(t))$ is equivalent to $\delta(\mathrm{d}L - \mu_k(t)\,\mathrm{d}\Phi^k)(\gamma(t)) = 0$ (theorem 4), the theorem follows. □

One can also think that the vakonomic Euler–Lagrange equations are not equations on curves in E but on curves in S_L. Then, we can just consider the projections of the solutions onto E.

The corresponding diagram is the following

$$\begin{tikzcd}
\mathsf{T}^*E \supset S_L \arrow[r, "\varepsilon"] & \mathsf{T}E^* \arrow[d, "\tau_{E^*}"] \\
S \arrow[u, "r_L"] \arrow[ur, dotted, "\Lambda_L"] \arrow[r, "\lambda_L"'] & E^*
\end{tikzcd} \tag{5.5}$$

where r_L is the relation which is the inverse of the projection $(\pi_E)_{|S_L} : S_L \to S$ and $\Lambda_L = \varepsilon \circ r_L$. Like in the non-constrained case, a curve γ in S satisfies the vakonomic E–L equation if it is related via Λ_L to an admissible curve in $\mathsf{T}E^*$.

From the above it should be clear that the phase space for the vakonomic constraint S is $\tau_{E^*}(\varepsilon(\pi^{-1}(S)))$ and the phase dynamic associated with the Lagrangian L is $\varepsilon(S_L)$. There is an obvious version of the above picture in the nonautonomous case.

Example 3 (Pontryagin's maximum principle)**.** For an algebroid (E, ε) over M consider the product algebroid $E_U = E \times \mathsf{T}U$. Considering an optimal control problem in which the manifold U plays the role of the set of control parameters and associated with

(1) an integrable constraint S defined by means of a U-dependent section $f : M \times U \to E$ of E by $(e, v) \in S \Leftrightarrow e = f(\tau(e), \tau_M(v))$ and

(2) a Lagrangian function $L : S \to \mathbb{R}$ depending only on the base, $L(e, v) = L(\tau(e), \tau_M(v))$.

In local coordinates $(x^a, u^\alpha, y^i, \dot{u}^\beta)$ in E_U and the adapted coordinates $(x, u, y, \dot{u}, p_x, p_u, \xi, \pi)$ in T^*E_U, the product algebroid morphism $\varepsilon_U = (\varepsilon, \varepsilon_M)$ reads

$$\varepsilon_U\,(x, u, y, \dot{u}, p_x, p_u, \xi, \pi) = \left(x, u, \xi, \pi, \rho^b_k(x) y^k, \dot{u}, c^k_{ij}(x) y^i \xi_k + \sigma^a_j(x) p_{x^a}, p_u\right).$$

The Lagrangian submanifold $S^*_L \subset \mathsf{T}^*E_U$ consists of points

$$\left(x, u, f(x,u), \dot{u}, \left(\frac{\partial L}{\partial x} - \xi \cdot \frac{\partial f}{\partial x}\right)(x,u), \left(\frac{\partial L}{\partial u} - \xi \cdot \frac{\partial f}{\partial u}\right)(x,u), \xi, 0\right),$$

so the phase (implicit) dynamics is given by $\varepsilon_U(S^*_L)$, which is the set of points

$$\Big(x, u, \xi, 0, \rho^b_k(x) f^k(x,u), \dot{u}, c^k_{ij}(x) f^i(x,u)\xi_k$$
$$+\sigma^a_j(x)\left(\frac{\partial L}{\partial x^a} - \xi \cdot \frac{\partial f}{\partial x^a}\right)(x,u), \left(\frac{\partial L}{\partial u} - \xi \cdot \frac{\partial f}{\partial u}\right)(x,u)\Big),$$

and the vakonomic Euler–Lagrange equations read

$$\frac{\mathrm{d}x^b}{\mathrm{d}t} = \rho^b_k(x) f^k(x,u), \tag{5.6}$$

$$\frac{\mathrm{d}\xi_j}{\mathrm{d}t} = c^k_{ij}(x) f^i(x,u)\xi_k + \sigma^a_j(x)\left(\frac{\partial L}{\partial x^a} - \xi_i \frac{\partial f^i}{\partial x^a}\right)(x,u), \tag{5.7}$$

$$\left(\frac{\partial L}{\partial u} - \xi_i \frac{\partial f^i}{\partial u}\right)(x,u) = 0. \tag{5.8}$$

