ELSEVIER

Journal of Geometry and Physics 18 (1996) 295–325

JOURNAL OF GEOMETRY AND PHYSICS

On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints

Franco Cardin *, Marco Favretti [1]

Dipartimento di Matematica Pura e Applicata, Universitá di Padova, via Belzoni, 7-35131 Padova, Italy

Received 29 July 1994; revised 27 February 1995

Abstract

The classical nonholonomic equations for a mechanical system subject to linear nonintegrable constraints are presented in Section 2. These are displayed in a geodesic form by the use of a suitable covariant derivative (due to Synge) in Section 3. We then express the Frobenius integrability of the constraint distribution by means of a zero torsion condition for the above Synge connection. Section 4 provides a self-contained derivation from a nonholonomic variational problem of the equations of motion for nonholonomic systems. These equations, which are nonequivalent to the previous ones, were first developed by Arnold and Kozlov and called vakonomic (vak) equations. Sections 5 and 6 are concerned with a geometrical interpretation of the terms occurring in the right-hand side of the vak equations. Under quite general assumptions, these latter can be described in terms of the curvature of an Ehresmann (local) connection whose horizontal subspace is precisely the constraint distribution. Furthermore, by introducing a suitable Lie group action on the configuration manifold, the local Ehresmann connection can be made into a global one which coincides with the mechanical connection of Smale–Marsden. Section 7 gives a motivation, in terms of Hopf–Rinow and Ambrose–Singer theorems, for the nonclassical requirement of the assignment of the reaction forces' values in the initial kinematical state in vakonomic mechanics. Section 8 develops a fundamental approach to the description of holonomic, i.e. geometrical constraints. We describe the reaction forces by using the Poincaré dual (a class of closed 1-forms) of the orientable constraint submanifold. As an instance of the construction developed in Sections 5 and 6, we consider in Section 9 the disk rolling without sliding on the plane.

Keywords: Nonholonomic systems; Vakonomic equations; Poincaré duality
1991 MSC: 49S05, 70F25, 53B10, 53C12, 93B03

* Corresponding author. E-mail: cardin@pdmat1.math.unipd.it. Fax: (39)(49) 8758596.
[1] E-mail: favretti@pdmat1.math.unipd.it.

SSDI 0393-0440(95)00016-X

1. Introduction

This paper presents an investigation of the dynamical equations for mechanical systems subject to nonholonomic constraints from both mechanical and geometrical points of view. Throughout this work, we describe the system in terms of the variables of the unconstrained system and introduce additional parameters to describe the reaction forces of the constraints. Although this approach allows one to look at the motion and the reaction forces simultaneously, it is rarely pursued in literature, mainly due to the difficulty of its geometrical globalization. The classical line of thought for the same problem (due to Lagrange) is to introduce coordinates adapted to the constraint submanifold and, by means of the Principle of Virtual Work, to get a resulting reduced system of equations where the unknown reaction forces are absent. The application of the above procedure to nonholonomic linear and nonlinear constraints is troublesome in that we are lacking a geometrically satisfactory definition of both the set of virtual displacements compatible with the constraints and that of the (workless) reaction forces. Nonetheless, one usually forces the holonomic construction to the nonholonomic case thereby obtaining a closed system of equations (herein called nonholonomic equations) by defining the virtual displacements according to the so-called nonholonomic condition (see Section 2).

An alternative way to tackle the problem is to drop the Principle of Virtual Work as a cornerstone of the theory and to adopt from the very beginning a variational axiomatic approach, that is to say, the equations of motion are the Euler–Lagrange equations related to a variational nonholonomic problem taking into account the constraints by Lagrange multipliers. This point of view, briefly reported in Section 3, is explained in the book [Arnold et al., 1988] where the authors coin for it the name of vakonomic dynamics (dynamics of variational axiomatic kind). For the sake of simplicity, we adopt here the same terminology. As a matter of fact, the two formulations are nonequivalent for nonintegrable nonholonomic constraints, so that the whole matter is still unsettled.

In this paper, the primary aim is to analyze the detailed geometrical structure of both theories in order to give selective criteria for (a choice between) them. To start with, in Section 2 we introduce a synthetic rational reconstruction of the mechanics of constrained systems and explain the nonholonomic conditions by a scheme close to the one in [Dazord, 1994]. In Section 3 we show that solutions of the nonholonomic equations are affine geodesics of a pseudo-connection introduced by Schouten, defined only on the constraint distribution (see also [Vershik, 1984; Koiller, 1992] for a critical discussion on this point), and then we derive the Synge canonical extension of the above connection. By means of it, we characterize the Frobenius integrability of the constraint linear distribution by a condition of zero *torsion* for the extended connection (Proposition 3.2). This characterization of a purely geometrical property of the system is meaningful from a mechanical point of view, since it uses the same object entering in the dynamical equations.

This feature has a counterpart in vakonomic dynamics in terms of *curvature*. In Section 5 we show in detail that given a constraint distribution $\mathcal{A}$, a Riemannian metric (e.g. induced by the kinetic energy), and if the distribution orthogonal to $\mathcal{A}$ is integrable, we can introduce an Ehresmann (locally defined) connection, whose horizontal subspace is precisely the assigned constraint distribution $\mathcal{A}$. It turns out that the reaction forces of the constraint along the motion, given by the right-hand side of vakonomic equations, can be expressed by means of the curvature of the connection. The same object characterizes the integrability of the constraint distribution by a zero *curvature* condition (Proposition 5.2).

In Section 6, we further pursue this construction. By supposing the distribution orthogonal to $\mathcal{A}$ involutive, hence integrable, we can introduce a group action whose orbits are precisely the leaves of the foliation related to the orthogonal distribution. A theorem of Arnold (Theorem 6.3) allows one under suitable hypotheses to make the above foliation into a fibration and to endow the configuration manifold with a principal bundle structure (Proposition 6.1). Then, we show that the Ehresmann connection previously introduced can be extended to a globally defined principal bundle connection, whose horizontal subspace coincides with $\mathcal{A}$. An instance of this construction, which is physically meaningful, is offered by the disk rolling without sliding on the plane, in Section 9. Moreover, the above connection is precisely the *mechanical connection* of Kummer–Smale in [Marsden, 1991]. In short, for a manifold which is the total space of a principal bundle, and equipped with a metric, the mechanical connection is defined through the related momentum map and the locked inertia tensor. Besides this geometrical analysis of vakonomic equations, Section 6 contains a hierarchical presentation of integrable constraints focussing on the global geometrical properties of the leaves of the constraint foliation.

Special attention is paid throughout the work to illustrate that the intimate structure of vakonomic (variational) equations is richer and more flexible than its nonvariational Chetaev counterpart. As a first instance of this, in Section 4 we show that, by adopting a different definition of varied path (Definition 4.3), the variational scheme generates a variational formulation of nonholonomic equations (Theorem 4.2). At the same time, in Section 4 it is stressed that the varied paths according to Chetaev conditions (necessarily) do not satisfy the constraint up to first order, unlike the classical variations. This latter fact may help to precise the geometrical nature of Chetaev conditions.

Moreover, in Section 7 we introduce the accessibility set of a nonintegrable constraint distribution by means of the classical Chow theorem and its version for principal bundles as given by the Ambrose–Singer theorem and we clarify the classical picture of nonholonomic constraints as "constraints not affecting the possible configurations" in terms of accessibility sets. A nonholonomic version of the Hopf–Rinow theorem in [Vershik and Gershkovich, 1994] (Theorem 7.3) gives the conditions for the existence of solutions of the variational vakonomic problem between two fixed mutually accessible configurations on a manifold equipped with a complete Riemannian metric and a completely nonintegrable constraint distribution. Since vakonomic equations can be given normal form with respect to the configuration variables and Lagrange multipliers, the (initial) Cauchy problem can be well-posed. It follows that vakonomic dynamics has two equivalent formulations: (i) as a variational problem with fixed boundary conditions ($2n$ parameters); (ii) as a Cauchy problem with assigned initial conditions ($2n - \dim \mathcal{A}$ parameters for the initial phase space point *plus* $\dim \mathcal{A}$ parameters for the initial Lagrange multipliers, that is $2n$ parameters). Such an equivalence motivates the quite unusual request to assign the value of the multipliers among the initial data, which we know to be equivalent to the specification of the reaction forces in the initial configuration. Unlike the above situation, the lack of a standard variational formulation of nonholonomic dynamics points out that the relation between dynamical and kinematical accessibility cannot be investigated by the aforementioned arguments.

Finally, Section 8 is concerned with a foundational approach to the Lagrangian holonomic dynamics by the use of the cohomological Poincaré dual of the constraint submanifold, within the vakonomic framework. As we remarked above, the Lagrangian description of the holonomic ideal constraints in the larger framework of the unconstrained manifold, that is $\mathcal{L}(x, \dot{x}, \lambda) = L(x, \dot{x}) + \lambda_\alpha \varphi^\alpha(x)$, is quite unsatisfactory due to the possibly indefinite tensorial character of $\mathcal{L}$. Moreover, it is known that the choice of a Lagrangian asymptotic procedure of realization of the holonomic constraint is highly arbitrary. These procedures are physically interesting because they determine the local full structure of the constraint in a whole sufficiently small neighbourhood of the aforementioned unconstrained manifold.

A survey of the topological obstructions to the aforementioned tensorial character of $\mathcal{L}$ leads us to place the problem at issue into the vakonomic framework. This procedure is performed by choosing, as a global representative of the holonomic distribution (locally given as kernel of the *exact* form $\mathrm{d}\varphi$), a suitable representative of the Poincaré dual class of the constraint submanifold (a *closed* form η). The *Localization Principle* allows us to choose the compact support of the Poincaré dual contained in an arbitrarily small tubular neighbourhood of the constraint manifold, which constitutes a *topological* realization of the holonomic constraint.

2. Constrained mechanical systems

For the sake of completeness, we give a brief outline of the theory of constrained mechanical systems. We suppose as given an inertial space Σ and n material points $M_i, i = 1, \ldots, n$, having masses $m_i, i = 1, \ldots, n$, respectively. We fix an inertial frame at a point O (origin) in Σ and hence describe the configurations of the mechanical system $\mathcal{S}$ of n points by n vectors $OP = (OP_i) = (OP_1, \ldots, OP_n) \in \mathbb{R}^{3n}$. Active forces acting on $\mathcal{S}$' points will be described by given functions $F_i : T\mathbb{R}^{3n} \cong \mathbb{R}^{3n} \times \mathbb{R}^{3n} \longrightarrow \mathbb{R}^{3n*}$, where $\mathbb{R}^{3n*}$ is the dual space of $\mathbb{R}^{3n}$, since we have in mind their characterization as semibasic forms working on virtual displacements (see, e.g. [Godbillon, 1969; Libermann and Marle, 1987]). We say that a constraint is imposed to $\mathcal{S}$ if (i) the positions and velocities of $\mathcal{S}$ are restricted to a submanifold of the tangent bundle $\mathcal{A} \subseteq T\mathbb{R}^{3n}$, eventually dependent on time, and (ii) to every admissible phase space element $(OP, \dot{OP}) \in \mathcal{A}$, this is realized physically by reaction forces $R_i, i = 1, \ldots, n$, belonging to an assigned set $\mathcal{R}_{(OP, \dot{OP})} \subseteq \mathbb{R}^{3n*}$.

To describe the dynamical evolution of $\mathcal{S}$ subject to constraints and active forces, we first define the *dynamically possible* motions $\mathbb{R} \supset I \ni t \mapsto OP_i(t)$ as those satisfying

$$(OP_i(t), \dot{OP}_i(t)) \in \mathcal{A}, \quad \forall t \in I, \tag{2.1}$$

i.e. kinematically possible, and (D'Alembert Principle)

$$m_i \ddot{OP}_i(t) - F_i(OP(t), \dot{OP}(t)) - R_i(t) = 0, \quad R_i(t) \in \mathcal{R}_{(OP, \dot{OP})}, \ \forall t \in I. \tag{2.2}$$

We use in (2.2) the identification between $\mathbb{R}^{3n}$ and $\mathbb{R}^{3n*}$ afforded by the canonical Euclidean isomorphism. Note that the system (2.1), (2.2) is largely undetermined. We can make a fuller use of (2.2) by means of the Principle of Virtual Work; this requires a deeper geometrical and dynamical description of the constraint than the one afforded by $\mathcal{A}$ and $\mathcal{R}$ respectively. Set, for simplicity, $\sigma := (OP_i, \dot{OP}_i) \in \mathcal{A}$, the typical phase space element allowed by the constraint. In the sequel, we follow the scheme developed in [Dazord, 1994].

From $\mathcal{A}$, consider the trivial vector bundle over $\mathcal{A}$, $\tilde{\mathcal{A}} := \mathcal{A} \times \mathbb{R}^{3n} \longrightarrow \mathcal{A}$, and let Φ,

$$\begin{aligned} &\Phi : \tilde{\mathcal{A}} \longrightarrow \mathbb{R}^k, && \Phi(\sigma, u) = \Phi_\sigma u, \\ &\Phi_\sigma \in L(\mathbb{R}^{3n}, \mathbb{R}^k), && \mathrm{rk}\Phi_\sigma = k, \quad \forall \sigma \in \mathcal{A}, \end{aligned} \tag{2.3}$$

be a surjective function, linear on the fibres of $\tilde{\mathcal{A}}$.

