The enigma of nonholonomic constraints

M. R. Flannery[a)]
School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332

(Received 16 February 2004; accepted 8 October 2004)

The problems associated with the modification of Hamilton's principle to cover nonholonomic constraints by the application of the multiplier theorem of variational calculus are discussed. The reason for the problems is subtle and is discussed, together with the reason why the proper account of nonholonomic constraints is outside the scope of Hamilton's variational principle. However, linear velocity constraints remain within the scope of D'Alembert's principle. A careful and comprehensive analysis facilitates the resolution of the puzzling features of nonholonomic constraints.
[DOI: 10.1119/1.1830501]

I. INTRODUCTION

The action integral,

$$S=\int_{t_1}^{t_2} L(\mathbf{q},\dot{\mathbf{q}},t)\,dt, \tag{1}$$

plays a central role in the dynamics of physical systems described by a Lagrangian L. Hamilton's principle states that the actual path $\mathbf{q}(t)$ of a particle is the path that makes the action S a minimum. It is well known that Hamilton's principle,

$$\delta S=\delta\int_{t_1}^{t_2} L(\mathbf{q},\dot{\mathbf{q}},t)\,dt=0, \quad \text{(Hamilton's principle)}, \tag{2}$$

when applied to problems involving c-holonomic constraints with the geometric form,

$$f_k(q_1,q_2,\ldots,q_n,t)=0, \quad (k=1,2,\ldots,c), \tag{3}$$

leads to Lagrange's equations of motion whose solution provides the time dependence of the $(n-c)$ independent generalized coordinates q_j for the unconstrained degrees of freedom.

For problems that require additional calculation of the forces Q_j^c of holonomic constraint, Hamilton's principle may be generalized to yield correct results simply by replacing L in Eq. (2) by

$$L^\dagger=L(\mathbf{q},\dot{\mathbf{q}},t)+\sum_{k=1}^{c}\lambda_k(t)f_k(\mathbf{q},t), \tag{4}$$

where the λ_k are Lagrange multipliers. Equation (2) is therefore replaced by Hamilton's generalized principle,

$$\delta S^\dagger=\delta\int_{t_1}^{t_2} L^\dagger(\boldsymbol{\eta},\dot{\boldsymbol{\eta}},t)\,dt=0, \quad \text{(Hamilton's generalized principle)}, \tag{5}$$

from which the Euler–Lagrange equations

$$\frac{d}{dt}\left(\frac{\partial L^\dagger}{\partial\dot{\eta}_j}\right)-\frac{\partial L^\dagger}{\partial\eta_j}=0, \quad (j=1,2,\ldots,n+c) \tag{6}$$

can be derived via *free* variations of the extended set $\boldsymbol{\eta}\equiv\{\mathbf{q}(q_1,q_2,\ldots,q_n),\boldsymbol{\lambda}(\lambda_1,\lambda_2,\ldots\lambda_c)\}$ of the $(n+c)$ variables involved in Eq. (5). Because $f_k(\mathbf{q},t)$ are independent of the generalized velocity $\dot{\mathbf{q}}$, the first n-equations of the Euler–Lagrange set (6) provide the correct equations of state. Because Eq. (4) is independent of $\dot{\lambda}_k$, the last c equations of the Euler–Lagrange set (6) for the λ_k $(k=1,2\ldots,c)$ simply reproduce the equations (3) of holonomic constraint.

A recurring theme[1–4] is whether Hamilton's principle (2) may be similarly generalized so as to treat nonholonomic (dynamic) constraints,

$$g_k(\mathbf{q},\dot{\mathbf{q}},t)=0, \tag{7}$$

which depend on generalized velocities $\dot{\mathbf{q}}$, simply by substituting

$$L^*=L+\sum_{k=1}^{c}\lambda_k(t)g_k(\mathbf{q},\dot{\mathbf{q}},t) \tag{8}$$

for L in Eq. (2). A theorem in the calculus of variations appears, at first sight, tailor-made for such a conjecture. The theorem[5–7] states that the path $\mathbf{q}(t)$ that makes the action Eq. (1) have an extremum under the side conditions (7) is the same as the path that makes the modified functional, $S^*=\int_{t_1}^{t_2}L^*(\boldsymbol{\eta},\dot{\boldsymbol{\eta}},t)\,dt$, an extremum, without any side conditions imposed. On the basis of this multiplier rule, the conjecture, the substitution of Eq. (8) in Eq. (2), was simply adopted without reservation for the general case (7) and equations of state were published.[1–3]

This conjecture becomes problematic, particularly because the multiplier rule does not yield the standard equations of state as obtained from D'Alembert's more basic principle for systems with less general nonholonomic constraints,

$$g_k^{(L)}(\mathbf{q},\dot{\mathbf{q}},t)=\sum_{j=1}^{n}A_{kj}(\mathbf{q},t)\dot{q}_j+B_k(\mathbf{q},t)=0, \tag{9}$$

which are now only linear in the velocities $\dot{q}_j$. Yet, the same multiplier rule[5–7] works for the holonomic constraints in Eq. (3).

