Acta Mech
https://doi.org/10.1007/s00707-024-04051-5

ORIGINAL PAPER

Federico Talamucci

Motion of general nonholonomic systems from the d'Alembert principle via an algebraic method

Received: 16 May 2024 / Revised: 18 July 2024 / Accepted: 23 July 2024

Abstract The aim of this study is to present an alternative way to deduce the equations of motion of general (i. e. also nonlinear) nonholonomic constrained systems starting from the d'Alembert principle and proceeding by an algebraic procedure. The two classical approaches in nonholonomic mechanics – Četaev method and vakonomic method – are treated on equal terms, avoiding integrations or other steps outside algebraic operations. In the second part of the work, we compare our results with the standard forms of the equations of motion associated with the two method and we discuss the role of the transpositional relation and of the commutation rule within the question of equivalence and compatibility of the Četaev and vakonomic methods for general nonholonomic systems.

1 Introduction

Our study concerns discrete mechanical systems subject to constraints involving the coordinates of the points and their velocities. We refer to this situation as nonholonomic systems even if the denomination, more generally, encompasses all situations complementary to the case of purely geometric restrictions, i. e. only on the coordinates of the points and possibly the time. Our approach is theoretical, that is, we deal with the mathematical formulation of the model and not with the physical realization of the constraints.

A fundamental point of reference for a comprehensive and historical review and for a systematic exposition of nonholonomic mechanics is [15]. For an update in the following decades and for an important analytical treatment of nonholonomic systems we quote [16]. The classical treatment of nonholonomic systems regards in the majority of cases linear constraints, i. e. the constraint equations are linear with respect to the kinetic variables. Among recent works that face the most general case of nonlinear kinematic constraints we refer to [2] for a comprehensive and general method for writing the dynamical equations through which various examples are analyzed. In [3] an example of a nonlinear system is presented, together with a review of the theory. A second series of works studying nonholonomic nonlinear constrained systems can be found in [24–26].

Our attention turns to constraints of a general type: one of the main objectives of the work is to explore the possibility of applying the d'Alembert principle to nonlinear nonholonomic systems, using elementary algebraic techniques. By "algebraic" we essentially intend to avoid the use of integrals with respect to time, a step very often recurring in most formulations. This possibility is known to exist for holonomic systems, constrained with only geometric restrictions. The idea to extend the d'Alembert principle to general nonholonomic systems requires only the definition of the virtual displacements (i. e. possible displacements) for which the virtual work of the constraint forces vanishes (ideal displacements).

In order to establish what types of displacements to consider, it is certain that we have to refer to the two main approaches prevailing in the theory of general constrained systems: we mean the displacements verifying

F. Talamucci (✉)
DIMAI Department of Mathematics and Informatics, University of Florence, Via di S. Marta 3, Florence 50139, Italy
E-mail: federico.talamucci@unifi.it

Published online: 07 August 2024

the Četaev condition, historically traceable to [4] on one side, and displacements complying the “vakonomic condition,” [8–10] (an interesting historical summary of the method is contained in [12]) for which we find in literature only integral (and not algebraic) approaches (this is indeed a motivation for the work).

The first part of the work (Sect. 2) is dedicated to recalling the D’Alembert principle in the most appropriate form and we develop simple linear algebra problems, coming from the declaration of two different classes of virtual displacements (class (A) and class (B)). The main result is stated in Proposition 1, where it is shown that the two systems of equations corresponding to the two different classes (A) and (B) are equivalent, if the displacements conditions verify a specific assumption (stated in (16)). We opt to start from assigning physical coordinates to the model and to develop the theory by using the radius vector of each point; then we show (Paragraph 2.4) that by passing to any arbitrary set of Lagrangian coordinates the formal structure of the problem and the results are identical.

In Sect. 3, the specific assignment (36) of the displacements conditions in terms of the constraint functions identifies the class (A) with the method based on the Četaev condition, class (B) with the formulation of the vakonomic mechanics. Proposition 3 which transfers the results of Sect. 2 to the mentioned application states that the two methods are equivalent if both the displacements conditions are assumed, as far as the equations of motion are obtained through algebraic considerations.

The final discussion in Sect. 4 compares the two set of nonholonomic equations (originating from (A) and (B)) with those present in literature and classified as Četaev systems or vakonomic systems. In particular, the vakonomic algebraic method is checked against the procedure of deducing the equations from a variational principle (Hamilton–Suslov principle). As a further issue, the possibility of making the conditions of the two kind valid simultaneously is explored. The question of equivalence of the two methods is indeed a current matter of debate and recently various works are dedicated to the study of the compatibility (or inequivalence) of the two methods [6,12,21–23].

A significant role in the question is played by the so called transpositional relation, which establishes a link between the two conditions on the displacements (Četaev and vakonomic types), in terms of the lagrangian derivatives of the constraint functions and of a quantity which vanishes if the commutation property between displacement and velocity is assumed (in simple terms: whether the velocity of the displacement is equal to the displacement of the velocity). Such a rule is another debated issue, starting from [15] up to the significant and exhaustive discussion in [5]. If on the one hand the validity of the commutation ule (tacitly or explicitly assumed in most of the vakonomic methods) entails the equivalence with the Četaev method only for holonomic systems [7,12,13], on the other hand the hypothesis of a non–zero coomutation rule offers new substance and significance to the vakonomic model [14,17]. The final observations of the work are developed precisely on the role of the transpositional relation assuming the simultaneous validity of the displacements conditions. Our comments are based mainly on the general vakonomic method presented in [14] and developed in [17].

2 Statement of the model

2.1 D’Alembert principle and virtual displacements

We consider a mechanical system consisting of N material points whose position is identified by the $\mathbb{R}^{3N}$ position vector $\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$, and subject to the nonholonomic constraints (linear or nonlinear)

$$\Psi_\nu(\mathbf{r}, \dot{\mathbf{r}}, t) = 0, \qquad \nu = 1, \dots, \kappa < 3N \tag{1}$$

where $\dot{\mathbf{r}} = (\dot{\mathbf{r}}_1, \dots, \dot{\mathbf{r}}_N)$ is the velocity vector of the system. The constraints are independent in the sense that

$$\text{the } \kappa \text{ vectors } \frac{\partial \Psi_\nu}{\partial \dot{\mathbf{r}}} \in \mathbb{R}^{3N}, \ \nu = 1, \dots, \kappa \text{ are linearly independent} \tag{2}$$

or, equivalently, the rank of the $(\kappa \times 3N)$ jacobian matrix $\dfrac{\partial \boldsymbol{\Psi}}{\partial \dot{\mathbf{r}}}$, $\boldsymbol{\Psi} = (\Psi_1, \dots, \Psi_\kappa)$ achieves its maximum value κ.

The Newtonian equations of motions for the systems can be written as

$$\dot{\mathbf{Q}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t) + \boldsymbol{R} \tag{3}$$

with $\mathbf{Q} = (m_1\dot{\mathbf{r}}_1, \ldots, m_N\dot{\mathbf{r}}_N)$ is the linear momentum of the system (m_i is the mass of each point), $\mathbf{F} = (\mathbf{F}_1, \ldots, \mathbf{F}_N)$ lists the forces acting on the points and $\mathbf{R} = (\mathbf{R}_1, \ldots, \mathbf{R}_N)$ are the constraint forces due to restrictions (1); the vector $\boldsymbol{R}$ is an unknown quantity of the problem.

The equations become operational if the mechanical behavior of the constraints is specified: following the d'Alembert Principle of virtual works, the constraints forces reveal the property

$$\boldsymbol{R} \cdot \delta\mathbf{r} = 0 \tag{4}$$

for the virtual displacements $\delta\mathbf{r} = (\delta\mathbf{r}_1, \ldots, \delta\mathbf{r}_N)$ of the system (i. e. the displacements performed at a fixed time anfd consistent with the constraints).