Let $V := \ker \Phi$ be the $3n - k$ subbundle of $\tilde{\mathcal{A}}$ whose typical fibre is $V_\sigma := \ker \Phi_\sigma \leq \mathbb{R}^{3n}$ and let V° be the annihilator of V, with typical fibre

$$V_\sigma^\circ := \{R \in \mathbb{R}^{3n*} \colon \langle R, w \rangle = 0, \ \forall w \in V_\sigma\}. \tag{2.4}$$

We call V_σ the space of *virtual* (reversible) *displacements compatible with the constraints* in the phase space configuration $\sigma \in \mathcal{A}$. Note that, unlike in the application to holonomic and nonholonomic constraint, we leave, at this level, the dimension of the constraint manifold $\dim \mathcal{A}$ completely independent from the dimension of the subbundle V of virtual displacements.

The constraint is *ideal* (workless) if and only if the set of reaction forces explicable by the constraints, $\mathcal{R}$, is the k-dimensional bundle over $\mathcal{A}$ of fibre

$$\mathcal{R}_\sigma := V_\sigma^\circ, \quad \forall \sigma \in \mathcal{A}. \tag{2.5}$$

As a straightforward consequence,

$$\ker \Phi_\sigma \subset \ker R, \quad \forall R \in \mathcal{R}_\sigma, \tag{2.6}$$

and, by a well-known theorem of homomorphism of vector spaces, to every $R \in \mathcal{R}_\sigma$, there exists a linear map $\Lambda = (\lambda^1, \ldots, \lambda^k) \in \mathbb{R}^{k*}$ such that $R = \Lambda \circ \Phi_\sigma$, and (2.2) becomes

$$m_i \ddot{OP}_i = F_i(\sigma(t)) + \Lambda(t) \circ \Phi_{\sigma(t)}. \tag{2.7}$$

In order to rewrite (2.1), we can, without loosing any generality, describe locally $\mathcal{A}$ as the union of smooth level sets of some function $f : U(\subset T\mathbb{R}^{3n}) \longrightarrow \mathbb{R}^m$, $f = (f_1, \ldots, f_m)$, and write (2.1) as

$$\mathcal{A}_{|U} = \{(OP_i, \dot{OP}_i) \colon f_\Gamma(OP_i, \dot{OP}_i) = 0\}, \quad \Gamma = 1, \ldots, m = 6n - \dim \mathcal{A}. \tag{2.8}$$

From (2.8), by deriving f_Γ with respect to time along any kinematically possible motion and substituting $\ddot{OP}$ as given by (2.7), we form the linear system

$$B_{\Gamma\alpha} \Lambda_\alpha = C_\Gamma, \tag{2.9}$$

where

$$B_{\Gamma\alpha} := \sum_{i=1}^{n} \frac{1}{m_i} \frac{\partial f_\Gamma}{\partial \dot{OP}_i} \cdot \Phi_{i\alpha} \tag{2.10}$$

and

$$C_\Gamma := -\sum_{i=1}^{n}\left(\frac{1}{m_i}\frac{\partial f_\Gamma}{\partial \dot{OP}_i}F_i + \frac{\partial f_\Gamma}{\partial OP_i}\cdot \dot{OP}_i\right). \tag{2.11}$$

The conditions to determine Λ uniquely can be derived by a simple algebra argument on the rank of B and C. As a consequence, if forces F_i are given and constraints $\mathcal{A}$ are assigned by local functions f_Γ, for every choice of $R \in \mathcal{R}_\sigma$, the Lagrange multipliers Λ can be given as constitutive functions of $\sigma \in \mathcal{A}$. We introduce now the classical choice of Φ for holonomic and nonholonomic (linear and nonlinear) constraints and make some considerations.

A smooth *holonomic*, i.e. purely geometrical, d-dimensional constraint is an embedded smooth submanifold S of $\mathbb{R}^{3n}$, $\dim S = d$, $\partial S = \emptyset$. In this case, $\mathcal{A} = TS$, the tangent bundle of S, and to express Φ we refer to a local representation of S, e.g. as union of level sets of smooth functions $\phi : \mathbb{R}^{3n} \longrightarrow \mathbb{R}^l$, $l = 3n - d$. In such a case, $S \supseteq U = \phi^{-1}(c)$ and $\forall OP \in U$, we state the Φ in (2.3) to be the following:

$$\Phi_\sigma = \Phi_{OP} := \mathrm{d}\phi(OP), \qquad V_{OP} := \ker \mathrm{d}\phi(OP) = T_{OP}U, \tag{2.12}$$

with k in (2.3) equal to l. The functions $f_\Gamma(OP, \dot{OP})$ in (2.8) are given now ($\Gamma = \alpha$) by $\sum_i (\partial\phi_\alpha/\partial OP_i)(OP_i)\dot{OP}_i$. It is easily seen that the $m \times k$ matrix $B_{\Gamma\alpha}$ becomes in this case a $l \times l$ matrix $B_{\alpha\beta}$ and that, to determine Λ uniquely, the above algebraic condition reduces to

$$\mathrm{rk}\sum_{i=1}^{n}\frac{\partial\phi_\alpha}{\partial OP_i}\cdot\frac{\partial\phi_\beta}{\partial OP_i} = \max = l,$$

which is always true if $\mathrm{rk}\,\mathrm{d}\phi = l$. However, the structure of the Lagrange multipliers Λ depends, to every fixed $R \in \mathcal{R}$, on the particular local function ϕ choosen, and this represents a serious drawback if S cannot be given globally as a level set of a single function. This point will be discussed at length in Section 8 where we will introduce a global description of holonomic codimension-one constraints that utilize the cohomological Poincaré dual class of S and in the general framework of vakonomic dynamics.

Nonholonomic (or kinematical) constraints are generally thought as additional constraints over a submanifold $M \subset \mathbb{R}^{3n}$, representing the configurations allowed by a holonomic constraints, that we refer to local coordinates x_i, $i = 1, \ldots, d = \dim M$; as before, we denote by $\sigma = (x, \dot{x})$ the typical point of TM in a fibred chart. For us, a $(d-k)$-dimensional linear nonholonomic constraint is a nonsingular smooth distribution (subbundle)

$$\mathcal{A} \subset TM, \qquad \mathcal{A}_x = \mathcal{A} \cap T_xM \leq T_xM, \qquad \dim \mathcal{A}_x = d - k, \quad \forall x \in M. \tag{2.13}$$

An integral manifold of $\mathcal{A}$ (*leaf*) through $x \in M$ is a $(d-k)$-dimensional submanifold S of M such that $T_xS = \mathcal{A}_x$, and $\mathcal{A}$ is called *integrable* if and only if at every $x \in M$ there exists a (unique) maximal integral manifold through x. The set of leaves $\mathcal{L}$, called *foliation*, forms a partition of M. Frobenius theorem below gives a necessary and sufficient condition of integrability of $\mathcal{A}$. Denote, with a little abuse of notation, by $\mathcal{X}(\mathcal{A})$ the module of smooth vector fields on M, sections of the bundle $\tau_M : \mathcal{A}(\subset TM) \longrightarrow M$.

Theorem (Frobenius). *$\mathcal{A}$ is a completely integrable distribution if and only if for every pair X, Y of differentiable sections of $\mathcal{A}$, $X, Y \in \mathcal{X}(\mathcal{A})$, their Lie bracket is a differentiable section of $\mathcal{A}$, $[X, Y] \in \mathcal{X}(\mathcal{A})$.*

A general (linear) $(d-k)$-dimensional distribution $\mathcal{A} \subset TM$ can be locally assigned as the kernel of k suitable linearly independent 1-forms $A_\alpha : TU \to \mathbb{R}$, i.e.

$$\mathcal{A}_{|U} = \{(x, \dot{x}) \in TU \colon A_{\alpha i}(x)\dot{x}^i = 0\}, \qquad \operatorname{rk} A_{\alpha i} = k, \tag{2.14}$$

and Lichnerowitz version of Frobenius theorem [Lichnerowicz, 1954, p.41] gives, whenever $\mathcal{A}$ is completely integrable, the following local representation for the forms A_α: there exists

(i) a matrix-valued function $a_{\alpha\beta} = a_{\alpha\beta}(x)$, $\det a \neq 0$, and
(ii) k real-valued functions $\phi_\alpha : U \subset M \longrightarrow \mathbb{R}$ such that

$$A_{\alpha i}(x) = a_{\alpha\beta}(x)\frac{\partial \phi_\beta}{\partial x^i}(x), \quad x \in U, \tag{2.15}$$

hence $\mathcal{A}$ can be rewritten as

$$\mathcal{A}_{|U} = \left\{(x, \dot{x}) \in TU \colon a_{\alpha\beta}(x)\frac{\partial \phi_\beta(x)}{\partial x^i}\dot{x}^i = 0\right\} \equiv \left\{(x, \dot{x}) \in TU \colon \frac{\partial \phi_\beta(x)}{\partial x^i}\dot{x}^i = 0\right\}. \tag{2.16}$$

It is now clear that U is the union of level sets (leaves of the foliation), i.e. there exists an open set $W \subset \mathbb{R}^k$, and

$$U = \bigcup_{c \in W} \phi^{-1}(c), \qquad \phi = (\phi_1, \ldots, \phi_k), \tag{2.17}$$

so that

$$\mathcal{A}_{|U} = \bigcup_{c \in W} T(\phi^{-1}(c)) = \bigcup_{c \in W} \{(x, \dot{x}) \in TU \colon x \in \phi^{-1}(c), \dot{x} \in T_x(\phi^{-1}(c))\}. \tag{2.18}$$

Condition (2.1) is locally given by $A_{\alpha i}(x(t))\dot{x}^i(t) = 0$, so the kinematically possible motions are the motions along a fixed leaf. Therefore, the previous choice for Φ still applies and, in particular,

$$\Phi_{(x,\dot{x})} = \Phi_x := (A_{\alpha i}(x)) = \left(\frac{\partial \phi_\alpha(x)}{\partial x^i}\right), \qquad V_x := \ker(A_{\alpha i}(x)) = \ker\left(\frac{\partial \phi_\alpha(x)}{\partial x^i}\right). \tag{2.19}$$

Note that, as before, we get a global description of the constraint $\mathcal{A}$ if and only if the related foliation is a fibration (see Section 5).

For nonholonomic *nonintegrable* linear constraints, even in the lack of a local integral manifold, and hence of an underlying geometrical interpretation of the set of virtual displacements, the natural assignment of Φ is to *identify* the subbundle V of virtual displacements precisely with the distribution $\mathcal{A}$, and to set again $\Phi := (A_{\alpha i})$. This is in agreement with the historical line of thought, as drawn in [Whittaker, 1944].

Finally, for *nonholonomic nonlinear* constraints, given as usual by local functions

$$\mathcal{A}_{|U} = \{(x,\dot{x}) \in TM\colon f_\alpha(x,\dot{x}) = 0, \alpha = 1,\ldots,k\}, \qquad \mathrm{rk}\left(\frac{\partial f_\alpha}{\partial \dot{x}^i}\right) = k, \tag{2.20}$$

it is largely acknowledged that the choice of Φ and V is given by the Appell-nonholonomic-Hamel conditions [Appell, 1904; Chetaev, 1932, 1933; Arnold et al., 1988, p.17]

$$\Phi_{(x,\dot{x})} := \left(\frac{\partial f_\alpha}{\partial \dot{x}^i}(x,\dot{x})\right) : \mathbb{R}^d \longrightarrow \mathbb{R}^k, \qquad V_{(x,\dot{x})} := \ker\left(\frac{\partial f_\alpha}{\partial \dot{x}^i}(x,\dot{x})\right) \tag{2.21}$$

which still allows one to determine uniquely the Lagrange multipliers Λ from (2.9). Note that this assignment of Φ is partially unsatisfactory for a number of reasons:

(a) The related equations of motion (2.22) and (2.23) below are not equivalent to the ones deriving from the variational principle of stationary action with constraints (see Section 3).

(b) The expression of multipliers Λ depends on the assignment of local functions f_α and the tensorial character of Λ with respect to changes of overlapping charts is not clear.

If active and inertial forces acting on $\mathcal{S}$ points are taken into account by a smooth Lagrangian function $L : TM \longrightarrow \mathbb{R}$, the dynamical equations of $\mathcal{S}$ subject to nonholonomic ideal constraints are – respectively in the linear and nonlinear case – the following analogues of (2.1), (2.2),

$$[L]_i = \lambda^\alpha A_{\alpha i}(x), \qquad A_{\alpha i}(x)\dot{x}^i = 0, \tag{2.22}$$

and

$$[L]_i = \lambda^\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i}(x,\dot{x}), \qquad f_\alpha(x,\dot{x}) = 0, \tag{2.23}$$

where

$$[L]_i := \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{x}^i}\right) - \frac{\partial L}{\partial x^i}.$$

In the following, we will restrict to systems subject to linear constraints and we will call for simplicity, Eq. (2.22) nonholonomic equations of motion.

3. A torsion free condition for the integrability of constraint distribution in nonholonomic dynamics

Let M be a smooth manifold referred to local coordinates $x_i, i = 1,\ldots,n$, and let $L : TM \longrightarrow \mathbb{R}$, $L := \frac{1}{2}g_{ij}(x)\dot{x}^i\dot{x}^j$ be the Lagrangian. M has naturally a structure of Riemannian manifold induced by the metric g_{ij}; furthermore, the motions of the free system (M, L) are geodesic curves with respect to g_{ij}, i.e. solutions of

$$\nabla_{\dot{x}}\dot{x} = [L] = 0, \tag{3.1}$$

where ∇ is the covariant derivative of the Levi-Civita torsion-free connection. Add to the system a linear constraint $\mathcal{A} \subset TM$, whose local expression is (2.14). In [Vershik, 1984], the author proves the following proposition.