The question of whether the use of Eq. (8) in Eq. (2) is a viable generalization of Hamilton's principle is of interest here, because Ref. 1 advocates its use and cites the equations of state derived from it.[3] However, this generalization had previously been acknowledged[4] as being incorrect because it did not reproduce the correct equations of state for systems under linear constraints in Eq. (9). Some textbooks[8–11] also have indicated the fallacy of using Eq. (8) in Eq. (2). However, the basic reason for its failure has remained obscure.

The multiplier rule[5–7] is indeed correct, as stated, so the fact that it works for holonomic constraints (3), but not for nonholonomic constraints (7) poses a dilemma.

Many examples can be given that explicitly illustrate that Eq. (8) does not provide the correct results as obtained from Newtonian mechanics.[12] In this paper, we search for the reason why the procedure fails and, in so doing, we also explain why the proper account of nonholonomic constraints given by Eqs. (7) and (9) is outside the scope of Hamilton's principle, even though the linear constraints in Eq. (9) remain within the scope of D'Alembert's principle. We will find the conditions that Eq. (8) must satisfy for valid substitution into Eq. (2). We also will indicate why the general nonholonomic constraints in Eq. (7) are outside the scope of a principle based on virtual displacements. Rather than beginning from Eq. (2) and showing, as has been done, that an application involving Eq. (7) or (9) leads to erroneous results,[4,8–12] more insight can be gained by tracing the various stages of development of the variational principle, Eq. (2), from the more fundamental principle of D'Alembert. The essential reason will then become apparent.

Because variational theorems and methods are essential tools of modern analytical dynamics and because various fallacies underlying their use are subtle and are not generally well appreciated, it is hoped that the following account will help illuminate their scope of application.

II. THEORY

We first outline some standard deductions of D'Alembert's principle, which is then expressed in a useful variational form that will provide a "royal road" from which Hamilton's principle can be easily extracted. The resolution of why the extended Lagrangian Eq. (4) works, while Eq. (8) does not, in Hamilton's principle, Eq. (2), will then become apparent via this approach.

A. Differential form of D'Alembert's principle

The motion of a system of particles, $i=1,2,\ldots,N$ of mass m_i located at $\mathbf{r}_i(t)$ in an inertial frame of reference is governed by Newton's equations,

$$\mathbf{F}_i+\mathbf{F}_i^c=m_i\ddot{\mathbf{r}}_i, \tag{10}$$

where the net force acting on each particle is decomposed into an active force $\mathbf{F}_i$ and a force $\mathbf{F}_i^c$ of constraint. A virtual displacement $\delta\mathbf{r}_i$ is an instantaneous variation from a given configuration $\mathbf{r}_i$ performed at a fixed time t and taken consistent with the constraints at that time. The summation convention, $a_{ij}q_j\equiv\Sigma_{j=1}^{n}a_{ij}q_j$ for repeated indices j will be adopted.

Assume that the total virtual work $\mathbf{F}_i^c\cdot\delta\mathbf{r}_i$ performed by all the constraining forces is zero. D'Alembert's principle, in both Newtonian $\mathbf{r}_i$ ($i=1,2,\ldots,N$) and generalized q_j ($j=1,2,\ldots,3N$) coordinate versions, states that[1,8–10,13]

$$(m_i\ddot{\mathbf{r}}_i-\mathbf{F}_i)\cdot\delta\mathbf{r}_i=\left[\frac{d}{dt}\left(\frac{\partial T}{\partial\dot{q}_j}\right)-\frac{\partial T}{\partial q_j}-Q_j\right]\delta q_j=0, \tag{11}$$

where the total kinetic energy $T=\frac{1}{2}m_i\dot{\mathbf{r}}_i^2(\mathbf{q},\dot{\mathbf{q}},t)$ is expressed in terms of the $n=3N$ generalized coordinates of all the particles. The generalized force,

$$Q_j\equiv\mathbf{F}_i\cdot\frac{\partial\mathbf{r}_i}{\partial q_j}, \tag{12}$$

is such that the virtual work $Q_j\delta q_j=\mathbf{F}_i\cdot\delta\mathbf{r}_i$ is equivalent in both representations and may be decomposed into a potential part,

$$Q_j^{(P)}(\mathbf{q},\dot{\mathbf{q}},t)\equiv\frac{d}{dt}\left(\frac{\partial U}{\partial\dot{q}_j}\right)-\frac{\partial U}{\partial q_j}, \tag{13}$$

derived from a generalized monogenic (the same for all particles) potential $U(\mathbf{q},\dot{\mathbf{q}},t)$ and a nonpotential part $Q_j^{\mathrm{NP}}=\mathbf{F}_i^{\mathrm{NP}}\cdot\partial\mathbf{r}_i/\partial q_j$. D'Alembert's principle is then