Equations (5.6) and (5.7) describe the phase dynamics on $E^* \times U$ associated with the Hamiltonian $H(x,u,\xi) = f^i(x,u)\xi_i - L(x,u)$ via the tensor Π_ε—trivially extended from E^* to $E^* \times U$. Equation (5.8), in turn, is the equation for critical points of this Hamiltonian with respect to the control variable u. In the classical case $E = \mathsf{T}M$, equations (5.7) and (5.8) read

$$\frac{\mathrm{d}p_a}{\mathrm{d}t} = \left(\frac{\partial L}{\partial x^a} - p_b \frac{\partial f^b}{\partial x^a}\right)(x,u), \qquad \left(\frac{\partial L}{\partial u^\alpha} - p_b \frac{\partial f^b}{\partial u^\alpha}\right)(x,u) = 0.$$

We recognize the Pontryagin's maximum principle in its normal differential form. For Lie algebroids, this principle was first proposed in [31].

5.3. Nonholonomic constraints—variational approach

A naive but instructive approach is to assume in this case that the constrained Euler–Lagrange equations describe admissible paths γ in S which are *critical points of W_L relative to the generalized distribution* $\mathcal{C}^{nh}_0(S)$, i.e. such that (5.1) is satisfied for all $\delta_\zeta\gamma \in \mathcal{C}^{nh}_0(S)$. Again, we will not discuss the problem how large is $\mathcal{C}^{nh}_0(S)$. Recall that $\delta_\zeta\gamma \in \mathcal{C}^{nh}(S)$ means that $\zeta(t) \in \mathsf{V}(S)$, where $\mathsf{V}(S) = \mathsf{T}S \bigcap \mathsf{V}E$ is the vertical part of $\mathsf{T}S$. If S is a submanifold and $\mathsf{V}(S)$ has a constant rank, then the annihilator $(\mathsf{V}(S))^0 \subset \mathsf{T}^*E_{|S}$ is a vector subbundle (over S) in T^*E. In this case, the quotient bundle $\mathsf{T}^*E_{|S}/(\mathsf{V}(S))^0$ is canonically isomorphic to the bundle $\mathsf{V}^*(S)$—dual to $\mathsf{V}(S)$. The latter, viewed as a subbundle in $pr_1 : E \oplus_M E \to E$ in an obvious way, is called the *bundle of virtual displacements* in [5, section 8]. Of course, $(\mathsf{V}(S))^0$ can be viewed in a similar way as a subbundle in $E \oplus_M E^* \to E$. In this interpretation, which we will generally use in the sequel, $(\mathsf{V}(S))^0_e \subset E^*_{\tau(e)}$, $e \in S$, is the annihilator of $(\mathsf{V}(S))_e \subset E_{\tau(e)}$ and $(\mathsf{V}^*(S))_e = E^*_{\tau(e)}/(\mathsf{V}(S))^0_e$.

It is obvious that (5.1) is satisfied for all $\delta_\zeta\gamma \in \mathcal{C}^{nh}_0(S)$ if (and not only if, in general)

$$\delta L(\mathrm{t}(\gamma)(t)) \in (\mathsf{V}(S))^0_{\gamma(t)}. \tag{5.9}$$

This equation we will view as the *constrained nonholonomic Euler–Lagrange equation*. Again, it is not exactly equivalent to the variational principle in general as it gives only a sufficient condition for a relative critical point of the action functional W_L. In local coordinates, if Φ^k are functions defining the constraint S via equations $\Phi^k(x,y) = 0$, then $\mathsf{V}(S)^0$ is generated by $\frac{\partial\Phi^k}{\partial y^i}$ at points of S, so the constrained nonholonomic Euler–Lagrange equation reads

$$\Phi^k(x,y) = 0, \qquad \frac{\mathrm{d}x^a}{\mathrm{d}t} = \rho^a_i(x) y^i, \tag{5.10}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^j}(x,y) - c^l_{ij}(x) y^i \frac{\partial L}{\partial y^l}(x,y) - \sigma^a_j(x)\frac{\partial L}{\partial x^a}(x,y) = \mu_k(t)\frac{\partial \Phi^k}{\partial y^j}(x,y). \tag{5.11}$$

For $E = \mathsf{T}M$ this is exactly the *Chetaev principle*, and for E being an arbitrary Lie algebroid equations (5.10) and (5.11) coincide with the equations associated with nonlinear nonholonomic constraints considered in [5, 25].