Remark 1 We follow the formulation given in [15] (as well as in recent articles, see [7,21]) where the variables which the vector $\boldsymbol{R}$ depends on are not explicitly indicated. Actually, one expects that the classical dependencies $\mathbf{r}$, $\dot{\mathbf{r}}$, t of the holonomic case are extended including $\ddot{\boldsymbol{r}}$ if the system is nonholonomic, but this may appear in conflict with the structure of the Newtonian equations. The discussion of this topic goes beyond the scope of the work and we refer to [15] for a broader overview; furthermore, by (46) the constraint forces are excluded.

Equations (3) together with conditions (4) provide

$$\left(\dot{\mathbf{Q}} - \mathbf{F}\right) \cdot \delta\mathbf{r} = 0 \tag{5}$$

from which the correct equations of motion will be deduced. The question therefore shifts to identifying the appropriate $\delta\mathbf{r}$: the selection must be compatible with the constraints (1) i. e. it must be expressed in terms of the functions Ψ_ν; if on the one hand in the case of holonomic constraints (that is in absence of $\dot{\mathbf{r}}$ or in case of integrable constraints) the answer is clear and unambiguous, in the case of general constraints we can say that the question is open, especially in the nonlinear case.

From a formal point of view for the moment, we indicate two general categories of displacements:

(A) the displacements verify

$$A(\mathbf{r}, \dot{\mathbf{r}}, t)\delta\mathbf{r} = \mathbf{0} \tag{6}$$

where A is a $\kappa \times 3N$–matrix with full rank κ.

(B) the displacements fulfill the condition

$$B(\mathbf{r}, \dot{\mathbf{r}}, t)\delta\mathbf{r} + C(\mathbf{r}, \dot{\mathbf{r}}, t)\delta\dot{\mathbf{r}} = \mathbf{0} \tag{7}$$

where B, C are $\kappa \times 3N$–matrices and $\delta\dot{\mathbf{r}}$ are the virtual variations of $\dot{\mathbf{r}}$ consistent with (1).

2.2 The mathematical problem

For the moment we are dealing with an abstract situation: the model must be completed by linking the matrices A, B anc C to the constraint functions (1). The problem is simply posed in these terms: once the position and the velocity of the system are fixed (by $\mathbf{r}$ and $\dot{\mathbf{r}}$), which displacements are compatible with the conditions (6) or (7)?

Case (A) can be expanded by a simple argument of linear algebra: for each fixed $\mathbf{r}$, $\dot{\mathbf{r}}$ and t the solution of (6) is the totality $\mathcal{W}$ of $\delta\mathbf{r}$ orthogonal to the row vectors $\mathbf{A}_\nu \in \mathbb{R}^{3N}$ of the matrix A, $\nu = 1, \ldots, \kappa$:

$$\delta\mathbf{r} \in \mathcal{W} = \langle \mathbf{A}_1, \ldots, \mathbf{A}_\kappa \rangle^\perp$$

Since the vectors are linearly independent, $\mathcal{W}$ is a vector space of dimension $n - \kappa$. Owing to (4), the constraint force $\mathbf{R}$ is orthogonal to all the vectors of $\mathcal{W}$, hence it must be $\mathbf{R} \in \mathcal{W}^\perp = \langle \mathbf{A}_1, \ldots, \mathbf{A}_\kappa \rangle$, namely $\mathbf{R} = \sum_{\nu=1}^{\kappa} \lambda_\nu \mathbf{A}_\nu$ for some coefficients. We conclude that Eq. (3) joined to (4), where the displacements have to verify (6), are equivalent to

$$\dot{\mathbf{Q}} = \mathbf{F} + \sum_{\nu=1}^{\kappa} \lambda_\nu \mathbf{A}_\nu \qquad case\ (A) \tag{8}$$

where the multipliers λ_ν are unknown quantities. The essential assumption is the full rank of A. Eq. (8) are coupled with (1) in order to form a system of $3N + \kappa$ equations in the unknowns $\mathbf{r}(t) \in \mathbb{R}^{3N}$ and $\lambda_1, \ldots, \lambda_\kappa$.

The essential final step (missing for the moment) is to establish the class (6) according to the restrictions (1), namely to set A as a function of Ψ_ν and their derivatives.

Let us now examine case (B): assuming that the rank of B is full, we see that (5) and (7) entail

$$\dot{\mathbf{Q}} - \mathbf{F} = \sum_{\nu=1}^{\kappa} \mu_\nu \mathbf{B}_\nu + \mathbf{y} \tag{9}$$

where the vector $\mathbf{y}$ is subject to the request

$$\mathbf{y} \cdot \delta\mathbf{r} = \sum_{\nu=1}^{\kappa} \mu_\nu \mathbf{C}_\nu \cdot \delta\dot{\mathbf{r}} \tag{10}$$

($\mathbf{B}_\nu, \mathbf{C}_\nu$ are the column of B, C and μ_ν are unknown mutliplying factors). We transform the previous expression by considering that

$$\mu_\nu \mathbf{C}_\nu \cdot \delta\dot{\mathbf{r}} = \frac{d}{dt}(\mu_\nu \mathbf{C}_\nu \cdot \delta\mathbf{r}) - \frac{d}{dt}(\mu_\nu \mathbf{C}_\nu) \cdot \delta\mathbf{r} + \mu_\nu \mathbf{C}_\nu \cdot \left(\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r}\right)$$

The presence of the last term is due to the uncertainty of the rule (we will deal with this question afterward)

$$\frac{d}{dt}\delta\mathbf{r} = \delta\dot{\mathbf{r}} \tag{11}$$

which would express the commutation of the operations $\frac{d}{dt}$ and δ (obviously $\dot{\mathbf{r}} = \frac{d}{dt}\mathbf{r}$).

Let us assume that it is possible to write (the hypothesis will be discussed later)

$$\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r} = W(\mathbf{r}, \dot{\mathbf{r}}, \ddot{r}, t)\delta\mathbf{r} \tag{12}$$

where W is a square matrix of order $3N$. The specified dependencies $\mathbf{r}, \dot{\mathbf{r}}, \ddot{r}, t$ are quite natural, since the variation $\delta\dot{\mathbf{r}}$ of possible velocities involves the accelerations of the system.

If (12) is the case, (10) reduces to

$$\mathbf{y} \cdot \delta\mathbf{r} = \sum_{\nu=1}^{\kappa} \left(-\frac{d}{dt}(\mu_\nu \mathbf{C}_\nu) + \mu_\nu W^T \mathbf{C}_\nu\right) \cdot \delta\mathbf{r} + \frac{d}{dt}\left(\sum_{\nu=1}^{\kappa} \mu_\nu \mathbf{C}_\nu \cdot \delta\mathbf{r}\right)$$

so that (9) can be written as

$$\begin{aligned} \dot{\mathbf{Q}} - \mathbf{F} &= \sum_{\nu=1}^{\kappa} \left(\mu_\nu \mathbf{B}_\nu - \frac{d}{dt}(\mu_\nu \mathbf{C}_\nu) + \mu_\nu W^T \mathbf{C}_\nu\right) + \mathbf{y}_1 \\ &= \sum_{\nu=1}^{\kappa} \mu_\nu \left(\mathbf{B}_\nu - \frac{d}{dt}(\mathbf{C}_\nu) + W^T \mathbf{C}_\nu\right) - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \mathbf{C}_\nu + \mathbf{y}_1 \qquad case\ (B) \end{aligned} \tag{13}$$

where the vector $\mathbf{y}_1$ must verify

$$\mathbf{y}_1 \cdot \delta\mathbf{r} = \frac{d}{dt}\left(\sum_{\nu=1}^{\kappa} \mu_\nu \mathbf{C}_\nu \cdot \delta\mathbf{r}\right). \tag{14}$$

For $C = \mathbb{O}$ (null square matrix of order N) the cases (A) and (B) are the same with $A = B$ and the corresponding Eqs. (8), (13) do coincide. As in the previous case, it is necessary to link the elements of B, C to the functions Ψ_ν and their derivatives. Regarding (14), we could say that the condition does not precisely define $\mathbf{y}_1$ therefore the equations are not closed; actually, the not explicit form of $\mathbf{y}_1$ in (13) except through $\delta\mathbf{r}$ makes the equations unusable unless other considerations are added. However, we are interested in the case when both conditions (6) and (7) hold and this allows us to eliminate $\mathbf{y}_1$, as we will see: basically, the physically interesting situation that we will consider corresponds to $A = C$, which allows us to suppress $\mathbf{y}_1$.