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_j}\right)-\frac{\partial L}{\partial q_j}-Q_j^{\mathrm{NP}}\right]\delta q_j=0,\quad\text{(D'Alembert's principle)}, \tag{14}$$

where the Lagrangian is

$$L(\mathbf{q},\dot{\mathbf{q}},t)=T(\mathbf{q},\dot{\mathbf{q}},t)-U(\mathbf{q},\dot{\mathbf{q}},t). \tag{15}$$

B. Holonomic constraints

When the c-constraint conditions in Eq. (3) are utilized to reduce the number of generalized coordinates from n to the minimum number $(n-c)$ of actual independent degrees of freedom, that is, when the constraints are embedded within the problem at the outset, then all the $(n-c)$ δq_j's in Eq. (14) are independent of each other. Because each displacement can take on any value at each t, the satisfaction of D'Alembert's principle, Eq. (14), demands that each coefficient of δq_j in Eq. (14) separately vanishes to yield Lagrange's equations,[1,8–10,13]

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_j}\right)-\frac{\partial L}{\partial q_j}=Q_j^{\mathrm{NP}}, \tag{16}$$

for the $(n-c)$ independent degrees of freedom.

When the holonomic constraints Eq. (3) are not used to reduce the set of generalized coordinates to this minimum number, that is, when they are instead "adjoined," then c of the δq_j's in Eq. (14) depend on the independent $(n-c)$ coordinates and are constrained by the c conditions,

$$\frac{\partial f_k}{\partial q_j}\delta q_j=0,\quad(k=1,2,\ldots,c) \tag{17}$$

which is obtained by differentiating Eq. (3) and keeping t fixed. The Lagrange multipliers $\lambda_k(t)$ can then be introduced by subtracting the quantity $\lambda_k(\partial f_k/\partial q_j)\delta q_j=0$ from the left-hand side of Eq. (14) to give

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_j}\right)-\frac{\partial L}{\partial q_j}-\lambda_k(t)\frac{\partial f_k(\mathbf{q},t)}{\partial q_j}-Q_j^{\mathrm{NP}}\right]\times\delta q_j(t)=0.\quad(j=1,2,\ldots,n). \tag{18}$$

Nonpotential forces Q_j^{NP} are included in Eq. (18). If we denote the $m=n-c$ independent (free) coordinates by $q_1,q_2,\ldots,q_m$ and the c-dependent ones by $q_{m+1},q_{m+2},\ldots,q_n$, then the previously unassigned c multipliers, λ_k, are now chosen to satisfy the c equations,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial\dot{q}_j}\right)-\frac{\partial L}{\partial q_j}=\lambda_k(t)\frac{\partial f_k(\mathbf{q},t)}{\partial q_j}+Q_j^{\mathrm{NP}}\quad(j=m+1,m+2,\ldots,n). \tag{19}$$

Equation (18) then reduces to

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)-\frac{\partial L}{\partial q_j}-\lambda_k(t)\frac{\partial f_k(\mathbf{q},t)}{\partial q_j}-Q_j^{\mathrm{NP}}\right]$$
$$\times\,\delta q_j(t)=0,\quad (j=1,2,\ldots,m) \tag{20}$$

for the free $m=n-c$ coordinates. Because the m δq_j's in Eq. (20) are all independent and arbitrary, each of the δq_j coefficients in Eq. (20) must separately vanish. The set,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)-\frac{\partial L}{\partial q_j}=\lambda_k(t)\frac{\partial f_k(\mathbf{q},t)}{\partial q_j}+Q_j^{\mathrm{NP}},\quad (j=1,2,\ldots,n), \tag{21}$$

therefore represents the equations of state for the full array of dependent and independent variables $q_1,q_2,\ldots,q_n$.

Now adjoin the constraint equations (3) to the Lagrangian set in Eq. (21) of n-equations to provide $n+c$ equations for the $n+c$ unknowns, the n q_j's and the c λ_k's, so that the sets $\mathbf{q}\equiv\{q_j\}$ and $\boldsymbol{\lambda}\equiv\{\lambda_k\}$ may in principle be determined. By comparing Eq. (21) with Eq. (16), it is seen that $Q_j^c=\lambda_k(\partial f_k/\partial q_j)$ are additional forces acting on the system. These Q_j^c must therefore be the forces of constraint which, because of Eq. (17), do no virtual work, as required for the validity of D'Alembert's principle. Although standard,[1,8–13] the above review will help provide the context to what now follows.