Example 4 (rolling ball)**.** Consider now the celebrated example of a ball rolling on a rotating table (cf [2, 5]), more precisely, of a homogeneous sphere of radius $r > 0$, mass m and inertia about any axis k^2, moving without sliding on a horizontal table which rotates with a constant angular velocity Ω. Like in example 1, we can recognize that the system lives on the Lie algebroid $\tau : \mathsf{T}\mathbb{R}^2 \times so(3,\mathbb{R}) \to \mathbb{R}^2$ with the product Lie algebroid structure and is ruled by the pure kinetic Lagrangian

$$L = \tfrac{1}{2}m\left(\dot{x}^2 + \dot{y}^2 + k^2\left(\omega_1^2 + \omega_2^2 + \omega_3^2\right)\right),$$

this time however with the presence of nonholonomic constraints

$$\Phi^1(x, y, \dot{x}, \dot{y}, \omega) = \dot{x} - r\omega_2 + \Omega y = 0,$$
$$\Phi^2(x, y, \dot{x}, \dot{y}, \omega) = \dot{y} + r\omega_1 - \Omega y = 0.$$

According to (5.10) and (5.11), we get the constrained nonholonomic Euler–Lagrange equation in the form

$$\dot{x} - r\omega_2 + \Omega y = 0, \qquad \dot{y} + r\omega_1 - \Omega y = 0, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}) = \mu_1, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{y}) = \mu_2,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}(mk^2\omega_1) = r\mu_2, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(mk^2\omega_2) = -r\mu_1, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(mk^2\omega_3) = 0$$

that easily implies

$$\ddot{x} + \frac{k^2\Omega}{r^2 + k^2}\dot{y} = 0, \qquad \ddot{y} - \frac{k^2\Omega}{r^2 + k^2}\dot{x} = 0.$$

If S is a linear constraint, i.e. S is a vector subbundle in E, then $(\mathsf{V}(S))_e$ can be identified with $S_{\tau(e)}$ and $(\mathsf{V}(S))_e^0$ with $S^0_{\tau(e)} \subset E^*_{\tau(e)}$. In this case, the constrained nonholonomic Euler–Lagrange equation (5.9) takes the form

$$\delta L(\mathsf{t}(\gamma)(t)) \in S^0_{\underline{\gamma}(t)}, \tag{5.12}$$

which is exactly the *d'Alembert's principle* of virtual work. The d'Alembert's principle for Lie algebroids was first proposed in [7].

More generally, assume that $S = A$ is an affine constraint, i.e. A is an affine subbundle in E. Then, $\mathsf{V}(A)_e$ can be canonically identified with the fiber $\mathsf{v}(A)_{\tau(e)} \subset E_{\tau(e)}$ of a vector bundle $\mathsf{v}(A)$ which serves as a model vector bundle of A. Hence, $(\mathsf{V}(S))_e^0$ can be identified with $\mathsf{v}(A)^0_{\tau(e)} \subset E^0_{\tau(e)}$ and the constrained nonholonomic Euler–Lagrange equation reads

$$\delta L(\mathsf{t}(\gamma)(t)) \in (\mathsf{v}(A))^0_{\underline{\gamma}(t)}. \tag{5.13}$$

5.4. *Affine nonholonomic constraints—geometric approach*

Let us assume that $S = A$ is an affine subbundle in E (over A_M) satisfying the integrability condition $\rho(A) \subset \mathsf{T}A_M$. In this case $(\mathsf{V}(A))_e$ is constant along fibers of A and coincides with $\mathsf{v}(A)_{\tau(e)}$. Let $\mathsf{v}(A)^0$ be the annihilator of $\mathsf{v}(A)$ which is a subbundle in E^* (over A_M). Let $i_{\mathsf{v}(A)} : \mathsf{v}(A) \hookrightarrow E$ be the inclusion of $\mathsf{v}(A)$ in E, let $i^*_{\mathsf{v}(A)} : E^*_{|A_M} \to \mathsf{v}(A)^*$ be the dual of $i_{\mathsf{v}(A)}$, and let