We finally establish a relation between the two expressions $A\delta\mathbf{r}$ and $B\delta\mathbf{r} + C\delta\dot{\mathbf{r}}$ (appearing in (6) and (7)) written in a way so that the difference appearing in (12):

$$B\delta\mathbf{r} + C\delta\dot{\mathbf{r}} - \frac{d}{dt}(A\delta\mathbf{r}) = A\left(\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r}\right) + \left(B - \frac{dA}{dt}\right)\delta\mathbf{r} + (C - A)\delta\dot{\mathbf{r}} \tag{15}$$

The check of (15) is immediate and we refer to it as the *transpositional relation*. In examining and employing (15) one does not necessarily have to assume that the two conditions (6) and (7) hold simultaneously (that is both the expressions vanish); rather, the right terms in the equality (15) allow us to investigate about the compatibility and the properties of the two classes of displacements (6) and (7).

2.3 The case $A = C$

Let us consider the special case

$$A = C \tag{16}$$

The just written condition seems artificial for the moment, but when the coefficients are specified in the physical context, a significant situation will be recognized.

In the next two Propositions, we assume that both conditions for displacements (6) and (7) hold and we check the mathematical implications. The sense of the investigation lies in the question (discussed later on) whether the two sets of Eqs. (8) and (13) are equivalent.

Proposition 1 *Assume that the displacements verify both* (6) *and* (7)*. If* (16) *holds, then the vector in* (14) *can be taken as*

$$\mathbf{y}_1 = \mathbf{0} \tag{17}$$

and Eqs. (8) *and* (13) *are equivalent.*

Proof The null vector (17) satisfies the requirement (14), since (6) holds with $C = A$. Moreover, if (6) and (7) are both in effective, relation (15) reduces to

$$0 = \mathbf{A}_\nu \cdot \left(\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r}\right) + \left(\mathbf{B}_\nu - \frac{d}{dt}\mathbf{A}_\nu\right) \cdot \delta\mathbf{r} \qquad \nu = 1, \ldots, \kappa \tag{18}$$

or equivalently, by (12) and (16),

$$\mathbf{C}_\nu \cdot W\delta\mathbf{r} + \left(\mathbf{B}_\nu - \frac{d}{dt}\mathbf{C}_\nu\right) \cdot \delta\mathbf{r} = \left(\mathbf{B}_\nu - \frac{d}{dt}\mathbf{C}_\nu + W^T\mathbf{C}_\nu\right) \cdot \delta\mathbf{r} = 0 \qquad \nu = 1, \ldots, \kappa.$$

Since the vectors $\mathbf{A}_\nu$, $\nu = 1, \ldots, \kappa$, are linearly independent and the set of displacements $\delta\mathbf{r}$ coincides with the orthogonal complement of the space generated by same vectors, one can write

$$\mathbf{B}_\nu - \frac{d}{dt}\mathbf{C}_\nu + W^T\mathbf{C}_\nu = \sum_{\sigma=1}^{\kappa} \rho_\sigma \mathbf{A}_\sigma$$

for some κ–uple $(\rho_1, \ldots, \rho_\kappa)$. We conclude that, once again by virtue of (16), Eq. (13) (second version) can be written as

$$\dot{\mathbf{Q}} - \mathbf{F} = \sum_{\nu,\sigma=1}^{\kappa} (\mu_\nu \rho_\sigma - \dot{\mu}_\sigma)\, \mathbf{A}_\sigma$$

which are equivalent to (8), simply with a different role of the multipliers. □

Proposition 2 *Again in the case* (16) *assume that* (6) *and* (7) *are valid simultaneously. Then*

$$\mathbf{C}_\nu \cdot \left(\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r}\right) = \mathbf{0} \tag{19}$$

if and only if

$$\mathbf{B}_\nu - \frac{d}{dt}\mathbf{A}_\nu = \sum_{\sigma=1}^{\kappa} \varrho_\sigma \mathbf{A}_\sigma, \qquad \nu = 1, \ldots, \kappa \tag{20}$$

for some real coefficients $\varrho_1, \ldots, \varrho_\kappa$.

Proof If (19) is true, also $(\mathbf{B}_\nu - \frac{d}{dt}\mathbf{A}_\nu) \cdot \delta\mathbf{r} = 0$ hence (20) must be verified for some coefficients $\varrho_1, \ldots, \varrho_\kappa$. Conversely, if (20) holds, then

$$\left(\mathbf{B}_\nu - \frac{d}{dt}\mathbf{A}_\nu\right) \cdot \delta\mathbf{r} = \sum_{\sigma=1}^{\kappa} \varrho_\sigma \mathbf{A}_\sigma \cdot \delta\mathbf{r} = 0$$

by virtue of (6) and (16). We conclude that (19) is true, due to (18). □

Corollary 1 *A necessary condition for the commutation rule* (11) *is the equality* (20).

Indeed, in that case (19) holds, which implies (20). Notice that the inverse statement is not true, that is if (20) is valid, the rule (11) does not necessarily have to hold.

2.4 Lagrangian coordinates

In terms of generalized coordinates $\mathbf{q} = (q_1, \ldots, q_n)$ and generalized velocities $\dot{\mathbf{q}} = (\dot{q}_1, \ldots, q_n)$, $n = 3N$ one has

$$\mathbf{r} = \mathbf{r}(\mathbf{q}, t), \qquad \dot{\mathbf{r}}(\mathbf{q}, \dot{\mathbf{q}}, t) = \sum_{j=1}^{n} \frac{\partial \mathbf{r}}{\partial q_j}\dot{q}_j + \frac{\partial \mathbf{r}}{\partial q_j} \tag{21}$$

where the jacobian matrix $\dfrac{\partial \mathbf{r}}{\partial \mathbf{q}} = \begin{pmatrix} \frac{\partial \mathbf{r}_1}{\partial q_1} & \ldots & \frac{\partial \mathbf{r}_1}{\partial q_n} \\ \ldots & \ldots & \ldots \\ \frac{\partial \mathbf{r}_N}{\partial q_1} & \ldots & \frac{\partial \mathbf{r}_N}{\partial q_n} \end{pmatrix}$ has maximum rank n.

Remark 2 When passing from the coordinates $\mathbf{r}$ to the generalized coordinates $\mathbf{q}$ we can assume either that no further geometric constraint occurs, i. e. $n = 3N$, or that additional holonomic conditions are juxtaposed: in the latter case, the selection of a smaller number of generalized coordinates $n < 3N$ does not lead to substantial changes in the reformulation of the problem (5).

The displacements are expressed in terms of lagrangian variables as

$$\delta\mathbf{r} = \sum_{j=1}^{n} \frac{\partial \mathbf{r}}{\partial q_j}\delta q_j \tag{22}$$

$$\delta\dot{\mathbf{r}} = \sum_{j=1}^{n} \frac{\partial \dot{\mathbf{r}}}{\partial q_j}\delta q_j + \sum_{j=1}^{n} \frac{\partial \dot{\mathbf{r}}}{\partial \dot{q}_j}\delta \dot{q}_j = \sum_{j=1}^{n} \frac{\partial \dot{\mathbf{r}}}{\partial q_j}\delta q_j + \sum_{j=1}^{n} \frac{\partial \mathbf{r}}{\partial q_j}\delta \dot{q}_j \tag{23}$$

and the problem (5) joined with (6) or with (7) is reformulated in the following way:

$$(A) \quad \begin{cases} \sum\limits_{j=1}^{n} (\dot{Q}^{(j)} - F^{(j)})\delta q_j = 0 \\ \sum\limits_{j=1}^{n} \alpha_{\nu,j}\delta q_j = 0 \end{cases} \quad or \ (B) \quad \begin{cases} \sum\limits_{j=1}^{n} (\dot{Q}^{(j)} - F^{(j)})\delta q_j = 0 \\ \sum\limits_{j=1}^{n} (\beta_{\nu,j}\delta q_j + \gamma_{\nu,j}\delta\dot{q}_j) = 0 \end{cases} \tag{24}$$