Proposition 3.1. *The equations of motions of the constrained system* $(M, L, \mathcal{A})$ *can be given the geodesic form* (3.1) *with respect to a suitable covariant derivative to be defined below by* (3.3) *and* (3.4).

Proof. By using the metric g_{ij} we define $\mathcal{A}^{\perp}$ as the orthogonal bundle of $\mathcal{A}$; then

$$\mathcal{A}_x^{\perp} := \mathrm{span}\{A^i{}_\alpha\} = \left\langle A^i{}_\alpha(x)\frac{\partial}{\partial x^i}\right\rangle, \tag{3.2}$$

where $A^i{}_\alpha := g^{ij}A_{\alpha j}$, and the orthogonal projector on $\mathcal{A}$

$$P : TM \longrightarrow TM, \qquad \ker P = \mathcal{A}^{\perp}, \qquad \mathrm{Im}\, P = \mathcal{A}. \tag{3.3}$$

On the fibred product $\mathcal{X}(\mathcal{A}) \times_M \mathcal{X}(\mathcal{A})$ define

$$\nabla^v := P\nabla, \qquad \nabla^v : \mathcal{X}(\mathcal{A}) \times_M \mathcal{X}(\mathcal{A}) \longrightarrow \mathcal{X}(\mathcal{A}). \tag{3.4}$$

Although ∇^v satisfies the formal properties of covariant derivative (see [Vershik, 1984, p.291]), it is not defined on the whole bundle (see also [Koiller, 1992, p.141]). However, recalling (2.22),

$$[L]^i = \nabla^i_{\dot{x}}\dot{x} = \lambda^\alpha A^i_\alpha(x),$$

by (3.4) we obtain

$$\nabla^v_{\dot{x}}\dot{x} = P\nabla_{\dot{x}}\dot{x} = P([L]) = P(\lambda^\alpha A^i_\alpha) = 0. \qquad \square \tag{3.5}$$

Remark. Solutions of (3.5) are affine geodesics, which are not solutions of a variational problem of minimum length because ∇^v is not inherited from some Riemannian metric. This point will be focussed in Section 4, making a comparison with dynamic equations for the same system derived from a nonholonomic variational problem (vak).

Now, we define a genuine connection on TM whose covariant derivative $\tilde{\nabla}$ is a proper extension of ∇^v to the whole bundle TM. This result, due to [Synge, 1928], is used in [Benenti, 1987]; we will employ it to characterize the Frobenius complete integrability of the constraint subbundle $\mathcal{A}$.

Let $Q : TM \longrightarrow TM$, $Q = I - P$ be the orthogonal projection operator on $\mathcal{A}^{\perp}$; $\ker Q = \mathcal{A}$, $\mathrm{Im}\, Q = \mathcal{A}^{\perp}$. From (3.4), for every vector field $Y \in \mathcal{X}(M)$

$$\begin{aligned} \nabla^v_i Y &= P\nabla_i Y = (I - Q)\nabla_i Y = \nabla_i Y - Q(\nabla_i Y) \\ &= \nabla_i Y - \nabla_i(QY) + (\nabla_i Q)Y = \tilde{\nabla}_i Y - \nabla_i(QY), \end{aligned} \tag{3.6}$$

where we have defined

$$(\tilde{\nabla}_i Y)^h := (\nabla_i Y + (\nabla_i Q)Y)^h = Y^h_{,i} + \Gamma^h_{ij}Y^j + (\nabla_i Q)^h_j Y^j = Y^h_{,i} + \tilde{\Gamma}^h_{ij}Y^j. \tag{3.7}$$

Since $(\nabla_i Q)^h_j$ is a (2, 1)-tensor, $\tilde{\Gamma}^h_{ij} := \Gamma^h_{ij} + (\nabla_i Q)^h_j$ are the Christoffel symbols of a genuine connection defined on the *whole* bundle; moreover, from (3.6)

$$\forall Z \in \mathcal{X}(\mathcal{A}), \quad QZ = 0 \Rightarrow \tilde{\nabla}_i Z = \nabla^v_i Z,$$

so $\tilde{\nabla}$ is a canonical extension of ∇^v, since it is independent of the particular choice of A_α defining $\mathcal{A}$. The local expression of $Q(= Q^2)$ is the following:

$$\begin{aligned} &Q^i_k = g^{ij} Q_{jk} = g^{ij} \Pi^{\alpha\beta} A_{\alpha j} A_{\beta k}, \\ &\Pi_{\alpha\beta} := g^{rs} A_{\alpha r} A_{\beta s}, \quad \det(\Pi_{\alpha\beta}) \neq 0, \quad \Pi^{\alpha\beta} = (\Pi^{-1})_{\alpha\beta}. \end{aligned} \tag{3.8}$$

When restricted to $\mathcal{X}(\mathcal{A})$, the symbols $\tilde{\Gamma}^h_{ij}$ have the special form (3.9)

$$\begin{aligned} (\tilde{\Gamma}^h_{is} - \Gamma^h_{is}) Z^s &= (\nabla_i Q)^h_s Z^s = -Q^h_s (\nabla_i Z)^s = -g^{hr} \Pi^{\alpha\beta} A_{\alpha r} A_{\beta s} (\nabla_i Z)^s \\ &= -g^{hr} \Pi^{\alpha\beta} A_{\alpha r} [\nabla_i (A_{\beta s} Z^s) - (\nabla_i A_\beta)_s Z^s] \\ &= g^{hr} \Pi^{\alpha\beta} A_{\alpha r} (\nabla_i A_\beta)_s Z^s. \end{aligned} \tag{3.9}$$

By introducing the deformation tensor of the vector field $A_\alpha \in \mathcal{X}(M)$,

$$D_{is\beta} = \tfrac{1}{2}[(\nabla_i A_\beta)_s + (\nabla_s A_\beta)_i], \tag{3.10}$$

we see that, along any kinematically possible motion, we can rewrite (3.5) as

$$(\nabla^v_{\dot{x}} \dot{x})^i = \ddot{x}^i + \tilde{\Gamma}^i_{jk} \dot{x}^j \dot{x}^k = 0, \tag{3.11}$$

$$(\nabla_{\dot{x}} \dot{x})^i + \Pi^{\alpha\beta} D_{rs\alpha} A^i_\beta \dot{x}^r \dot{x}^s = 0, \tag{3.12}$$

and find the expression of the reaction forces of the constraint along the motion as a constitutive function of $(x, \dot{x})$ by comparison with (2.22), that is

$$\lambda^\alpha A^i_\alpha = R^i = -\Pi^{\alpha\beta} D_{rs\alpha} A^i_\beta \dot{x}^r \dot{x}^s. \tag{3.13}$$

Note that [Benenti, 1987, p.9]

$$D_{is\beta} \equiv 0 \Leftrightarrow A_\beta \text{ is a } \textit{Killing vector} \text{ of the metric } g_{ij}.$$

The connection $\tilde{\Gamma}^h_{is}$ introduced above allows us to characterize the Frobenius integrability of the constraint distribution $\mathcal{A} = \ker(A_{\alpha i})$. We prove the following proposition.

Proposition 3.2. *$\mathcal{A}$ is a completely integrable subbundle of TM $\Leftrightarrow$ the 2-form $\mathrm{d}A_\alpha$ verifies $\mathrm{d}A_\alpha(X, Y) = 0$, $\forall X, Y \in \mathcal{X}(\mathcal{A})$ $\Leftrightarrow$ $\tilde{T}(X, Y) = 0$, $\forall X, Y \in \mathcal{X}(\mathcal{A})$, where $\tilde{T}$ is the torsion of the extended connection $\tilde{\Gamma}$.*

Proof. Define as usual

$$\tilde{T}(X, Y) = \tilde{\nabla}_X Y - \tilde{\nabla}_Y X - [X, Y]. \tag{3.14}$$

Since, if $X, Y \in \mathcal{X}(\mathcal{A})$, $\tilde{\nabla}_X Y = \nabla^v_X Y = P \nabla_X Y = (I - Q) \nabla_X Y$, we find

$$\begin{aligned}\tilde{T}(X,Y) &= P\nabla_X Y - P\nabla_Y X - [X,Y] \\ &= P(\nabla_X Y - \nabla_Y X) - [X,Y] = P([X,Y]) - [X,Y]\end{aligned}$$

so

$$\tilde{T}(X,Y) = 0, \forall X,Y \in \mathcal{X}(\mathcal{A}) \Leftrightarrow P([X,Y]) = [X,Y],$$

i.e. if and only if $\mathcal{A}$ is an integrable distribution according to Frobenius theorem. From (3.14), we obtain $\tilde{T}$ referred to coordinate fields

$$\tilde{T}^i_{jk} = \tilde{\Gamma}^i_{jk} - \tilde{\Gamma}^i_{kj} = (\nabla_j Q)^i_k - (\nabla_k Q)^i_j,$$

so, if $\forall X, Y \in \mathcal{X}(\mathcal{A})$, recalling (3.9), we have

$$\begin{aligned}\tilde{T}^i_{jk} &= g^{ir}\Pi^{\alpha\beta}[(\nabla_j A_\beta)_k - (\nabla_k A_\beta)_j]A_{\alpha r} \\ &= g^{ir}\Pi^{\alpha\beta}[A_{\beta k,j} - A_{\beta j,k}]A_{\alpha r} = \Pi^{\alpha\beta}(\mathrm{d}A_\beta)_{jk}A^i_\alpha\end{aligned} \tag{3.15}$$

and

$$\tilde{T}(X,Y) = \Pi^{\alpha\beta}\,\mathrm{d}A_\beta(X,Y)A_\alpha = 0, \forall X,Y \in \mathcal{X}(\mathcal{A}) \Leftrightarrow \mathrm{d}A_{\beta|\mathcal{A}} \equiv 0. \qquad \square \tag{3.16}$$

Remark. Note that the projected covariant derivative ∇^v does not allow to define torsion by the usual formula if $\mathcal{A}$ is not integrable since, for the natural candidate torsion T^v we have, for $X, Y \in \mathcal{X}(\mathcal{A})$,

$$\begin{aligned}\tilde{T}(X,Y) &:= \nabla^v_X Y - \nabla^v_Y X - [X,Y] = P(\nabla_X Y - \nabla_Y X) - [X,Y] \\ &= P([X,Y]) - [X,Y] = -Q([X,Y]) \in \mathcal{A}^\perp,\end{aligned}$$

that is T^v does not work into $\mathcal{X}(\mathcal{A})$, unlike (3.4). This difficulty, pointed out by Vershik [Vershik, 1984, p.292] is here removed by the very use of the canonically extended connection $\tilde{\Gamma}$.

4. Equations of motion of variational axiomatic kind (vak) for constrained mechanical systems

The framework of the Virtual Work Principle to draw the equations of motions in presence of ideal constraints, is not the only rational scheme ensued by theorists on mechanics. Actually, one can take as starting point for a theory of constrained mechanical systems the Hamilton (or stationary action) principle, stating that the dynamically possible motion is the one which extremizes a chosen functional F in a suitable class of kinematically possible paths joining two fixed points. This amounts to say that the equations of motion are the Euler–Lagrange equations of a variational Lagrange problem of calculus of variations with differential constraints. For Lagrangian mechanical systems subject to holonomic or linear integrable constraints, the two principles are perfectly equivalent in that they do provide the same set of dynamic equations, but for a more general choice of constraint, i.e. nonlinear or nonintegrable ones, it is acknowledged, at least from the mid fifties works of [Capon, 1952; Jeffrey, 1954; Pars, 1954] that this equivalence is lost and the two formulations split up. It is clear that the problem attains to the very foundations of analytical mechanics and that it can be settled only by looking at the ability of both theories to deal with concrete examples. In fact, most of the literature on the subject, known to us, deals with nonholonomic equations of motion especially in order to give an intrinsic geometrical formulation of them [Vershik and Faddeev, 1981] and of nonholonomic conditions [Massa and Pagani, 1991], but there are a number of works [Rumyantsev, 1981; Kozlov, 1982a, 1982b, 1983; Arnold et al., 1988] devoted to the study of the equations of motion, for nonholonomic constrained systems, deriving from a variational nonholonomic problem or, as the authors [Arnold et al., 1988] say, to the study of vakonomic mechanics (mechanics of variational axiomatic kind).

Up to now, the problem at issue, especially as far as concrete examples are concerned, is far from resolved. However, the study of variational problems with nonholonomic constraints, initiated by [Vranceanu, 1928; Synge, 1928; Vagner, 1940] and continued in the works of [Chow, 1939] (see [Vershik and Gershkovich, 1994] for a historical survey) has made clear, in a rigorous way, the difference between "straightest" curves, i.e. the mechanics of nonholonomic systems, and "shortest", i.e. the variational theory of nonholonomic systems, a terminology introduced by Hertz. In this section we first introduce the dynamics of nonholonomic constrained systems according to [Arnold et al., 1988]. More in detail, we lay down a unifying variational framework; within this, we gain the main result of vak dynamics (Theorem 4.1). Then, in the same framework, we show that (i) the nonholonomic equations of motion can be given as solutions of a conditional extremum problem, (ii) once (i) is achieved, the difference between these equations and vakonomic ones (see below (4.3)) amounts to a different definition for the admissible varied paths, (iii) the equations do coincide for holonomic and linear integrable constraints, and (iv) the reaction forces of the constraints in the vakonomic formulation are gyroscopic in the general case.