$$\mathsf{T}i^*_{\mathsf{v}(A)} : \mathsf{T}(E^*_{|A_M}) \to \mathsf{T}\mathsf{v}(A)^*$$

be its tangent prolongation. According to the integrability condition $\rho(A) \subset \mathsf{T}A_M$, the image $\varepsilon(\mathsf{T}^*E_{|A})$ lies in $\mathsf{T}(E^*_{|A_M})$ and the corresponding diagram is the following:

$$
\begin{tikzcd}
\mathsf{T}^*E_{|A} \arrow[r, "\varepsilon"] & \mathsf{T}(E^*_{|A_M}) \arrow[r, "\mathsf{T}i^*_{\mathsf{v}(A)}"] & \mathsf{T}\mathsf{v}(A)^* \arrow[d, "\tau_{\mathsf{v}(A)^*}"] \\
E \supset A \arrow[u, "\mathrm{d}L"] \arrow[r, "\lambda_L"] \arrow[urr, dotted, "\Lambda_L^A"] \arrow[rr, dotted, bend right, "\lambda_L^A"'] & E^*_{|A_M} \arrow[r, "i^*_{\mathsf{v}(A)}"] & \mathsf{v}(A)^*
\end{tikzcd}
\tag{5.14}
$$

The space $\mathsf{v}(A)^*$ is the phase space for the nonholonomic constraint A with $\lambda_L^A : A \to \mathsf{v}(A)^*$, $\lambda_L^A = i^*_{\mathsf{v}(A)} \circ \lambda_L$, as the *constrained Legendre map*, and $\Lambda_L^A : A \to \mathsf{T}\mathsf{v}(A)^*$, with $\Lambda_L^A = \mathsf{T}i^*_{\mathsf{v}(A)} \circ \varepsilon \circ \mathrm{d}L$, as the *constrained Tulczyjew differential*. The set $\Lambda_L^A(A) \subset \mathsf{T}\mathsf{v}(A)^*$ is the phase dynamics associated with the Lagrangian L. The nonholonomic Euler–Lagrange equation is described as follows.

Theorem 6. *A curve $\gamma : \mathbb{R} \to A$ satisfies the nonholonomic Euler–Lagrange equation $\delta L(\mathsf{t}(\gamma)(t)) \in \mathsf{v}(A)^0_{\underline{\gamma}(t)}$ if and only if the curve $\Lambda_L^A(\gamma(t))$ in $\mathsf{T}\mathsf{v}(A)^*$ is admissible (is the tangent prolongation of a curve in $\mathsf{v}(A)^*$).*

Proof. Consider local coordinates $(x^I) = (x^i, x^\iota)$ on an open set U of M such that A_M is determined by the constraint $x^\iota = 0$. A local basis $\{e_a\}_{a=1,\dots,n-r}$ of sections of $\mathsf{v}(A)$ together with a section e_0 of A we can extend to local sections of E and complete them to a local basis of sections $\{e_0, e_a, e_\alpha\}$ of the vector bundle E. Then, in coordinates $(x^I, y^A) = (x^i, x^\iota, y^0, y^a, y^\alpha)$ adapted to this bases, the local equations defining the constrained subbundle A as an affine subbundle of E over A_M are $x^\iota = 0, y^0 = 1, y^\alpha = 0$, so points of A have coordinates $(x^i, 0, 1, y^a, 0)$. Note that integrability of the constraint A means that $\rho_0^\iota(x) = 0$ and $\rho_a^\iota(x) = 0$ at points $x \in A_M$.

Taking local coordinates (x^i, y^a) on $\mathsf{v}(A)$ we may write $i_{\mathsf{v}(A)} : \mathsf{v}(A) \hookrightarrow E$ as $i_{\mathsf{v}(A)}(x^i, y^a) = (x^i, 0, 0, y^a, 0)$ and $i^*_{\mathsf{v}(A)}(x^i, 0, \xi_0, \xi_a, \xi_\alpha) = (x^i, \xi_a)$, so

$$\mathsf{T}i^*_{\mathsf{v}(A)}(x^i, 0, \xi_A, \dot{x}^j, 0, y^B) = (x^i, \xi_a, \dot{x}^j, \dot{\xi}_a).$$