where $\nu = 1, \ldots, \kappa$ and

- $v^{(j)} = \mathbf{v} \cdot \dfrac{\partial \mathbf{r}}{\partial q_j}$, $j = 1, \dots, n$, indicates the j–lagrangian component of a $3N$–vector $\mathbf{v}$; in particular, the relation $\dot{\mathbf{Q}}^{(j)} = \dfrac{d}{dt}\dfrac{\partial T}{\partial \dot{q}_j} - \dfrac{\partial T}{\partial q_j}$ is well known for each $j = 1, \dots, n$ and that the presence of active forces coming from a generalized potential, that is

$$F_P^{(j)} = \frac{\partial U}{\partial q_j} - \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_j}, \qquad j = 1, \dots, n$$

for some function $U(\mathbf{q}, \dot{\mathbf{q}}, t)$ makes us write the upper set of equations in (A) or in (B) as

$$\sum_{j=1}^{n} \left(\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} - F_{NP}^{(j)} \right) \delta q_j = 0 \tag{25}$$

where $\mathcal{L}(\mathbf{q}, \dot{\mathbf{q}}, t) = T + U$ is the Lagrangian function and the term $F_{NP}^{(j)}$ takes into account the remaining active forces not deriving from a potential;
- the coefficients are related to those appearing in (6) and (7) by means of

$$\alpha_{\nu,j} = \mathbf{A}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j}, \ \beta_{\nu,j} = \mathbf{B}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} + \mathbf{C}_\nu \cdot \frac{\partial \dot{\mathbf{r}}}{\partial q_j}, \ \gamma_{\nu,j} = \mathbf{C}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} \tag{26}$$

and they depend on the variables $(\mathbf{q}, \dot{\mathbf{q}}, t)$ by virtue of the replacements (21).

In (25), we recognize the ordinary way to formulate the D'Alembert principle in generalized coordinates version.

Equations (8) (case (A)) converted to lagrangian variables are immediate:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} - F_{NP}^{(j)} = \sum_{\nu=1}^{\kappa} \lambda_\nu \alpha_{\nu,j} \quad j = 1, \dots, n \qquad case\ (A) \tag{27}$$

Equations (27) can be achieved (i) either using the same technique of linear spaces as performed in the previous analysis, (ii) or by calculating the Lagrangian components (by means of the scalar product $\cdot \frac{\partial \mathbf{r}}{\partial q_j}$, $j = 1, \dots, n$) of (8) and (13). In regard to (i), we see that the matrix $(\alpha_{\nu,j})$ has maximum rank κ, since $(\alpha_{\nu,j}) = A(J_{\mathbf{r}}\mathbf{q})$ (see (6)), hence the formal procedure is identical: the equations (A) in (24) imply that it must be $\dot{Q}^{(j)} - F^{(j)} = \sum_{\nu=1}^{\kappa} \lambda_\nu \alpha_{\nu,j}$ for suitable coefficients, hence (27).

Concerning case (B), we start by stating the following

Property 1 *The commutation property* $\delta \dot{\mathbf{r}} = \dfrac{d}{dt}\delta \mathbf{r}$ *holds if and only if*

$$\frac{d}{dt}(\delta q_j) = \delta \left(\frac{d}{dt} q_j \right) \tag{28}$$

holds for any set of independent lagrangian coordinates $\mathbf{q} = (q_!, \dots, q_n)$.

Proof We see that

$$\begin{aligned} \delta \dot{\mathbf{r}} - \frac{d}{dt}\delta \mathbf{r} &= \sum_{j=1}^{n} \left(\cancel{\frac{\partial \dot{\mathbf{r}}}{\partial q_j}\delta q_j} + \frac{\partial \mathbf{r}}{\partial q_j}\delta \dot{q}_j \right) - \sum_{j=1}^{n} \left(\frac{\partial \mathbf{r}}{\partial q_j}\frac{d}{dt}(\delta q_j) + \cancel{\frac{\partial \dot{\mathbf{r}}}{\partial q_j}\delta q_j} \right) \\ &= \sum_{j=1}^{n} \frac{\partial \mathbf{r}}{\partial q_j} \left(\delta \dot{q}_j - \frac{d}{dt}\delta q_j \right) \end{aligned} \tag{29}$$

Since the vectors $\frac{\partial \mathbf{r}}{\partial q_j}$, $j = 1, \dots, n$ are independent, the commutation (11) holds if and only if (28) holds. □

Assumption (12) can be assigned in terms of lagrangian coordinates as

$$\delta \dot{q}_j - \frac{d}{dt}(\delta q_j) = \sum_{h=1}^{n} \omega_{j,h} \delta q_h. \tag{30}$$

Owing to (29), the terms $\omega_{j,h}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{q}, t)$, $j, h = 1, \ldots, n$, are related to the entries $w_{r,s}$, $r, s = 1, \ldots, 3N$ of W in (12) by means of

$$\sum_{s=1}^{3N} w_{r,s} \frac{\partial \xi_s}{\partial q_j} = \sum_{h=1}^{n} \frac{\partial \xi_r}{\partial q_h} \omega_{h,j} \quad r = 1, \ldots, 3n, \quad j = 1, \ldots, n, \quad (\xi_1, \ldots, \xi_{3N}) = \mathbf{r}$$

that is, in terms of W and the matrix $\Omega = (\omega_{j,h})_{j,h=1,\ldots,n}$:

$$W \frac{\partial \mathbf{r}}{\partial \mathbf{q}} = \frac{\partial \mathbf{r}}{\partial \mathbf{q}} \Omega, \qquad \left(\frac{\partial \mathbf{r}}{\partial \mathbf{q}} \right)_{(r,j)} = \frac{\partial \xi_r}{\partial q_j}, \quad r = 1, \ldots, 3N, \quad j = 1, \ldots, n. \tag{31}$$

Roughly speaking, the physical role of W (hence of Ω) is to compensate for the lack of interchange of the two operations δ and d/dt: the opinion that the presence of nonholonomic constraints makes the operations d (time derivative) and δ (variation) non–commutative [5] is more and more common and we refer to [14] for a general overview on that issue; in [17] an extended discussion on how to determine the matrix Ω is performed.

Equations (13) in terms of the coefficients (26) are

$$\begin{aligned} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}_j} - \frac{\partial \mathcal{L}}{\partial q_j} - F_{NP}^{(j)} &= \sum_{\nu=1}^{\kappa} \left(\mu_\nu \beta_{\nu,j} - \frac{d}{dt}(\mu_\nu \gamma_{\nu,j}) + \mu_\nu \sum_{h=1}^{n} \gamma_{\nu,h} \omega_{h,j} \right) + y_j \\ &= \sum_{\nu=1}^{\kappa} \mu_\nu \left(\beta_{\nu,j} - \frac{d}{dt}(\gamma_{\nu,j}) + \sum_{h=1}^{n} \gamma_{\nu,h} \omega_{h,j} \right) - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \gamma_{\nu,j} + y_j \\ & j = 1, \ldots, n \qquad case\ (B) \end{aligned} \tag{32}$$

The terms containing $\omega_{h,j}$ come from (see (31))

$$\begin{aligned} W^T \mathbf{C}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} &= \mathbf{C}_\nu \cdot W \frac{\partial \mathbf{r}}{\partial q_j} = \mathbf{C}_\nu \cdot \frac{\partial \mathbf{r}}{\partial \mathbf{q}} \begin{pmatrix} \omega_{1,j} \\ \cdots \\ \omega_{n,j} \end{pmatrix} = \left(\frac{\partial \mathbf{r}}{\partial \mathbf{q}} \right)^T \mathbf{C}_\nu \cdot \begin{pmatrix} \omega_{1,j} \\ \cdots \\ \omega_{n,j} \end{pmatrix} \\ &= \begin{pmatrix} \gamma_{\nu,1} \\ \cdots \\ \gamma_{\nu,n} \end{pmatrix} \cdot \begin{pmatrix} \omega_{1,j} \\ \cdots \\ \omega_{n,j} \end{pmatrix} \end{aligned}$$

and $y_j = \mathbf{y}_1 \cdot \dfrac{\partial \mathbf{r}}{\partial q_j}$ (see (14)) is such that

$$y_j \delta q_j = \frac{d}{dt} \left(\sum_{\nu=1}^{\kappa} \gamma_{\nu,j} \delta q_j \right), \qquad j = 1, \ldots, n. \tag{33}$$