Let M be a smooth manifold, $L : TM \longrightarrow \mathbb{R}$ a smooth Lagrangian function, $f_\alpha : TM \longrightarrow \mathbb{R}$ k functions whose covectors $\partial f_\alpha/\partial \dot{x}^i$ are linearly independent, i.e. $\mathrm{rk}(\partial f_\alpha/\partial \dot{x}^i) = k$, as in (2.21). A smooth path $x : [t_1, t_2] \longrightarrow M$ is *admissible* if and only if it satisfies the constraint $f_\alpha(x, \dot{x})$. We first define the variations of admissible paths.

Definition 4.1. A variation of the admissible path $x : [t_1, t_2] \longrightarrow M$ is a smooth homotopy $z : [-c, c] \times [t_1, t_2] \longrightarrow M$ such that

(i) $z(0, t) = x(t), \quad \forall x \in [t_1, t_2]$,

(ii) $z(\epsilon, t_1) = x_1, \quad z(\epsilon, t_2) = x_2, \quad \forall \epsilon \in [-c, c]$,

(iii) $z(\epsilon, t)$ satisfies the constraint f_α to the *first order* in ϵ.

Let $\gamma := x([t_1, t_2])$. A variation of $x(\cdot)$ defines a *variation vector field*

$$W : \gamma \subset M \to TM, \qquad t \mapsto (x(t), w(t)), \qquad w(t) := \frac{\partial z}{\partial \epsilon}(0, t) \tag{4.1}$$

and

$$\tau_M \circ W = \mathrm{id}_\gamma. \tag{4.2}$$

As a straightforward consequence of Definition 4.1 we have the following proposition.

Proposition 4.1. *A smooth section W as in* (4.2) *is a variation vector field* (4.1) *along a fixed $x(\cdot)$ if and only if*
(1) $w(t_1) = w(t_2) = 0$,
(2) $\mathcal{F}^{\mathrm{vak}}(w) = 0$, $\mathcal{F}^{\mathrm{vak}} = (\mathcal{F}_1^{\mathrm{vak}}, \dots, \mathcal{F}_k^{\mathrm{vak}})$, *where*

$$\mathcal{F}_\alpha^{\mathrm{vak}}(w) := \frac{\partial f_\alpha}{\partial x^i}(x,\dot{x})w^i + \frac{\partial f_\alpha}{\partial \dot{x}^i}(x,\dot{x})\dot{w}^i = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_\alpha}{\partial \dot{x}^i}w^i\right) - [f_\alpha]_i w^i. \tag{4.3}$$

$\mathcal{F}_\alpha$ is a linear functional between the linear space $X := \{t \mapsto (x(t), w(t)) \colon (x,w) \in C^\infty([t_1,t_2], TM), w(t_1) = w(t_2) = 0\}$ and $C^\infty([t_1,t_2], \mathbb{R}^k)$. Moreover, given a variation vector field w, it uniquely defines a smooth variation, whose expression in a local chart is $z(\epsilon, t) = x(t) + w(t)\epsilon$. We can now state [Arnold et al., 1988, p.32]:

Definition 4.2 (VAK). The admissible path $x(\cdot)$ is a dynamically possible motion if and only if it is an extremal of the functional

$$F = \int_{t_1}^{t_2} L(x,\dot{x})\,\mathrm{d}t \tag{4.4}$$

with respect to every conditioned variation according to Definition 4.1.

The following theorem of calculus of variations characterizes the dynamically possible motions $x(\cdot)$ according to Definition 4.2.

Theorem 4.1. *The admissible path $x(\cdot) : [t_1,t_2] \to M$ is a conditional extremal for the functional F if and only if there exist k smooth functions $\lambda^\alpha : [t_1,t_2] \to \mathbb{R}$ such that $x(\cdot)$ verifies*

$$[L]_i = \lambda^\alpha [f_\alpha]_i + \dot{\lambda}^\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i}(x,\dot{x}). \tag{4.5}$$

Remark. Eqs. (4.5) are called equations of motion of *vakonomic* type. It is important to stress that vakonomic equations (4.5) can be derived as Euler–Lagrange equations of an unconditioned variational problem for a functional of type (4.4) with Lagrangian

$$\mathcal{L}(x,\dot{x},\lambda) = L(x,\dot{x}) - \lambda^\alpha f_\alpha(x,\dot{x}), \tag{4.6}$$

where the Lagrange multipliers λ are added independent parameters and

$$[\mathcal{L}]_{\lambda_\alpha} = f_\alpha(x,\dot{x}) = 0$$

are precisely the constraint equations (2.23).

Proof of Theorem 4.1. Let $x(\cdot)$ be an admissible path, $z(\epsilon, t)$ a variation of $x(\cdot)$. Define

$$F(\epsilon) := \int_{t_1}^{t_2} L(z(\epsilon,t), \dot{z}(\epsilon,t))\, \mathrm{d}t. \tag{4.7}$$

It is a well-known result that

$$\mathrm{d}F(w) := \left.\frac{\mathrm{d}F(\epsilon)}{\mathrm{d}\epsilon}\right|_{\epsilon=0} w = -\int_{t_1}^{t_2} [L]_i w^i \,\mathrm{d}t$$

is a linear functional on X. The path $x(\cdot)$ is an extremal for F, according to Definition 4.2 (VAK), if and only if

$$\ker \mathcal{F}^{\mathrm{vak}} \subseteq \ker \mathrm{d}F, \tag{4.8}$$

where $\mathcal{F}^{\mathrm{vak}}$ is defined by (4.3). By a well-known theorem of homomorphism of vector spaces (see [Jacobson, 1965, p.61]) there exists $\Lambda : C^{\mathrm{vak}}(:= \mathrm{Im}\, \mathcal{F}^{\mathrm{vak}}) \to \mathbb{R}$, linear, such that $\mathrm{d}F = \Lambda \circ \mathcal{F}^{\mathrm{vak}}$. By Riesz' Representation Theorem on $C^\infty([t_1,t_2], \mathbb{R}^k)$, endowed with the scalar product $(f, g) = \int_{t_1}^{t_2} f^\alpha g_\alpha \,\mathrm{d}t$, here $f^\alpha \equiv f_\alpha$, we can single out k smooth functions $\lambda^\alpha : [t_1, t_2] \to \mathbb{R}$ such that, $\forall y \in C^{\mathrm{vak}}$

$$\Lambda(y) = \int_{t_1}^{t_2} \lambda^\alpha y_\alpha \,\mathrm{d}t,$$

so, to every $W = (x, w) \in X$,

$$\mathrm{d}F(w) = \Lambda \circ \mathcal{F}^{\mathrm{vak}}(w),$$

$$\mathrm{d}F(w) = -\int_{t_1}^{t_2} [L]_i w^i \mathrm{d}t = \Lambda \circ \mathcal{F}^{\mathrm{vak}}(w) = \int_{t_1}^{t_2} \lambda^\alpha \left\{ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial f_\alpha}{\partial \dot{x}^i} w^i \right) - [f_\alpha]_i w^i \right\} \mathrm{d}t$$

$$= \int_{t_1}^{t_2} -\left\{ \dot{\lambda}^\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i} + \lambda^\alpha [f_\alpha]_i \right\} w^i \mathrm{d}t + \left[\lambda^\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i} w^i \right]_{t=t_1}^{t=t_2},$$

and thesis (4.5) follows from the application of the fundamental lemma of calculus of variations. □

The fundamental entry of the above demonstration is that $\mathcal{F}^{\mathrm{vak}}$ is a *linear* functional. Bearing this in mind, we now show that nonholonomic equations (2.23) can be obtained as solutions of a variational problem (4.4) for a suitable choice of $\mathcal{F}$. We first state

Definition 4.3. A smooth vector field W as in (4.2) is a nonholonomic variation vector field along a fixed γ if and only if

(i) $w(t_1) = w(t_2) = 0$,

(ii) $\mathcal{F}^{\mathrm{Ch}}(w) = 0, \mathcal{F}^{\mathrm{Ch}} = (\mathcal{F}_1^{\mathrm{Ch}}, \dots, \mathcal{F}_k^{\mathrm{Ch}})$,
where

$$\mathcal{F}^{\mathrm{Ch}}(w) := \frac{\partial f_\alpha}{\partial \dot{x}^i}(x, \dot{x}) w^i. \tag{4.9}$$

We recognize in (ii) the variational analogue of Chetaev' conditions (2.21). As an immediate consequence of Definition 4.3, we have the following theorem.

Theorem 4.2. *The admissible path* $x(\cdot)$ *is a conditional extremal for* (4.4), *where nonholonomic variations according to Definition* 4.3 *are taken into account, if and only if there exist k smooth functions* $\mu^\alpha : [t_1, t_2] \to M$ *such that*

$$[L]_i = \mu_\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i}(x, \dot{x}). \tag{4.10}$$

Proof. Utilize $\mathcal{F}^{\mathrm{Ch}}$ instead of $\mathcal{F}^{\mathrm{vak}}$ in the proof of Theorem 4.1. Then, there exists $\hat{\Lambda} = (\mu^1, \dots, \mu^k) : C^{\mathrm{Ch}}(:= \mathrm{Im}\, \mathcal{F}^{\mathrm{Ch}}) \to \mathbb{R}$ such that

$$\mathrm{d}F = \hat{\Lambda} \circ \mathcal{F}^{\mathrm{Ch}},$$

and, $\forall\, W = (x, w) \in X$,

$$\mathrm{d}F(w) = \int_{t_1}^{t_2} [L]_i w^i \, \mathrm{d}t = \hat{\Lambda} \circ \mathcal{F}^{\mathrm{Ch}}(w) = \int_{t_1}^{t_2} \mu^\alpha \frac{\partial f_\alpha}{\partial \dot{x}^i} w^i \, \mathrm{d}t.$$

We conclude by using again the fundamental lemma. □

We now particularize the expression of $\mathcal{F}^{\mathrm{vak}}$ and $\mathcal{F}^{\mathrm{Ch}}$ for holonomic and nonholonomic linear integrable constraints.

Holonomic constraints. Their local expression is $\phi_\beta(x) = 0, \beta = 1, \dots, k$, so $f_\beta(x, \dot{x}) = \phi_{\beta,i}(x)\dot{x}^i = 0$ is the kinematical constraint associated to ϕ_β. Then, from (4.10) and (4.5) respectively, since $[f_\beta]_i = 0$, we have that

$$\begin{aligned} [L]_i &= \mu^\beta \phi_{\beta,i} \quad \text{are the nonholonomic equations,} \\ [L]_i &= \dot{\lambda}^\beta \phi_{\beta,i} \quad \text{are the vakonomic equations.} \end{aligned} \tag{4.11}$$

They generate the same dynamics under the identification $\mu^\beta = \dot{\lambda}^\beta$.

Integrable linear constraints. Their local expression is (see (2.15))

$$f_\alpha(x, \dot{x}) = a_\alpha^\beta(x) \phi_{\beta,i}(x) \dot{x}^i = 0.$$

By a straightforward computation, we have in the vakonomic and nonholonomic case respectively

$$[L]_i = \frac{\mathrm{d}}{\mathrm{d}t}(\lambda^\alpha a_\alpha^\beta)\phi_{\beta,i} = l^\beta \phi_{\beta,i},$$

$$[L]_i = \mu^\alpha a_\alpha^\beta \phi_{\beta,i} = m^\beta \phi_{\beta,i}.$$

In both the holonomic and linear integrable cases, the redefinition of Lagrange multipliers by time derivative or by means of the matrix a_α^β, does not alter the structure of the dynamical equations. Finally, for linear nonintegrable constraints, the vakonomic equations are

$$[L]_i = \dot\lambda^\alpha A_{\alpha i}(x) + \lambda^\alpha \, \mathrm{d}A_\alpha(\dot x, \cdot)_i . \tag{4.12}$$

From $A_{\alpha i}(x)\dot x^i = 0, \mathrm{d}A_\alpha(\dot x, \dot x) = 0$, it is immediate to recognize their gyroscopic character. In Section 5, we will continue the study of linear nonintegrable constraints.

5. A zero curvature condition for the integrability of constraint distribution in vakonomic dynamics

Let $(M, L, \mathcal{A})$ be a Lagrangian system subject to linear nonintegrable constraints given by a k-codimensional distribution $\mathcal{A}$ on TM, whose local expression is $A_{\alpha i}(x)\dot x^i = 0$, $\alpha = 1, \dots, k$. Define $A_\alpha^i := g^{ij}A_{\alpha j}$, where g is a Riemannian metric, e.g. induced by the kinetic energy component of L, and consider the k-dimensional orthogonal distribution

$$x \in M, \quad x \mapsto \mathcal{A}_x^\perp := \mathrm{span}\{A_\alpha\} = \langle A_\alpha^i \frac{\partial}{\partial x^i}\rangle . \tag{5.1}$$

Note that $\forall x \in M, \mathcal{A}_x^\perp \oplus \mathcal{A}_x = T_xM$. We make the following hypotheses:

(H.1) The smooth distribution $\mathcal{A}^\perp$ is Frobenius integrable.

Since $\dim \mathcal{A}^\perp = \operatorname{codim} \mathcal{A}$, hypothesis (H.1) is very reasonable when the system (M, L) is weakly constrained: in the limiting case $\operatorname{codim} \mathcal{A} = 1$, (H.1) is trivially satisfied.