For the adapted local coordinates $(x^i, x^\iota, y^0, y^a, y^\alpha, p_i, p_\iota, \xi_0, \xi_a, \xi_\alpha)$ in T^*E, the map ε reduced to $(\mathsf{T}^*E)_{|A}$ takes values in $\mathsf{T}(E_{|A_M})$ (integrability) and reads

$$\begin{aligned}
\varepsilon(x^i, 0, 1, y^a, 0, p_I, \xi_A) = \big(&x^i, 0, \xi_A, \rho_e^j(x^i, 0)y^e \\
&+ \rho_0^j(x^i, 0), 0, \big(c_{eB}^D(x^i, 0)y^e + c_{0B}^D(x^i, 0)\big)\xi_D + \sigma_B^I(x^i, 0)p_I\big).
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathsf{T}i^*_{\mathsf{v}(A)} \circ \varepsilon(x^i, 0, 1, y^a, 0, p_I, \xi_A) = \big(&x^i, \xi_a, \rho_e^j(x^i, 0)y^e \\
&+ \rho_0^j(x^i, 0), \big(c_{eb}^D(x^i, 0)y^e + c_{0b}^D(x^i, 0)\big)\xi_D + \sigma_b^I(x^i, 0)p_I\big)
\end{aligned} \tag{5.15}$$

and

$$\begin{aligned}
\Lambda_L^A(x^i, y^a) &= Ti^*_{\mathsf{v}(A)} \circ \epsilon\left(x^i, 0, 1, y^a, 0, \frac{\partial L}{\partial x^I}(x^j, 0, 1, y^a, 0), \frac{\partial L}{\partial y^A}(x^j, 0, 1, y^a, 0)\right) \\
&= \left(x^i, \frac{\partial L}{\partial y^b}(x^j, 0, 1, y^a, 0), \rho_e^j(x^i, 0)y^e + \rho^0(x^i, 0), \big(c_{eb}^D(x^i, 0)y^e + c_{0b}^D(x^i, 0)\big)\right. \\
&\quad \left. \times \frac{\partial L}{\partial y^D}(x^j, 0, 1, y^a, 0) + \sigma_b^I(x)\frac{\partial L}{\partial x^I}(x^j, 0, 1, y^a, 0)\right).
\end{aligned}$$

Therefore, locally, the nonholonomic Euler–Lagrange equations read

$$x^\iota = 0, \qquad y^0 = 1, \qquad y^\alpha = 0, \qquad \frac{\mathrm{d}x^j}{\mathrm{d}t} = \rho^j_e(x^i, 0)y^e + \rho^j_0(x^i, 0), \tag{5.16}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial y^b}(x^i, 0, 1, y^a, 0) = \left(c^D_{eb}(x^i, 0)y^e + c^D_{0b}(x^i, 0)\right)\frac{\partial L}{\partial y^D}(x^i, 0, 1, y^a, 0) \\ + \sigma^I_b(x^i, 0)\frac{\partial L}{\partial x^I}(x^i, 0, 1, y^a, 0). \tag{5.17}$$

On the other hand, (5.17) means that $\delta L_b = 0$ for all b, i.e. $\delta L \in \mathsf{v}(A)^0$. □

In the case of a Lie algebroid and linear constraints $A = \mathsf{v}(A)$ covering the whole M, when we have in particular $A_M = M$, $\sigma^i_e = \rho^i_e$, the previous equations are precisely the nonholonomic equations obtained in [5] (see equations (3.8)).

Again, there is an obvious version of the above picture for a time-dependent Lagrangian based on (3.6). In the nonholonomic case, however, we cannot restrict the Lagrangian function to the constraint, except for the case which is, in fact, *holonomic*.

5.5. Holonomic constraints

In the nonholonomic case, we can restrict the Lagrangian L to the constraint S if the geometric constraint is holonomic. Note, however, that this does not imply automatically that the corresponding vakonomic and nonholonomic Euler–Lagrange equations are the same, since the equations are not precisely variational (they describe only sufficient conditions that the variational principle holds true) and they are obtained in different ways. On the other hand, in the linear case, holonomicity means that the vector subbundle S is closed with respect to the algebroid bracket. Since the constraints are assumed to be integrable, for the canonical Lie algebroid $E = \mathsf{T}M$ this means, in turn, that $S = \mathsf{T}S_M$, so the constraints are holonomic in the classical sense. More generally, assume that A is an affine constraint, i.e. A is an affine subbundle in E.