It should be noticed that despite the non–complete formal adherence between the Greek and Latin functions in (26) (in fact in $\beta_{\nu,j}$ there is an additional $\mathbf{C}_\nu$), the formal correspondence between the Newtonian equations (13) and the lagrangian ones (32) is respected, in the sense that "greek" and "latin" terms are present in the same role. This is true by virtue of the cancelation which occurs calculating the following lagrangian components in (13):

$$\begin{aligned} \mathbf{B}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} - \frac{d}{dt}(\mathbf{C}_\nu) \cdot \frac{\partial \mathbf{r}}{\partial q_j} &= \left(\beta_{\nu,j} - \cancel{\mathbf{C}_\nu \cdot \frac{\partial \dot{\mathbf{r}}}{\partial q_j}} \right) - \left(\frac{d}{dt} \left(\mathbf{C}_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} \right) - \cancel{\mathbf{C}_\nu \cdot \frac{\partial \dot{\mathbf{r}}}{\partial q_j}} \right) \\ &= \beta_{\nu,j} - \frac{d}{dt} \gamma_{\nu,j}. \end{aligned}$$

The same formal analogy is present also if we write the transpositional relation (15) using lagrangian coordinates, which is

$$\sum_{j=1}^{n}\left(\beta_{\nu,j}\delta q_j+\gamma_{\nu,j}\delta\dot{q}_j\right)-\frac{d}{dt}\left(\sum_{j=1}^{n}\alpha_{\nu,j}\delta q_j\right)$$
$$=\sum_{j=1}^{n}\alpha_{\nu,j}\left(\delta\dot{q}_j-\frac{d}{dt}\delta q_j\right)+\sum_{j=1}^{n}\left(\beta_{\nu,j}-\frac{d\alpha_{\nu,j}}{dt}\right)\delta q_j+\sum_{j=1}^{n}\left(\gamma_{\nu,j}-\alpha_{\nu,j}\right)\delta\dot{q}_j \tag{34}$$

The relation can be obtained either rearranging directly the terms in the left–hand side, or replacing (22), (23), (26) in (15); even in this case, cancelations and transfers of terms make the formal structure of (34) in agreement with (15).

Finally, we see that the case $A=C$ (see (16)) opportunely overlaps the following relations:

$$\alpha_{\nu,j}=\gamma_{\nu,j},\qquad \nu=1,\ldots,\kappa,\quad j=1,\ldots,n \tag{35}$$

and the effects are the same as those outlined in Paragraph. In particular:

- the terms y_j for each $j=1,\ldots,n$ in (32) can be eliminated,
- the last summation in (34) vanishes.

3 A significant application

The correlation between the matrices A, B, C and the constraint functions (1) is now specified as follows:

$$\mathbf{A}_\nu(\mathbf{r},\dot{\mathbf{r}},t)=\mathbf{C}_\nu(\mathbf{r},\dot{\mathbf{r}},t)=\nabla_{\dot{\mathbf{r}}}\Psi_\nu,\qquad \mathbf{B}_\nu(\mathbf{r},\dot{\mathbf{r}},t)=\nabla_{\mathbf{r}}\Psi_\nu \tag{36}$$

where $\nabla_{\mathbf{r}}=(\frac{\partial}{\partial\xi_1},\ldots,\frac{\partial}{\partial\xi_{3N}})$, setting $\mathbf{r}=(\xi_1,\ldots,\xi_{3N})$. It is evident that the assignment (36) satisfies the requirement (16).

We indicate by $\delta^{(v)}$ and $\delta^{(c)}$ respectively the operations appearing to (6) and (7):

$$\delta^{(c)}\Psi_\nu=\nabla_{\dot{\mathbf{r}}}\Psi_\nu\cdot\delta\mathbf{r},\qquad \delta^{(v)}\Psi_\nu=\nabla_{\mathbf{r}}\Psi_\nu\cdot\delta\mathbf{r}+\nabla_{\dot{\mathbf{r}}}\Psi_\nu\cdot\delta\dot{\mathbf{r}} \tag{37}$$

Definitions (37) discover the virtual displacement condition according to the Četaev theory (first equation) and to the vakonomic model (second one).

As it is known, the debate about the appropriate choice in (37) for displacements and their consistency with models is a crucial aspect in the theory of nonholonomic systems. Although the two types of displacements (37) appear in contradition, a part of literature is dedicated to the study of their compatibility or inequivalency [5,12,21–23]. A key role is played by the possible interchangeability of the operators δ (variation) and d/dt (time derivative): if one assumes that the commutative rule (11) is valid, then the equivalency of the two classes of displacements is claimed to exist only for holonomic systems [12]. Hence, the attempt to reconcile the two methods must necessarily pass through a non–commutative statement like (12) and the transpositional relation is the main reference formula, as we will see further on.

The equations of motion (8) and (13) are now

$$\dot{\mathbf{Q}}=\mathbf{F}+\sum_{\nu=1}^{\kappa}\lambda_\nu\nabla_{\dot{\mathbf{r}}}\Psi_\nu \tag{38}$$

$$\dot{\mathbf{Q}}-\mathbf{F}=\sum_{\nu=1}^{\kappa}\left(\mu_\nu\nabla_{\mathbf{r}}\Psi_\nu-\frac{d}{dt}(\mu_\nu\nabla_{\dot{\mathbf{r}}}\Psi_\nu)+\mu_\nu W^T\nabla_{\dot{\mathbf{r}}}\Psi_\nu\right)$$
$$=\sum_{\nu=1}^{\kappa}\mu_\nu\left(\nabla_{\mathbf{r}}\Psi_\nu-\frac{d}{dt}(\nabla_{\dot{\mathbf{r}}}\Psi_\nu)+W^T\nabla_{\dot{\mathbf{r}}}\Psi_\nu\right)-\sum_{\nu=1}^{\kappa}\dot{\mu}_\nu\nabla_{\dot{\mathbf{r}}}\Psi_\nu+\mathbf{y}_1 \tag{39}$$

where the vector $\mathbf{y}_1$ is such that (see (14))

$$\mathbf{y}_1 \cdot \delta \mathbf{r} = \frac{d}{dt}\left(\sum_{\nu=1}^{\kappa} \mu_\nu \nabla_{\dot{\mathbf{r}}} \Psi_\nu \cdot \delta \mathbf{r} \right).$$

The transpositional relation (15) takes the form

$$\delta^{(v)}\Psi_\nu - \frac{d}{dt}\left(\delta^{(c)}\Psi_\nu\right) = \nabla_{\dot{\mathbf{r}}}\Psi_\nu \cdot \left(\delta\dot{\mathbf{r}} - \frac{d}{dt}\delta\mathbf{r}\right) - \mathcal{D}^{(\mathbf{r})}\Psi_\nu \cdot \delta\mathbf{r} \tag{40}$$

where

$$\mathcal{D}^{(\mathbf{r})} F = \frac{d}{dt}\left(\nabla_{\dot{\mathbf{r}}} F\right) - \nabla_{\mathbf{r}} F \tag{41}$$

is the lagrangian derivative of a function $F(\mathbf{r}, \dot{\mathbf{r}}, t)$.

A clear difference between the two equations of motion (38) and (39) is the presence of the time derivative $\dot{\mu}_\nu$ of the Lagrange multipliers in the second set of equations, whereas in (38) the coefficients λ_ν appear only in entire form. This presupposes that it is necessary to prescribe the initial conditions $\mu_\nu(0)$, $\nu = 1, \dots, \kappa$ in order to solve the problem, introducing in this way an unclear element to disentangle, whenever Eq. (39) are used. The need of assigning initial conditions for the Lagrange multipliers is indeed an inconvenience related to Vakonomic equations Dealing with this issue is outside our aim and we refer to [12], [17] for some discussion.