Denote with $M/\mathcal{A}^\perp$ the set of leaves of the foliation associated to the distribution $\mathcal{A}^\perp$ and let $\pi : M \to M/\mathcal{A}^\perp$ be the map $x \mapsto \{\text{leaf through } x\}$. Proposition 5.1 shows the local structure of the foliation.

Proposition 5.1 ([Molino, 1988, p.15]). *To every $x \in M$ there exists a neighbourhood U of x which is a simple set of the foliation, that is,* (i) *$\pi : U \to U/\mathcal{A}^\perp$ is a differentiable surjective submersion,* (ii) *the leaves (level sets of π) are connected,* (iii) *$U/\mathcal{A}^\perp$ is a $(n-k)$-differentiable manifold.*

This, together with the metric g, is all we need to define an Ehresmann connection [Ehresmann, 1950] on the fibre bundle $\pi : U \to U/\mathcal{A}^\perp$. We proceed as in [Marsden et al., 1990]. The vertical subspace is $VU := \ker T\pi = \mathcal{A}^\perp = \mathrm{span}\{A_\alpha\}$. To define a connection, we define a *horizontal* subspace HU, supplementary to VU by means of g,

$$H_xU := (\mathcal{A}_x^\perp)^\perp = \mathcal{A}_x = \ker\{A_\alpha\}, \qquad H_xU \oplus V_xU = T_xU, \tag{5.2}$$

so $\forall X \in T_xU, X = X^h + X^v$, where $X^h := h(X)$, and $X^v := v(X)$ are the horizontal, respectively, vertical projections of X. The vertical projector, $v : TU \longrightarrow VU$, $\ker v = HU$, is the *connection* 1*-form*; to express v components, instead of using the base of T_xM in a local chart adapted to the submersion, we use the bases $\{\partial/\partial x^i\}$ and $\{A_\alpha\}$ of TU and

$VU = \mathcal{A}^{\perp}$ respectively. By comparison with the projector Q, introduced in (3.8), we derive immediately

$$v_i^{\alpha}(x) := Q_i^{\alpha} = \Pi^{\alpha\beta} A_{\beta i}. \tag{5.3}$$

Then, by using the well-known Cartan formula [Kobayashi and Nomizu, 1963, p.77] we define the *curvature 2-form* of the connection v: $\forall X, Y \in \mathcal{X}(M)$

$$\Omega^{\alpha}(X, Y) := \mathrm{d}v^{\alpha}(X^h, Y^h) = -v^{\alpha}([X^h, Y^h]) = \Pi^{\alpha\beta}\, \mathrm{d}A_{\beta}(X^h, Y^h). \tag{5.4}$$

By means of v and Ω, we rewrite vakonomic equations (4.12) of last section and the constraint equations as follows:

$$A_{\alpha i}(x)\dot{x}^i = 0 \Leftrightarrow \dot{x} \in \mathcal{A}_x = H_x U \Leftrightarrow v(x)\dot{x} = 0, \tag{5.5}$$

$$[L](\cdot) = \langle \dot{\mu}, v(\cdot)\rangle + \langle \mu, \Omega\rangle(\dot{x}, h(\cdot)) + \langle \mu, \mathrm{d}A\rangle(\dot{x}, v(\cdot)), \tag{5.6}$$

where $\mu_\beta := \Pi_{\alpha\beta}\lambda^{\alpha}$, $\dot{\mu}_\beta := \Pi_{\alpha\beta}\dot{\lambda}^{\alpha}$, $\langle \dot{\mu}, v(\cdot)\rangle_i := \dot{\lambda}^{\alpha}\Pi_{\alpha\beta}\Pi^{\delta\beta}A_{\delta i} = \dot{\lambda}^{\alpha}A_{\alpha i}$ and $\langle\cdot,\cdot\rangle$ is the pairing between $\mathbb{R}^k$ and $\mathbb{R}^{k*}$.

On the one hand, the horizontal distribution $x \mapsto H_x U$, locally defined in (5.1), does coincide on U with the smooth distribution $\mathcal{A}$; since the intersection of two simple sets is a simple set, it is correctly defined on the whole manifold M. On the other hand, since the diffeomorphism π between HU and $U/\mathcal{A}^{\perp}$ is merely a local one, we can only lift paths that lies wholly inside $U/\mathcal{A}^{\perp}$. However, thanks to its local character, we can characterize the integrability of the distribution $\mathcal{A}$ in terms of the curvature Ω (strictly speaking, $\Omega_{|U}$) defined above.

By invoking the Frobenius criterion of integrability, we lay down the following proposition.

Proposition 5.2. *$\mathcal{A} = HM$ is an integrable distribution $\Leftrightarrow$ $\forall X, Y \in \mathcal{X}(\mathcal{A}), [X, Y] \in \mathcal{X}(\mathcal{A})$, i.e. $\forall X, Y$ horizontal vector fields, their Lie bracket is a horizontal vector field $\Leftrightarrow$ $\forall X, Y \in \mathcal{X}(\mathcal{A}),\ -v([X, Y]) = \Omega(X, Y) = 0 \Leftrightarrow \Omega \equiv 0$.*

Recall the projected connection $\tilde{\Gamma}$ on M defined in Section 2 and the associated torsion $\tilde{T}$. Putting together Propositions 5.2 and 3.2 we have the following proposition.

Proposition 5.3. *$\mathcal{A}$ is an integrable distribution $\Leftrightarrow$ $\mathrm{d}A_{\alpha|\mathcal{A}} = 0$, $\alpha = 1, \ldots, k$ $\Leftrightarrow$ $\tilde{T}_{|\mathcal{A}} = 0 \Leftrightarrow \Omega = 0$.*

Note that, in order to characterize the integrability of $\mathcal{A}$, the torsion $\tilde{T}$ of $\tilde{\Gamma}$ which is a linear connection defined on the linear frame bundle $F(M) \to M$ with structure group $G = GL(n, \mathbb{R})$, plays the *same role* of the curvature Ω of v which is a locally defined connection on the fibre bundle $U \to U/\mathcal{A}^{\perp}$, $U \subset M$.

In Section 6, by adding suitable hypotheses on distribution $\mathcal{A}^{\perp}$, we will be able to make the local fibration $U \to U/\mathcal{A}^{\perp}$ into a global fibration $M \to M/\mathcal{A}^{\perp}$ which does coincide with the one $M \to M/G$, to be defined through a suitable abelian Lie group action Φ :

$G \times M \to M$. Then $(M, M/G, G, \pi)$ is a principal bundle and we can introduce notions as the accessibility set of the constraint *horizontal* distribution, $\mathcal{A} = HM$, or the holonomy of a kinematically possible, i.e. horizontal, path. These are useful tools in the study of global properties of nonholonomic linear nonintegrable constraints.

6. Principal bundle structure for constraints whose orthogonal distribution is involutive

Let $(M, L, \mathcal{A})$ be a Lagrangian mechanical system, subject to linear nonintegrable constraints $\mathcal{A}$. By hypothesis (H.1) of Section 5, we regard the constraint distribution $\mathcal{A}$ as the horizontal subspace of an Ehresmann connection v and we are able to describe the reaction forces of the constraints by the very use of the related curvature Ω. Note that this analysis of the constraints is local, since we are able to describe the geometrical and mechanical behaviour of the constraints only in a neighbourhood of a point x in M. The following theorem allows one to study the global properties of a leaf of the constraint foliation.

Theorem 6.1 ([Guillemin and Sternberg, 1984, p.209]). *Let F be a smooth foliation and let L be a compact leaf whose holonomy group is trivial. Then there exists a neighbourhood U of L which is the union of leaves of F, and a disk D in $\mathbb{R}^k$, where k is the codimension of F, and a diffeomorphism $f : L \times D \to U$ such that $f_{|L\times\{0\}} = \mathrm{id}_L$.*

In particular, if all the leaves of F are compact and have trivial holonomy, (for example if they are simply connected) then F is a *fibration* (see Definition 6.1).

Definition 6.1 ([Dieudonné, 1974, p.74]). A smooth fibration is a triple (M, B, π) where M and B are smooth manifolds, π a smooth surjective map from M onto B verifying the following condition of *local triviality*:

$\forall b \in B$ there exists an open neighbourhood U of b in B, a smooth manifold F and a diffeomorphism $\varphi : U \times F \to \pi^{-1}(U)$ such that $\pi(\varphi(y, z)) = y$ for every $y \in U$ and $z \in F$.

The hypothesis of compact leaves is necessary to measure the distance between two neighbouring leaves by using geodesics of some Riemannian metric. It can be dropped if M is endowed with a *bundle-like* metric [Reinhart, 1959, p.122], that is a metric g on TM whose expression in a local chart $x^i = (z^1, \ldots, z^{d-k}, y^1, \ldots, y^k)$ (leaves are y^α = const.) is

$$\mathrm{d}s^2 = g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j = g_{\Gamma\Delta}(z, y)\omega^\Gamma\omega^\Delta + g_{\alpha\beta}(y)\,\mathrm{d}y^\alpha\,\mathrm{d}y^\beta, \tag{6.1}$$

where $\{\omega^\Gamma(z, y), \mathrm{d}y^\alpha\}$ is a base of $T^*_{(z,y)}M$. In this case, the following theorem holds.

Theorem 6.2 ([Hermann, 1960, p.445]). *Let M be a smooth manifold with a complete Riemannian metric which is bundle-like with respect to a foliation F. If all the leaves are closed,* *then M/F is a metric space and the application $x \mapsto \phi(x) = \{\text{leaf through } x\}$ is an open and continuous map. If all the leaves have trivial holonomy, then M/F is a smooth manifold and ϕ is a surjective submersion. Hence $(M, M/F, \phi)$ is a fibration.*

As an immediate outcome on the description of the trajectories of constrained systems, we see that if the constraint distribution gives rise to a foliation, then we have only a *local* equivalence with a holonomic system, because in the local fibrations (see Proposition 5.1) which describe the foliation at any point $x \in M$ we are not able to decide whether two local leaves belong to the same global leaf or not. On the contrary, if we deal with a fibration, related to the constraint distribution, then the kinematically possible motions are precisely the ones along a fixed connected fibre, which is an embedded submanifold L of M, that we can express as a level set of the submersion π. The last property is equivalent to say that the *normal bundle* to L, defined with the metric g, which we know to be the bundle of reaction forces, is *trivial* [Guillemin and Pollack, 1974, p.97] (see also Section 8). It is worthwhile to notice that by using the above outlined scheme of fibrations, Marle formulates a theory of active constraints [Marle, 1991].

We now introduce a global group action $G \times M \to M$, where G is abelian, such that $(M, M/G, G, \pi)$ is a principal fibre bundle, and a connection on it whose horizontal space is precisely the constraint distribution $\mathcal{A}$ coinciding with the *mechanical connection* in [Marsden, 1991, p.50]. Let $(M, L, \mathcal{A})$ be the Lagrangian constrained mechanical system introduced above and $\mathcal{A}_x^{\perp} = \operatorname{span}\{A_\alpha(x)\}$, where $A_\alpha^i(x) = g^{ij}A_{\alpha j}(x)$, a local representation of the orthogonal distribution (5.1). Besides hypothesis (H.1), we make the stronger

(H.2) The k smooth vector fields A_α are commuting, i.e. $[A_\alpha, A_\beta] = 0$.

Obviously, (H.2) is a special case of (H.1). Moreover, (H.2) is trivially satisfied if $k = 1$, and it is a mild request if the codimension of the system is low. Let $\Phi^\alpha : I \subset \mathbb{R} \times M \to M$ be the flux of the o.d.e.

$$\frac{\mathrm{d}x^i}{\mathrm{d}\tau^\alpha} = A_\alpha^i(x). \tag{6.2}$$

By (H.2), they are commuting, i.e. $\Phi^\alpha_{\tau^\alpha} \circ \Phi^\beta_{\tau^\beta} = \Phi^\beta_{\tau^\beta} \circ \Phi^\alpha_{\tau^\alpha}$. Now we suppose that

(H.3) Solutions of o.d.e. (6.2) have maximal extension to the whole $\mathbb{R}$.

Note that (H.3) is satisfied if, for instance, M is a compact manifold. (H.2) and (H.3) together allow to define a global smooth action of $(\mathbb{R}^k, +)$, abelian, on M

$$\begin{aligned} &\Phi(\tau, x) := \Phi^1_{\tau^1} \circ \cdots \circ \Phi^k_{\tau^k}(x), \qquad \Phi : \mathbb{R}^k \times M \to M, \\ &(\tau, x) \mapsto \Phi_\tau(x) = \Phi(\tau, x), \end{aligned} \tag{6.3}$$

since we have trivially $\Phi_0 = \mathrm{id}_M$, $\Phi(\tau + \rho, x) = \Phi(\tau, \Phi(\rho, x))$. Moreover, we identify the leaf of $\mathcal{A}^\perp$ through x with the orbit $\mathbb{R}^k x := \{\Phi(\tau, x) : \tau \in \mathbb{R}^k\}$ through x and $\mathcal{A}_x^\perp$ with $T_x(\mathbb{R}^k x)$. By definition, Φ acts *transitively* on M but not *freely* in the general case. However, starting from Φ, we can define a free group action by means of the following theorem.