Theorem 7. *An affine constraint A in a quasi-Lie algebroid E is holonomic if and only if the algebroid bracket of sections of A is a section of* $\mathsf{v}(A)$.

Proof. Let us choose a basis of sections e_i and the corresponding linear coordinates (x^a, y^i) in E such that A is locally defined by equations $y^i = 0, i > r + 1$, and $y^{r+1} = 1$ and let $\gamma(t) = (x(t), y(t))$ be an admissible path in A. Then, $y^i(t) = 0$ for $i > r + 1$ and $y^{r+1} = 1$. Moreover, ζ is a vertical variation of γ, $\zeta(x(t), y(t)) = f^i(t)\partial_{y^i}$, if and only if $f^i = 0$ for $i > r$. In view of (2.10), $\delta_\zeta\gamma$ is tangent to A only if

$$\frac{\mathrm{d}f^k}{\mathrm{d}t}(t) + c^k_{ij}(x(t))y^i(t)f^j(t) = 0$$

for $k > r$. But for any $k > r$

$$\frac{\mathrm{d}f^k}{\mathrm{d}t}(t) + \sum_{i,j} c^k_{ij}(x(t))y^i(t)f^j(t) = \sum_{j\leqslant r} c^k_{(r+1)j}(x(t))f^j(t) + \sum_{i,j\leqslant r} c^k_{ij}(x(t))y^i(t)f^j(t).$$

As $y^i(t), f^j(t)$ for $i, j \leqslant r$ are arbitrary, $c^k_{ij} = 0$ for $k > r$ and $i \leqslant r + 1, j \leqslant r$. Since $c^k_{ij} = -c^k_{ji}$, they vanish also for $k > r$ and for all $i, j \leqslant r + 1$. This means that the bracket of local sections $[e_i, e_j]_\varepsilon, i, j \leqslant r + 1$ belongs to the span of $\{e_1, \ldots, e_r\}$, i.e. is a section of

$\mathsf{v}(A)$. But sections of A are of the form $e_{r+1} + \sum_{i\leqslant r} \varphi_i(x)e_i$, so their brackets are sections of $\mathsf{v}(A)$. The converse is obvious. $\square$

According to the terminology of [10–12], one can say that affine holonomic constraints in a Lie algebroid are *Lie affgebroids*. A correct geometric description of time-dependent systems and other systems, based on the idea of Lie affgebroid was first proposed in [34, 37] and developed in [10–12].

5.6. Affine holonomic constraints—geometric approach

If A is a holonomic affine constraint, then, using local coordinates as above, we can prove analogously to the proof of theorem 7 that $c^D_{eb}(x^i, 0)$ and $c^D_{0b}(x^i, 0)$ can be nonzero only for D indexing a section of $\mathsf{v}(A)$, symbolically $D = d$, and that $\sigma^\iota_b(x^i, 0) = 0$. Now, using the local form (5.15) of $\mathsf{T}i^*_{\mathsf{v}(A)} \circ \varepsilon$, we conclude that $\mathsf{T}i^*_{\mathsf{v}(A)} \circ \varepsilon$ vanishes on the annihilator of $\mathsf{T}A$. Hence, $\mathsf{T}i^*_{\mathsf{v}(A)} \circ \varepsilon$ defines a map $\varepsilon^A : \mathsf{T}^*A \to \mathsf{T}\mathsf{v}(A)^*$ and diagram (5.14) reduces to the following:

$$\begin{tikzcd} \mathsf{T}^*A \arrow[rr, "\varepsilon^A"] & & \mathsf{T}\mathsf{v}(A)^* \arrow[d, "\tau_{\mathsf{v}(A)^*}"] \\ A \arrow[u, "\mathrm{d}L"] \arrow[urr, dotted, "\Lambda^A_L"] \arrow[rr, "\lambda^A_L"'] & & \mathsf{v}(A)^* \end{tikzcd} \quad . \tag{5.18}$$