3.1 Displacements and constrained systems

Let us now place the constraints (1) into the lagrangian formalism: we consider

$$g_\nu(\mathbf{q}, \dot{\mathbf{q}}, t) = \Psi_\nu(\mathbf{r}(\mathbf{q}, \mathbf{t}), \dot{\mathbf{r}}(\mathbf{q}, \dot{\mathbf{q}}, t), t) = 0, \qquad \nu = 1, \dots, \kappa \tag{42}$$

which correspond to (1) rewritten via (21). The jacobian matrix $\dfrac{\partial \mathbf{g}}{\partial \dot{\mathbf{q}}} = \dfrac{\partial \Psi}{\partial \dot{\mathbf{r}}}\dfrac{\partial \dot{\mathbf{r}}}{\partial \dot{\mathbf{q}}} = \dfrac{\partial \Psi}{\partial \dot{\mathbf{r}}}\dfrac{\partial \mathbf{r}}{\partial \mathbf{q}}$, where $\mathbf{g} = (g_1, \dots, g_k)$, has full rank κ, by virtue of (2) and the non–singularity of $\partial \mathbf{r}/\partial \mathbf{q}$, since the generalized coordinates $\mathbf{q}$ are independent.

The choice (36) corresponds to

$$\alpha_{\nu,j} = \gamma_{\nu,j} = \frac{\partial g_\nu}{\partial \dot{q}_j}, \quad \beta_{\nu,j} = \frac{\partial g_\nu}{\partial q_j}, \qquad \nu = 1, \dots, \kappa, \quad j = 1, \dots, n \tag{43}$$

and that the displacements (37) take the form, for each $\nu = 1, \dots, \kappa$:

$$\delta^{(c)}\Psi_\nu = \delta^{(c)} g_\nu = \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\delta q_j, \quad \delta^{(v)}\Psi_\nu = \delta^{(v)} g_\nu = \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial q_j}\delta q_j + \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\delta \dot{q}_j. \tag{44}$$

In order to verify (43) and (44) it suffices to take into account (22), (23) and the relations

$$\begin{cases} \dfrac{\partial g_\nu}{\partial q_j} = \displaystyle\sum_{i=1}^{n} \frac{\partial \Psi_\nu}{\partial \xi_i}\frac{\partial \xi_i}{\partial q_j} + \sum_{i=1}^{n} \frac{\partial \Psi_\nu}{\partial \dot{\xi}_i}\frac{\partial \dot{\xi}_i}{\partial q_j} = \nabla_{\mathbf{r}}\Psi_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j} + \nabla_{\dot{\mathbf{r}}}\Psi_\nu \cdot \frac{\partial \dot{\mathbf{r}}}{\partial q_j} \\ \dfrac{\partial g_\nu}{\partial \dot{q}_j} = \displaystyle\sum_{i=1}^{n} \frac{\partial \Psi_\nu}{\partial \dot{\xi}_i}\frac{\partial \dot{\xi}_i}{\partial \dot{q}_j} = \sum_{i=1}^{n} \frac{\partial \Psi_\nu}{\partial \dot{\xi}_i}\frac{\partial \xi_i}{\partial q_j} = \nabla_{\dot{\mathbf{r}}}\Psi_\nu \cdot \frac{\partial \mathbf{r}}{\partial q_j}, \end{cases} \tag{45}$$

where we indexed $\mathbf{r} = (\xi_1, \dots, \xi_{3N})$.

Remark 3 In defiance of the check, the conclusion could only be (44) if we want the operations $\delta^{(c)}$ and $\delta^{(v)}$ to have a meaning independent of the choice of variables. That is, if we had started with generic Lagrangian coordinates $\mathbf{q}$, the virtual displacements conditions have necessarily the form (44).

At this stage, we can summarize the path through the equations and definitions (1), (5), (6), (7), (24), (27), (32), (36), (38), (39), (42), (43), (44) by presenting the following scheme:

$$\sum_{j=1}^{n}(\dot{Q}^{(j)} - F^{(j)})\delta q_j = 0, \quad \begin{cases} g_\nu = 0 \\ \nu = 1, \dots, \kappa \end{cases} \quad \begin{array}{l} \text{D'Alembert principle} \\ \text{+constraints} \end{array}$$

$$\text{Č etaev condition} \swarrow (A) \qquad (B) \searrow \text{vakonomic condition} \tag{46}$$

$$\delta^{(c)} g_\nu = \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\delta q_j = 0 \qquad \delta^{(v)} g_\nu = \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial q_j}\delta q_j + \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\delta \dot{q}_j = 0$$

$$\Downarrow \qquad\qquad \Downarrow \tag{47}$$

$$\mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = \sum_{\nu=1}^{\kappa} \lambda_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \qquad \begin{cases} \mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = -\sum_{\nu=1}^{\kappa} \mu_\nu \mathcal{D}^{(q_j)} g_\nu \\ -\sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} + \sum_{\nu=1}^{\kappa} \mu_\nu \sum_{h=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_h}\omega_{h,j} + y_j \end{cases} \tag{48}$$

where y_j is defined by (33) and vanishes whenever (47) first condition holds,

$$\mathcal{D}^{(q_j)} f(\mathbf{q}, \dot{\mathbf{q}}, t) = \frac{d}{dt}\frac{\partial f}{\partial \dot{q}_j} - \frac{\partial f}{\partial q_j} \tag{49}$$

is the lagrangian derivative with respect to the variable q_j of a function f. In (46) and (47) the index ν takes each of the values $\nu = 1, \dots, \kappa$; the letters (A) and (B) refer to the two categories of displacements (6) and (7); the equations in (48) are n for each type, $j = 1, \dots, n$.

We now transfer the content of Proposition 1 (which is legitimate because we are in the case (16), as it is evident from(35) and (43)), to state the main result:

Proposition 3 *Let the dynamics of the constrained system governed by the principle* (46)*, where the displacements are related to the constraint functions by* (47)*, first equality, or second equality. Then, the equations of motion are those written in* (48)*. Furthermore, assume that the displacements verify both* (47)*. Then y_j defined in* (33) *can be taken as zero and the two groups of equations in* (48) *are equivalent.*

We link the equations of motion with the transpositional relation we developed through the formulas (15), (40) and (34), which we write again, in light of (43) and (44), as

$$\delta^{(v)} g_\nu - \frac{d}{dt}\left(\delta^{(c)} g_\nu\right) = \sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\left(\delta \dot{q}_j - \frac{d}{dt}\delta q_j\right) - \sum_{j=1}^{n} D^{(q_j)} g_\nu \delta q_j \tag{50}$$

where the derivative $D^{(q_j)}$ is defined in (49). If (30) is assumed to hold, then the relation takes the form

$$\delta^{(v)} g_\nu - \frac{d}{dt}\left(\delta^{(c)} g_\nu\right) = \sum_{h,j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\omega_{j,h}\delta q_h - \sum_{j=1}^{n} D^{(q_j)} g_\nu \delta q_j. \tag{51}$$

We also transfer the content of Proposition 2 to the relation (50) and we state the following

Proposition 4 *Assume that* (47) *are valid simultaneously. Then*

$$\sum_{j=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_j}\left(\delta \dot{q}_j - \frac{d}{dt}\delta q_j\right) = 0 \tag{52}$$

if and only if

$$\mathcal{D}^{(q_j)} g_\nu = \sum_{\sigma=1}^{\kappa} \varrho_\sigma \frac{\partial g_\sigma}{\partial \dot{q}_j}, \qquad \nu = 1, \dots, \kappa \tag{53}$$

for some real coefficients $\varrho_1, \dots, \varrho_\kappa$.