Theorem 6.3 ([Arnold, 1978]). *Let M be a smooth manifold, and $\Phi : R^k \times M \to M$ be a smooth nonfree action of $(\mathbb{R}^k, +)$ on M. Then*
(a) *the isotropy group Γ_y of Φ is the same $\forall y \in \mathbb{R}^k x$,*
(b) *if $\dim \Gamma_y = r$, $\mathbb{R}^k x$ is diffeomorphic to $\mathbb{T}^r \times \mathbb{R}^{k-r} := G$ (abelian) where $\mathbb{T}^r$ is the r-dimensional torus,*
(c) *if $\dim \Gamma_x = r \ \forall x \in M$, there exists a free action $\Psi : G \times M \to M$ whose orbits do coincide with the ones of Φ.*

Now we have made $\pi : M \to M/G$ into a fibre bundle, with G acting freely and transitively on the orbits. To make $\pi : M \to M/G$ into a principal bundle (see [Kobayashi and Nomizu, 1963, p.50]) we need the local trivializability, that is to say $\forall U \subset M/G$, $\pi^{-1}(U)$ is diffeomorphic to $U \times G$. This is afforded by the following theorem.

Theorem 6.4 ([Bourbaki, 1958, p.63]). *Let G be a Lie group and $G \times M \to M$ be a group action. If*
(H.4) *G acts properly and freely on M and*
(H.5) *$\forall x \in M$, the map $g \mapsto gx$ is an immersion of G in M,*
then $(M, M/G, G, \pi)$ is a principal fibre bundle.

We are ready to lay down the following proposition.

Proposition 6.1. *Suppose hypotheses* (H.2)–(H.5) *to hold. Then, with reference to the constraint distribution $\mathcal{A}$, we have the following chain of equalities*:

$$\mathcal{A}_x^{\perp} = \ker T_x\pi = T_x(Gx) = \{X_a(x) : a \in \mathcal{G}\} = V_x M, \tag{6.4}$$

where $\mathcal{G}$ is the Lie algebra of G and $X_a(x) \in T_x M$ is the vector field whose infinitesimal generator is $a \in \mathcal{G}$:

$$X_a^i := \frac{\mathrm{d}}{\mathrm{d}t}\left(\Psi_{\mathrm{e}^{at}}(x)\right)^i\Big|_{t=0} = A_\alpha^i a^\alpha, \quad \alpha = 1, \ldots, k, \ a \in \mathcal{G}(= \mathbb{R}^k). \tag{6.5}$$

Recall that $\mathcal{A}_x \oplus \mathcal{A}_x^{\perp} = T_x M$. If we define

$$H_x M := \mathcal{A}_x = \ker\{A_{\alpha i}(x)\}, \qquad T_x M = H_x M \oplus V_x M, \tag{6.6}$$

then $\mathcal{A}$ is the horizontal subspace of a *connection* on the principal bundle π, which coincides with the local one introduced in (5.3), if and only if (see [Kobayashi and Nomizu, 1963, p.63])

$$T_x\Psi_\tau(\mathcal{A}_x) = \mathcal{A}_{\Psi_\tau(x)}, \quad \forall x \in M, \ \tau \in G. \tag{6.7}$$

In our setting, condition (6.7) reads

(H.6) $T_x\Psi_\tau(\mathcal{A}_x) = (\mathcal{A}^{\perp}_{\Psi_\tau(x)})^{\perp}$.

Such a hypothesis (H.6), which involves the group action, the metric and the constraint, can also be written as

$$\Psi_{\tau *}(x)A_\alpha(x) = \Pi_{\alpha\rho}(\Psi_\tau(x))\Pi^{\rho\beta}(x)A_\beta(x). \tag{6.8}$$

Remark. The connection (6.6) is precisely the *mechanical* connection of Marsden [Marsden, 1991, p.50] on (M, g) whose construction (sketch) is reported below:

Define a *connection* on M by a connection 1-form $\alpha : TM \to \mathcal{G}$ by

$$\alpha := \Pi^{-1} \circ J \circ FL \tag{6.9}$$

where

$$FL : TM \to T^*M, \qquad FL(x, \dot{x}) = (x, p), \tag{6.10}$$

is the Legendre *transformation*,

$$J : T^*M \to \mathcal{G}^*, \qquad \langle J(x, p), a\rangle = p(X_a) = p_i A^i_\alpha a^\alpha, \quad \forall a \in \mathcal{G} \tag{6.11}$$

is the *momentum map* [Marsden, 1991, p.52],

$$\Pi : \mathcal{G} \to \mathcal{G}^*, \qquad \langle \Pi(x)a, b\rangle = g(X_a, Y_b) = g_{ij} A^i_\alpha a^\alpha A^j_\beta b^\beta = \Pi_{\alpha\beta} a^\alpha b^\beta, \tag{6.12}$$

is the *locked inertia tensor* (that coincide with $\Pi_{\alpha\beta}$ introduced in (3.8)). Finally, by the above formulas, we have the local expression of α

$$(\alpha(x))^\rho := \Pi^{\rho\sigma} A_{\sigma i} \; (= v^\rho_i) \tag{6.13}$$

that coincides with (5.3).

We refer to textbooks of differential geometry for the notion of *holonomy* of a connection. We simply recall here its construction to look at its mechanical interpretation for the case at issue. The assignation of a connection 1-form $\alpha : TM \to \mathcal{G}$ uniquely defines a global *horizontal lift* $l : TB \to HM$ where $(B = M/G)$. We can then lift a path γ in B (eventually closed) to a path $\tilde{\gamma} = l(\gamma)$ which is horizontal, i.e. $\dot{\tilde{\gamma}}(s) \in H_{\tilde{\gamma}(s)}M$, that is kinetically possible. If $\gamma : [0, 1] \to B$ is a loop, $x_0 = \tilde{\gamma}(0)$ and $x_1 = \tilde{\gamma}(1)$ are two, not necessarily coincident, configurations of the system allowed by constraints that lie on the same fibre $\pi^{-1}(\gamma(0))$, which is an orbit of Ψ. There exists therefore a $\tau \in G$ such that $\Psi(\tau, x_0) = x_1$. The element $\tau \in G$ is called the *holonomy of the loop* γ in x_0.

There is an increase in literature devoted to the study of applications of this subject; let us quote, among others, the isoholonomic problem, i.e. to find extremals of a given cost functional (e.g. the kinetic energy) on the set of admissible (i.e. horizontal) paths with fixed holonomy (cf. [Montgomery, 1990] and the bibliography quoted therein).

7. A comparison between nonholonomic and vakonomic dynamics in the light of Chow, Ambrose–Singer and Hopf–Rinow theorems

Solutions of vakonomic equations (4.5) or (5.6) are extremals of a variational constrained problem with fixed initial and final configurations, say x and y. Of course, the problem is well-posed when, among other requirements, y belongs to the *accessibility set* of x; more precisely, let M be a smooth manifold, and $\mathcal{A}$ be the constraint linear distribution. We state

Definition 7.1. To every $x \in M$, the accessibility set of x is

$$M(x) = \{y \in M\colon \exists\, \gamma : [0,1] \to M, \gamma \in KC^1, \gamma(0) = x, \gamma(1) = y, \dot{\gamma} \in \mathcal{A}\}. \quad (7.1)$$

Denote with $\mathcal{L}(\mathcal{A})$ the subalgebra generated by $\mathcal{A}$, that is the minimum subspace of $\mathcal{X}(M)$ containing $\mathcal{X}(\mathcal{A})$ and all the finitely iterated Lie products of vector fields in $\mathcal{X}(\mathcal{A})$. $\mathcal{L}(\mathcal{A})$ defines an involutive distribution, $\mathcal{D}(\mathcal{A})$, possibly of nonconstant dimension. If dimension of $\mathcal{D}(\mathcal{A})$ is constant, by Frobenius theorem (Section 2), it is integrable, and $M(x)$ is precisely the maximal integral manifold of $\mathcal{D}(\mathcal{A})$ through x [Sussmann, 1973]. The accessibility set of a typical distribution, even singular, can be investigated by the following theorem [Chow, 1939; Sussmann, 1973].

Theorem 7.1 (Chow). *Let M be an n-dimensional smooth manifold, $\mathcal{A}$ a possibly singular distribution. If* $\dim \mathcal{D}(\mathcal{A})(x) = n$, *then $M(x)$ is an open set containing x; if moreover* $\dim \mathcal{D}(\mathcal{A})(x) = n$, $\forall x \in M$, *then $M(x) = M$ (and $\mathcal{A}$ is called a completely nonholonomic distribution).*

It is now clear that the classical picture of nonholonomic constraints as constraints which do not affect the possible configurations of the system can be made rigorous in terms of the accessibility sets of the constraint. If the configuration space M has a principal bundle structure, and the constraint distribution is the horizontal distribution $\mathcal{A} = HM$ of a connection, we can easily compute $\mathcal{D}(\mathcal{A})$ by using the classical Ambrose–Singer theorem.

Theorem 7.2 (Ambrose–Singer [Kobayashi and Nomizu, 1963, p.89]). *Let $(M, M/G, G, \pi)$ be a principal bundle, with M/G a connected and paracompact manifold; denote with $\mathcal{A} = HM$ the horizontal subbundle of a connection, $G(x)$ the holonomy group at x of the connection, $M(x)$ the holonomy bundle (which is precisely the accessibility set of $\mathcal{A}$ in x). Let $\mathcal{G}_x$ be the Lie algebra of $G(x)$. Then $\mathcal{G}_x$ is the following image of the curvature 2-form when restricted to the horizontal distribution in the accessibility set of x, that is,*

$$\mathcal{G}_x = \{\Omega(y)(X, Y)\colon y \in M(x), X, Y \in H_y M\} \leq \mathcal{G}. \quad (7.2)$$

From the proof of Ambrose–Singer theorem [Kobayashi and Nomizu, 1963, p.89] we easily derive:

$$\mathcal{D}(HM)(y) = H_y M \oplus \sigma(\mathcal{G}_y) = T_y(M(x)), \qquad \forall y \in M(x), \quad (7.3)$$

where $\sigma(\mathcal{G}_y) = \{X_a(y)\colon a \in \mathcal{G}_y\}$ and X_a is the vector field whose infinitesimal generator is $a \in \mathcal{G}_y$. Moreover, since M is a disjoint union of holonomy bundles $M(x)$, which are all isomorphic, we have $\dim M(x) = \text{const.}\ \forall x \in M$. Then

$$\dim \mathcal{D}(HM)(x) = \dim (M/G) + \dim \mathcal{G}_x = \dim T_x(M(x)) = \text{const.} \quad \forall x \in M,$$

and $\mathcal{D}(\mathcal{A})$ is an involutive distribution of constant dimension, hence Frobenius integrable. The following corollary is a straightforward application of Chow's theorem.

Corollary. *If* $\sigma(\mathcal{G}_x) = V_x M$, *that is if* $\mathcal{G}_x = \mathcal{G}$, *then* $\dim \mathcal{D}(\mathcal{A}) = n$ *and* $M(x) = M$.

We are now concerned with the *existence* of solutions γ of the variational problem between mutually accessible, fixed configurations $x, y \in M$:

$$\inf \int_{t_0}^{t_1} L\,\mathrm{d}t, \quad L = g(\dot\gamma, \dot\gamma)^{1/2}, \quad \gamma(t_0) = x,\ \gamma(t_1) = y, \quad \dot\gamma \in \mathcal{A}. \tag{7.4}$$

For nonholonomic constraint distributions, this is stated in terms of the following version of the classical Hopf–Rinow theorem.

Theorem 7.3 (Hopf–Rinow [Vershik and Gershkovich, 1994, p.37]). *Let M be a smooth complete Riemannian manifold, $\mathcal{A}$ a completely nonholonomic distribution, i.e.* $\dim \mathcal{D}(\mathcal{A}) = n$. *Then, to every pair $x, y \in M$ there is a piecewise smooth geodesic with $\dot\gamma \in \mathcal{A}$ joining them.*

Remind that the above *geometrical* variational problem (7.4) is equivalent, up to reparametrizations, to the *mechanical* one with $L = \frac{1}{2} g(\dot\gamma, \dot\gamma)$. Therefore, there exists the mechanical solution too and we can gain smoothness for it by restricting ourselves to a suitably small neighbourhood of x. This attention avoids also the occurrence of conjugate points phenomena and consequent multiplicity of solutions, a matter explained by the Morse theory. The following considerations concern the vakonomic and nonholonomic formulations of dynamical equations under the hypothesis of *completely nonholonomic* constraint distribution $\mathcal{A}$;

(A) As seen in Section 4, vakonomic equations are the Euler–Lagrange equations of an unconditioned variational problem. The existence of solutions between accessible configurations, given by Hopf–Rinow theorem, allows us to give two *equivalent* formulations of the equations of motion:
  (1) as a *variational problem*, see (4.4), with *boundary conditions* x, y assigned, that is $2n$ parameters,
  (2) as a *Cauchy problem*, see (4.5), for the Euler–Lagrange equations deriving from (1) with assigned initial conditions $x(t_0)$, $\dot x(t_0) \in \mathcal{A}_{x(t_0)}$, that is $2n - k$ parameters *plus k initial values of the Lagrange multipliers* $\lambda(t_0)$, hence $2n$ parameters as in (1).

Actually, vakonomic equations (4.5) can be given the normal form with respect to x and λ (see [Vershik and Gershkovich, 1994, p.39]), and the above assignment (2) means to fix the reaction forces of the constraint in the initial phase space point $x(t_0), \dot x(t_0)$. In a different way, in nonholonomic nonvariational formulation of dynamics, the Lagrange multipliers can be expressed as constitutive functions of $(x, \dot x)$ as shown in Section 2, so the Cauchy initial data assignment involves *only* $x(t_0)$, $\dot x(t_0) \in \mathcal{A}_{x(t_0)}$.