This time, however, only the restriction of L to A does matter. The phase space is $\mathsf{v}(A)^*$, the phase dynamics is implicitly defined as $\Lambda^A_L(A) = \varepsilon^A \circ \mathrm{d}L(A) \subset \mathsf{T}\mathsf{v}(A)^*$, and the Euler–Lagrange equation for a curve γ in A reads

$$\Lambda^A_L \circ \gamma = \mathsf{t}\big(\lambda^A_L \circ \gamma\big).$$

In local coordinates,

$$x^\iota = 0, \qquad y^0 = 1, \qquad y^\alpha = 0, \qquad \frac{\mathrm{d}x^j}{\mathrm{d}t} = \rho^j_e(x^i, 0)y^e + \rho^j_0(x^i, 0), \tag{5.19}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial y^b}(x^i, y^a)\right) = \big(c^d_{eb}(x^i, 0)y^e + c^d_{0b}(x^i, 0)\big)\frac{\partial L}{\partial y^d}(x^i, y^a) + \sigma^j_b(x^i, 0)\frac{\partial L}{\partial x^j}(x^i, y^a). \tag{5.20}$$

The above equations (canonically reduced to A) are exactly the Euler–Lagrange equations for a (Lie) affgebroid obtained in [13, 19]. One can say that geometrical mechanics on a (Lie) affgebroid is just geometrical mechanics on (Lie) algebroid with a holonomic affine constraint.

References

[1] Batlle C, Gomis J, Pons J M and Román-Roy N 1986 Equivalence between the Lagrangian and Hamiltonian formalism for constrained systems *J. Math. Phys.* **27** 2953–62

[2] Bloch A M, Krishnaprasad P S, Marsden J E and Murray R M 1996 Nonholonomic mechanical systems with symmetry *Arch. Ration. Mech. Anal.* **136** 21–99

[3] Brown R and Mackenzie K C 1992 Determination of a double groupoid by its core diagram *J. Pure Appl. Algebra* **80** 237–72

[4] Cortés J, de Léon M, Marrero J C, Martín de Diego D and Martínez E 2006 A survey of Lagrangian mechanics and control on Lie algebroids and groupoids *Int. J. Geom. Methods Mod. Phys.* **3** 509–58

[5] Cortés J, de León M, Marrero J C and Martínez E 2005 Nonholonomic Lagrangian systems on Lie algebroids *Preprint* math-ph/0512003

[6] Cortés J, de Léon M, Martín de Diego D and Martínez E 2002 Geometric description of vakonomic and nonholonomic dynamics: comparison of solutions *SIAM J. Control Optim.* **41** 1389–412 (electronic)

[7] Cortés J and Martínez E 2004 Mechanical control systems on Lie algebroids *IMA J. Math. Control Inform.* **21** 457–92

[8] Crainic M and Fernandes R L 2003 Integrability of Lie brackets *Ann. Math.* **157** 575–620

[9] Dufour J-P 1991 Introduction aux tissus *Séminaire GETODIM* 55–76 (*Preprint*)

[10] Grabowska K, Grabowski J and Urbański P 2003 Lie brackets on affine bundles *Ann. Global Anal. Geom.* **24** 101–30

[11] Grabowska K, Grabowski J and Urbański P 2004 AV-differential geometry: Poisson and Jacobi structures *J. Geom. Phys.* **52** 398–446

[12] Grabowska K, Grabowski J and Urbański P 2006 Geometrical mechanics on algebroids *Int. J. Geom. Methods Mod. Phys.* **3** 559–75

[13] Grabowska K, Grabowski J and Urbański P 2007 AV-differential geometry: Euler–Lagrange equations *J. Geom. Phys.* **57** 1984–98

[14] Grabowski J 2003 Quasi-derivations and QD-algebroids *Rep. Math. Phys.* **32** 445–51

[15] Grabowski J and Rotkiewicz M 2007 Higher vector bundles and multi-graded symplectic manifolds *Preprint* math.DG/0702772

[16] Grabowski J and Urbański P 1997 Lie algebroids and Poisson–Nijenhuis structures *Rep. Math. Phys.* **40** 195–208

[17] Grabowski J and Urbański P 1999 Algebroids—general differential calculi on vector bundles *J. Geom. Phys.* **31** 111–41

[18] Iglesias D, Marrero J C, Martín de Diego D and Sosa D 2007 Singular Lagrangian systems and variational constrained mechanics on Lie algebroids *Preprint* 0706.2789v1

[19] Iglesias D, Marrero J C, Padrón E and Sosa D 2006 Lagrangian submanifolds and dynamics on Lie affgebroids *Rep. Math. Phys.* **57** 385–436