The proof is identical to that exhibited in Proposition 2. As we remarked through (28), the commutation property (i. e. (19) vanishes) is independent of the set of variables are chosen. The necessary condition of Corollary 1 for the commutation (28) becomes now (53). In [20] it is shown that exact constraints (that is $g_\nu = \dot{f}_\nu(\mathbf{q}, t)$ for some f_ν) and constraint functions admitting an integrating factor (that is $\phi_\nu(\mathbf{q}, t) g_\nu = \dot{f}_\nu(\mathbf{q}, t)$ for some factor ϕ_ν) so that the commutation (28) is consistent with these categories. On the other hand, even simple examples of linear constraints (i. e. $g_\nu = \sum_{j=1}^n \ell_{\nu,j}(\mathbf{q}, \mathbf{t})\dot{q}_j + p_\nu(\mathbf{q}, t)$) for which (53) does not hold, so that the commutation cannot be assumed as valid.

4 Discussion and conclusion

The two conditions $\delta^{(v)} g_\nu = 0$ and $\delta^{(c)} g_\nu = 0$ express two different points of view and their application involves various points of discussion: wondering whether they are valid separately, both are valid, one or the other should be applied only to specific classes of mechanical systems are all current topics on the subject.

The debate about the simultaneous validity of both conditions in (47) – or better their compatibility – can profitably look at (50), where it is evident that the question intersects the also relevant and controversial issue about the commutative property of the two operators δ and d/dt (an useful and extensive discussion on this issue can be found in [5]).

First of all a comparison with the motion equations of the same kind present in the literature is necessary. Concerning the class (A), Eq. (48), first group, do represent the classic equations of motion for nonholonomic systems deriving from the Četaev condition (47), first equality. This last condition extends in a simple way a fundamental point in the theory of nonholonomic constraints, which consists of assigning to the linear kinematic constraint $g_\nu = \sum_{j=1}^n \ell_{\nu,j}(\mathbf{q}, \mathbf{t})\dot{q}_j + p_\nu(\mathbf{q}, t) = 0$ the displacements δq_j such that $\sum_{j=1}^n \ell_{\nu,j}\delta q_j = 0$ (we refer, among others, to [11], [15]); clearly $\delta^{(c)} g_\nu = 0$ reduces to the previous condition, if g_ν is linear with respect to the generalized velocities.

Although a rigorous derivation from a physical principle or a theoretical validation of the Četaev condition are uncertain (an extensive discussion on this can be found in [5]), the method (A) shows two remarkable advantages:

- (i) it does not introduce the question of the commutation rule, since $\delta\dot{q}_j$, $j = 1, \ldots, n$, are absent in the definition $\delta^{(c)} g_\nu = 0$ of virtual displacements; Eq. (8) of case (A) can be formulated without making any pronouncements on (11), that is the validity or not of (28).
- (ii) it leads to the same equations even by switching the theoretical approach from the d'Alembert principle (as we performed in an algebraic way) to the Hamilton principle via a variational approach [18,19].

Let us move now on to case (B) and make the following distinction:

- $(B1)$ the hypothesis (11), namely $\delta\dot{\mathbf{r}} = \frac{d}{dt}\delta\mathbf{r}$ is assumed to hold, hence $\frac{d}{dt}\delta q_j = \delta\dot{q}_j$, $j = 1, \ldots, n$, for any set of lagrangian coordinates (see (28)),
- $(B2)$ (11) does not hold and the non–commutative operations is expressed by the formulation of (12), namely $\delta\dot{\mathbf{r}} = \frac{d}{dt}\delta\mathbf{r} = W\delta\mathbf{r}$ (where the matrix W needs to be specified), or equivalently (30).

The hypothesis $(B1)$ reduces Eq. (48), second group, to (we are also assuming both (47), hence $y_j = 0$)

$$\mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = -\sum_{\nu=1}^{\kappa} \mu_\nu \mathcal{D}^{(q_j)} g_\nu - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \tag{54}$$

and they do coincide with the vakonomic equations of motion which are obtained either by the integration (as in [6,21])

$$\int_{t_0}^{t_1} \sum_{j=1}^{n} \left(\mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)}\right)\delta q_j dt = 0$$

or by the constrained Hamilton principle [1]

$$\delta \int_{t_0}^{t_1} \left(\mathcal{L} - F_{NP}^{(j)} + \sum_{\nu=1}^{\kappa} \mu_\nu g_\nu\right)\delta q_j dt = 0. \tag{55}$$

Remark 4 The treatment by integration needs the additional conditions of both ends fixed

$$\delta \mathbf{r}|_{t=t_0} = \delta \mathbf{r}|_{t=t_1} = \mathbf{0}$$

(or equivalently on δq_j)that further specifies the class of displacements. Our method based on a common algebraic technique for both cases (A) and (B) does not require such a request on the displacements, that is the displacements are treated on equal terms.

The vakonomic model has been decisively criticized over the past few years either for the aspect of specific testing problems showing inconsistency (as in [12] and in [22]) and from the theoretical perspective. Regarding the latter, we glean two conclusions from the recent literature (as [6,12,13]):

- the vakonomic equations $(B1)$ do not reproduce the correct physical state even in the case of linear or homogeneous kinematic constraints,
- the first system in (48) and (54) are equivalent if and only if the constraints are holonomic, that is $g_\nu = \dot{f}_\nu$.

As far as we understand, the result in Proposition 3 would seem to offer a broader perspective (i. e. even more than the holonomic case) regarding the simultaneous validity of models (A) and (B) (possibly reduced to $(B1)$), attributable to the role of multipliers.

The case $(B2)$ offers an interesting and recent proposal to reconcile the two cases (A) and (B) through the action of the matrix W or Ω (defined in (12) and (30)) which regulates the non–commutativity of the operations. We mainly refer to [14,17] in order to compare our results with the ones obtained via a variational approach, summarized by (55). The equations of motion achieved at a first stage in [14] (and revisited in [17] as "modified vakonomic method") correspond, in our notation, to

$$\begin{aligned} \mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = &- \sum_{\nu=1}^{\kappa} \mu_\nu \mathcal{D}^{(q_j)} g_\nu - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \\ &+ \sum_{\nu=1}^{\kappa} \mu_\nu \sum_{h=1}^{n} \frac{\partial g_\nu}{\partial \dot{q}_h} \omega_{h,j} + \sum_{h=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_h} \omega_{h,j}, \quad j = 1, \ldots, n \end{aligned} \tag{56}$$

and actually they do not match precisely with (48), second group, because of the terms in (56) $\sum_{h=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_h} \omega_{h,j}$.

Remark 5 The presence of the additional sum can be understood, if we think about (55) and we rearrange (56) in the form

$$\mathcal{D}_1^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = - \sum_{\nu=1}^{\kappa} \mu_\nu \mathcal{D}_1^{(q_j)} g_\nu - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \qquad j = 1, \ldots, n \tag{57}$$

where $\mathcal{D}_1^{(j)}$ is the operator

$$\mathcal{D}_1^{(q_j)} = \mathcal{D}^{(q_j)} - \sum_{h=1}^{n} \omega_{h,j} \frac{\partial}{\partial \dot{q}_h} \quad j = 1, \ldots, n. \tag{58}$$

The disagreement between (57) and (48) second group, which can be written by means of (58) as

$$\mathcal{D}^{(q_j)}\mathcal{L} - F_{NP}^{(j)} = - \sum_{\nu=1}^{\kappa} \mu_\nu \mathcal{D}_1^{(q_j)} g_\nu - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \qquad j = 1, \ldots, n \tag{59}$$

is then overcome if the hypothesis

$$\sum_{h=1}^{n} \frac{\partial \mathcal{L}}{\partial \dot{q}_h} \omega_{h,j} = 0 \ \text{ for each } \ j = 1, \ldots, n \tag{60}$$

is assumed. The same condition (60) plays a special role in [17], in the sense we are going to explain. The transpositional condition (51) can be written by (58) and by assuming both (47) simply as

$$0 = \sum_{h,j=1}^{n} \omega_{h,j} \frac{\partial g_\nu}{\partial \dot{q}_h} \delta q_j - \sum_{j=1}^{n} D^{(q_j)} g_\nu \delta q_j = - \sum_{j=1}^{n} D_1^{(q_j)} g_\nu \delta q_j. \qquad (61)$$

The just written relation is used to claim $D_1^{(q_j)} g_\nu = 0$ for each $j = 1, \ldots, n$ singularly and to write Eq. (57) as in [17], recalling [14]:

$$\mathcal{D}_1^{(q_j)} \mathcal{L} - F_{NP}^{(j)} = - \sum_{\nu=1}^{\kappa} \dot{\mu}_\nu \frac{\partial g_\nu}{\partial \dot{q}_j} \qquad j = 1, \ldots, n \qquad (62)$$

At this point it is evident that (60) makes these equations equivalent to the Četaev Eq. (48), first group, simply by setting $\lambda_\nu = -\dot{\mu}_\nu$ (actually $\mathcal{D}_1^{(q_j)} \mathcal{L} = \mathcal{D}^{(q_j)} \mathcal{L}$ if (60) holds). Hence in [17] the assumption (60) is claimed as sufficient condition for the equivalence of the Četaev and vakonomic equations.