(B) The lack of a standard variational formulation of nonholonomic equations (2.23), implies that the relation between dynamical accessibility, i.e. dynamically possible motions and geometrical accessibility, i.e. kinematically possible motions, cannot be investigated with the aforementioned theorems. Recall that nonholonomic equations are *geodesic* of the projected connection (4.6) which is not Riemann-metrizable generally, so the equations are not identifiable with those coming from a variational problem of minimum length (they are only affine geodesic). This is confirmed by the fact that, in the variational problem of Theorem 4.2 of Section 4, which gives rise to nonholonomic equations, the varied paths do not satisfy the constraint.

For completeness, we rephrase here a theorem in [Vershik and Gershkovich, 1988] which shows that solutions of nonholonomic and vakonomic equations are in the general case totally different. Let (M, L) be as in Section 3 and let γ_ν be a geodesic of the projected connection, i.e. solution of nonholonomic equations (3.5), where $\nu \in S_x$ is an initial kinematically possible unit vector, $S_x = \{\nu \in \mathcal{A}_x, \|\nu\| = 1\}$. Let $\gamma_{\nu,\omega}$ be the solution of vakonomic equations (4.12) for $\nu \in S_x$ and let $\omega = \lambda^\alpha A_\alpha \in \mathcal{A}^\perp$ for a given choice of the initial value of the Lagrange multipliers. Denote by B the set of points $\nu \in S_x$ for which there is a choice of ω such that at x the germ of γ_ν coincides with the germ of $\gamma_{\nu,\omega}$.

Theorem 7.4. *Suppose $\mathcal{A}$ is a distribution in general position. Then, for some open, everywhere dense subset of the space of (kinetic energy) metrics, B is empty.*

8. A global geometrical setting of holonomic constraints in the framework of vakonomic dynamics by means of Poincaré duality

This section is concerned with a foundational approach to the description of holonomic constraints. Consider a Lagrangian holonomic system $L(x, \dot{x})$ on TM, where M is a smooth manifold. If a new holonomic ideal constraint $j : S \hookrightarrow M$ is added, the standard way to deal with it is via the classical line of thought (due to Lagrange): the resulting system is still Lagrangian, where

$$L_S := L \circ Tj : TS \to \mathbb{R}. \tag{8.1}$$

In such a formulation, the description of reaction forces is obviously absent. Since we wish to describe them , we have turned to the alternative scheme of Lagrange multipliers. Briefly, as before, the constraint is geometrically defined by the level set of suitable smooth local functions $\varphi_\alpha(x_i)$, $\alpha = 1, \ldots, k$, $i = 1, \ldots, n$; the parameters describing the mechanical system are the old variables x plus the Lagrange multipliers λ, and the new Lagrange function is

$$\mathcal{L}(x, \lambda, \dot{x}, \dot{\lambda}) := L(x, \dot{x}) + \lambda^\alpha \varphi_\alpha(x). \tag{8.2}$$

Although the Lagrangian function (8.2) could be interpreted as a generalized Lagrangian generating function (Morse family) on TM in the Tulczyjew sense (see [Benenti, 1982] and the bibliography quoted therein), we have already remarked that neither the global meaning of (8.2) nor the tensorial character of λ^α is clear. In other words, if S cannot be given as a level set of a global function defined in some open neighbourhood of S in M, then $\mathcal{L}$ in (8.2) is not a scalar invariant function; this requirement for S, i.e. $S = \varphi^{-1}(0)$, is a severe topological condition; it holds if and only if the *normal bundle* of S (for some Riemannian metric on M, e.g. the kinetic energy one) is *trivial* (see [Guillemin and Pollack, 1974, p.77]). A generalization of the above hypothesis on S leads us into the hard task of the topological characterization of trivial normal bundles. This difficulty is illustrated by the following classical theorem on the subject.

Theorem 8.1 ([Hirsh, 1988, p.79]). *Every vector bundle over a contractible paracompact space is trivial.*

We might enlighten the latter *homotopic* hypothesis by a *homological* argument. In particular, we have in the mind to consider constraint manifold S whose homology is not trivial. This choice seems to be adequate for a number of reasons:

(a) If, for example $S = \varphi^{-1}(0)$, S orientable, $\dim S = \dim M - k$, where φ is defined on the *whole* manifold M, then S is a *boundary* of some Ω, e.g. setting $\Omega = \{x \in M: \varphi_\alpha(x) > 0,\ \alpha = 1, \ldots, k\}$ we have $S = \partial\Omega$, and, as a straightforward consequence, the Poincaré dual cohomological class of S (see Appendix A) is *trivial*. Indeed, let the closed k-form η on M be a representative of the Poincaré dual class of S, $\eta \in [\eta_S] \in H^k(M)$; By definition, for every $n-k$ closed form $\omega \in H_c^{n-k}(M)$ one has

$$\int_M \omega \wedge \eta = \int_S j^*\omega. \tag{8.3}$$

Now, by using Stokes theorem

$$\int_S j^*\omega = \int_{\partial\Omega} j^*\omega = \int_\Omega \mathrm{d}(j^*\omega) = \int_\Omega j^*\,\mathrm{d}\omega = 0, \tag{8.4}$$

and we have that $[\eta_S]$ vanishes; in the case $k = 1$, $[\eta_S]$ contains the *exact* form $\mathrm{d}\varphi$, which enters in the description of reaction forces as given by the Euler–Lagrange equations for $\mathcal{L}$ in (8.2).

(b) The above point (a) suggests that the Poincaré dual of a homologically nontrivial one-codimensional constraint S is a tool that might allow us to describe the reaction forces by means of *closed* 1-forms η_S *globally* defined and with compact support contained in a tubular neighbourhood of S, which generalize the *exact* local 1-forms $\mathrm{d}\varphi$. Moreover, when we describe the reaction forces of a constraint by using a nontrivial Poincaré class, we are sure to deal with a manifold which is not a boundary, and therefore that cannot be the level set of a unique function φ defined on the whole M. The last statement follows from de Rham and Poincaré dualities.

(c) The description of reaction forces by the 1-form $\eta \in [\eta_S]$ has a precise global *geometrical* sense, and the choice of a particular representative η, together with its compact support, corresponds to the assignation of the physical zone of influence of the constraint.

The latter consideration is a promising starting point for a theory of realization of constraints. Briefly, one generally considers certain *asymptotic* procedures of realization of the new holonomic constraint S, e.g. by means of a sort of penalty method:

$$\mathcal{L}_\epsilon(x,\dot{x}) := L(x,\dot{x}) - (1/2\epsilon)\langle\varphi(x),\varphi(x)\rangle, \quad \epsilon \to 0^+, \tag{8.5}$$

where $\langle,\rangle$ is some nondegenerate bilinear form on $\mathbb{R}^k$ and $S = \varphi^{-1}(0)$ as above.

Note that a geometrical desingularization of the Euler–Lagrange system related to $\mathcal{L}_\epsilon$ is given by (see [Cardin, 1991]):

$$\mathbf{L}_\epsilon(x,\mu,\dot{x},\dot{\mu}) := L(x,\dot{x}) + \tfrac{1}{2}\epsilon\langle\mu,\mu\rangle - \langle\mu,\varphi(x)\rangle. \tag{8.6}$$

Indeed, on one hand,

$$[\mathcal{L}_\epsilon]_i = [L]_i + \frac{1}{\epsilon}\left\langle\frac{\partial\varphi(x)}{\partial x^i},\varphi(x)\right\rangle = 0, \tag{8.7}$$

on the other hand

$$[\mathbf{L}_\epsilon]_i = [L]_i + \left\langle\mu,\frac{\partial\varphi(x)}{\partial x^i}\right\rangle = 0,$$

$$[\mathbf{L}_\epsilon]_\alpha = \epsilon\mu_\alpha - \varphi_\alpha(x) = 0, \tag{8.8}$$

hence we obtain precisely (8.7). Furthermore,

$$\lim_{\epsilon\to 0^+} \mathbf{L}_\epsilon = \mathbf{L}_0 = \mathcal{L}, \tag{8.9}$$

where the last equality holds by recognizing the multiplier μ of $\mathbf{L}_0$ to be the multiplier λ up a sign.

The *analytical* problem – that is when and in which sense the solutions of Euler–Lagrange equations of $\mathcal{L}_\epsilon$ tend to the solutions of $\mathcal{L}$, as $\epsilon \to 0^+$ – of the realization of constraint by asymptotic procedures is widely treated (see [Benettin et al., 1987, 1989] and the literature quoted therein).

By reconsidering the above asymptotic procedure, we notice that the choice of the *confinement potential* $(1/2\epsilon)\langle\varphi(x),\varphi(x)\rangle$ is highly arbitrary and therefore it is natural to look for a geometrical, global and intrinsic class of asymptotic realizations.

This aim can be pursued, in our opinion, by the very use of the vakonomic dynamic approach coupled with the use of the cohomological Poincaré class $[\eta_S]$ in order to assign, in a physically meaningful way, the reaction forces of the holonomic ideal constraint S, with codim $S = 1$.

We proceed in a axiomatic way: to the smooth orientable manifold S we associate its Poincaré dual $[\eta_S]$, which we assume to be nontrivial, since we left aside the trivial case when S is a boundary. We postulate that, for a fixed choice of the 1-form $\eta \in [\eta_S]$, the constraint is described by

$$\eta_i(x)\dot{x}^i = 0, \tag{8.10}$$

and the dynamics is the one defined by the Lagrangian

$$\hat{\mathcal{L}}(x, \dot{x}, \zeta) = L(x, \dot{x}) - \zeta \eta_i(x)\dot{x}^i, \tag{8.11}$$

so the corresponding dynamical equations are ($\mathrm{d}\eta = 0$)

$$\begin{aligned} &[\hat{\mathcal{L}}]_i = 0\colon [L]_i = \dot{\zeta}\eta_i(x), \\ &[\hat{\mathcal{L}}]_\zeta = 0\colon \eta_i(x)\dot{x}^i = 0. \end{aligned} \tag{8.12}$$

This formulation has the following remarkable features:

(i) $\hat{\mathcal{L}}$ is globally defined on M in a rigorous geometrical way.

(ii) When $S = \varphi^{-1}(0)$ for some globally defined φ on M, then we obviously choose $\eta = \mathrm{d}\varphi$ and get back the complete description of vakonomic dynamics as in (4.11) by identifying λ with $\dot{\zeta}$.

(iii) In the general case (possibly $[\eta_S] \neq 0$), we fix an arbitrarily small tubular neighbourhood N of S in M (if we are given a Riemannian metric), and we can find a representative η of $[\eta_S]$ whose support is contained in N; this gives a concrete meaning to the *Localization Principle*. Moreover, it represents the *topological* equivalent of the analytical asymptotic realization procedure, since we can shrink the tubular neighbourhood N to S, finding always a suitable representative. Pictorially, in $M\backslash\{\operatorname{supp}\eta\}$, the system behaves as an unconstrained one, whose trajectory $x(t)$ verifies $[L]_i = 0$, but inside $\{\operatorname{supp}\eta\}$ the reaction force of the constraint, given by the right-hand side of (8.12), bends the trajectory (see (8.10)) to stay inside the fixed neighbourhood N. By shrinking N to S, we have the asymptotic topological realization of S.

9. An Ehresmann connection on a principal bundle for the vertical rolling disk

Consider a homogeneous disk D of unit mass and radius r, rolling without sliding and leaning on a plane. Let O be the origin in the plane and refer the system to coordinates $z = (x, y, \varphi, \theta)$, where $(x, y) \in \mathbb{R}^2$ are the coordinates of the point P of contact between the disk and the plane, and (φ, θ) are respectively the angle between the x-axis and the vector OP, and the angle between the vertical axis, orthogonal to the plane, and a fixed radius of the disk. The configuration manifold is then $M = \mathbb{R}^2 \times \mathbb{T}^2$. Let C be the centre of the disk and ω the angular velocity of the rigid body D. The condition of rolling without sliding imposes to the system the following nonholonomic linear *nonintegrable* constraint

$$V_P = V_C + \omega \wedge CP = 0, \tag{9.1}$$

equivalent to the system of two equations, linear in $\dot{z}$,

$$\begin{aligned} A_{1i}(z)\dot{z}^i &= \dot{x} + r\dot{\theta}\cos\varphi = 0, \\ A_{2i}(z)\dot{z}^i &= \dot{y} + r\dot{\theta}\sin\varphi = 0, \end{aligned} \tag{9.2}$$

where $A_1 = (1, 0, 0, r\cos\varphi) \in T^*M$, $A_2 = (0, 1, 0, r\sin\varphi) \in T^*M$ are two linearly independent 1-forms, globally defined on TM, describing the constraint as in (2.14). Denote by I the inertia tensor of D; the Lagrangian of the disk is

$$L = T = \tfrac{1}{2}\{V_C \cdot V_C + \omega \cdot I\omega\}$$
$$= \tfrac{1}{2}\{\dot{x}^2 + \dot{y}^2 + \tfrac{1}{4}r^2(\dot{\varphi}^2 + 2\dot{\theta}^2)\} = \tfrac{1}{2}g(z)_{ij}\dot{z}^i\dot{z}^j, \tag{9.3}$$

where $\cdot$ is the scalar product in $\mathbb{R}^3$, and $g(z) = \mathrm{Diag}\{1, 1, r^2/4, r^2/2\}$ is the (diagonal) kinetic energy matrix. In the sequel we endow M with the metric g instead of the standard Euclidean one. Note that (9.2) defines a 2-dimensional distribution (of virtual displacements) $\mathcal{A} = \ker\{A_1, A_2\}$. Proceeding as in Section 5, we define the orthogonal distribution $\mathcal{A}^\perp$ by

$$z \mapsto \mathcal{A}_z^\perp := \mathrm{span}\{A^1(z), A^2(z)\} = \langle A^{\alpha i}(z)\frac{\partial}{\partial z^i}\rangle, \tag{9.4}$$

where $A^{\alpha i}(z) := g^{ij}(z)A_{\alpha j}(z)$, $\alpha = 1, 2$ and $g^{ij} = g_{ij}^{-1}$. In particular,

$$A^1 = (1, 0, 0, 2/r\cos\varphi), \qquad A^2 = (0, 1, 0, 2/r\sin\varphi).$$

By direct calculation, one proves the following proposition.