[20] Kamimura K 1982 Singular Lagrangians and constrained Hamiltonian systems, generalized canonical formalism *Nuovo Cimento* B **69** 33–54

[21] Klein J 1962 Espaces variationelles et mécanique *Ann. Inst. Fourier* **12** 1–124

[22] Konieczna K and Urbański P 1999 Double vector bundles and duality *Arch. Math. (Brno)* **35** 59–95

[23] Libermann P 1996 Lie algebroids and mechanics *Arch. Math.* **32** 147–62

[24] de León M, Marrero J C and Martínez E 2005 Lagrangian submanifolds and dynamics on Lie algebroids *J. Phys. A: Math. Gen.* **38** R241–308

[25] de León M, Marrero J C and Martín de Diego D 1997 Mechanical systems with nonlinear constraints *Int. J. Theor. Phys.* **36** 979–95

[26] Mackenzie K C H 1995 Lie algebroids and Lie pseudoalgebras *Bull. Lond. Math. Soc.* **27** 97–147

[27] Mackenzie K C H 2005 *General Theory of Lie Groupoids and Lie Algebroids* (Cambridge: Cambridge University Press)

[28] Marle C-M 1998 Various approaches to conservative and nonconservative nonholonomic systems *Rep. Math. Phys.* **42** 211–29

[29] Martínez E 2001 Lagrangian mechanics on Lie algebroids *Acta Appl. Math.* **67** 295–320

[30] Martínez E 2001 Geometric formulation of mechanics on Lie algebroids *Proc. VIII Fall Workshop on Geometry and Physics (Medina del Campo, 1999) Publicaciones de la RSME* vol 2 pp 209–22

[31] Martínez E 2004 Reduction in optimal control theory *Rep. Math. Phys.* **53** 79–90

[32] Martínez E 2005 Classical field theory on Lie algebroids: variational aspects *J. Phys. A: Math. Gen.* **38** 7145–60

[33] Martínez E 2006 Variational calculus on Lie algebroids *Preprint* math-ph/0603028v2

[34] Martí nez E, Mestdag T and Sarlet W 2002 Lie algebroid structures and Lagrangian systems on affine bundles *J. Geom. Phys.* **44** 70–95

[35] Pradines J 1974 Fibrés vectoriels doubles et calcul des jets non holonomes *Notes polycopiées, Amiens* (in French)

[36] Pradines J 1974 Représentation des jets non holonomes par des morphismes vectoriels doubles soudés *C. R. Acad. Sci. Paris* A **278** 1523–6 (in French)

[37] Sarlet W, Mestdag T and Martínez E 2002 Lie algebroid structures on a class of affine bundles *J. Math. Phys.* **43** 5654–74

[38] Tulczyjew W M 1974 Hamiltonian systems, Lagrangian systems, and the Legendre transformation *Symp. Math.* **14** 101–14

[39] Tulczyjew W M 1976 Les sous-variétés Lagrangiennes et la dynamique Lagrangienne *C. R. Acad. Sci. Paris* A–B **283** A675–8 (in French)

[40] Tulczyjew W M 2003 A note on holonomic constraint *Boston Stud. Philos. Sci.* **234** 403–19

[41] Tulczyjew W M and Urbański P 1999 A slow and careful Legendre transformation for singular Lagrangians *(The Infeld Centennial Meeting (Warsaw, 1998)) Acta Phys. Pol.* B **30** 2909–78

[42] Urbański P 1996 Double vector bundles in classical mechanics *Rend. Sem. Mat. Univ. Pol. Torino* **54** 405–21

[43] Urbański P 2003 An affine framework for analytical mechanics *Classical and Quantum Integrability* vol 59 ed J Grabowski (Polish Academy of Science, Warsaw: Banach Center Publications) pp 257–79

[44] van der Schaft A J and Maschke B *et al* 1994 On the Hamiltonian formulation of nonholonomic mechanical systems *Rep. Math. Phys.* **34** 225–33

[45] Weinstein A 1996 Lagrangian mechanics and grupoids *Fields Inst. Commun.* **7** 207–31

[46] Zakrzewski S 1990 Quantum and classical pseudogroups: I *Commun. Math. Phys.* **134** 347–70
Zakrzewski S 1990 Quantum and classical pseudogroups: II *Commun. Math. Phys.* **134** 371–95