Let us comment this conclusion: as far as we are concerned, the relation (59) does not imply *logically* the single conditions $D_1^{(q_j)} g_\nu = 0$, $j = 1, \ldots, n$; hence, we understand that the latter conditions are *imposed* in order to determine W, and they are consistent with (59).

In any case, the same condition (60) turns out to be essential also for our presentation, not really for the question of the equivalence of the two types of equations (that we assert to exist in general, independently of (60)), but for the consistency of the vakonomic equations following the two different deductions of algebraic or variational type.

Our interest in analyzing deeper the topics introduced will focus mainly on the following points:

- to investigate condition (53), in order to define the class of constraints for which the terms containing $\delta \dot{q}_j - \frac{d}{dt} \delta q_j$ can be eliminated from the transpositional relation (50), owing to (52), independently of the validity of (28),
- still concerning (50), to determine exactly the validity of the model $(B1)$, that is the entire set of functions g_ν verifying $\sum_{j=1}^{n} D^{(q_j)} g_\nu \delta q_j = 0$ (the sum, not singularly), in particular whether it coincides with the class of functions that admitting integrating factor,
- what and how many conditions allow us to determine the matrix W or Ω – defined in (12) and (30) – and whether we actually always arrive at a closed system.

The last question is a delicate point and it is crucial in order to close the problem, by fixing the $3N \times 3N$ entries of W, or the $n \times n$ entries of Ω. One aspect of the question has already emerged regarding the validity (which we have questioned) of $\mathcal{D}_1^{(q_j)} g_\nu = 0$ for each $j = 1, \ldots, n$: this would provide $\kappa \times n$ conditions on the n^2 entries of Ω. On the contrary, the (only one) condition (61) supplies just $3N - \kappa$ conditions, from our point of view.

Finally, we remark that another source of information for the matrix W or Ω may come from the particular structure of the problem: for instance, if $\bar{q}_j$ is an ignorable lagrangian coordinate in $\mathcal{L}$, that is $\frac{\partial \mathcal{L}}{\partial \bar{q}_j} = 0$. Let us add the hypotheses $\frac{\partial g_\nu}{\partial \dot{\bar{q}}_j} = 0$ for any $\nu = 1, \ldots, \kappa$ (the circumstance is not uncommon, as in the examples studied in [17]): then the corresponding equation of motion (48), first group states that

$$\frac{\partial \mathcal{L}}{\partial \dot{\bar{q}}_j} = \bar{p}_j = costant$$

and transferring this information to the corresponding equation on the right one gets the condition on $\omega_{h,j}$

$$\bar{p}_j = \sum_{h=1}^{n} \left(g_\nu - \sum_{h=1}^{\kappa} \frac{\partial g_\nu}{\partial \dot{q}_h} \omega_{h,j} \right)$$

On the contrary, if (62) are claimed to hold, then the deduced condition confirms (60) for the index j corresponding to the ignorable variable $\bar{q}_j$.

Funding Open access funding provided by Università degli Studi di Firenze within the CRUI-CARE Agreement. The authors have no relevant financial or non-financial interests to disclose. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

Declarations

Conflict of interest The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Arnold, V.I., Kozlov, V.V., Nejshtadt, A.I.: Mathematical aspects of classical and celestial mechanics. Dynamical System III, Encyclopaedia of Mathematical Sciences, Springer, Berlin (1988)
2. Benenti, S.: A general method for writing the dynamical equations of nonholonomic systems with ideal constraints. Regul. Chaot. Dyn. **13**(4), 283–315 (2008)
3. Benenti, S.: The non-holonomic double pendulum, an example of non-linear non-holonomic system. Regul. Chaot. Dyn. **1**(5), 417–442 (2011)
4. Četaev, N.G.: On the Gauss principles. Papers on analytical mechanics **323**, Science Academy (1962)
5. Flannery, M.R.: d'Alembert-Lagrange analytical dynamics for nonholonomic systems. J. Math. Phys. **52**(032705), 1–29 (2011)
6. Flannery, M.R.: The elusive d-Alembert-Lagrange dynamics of nonholonomic systems. Am. J. Phys. **79**(9), 932–94 (2011)
7. Flannery, M.R.: The enigma of nonholonomic constraints. Am. J. Phys. **73**, 265–272 (2005)
8. Kozlov, V.V.: Dynamics of systems with nonintegrable constraints I. Mosc. Univ. Mech. Bull. **37**, 27–34 (1982)
9. Kozlov, V.V.: Dynamics of systems with nonintegrable constraints II. Mosc. Univ. Mech. Bull. **37**, 74–80 (1982)
10. Kozlov, V.V.: Dynamics of systems with nonintegrable constraints III. Mosc. Univ. Mech. Bull. **38**, 40–51 (1983)
11. Lemos, N. A.: Analytical mechanics. Cambridge University Press (2018)
12. Lemos, N.A.: Complete inequivalence of nonholonomic and vakonomic mechanics. Acta Mech. **233**, 47–56 (2022)
13. Lewis, A.D., Murray, R.M.: Variational principles for constrained systems: theory and experiment. Int. J. Non-Linear Mech. **30**(6), 793–815 (1995)
14. Llibre, J., Ramírez, R., Sadovaskaia, N.: A new approach to vakonomic mechanics. Nonlinear Dyn. **78**, 2219–2247 (2014)
15. Neĭmark, Ju. I., Fufaev, N. A.: Dynamics of Nonholonomic Systems. Translations of Mathematical Monographs **33**, Providence: American Mathematical Society (1972)
16. Papastavridis, J.G.: Analytical Mechanics: a comprehensive treatise on the dynamics of constrained systems. World Scientific (2014)
17. Pastore, A., Giammarini, A., Grillo, A.: Reconciling Kozlov's vakonomic method with the traditional non-holonomic method: solution of two benchmark problems. Acta Mech. **235**, 2341–2379 (2024)
18. Rund, H.: The Hamilton-Jacobi theory in the calculus of variations. Van Nostrand, London (1966)
19. Saleta, E.J., Cromer, A.H.: A variational principle for nonholonomic systems. Am. J. Phys. **38**, 892–897 (1970)
20. Talamucci, F.: On the transpositional relation for nonholonomic systems. *to appear* (2024)
21. Xu, S.-S., Berakdar, J.: On the validity of the vakonomic model and the Čhetaev model for constraint dynamical systems. Rep. Math. Phys. **60**(1), 107–116 (2007)
22. Xu, S.-S., Li, S.-M., Berakdar, J.: Appell-Hamel dynamical system: a nonlinear test of the Chetaev and the vakonomic model. ZAMM Z. Angew. Math. Mech. **87**(10), 692–697 (2007)
23. Zampieri, G.: Nonholonomic versus vakonomic dynamics. J. Differ. Eq. **163**, 335–347 (2000)
24. Zeković, D.N.: Dynamics of mechanical systems with nonlinear nonholonomic constraints I The history of solving the problem of a material realization of a nonlinear nonholonomic constraint. Z. Angew. Math. Mech. **91**(11), 883–898 (2011)
25. Zeković, D.N.: Dynamics of mechanical systems with nonlinear nonholonomic constraints II Differential equations of motion. Z. Angew. Math. Mech **91**(11), 899–922 (2011)
26. Zeković, D.N.: Dynamics of mechanical systems with nonlinear nonholonomic constraints III. Z. Angew. Math. Mech **93**(8), 550–574 (2013)