Proposition 9.1. *The vector fields $A^\alpha \in \mathcal{X}(M)$ are commuting, i.e.*

$$[A^\alpha, A^\beta] = 0, \quad \alpha, \beta = 1, 2.$$

As a consequence, the distribution $\mathcal{A}^\perp$ is obviously Frobenius integrable and the system $(M, g, L, \mathcal{A})$ at issue *verifies hypothesis* (H.2) of Section 6. Consider then the flux of the o.d.e. (6.2) for the present case. It is easy to see that (6.3) defines a group action Φ : $(\mathbb{R}^2, +) \times M \to M$, $(\tau, \hat{z}) \mapsto \Phi_\tau(\hat{z}) = z$, whose expression is

$$\begin{cases} x := \hat{x} + \tau^1, \\ y := \hat{y} + \tau^2, \\ \varphi := \hat{\varphi}, \\ \theta := \hat{\theta} + (2/r)(\tau^1\cos\hat{\varphi} + \tau^2\sin\hat{\varphi}). \end{cases} \tag{9.5}$$

The orbit of Φ through $\hat{z} = (\hat{x}, \hat{y}, \hat{\varphi}, \hat{\theta}) \in M$ is the 2-dimensional submanifold

$$\mathbb{R}^2\hat{z} = \{(x, y, \hat{\varphi}, \theta(x, y)) \colon (x, y) \in \mathbb{R}^2,$$
$$\theta(x, y) = \hat{\theta} + (2/r)[(x - \hat{x})\cos\hat{\varphi} + (y - \hat{y})\sin\hat{\varphi}]\} \tag{9.6}$$

which is diffeomorphic to $\mathbb{R}^2$. The orbits of Φ being all diffeomorphic, they define a partition of M. By a simple check, we have also the following proposition.

Proposition 9.2. *The smooth group action defined by* (9.5) *is free, transitive on the orbit, proper and, to every $z \in M$, $\mathbb{R}^2 z$ is an immersion of $\mathbb{R}^2$ in M.*

As a consequence, Φ *verifies hypotheses* (H.2) to (H.5) of Section 6. Then the partition of M defined by Φ is a *fibration*; moreover, since $M/\mathbb{R}^2 = \mathbb{R}^2 \times \mathbb{T}^2/\mathbb{R}^2 = \mathbb{T}^2, \pi : M \to M/G$ is a *trivial* principal bundle. We will use (φ, θ) as coordinates on the quotient space and (x, y) as coordinates on the fibres. By Proposition 6.1 (see (6.4)) $\mathcal{A}^\perp = \mathrm{span}\{A^1, A^2\}$ is the vertical subspace. On the other hand, hypothesis (H.6) of Section 6 fails to hold in this example, so the constraint distribution $\mathcal{A}$ cannot define a principal connection on the above bundle. Following the construction displayed in Section 5 (formulae (5.2)–(5.4)) we can put an Ehresmann connection on $\pi : M \to M/G$. After some calculations, one obtains

$$\alpha = \alpha(z) = \alpha(\varphi) = \frac{1}{3}\begin{pmatrix} 1+2\sin^2\varphi & -2\sin\varphi\cos\varphi & 0 & r\cos\varphi \\ -2\sin\varphi\cos\varphi & 1+2\cos^2\varphi & 0 & r\sin\varphi \end{pmatrix}. \tag{9.7}$$

The expression of the curvature 2-form (see (5.4)) is $\Omega = (\Omega^1, \Omega^2) : TM \times TM \to \mathcal{G}$

$$\Omega^1 = \Omega^1(\varphi) = \tfrac{2}{3}\sin\varphi(\cos\varphi\, \mathrm{d}\varphi \wedge \mathrm{d}x + \sin\varphi\, \mathrm{d}\varphi \wedge \mathrm{d}y + \tfrac{1}{2}r\, \mathrm{d}\varphi \wedge \mathrm{d}\theta) \tag{9.8}$$

$$\Omega^2 = \Omega^2(\varphi) = -\tfrac{2}{3}\cos\varphi(\cos\varphi\, \mathrm{d}\varphi \wedge \mathrm{d}x + \sin\varphi\, \mathrm{d}\varphi \wedge \mathrm{d}y + \tfrac{1}{2}r\, \mathrm{d}\varphi \wedge \mathrm{d}\theta).$$

Acknowledgements

The authors wish to thank C.-M. Marle and M. Spera for valuable remarks and suggestions on Sections 2 and 8 respectively. They are indebted to the referee for useful comments.

Appendix A. Poincaré duality

Let M be an n-dimensional oriented smooth manifold. Denote by $H^k(M)$ the k-dimensional space of the de Rham cohomology. The Poincaré duality, produced by the nondegenerate pairing

$$H^{n-k}(M) \times H_c^k(M) \ni ([\alpha],[\beta]) \mapsto \int_M \alpha \wedge \beta \in \mathbb{R},$$

c: compact support, states an isomorphism between

$$H^{n-k}(M) \simeq (H_c^k(M))^*,$$

see [Bott and Tu, 1982, p.44]. Let S be a k-dimensional oriented submanifold of M, $j : S \hookrightarrow M$. By Stokes's theorem, the map

$$H_c^k(M) \ni [\omega^{(k)}] \mapsto \int_S j^*\omega^{(k)} \in \mathbb{R}$$

is a linear functional on $H_c^k(M)$; hence, by the above duality, a cohomological class $[\eta_S] \in H^{n-k}(M)$ is associated to S, the so-called Poincaré dual of S:

$$\int_M \omega^{(k)} \wedge \eta_S = \int_S j^*\omega^{(k)}, \quad \forall \omega^{(k)} \in H_c^k(M).$$

Localization Principle [Bott and Tu, 1982, pp. 53, 67] agrees the Poincaré dual to have support as small as we like into any open tubular neighbourhood of S.

References

P. Appell, *Traité de mécanique rationelle, Tome II* (Gauthier-Villars, Paris, 1904).

V.I. Arnold, Mathematical methods of classical mechanics, in: *Graduate Texts in Math.*, Vol. 60 (Springer, Berlin, 1978).

V.I. Arnold, V.V. Kozlov and A.I. Nejshtadt, Mathematical aspects of classical and celestial mechanics, in: *Dynamical Systems III. Encyclopaedia of Mathematical Sciences* (Springer, Berlin, 1988).

S. Benenti, Geometrical aspects of the dynamics of nonholonomic systems, Journées Relativistes (1987) Univ. de Chambery, 14–16 Mai 1987.

S. Benenti, Symplectic relations in analytical mechanics, *Proc. IUTAMM-ISIMM Symposium on "Modern developments in Analytical Mechanics"* (Torino, 1982).

G. Benettin, L. Galgani and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, Part I, Comm. Math. Phys. 113 (1987) 87–103.

G. Benettin, L. Galgani and A. Giorgilli, Realization of holonomic constraints and freezing of high frequency degrees of freedom in the light of classical perturbation theory, Part II, Comm. Math. Phys. 121 (1989) 557–601.

R. Bott and L.W. Tu, *Differential Forms in Algebraic Topology* (Springer GTM 82, Berlin, New York, 1982).

N. Bourbaki, Variétées differentielles et analytiques, in: *Éléments de mathématiques XXXIII* (Hermann, Paris, 1958).

R. Capon, Hamilton's principle in relation to nonholonomic mechanical systems, Quart. J. Mech. Appl. Math. V(4) (1952) 472–480.

F. Cardin, Morse families and constrained mechanical systems. Generalized hyperelastic materials, Meccanica 26 (1991) 161–167.

N.G. Chetaev, On the Gauss principle, Izv. Physico–Mathematical Society of Kazan University s.3, 6 (1932, 1933) 62–67.

W.L. Chow, Systeme von linearen partiellen Differential–gleichungen erster ordnung, Math. Ann. 117 (1939) 98–105.

P. Dazord, Mécanique hamiltonienne en presence de contraintes, Illinois J. Math. 38(1) (1994) 148–175.

J. Dieudonné, *Éléments d'analyse, Tome III* (Gauthier-Villars, Paris, 1974).

C. Ehresmann, Les connexions infinitesimales dans un espace fibré differentiable, Coll. de Topologie Bruxelles CBRM (1950) 29–55.

C. Godbillon, Géometrie Différentielle et Mécanique Analytique, in: *Collection Méthodes* (Hermann, Paris, 1969).

V. Guillemin and A. Pollack, *Differential Topology* (Prentice-Hall, Englewood Cliffs, NJ, 1974).

V. Guillemin and S. Sternberg, *Symplectic Tecniques in Physics* (Cambridge University Press, Cambridge, 1984).

R. Hermann, On the differential geometry of foliations, Ann. of Math. 72(3) (1960) 445–457.

M. Hirsh, Differential topology, in: *Graduate Texts in Math.* (Springer, Berlin, 1988).

R. Jacobson, *Basic Algebra* (Freeman, New York, 1965).

H. Jeffreys, What is Hamilton principle?, Quart. J. Mech. Appl. Math. VII(3) (1954) 335–337.

S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. I (Interscience, New York, 1963).

J. Koiller, Reduction of some classical nonholonomic systems with simmetry, Arch. Rational Mech. Anal. 118 (1992) 113–148.

V.V. Kozlov, Dynamics of systems with nonintegrable constraints I, Moscow Univ. Mech. Bull. 37(3, 4) (1982) 27–34.

V.V. Kozlov, Dynamics of systems with nonintegrable constraints II, Moscow Univ. Mech. Bull. 37(4) (1982) 70–76.

V.V. Kozlov, Dynamics of systems with nonintegrable constraints III, Moscow Univ. Mech. Bull. 38(3) (1983) 102–111.

P. Libermann and C.M. Marle, *Symplectic Geometry and Analytical Mechanics* (Dreidel Publ. Company, Dordrecht, 1987).

A. Lichnerowicz, *Theorie globale des connexions et des groupes d'holonomie*, Ed. Cremone- se Roma (1954).

C.M. Marle, Géometrie des systèmes mécaniques à liason actives, in: *Symplectic Geometry and Mathematical Physics*, Actes du colloque en l'honneur de Jean–Marie Soriau, (Birkhauser, Boston, 1991).

J. Marsden, *Lectures on Mechanics* (Cambridge University Press, Cambridge, 1991).
J. Marsden, R. Montgomery and T. Ratiu, *Reduction, Symmetry and Phases in Mechanics* (Cambridge University Press, Cambridge, 1990).
E. Massa and E. Pagani, Classical dynamics of non-holonomic systems : A geometric approach, Ann. Inst. H. Poincaré Phys. Théor. 55(1) (1991) 511–544.
P. Molino *Riemannian Foliations* (Birkhauser, Boston, 1988).
M. Montgomery, Isoholonomic problems and some applications, Comm. Math. Phys. 128 (1990) 565–592.
L.A. Pars, Variation principles in dynamics, Quart. J. Mech. Appl. Math. VII(3) (1954) 338–351.
B. Reinhart, Foliated manifolds with bundle-like metrics, Ann. of Math, 69 (1959) 119–132.
V.V. Rumyantsev, Certain variational principles of mechanics, in: *Advances in Theoretical and Applied Mechanics* (MIR, Moscow, 1981).
H.J. Sussmann, Orbits of families of vector fields and integrability of distributions, Trans. Amer. Math. Soc. 180 (1973) 171–188.
J.L. Synge, Geodesics in nonholonomic geometry, Math. Ann. 99 (1928) 738–751.
V.V. Vagner, Differential geometry of nonholonomic manifolds, The VIII Internat. Competition for the N.I. Lobatschewski Prize, Report of The Physico-Mathematical Society of Kazan (1940).
A.M. Versihk, Classical and nonclassical dynamics with constraints, *Lecture Notes in Math.*, 1108 (1984) 278–301.
A.M. Versihk and L.D. Faddeev, Lagrangian mechanics in invariant form, Selecta Math. Soviet 4(1) (1981) 339–350.
A.M. Vershik and V.Ya. Gershkovich, Nonholonomic problems and geometry of distributions, Acta Appl. Math. 12 (1988) 181–209.
A.M. Vershik and V.Ya. Gershkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems, in: *Dynamical Systems VII. Encyclopaedia of Mathematical Sciences* (Springer, Berlin, 1994) pp. 4–79.
G. Vranceanu, Parallèlisme et courbure dans une variété non holonome, Atti del congresso Internaz. dei matematici Bologna 5 (1928) 63–75, Ed. Zanichelli Bologna.
E.T. Whittaker, *A Treatise on Analytical Dynamics of Particles and Rigid Bodies* (Dover, New York, 1944).