Because f_k is independent of the velocities $\dot{\mathbf{q}}$, a generalized D'Alembert principle,

$$\left[\frac{d}{dt}\left(\frac{\partial L^\dagger}{\partial \dot{\eta}_j}\right)-\frac{\partial L^\dagger}{\partial \eta_j}-Q_j^{\mathrm{NP}}\right]\delta\eta_j=0,\quad (j=1,2,\ldots,n+c)$$
$$(\text{D'Alembert generalized principle}), \tag{22}$$

can therefore be introduced where $L^\dagger(\dot{\boldsymbol{\eta}},\boldsymbol{\eta},t)=L+\lambda_k(t)f_k(\mathbf{q},t)$ is an augmented Lagrangian over an extended set of coordinates $\boldsymbol{\eta}\equiv(\mathbf{q},\boldsymbol{\lambda})$. On regarding all η_j as free, then

$$\frac{d}{dt}\left[\frac{\partial(L+\lambda_k f_k)}{\partial \dot{\eta}_j}\right]-\frac{\partial(L+\lambda_k f_k)}{\partial \eta_j}=Q_j^{\mathrm{NP}},\quad (j=1,2,\ldots,n+c) \tag{23}$$

are the generalized Lagrange equations for the extended set η_j. The first n equations of Eq. (23) reproduce the correct equations of state, (21), and the last c equations reproduce the constraint equations, $f_k=0$. Hence, D'Alembert's principle in Eq. (14), with the displacements δq_j subject to the c conditions in Eq. (17), is equivalent to the generalized principle, Eq. (22), with all coordinates η_j free. The replacement of the basic principle Eq. (14) with the subsidiary conditions Eq. (17) by the generalized principle Eq. (22) without subsidiary conditions is the Lagrange multiplier rule. Both principles provide identical equations of state, Eq. (21), and the multiplier rule in Eq. (22) provides the shortcut.

It is important to note that the displaced paths $q_j+\delta q_j$, not only comply with the essential conditions in Eq. (17) for the displacements, but also satisfy the equations of constraint,

$$f_k(\mathbf{q}+\delta\mathbf{q},t)=f_k(\mathbf{q},t)+\delta f_k(\mathbf{q},t)=0, \tag{24}$$

because there is no change $\delta f_k=(\partial f_k/\partial q_j)\delta q_j=0$ to the constraint Eq. (3). The displaced paths are therefore all geometrically possible because they all conform to Eq. (24). The key requirement for application of the multiplier rule is that the displaced paths must be geometrically possible by satisfying the equations (24) of constraint. As will be shown next, this condition is violated, in general, by nonholonomic constraints.

C. Nonholonomic constraints

The virtual displacements δq_j for nonholonomic systems with c linear constraints,

$$g_k^{(L)}(\mathbf{q},\dot{\mathbf{q}},t)=A_{kj}(\mathbf{q},t)\dot{q}_j+B_k(\mathbf{q},t)=0, \tag{25}$$

obeyed by the actual path, are themselves constrained to obey c instantaneous conditions

$$A_{kj}(\mathbf{q},t)\,\delta q_j=0,\quad (k=1,2,\ldots,c) \tag{26}$$

obtained by first writing Eq. (25) in differential form as

$$g_k^{(L)}dt=A_{kj}(\mathbf{q},t)dq_j+B_k(\mathbf{q},t)dt, \tag{27}$$

and then by setting $dt=0$ and $dq_j=\delta q_j$ as prescribed. As with Eq. (17), the linear conditions (26) also may be absorbed in D'Alembert's principle because Eq. (14) is linear in δq_j. By adding $\lambda_k A_{kj}\delta q_j=0$ to the right-hand side of Eq. (14), and by proceeding as before in Sec. II B, the equations of state under the linear constraints in Eq. (25) are obtained in the form

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)-\frac{\partial L}{\partial q_j}=\lambda_k(\mathbf{q},t)A_{kj}(\mathbf{q},t)+Q_j^{\mathrm{NP}},\quad (j=1,2,\ldots,n) \tag{28}$$

for all the coordinates. We now examine the validity of D'Alembert's generalized principle

$$\left\{\frac{d}{dt}\left[\frac{\partial(L+\mu_k g_k)}{\partial \dot{\eta}_j}\right]-\frac{\partial(L+\mu_k g_k)}{\partial \eta_j}-Q_j^{\mathrm{NP}}\right\}\delta\eta_j=0,$$
$$(j=1,2,\ldots,n+c), \tag{29}$$

applied to nonholonomic constraints Eq. (7), where $\mu_k(t)$ are a different set of multipliers and where all $\delta\eta_j$ are regarded as free. On introducing G_{kj}, where

$$G_{kj}=\left[\frac{d}{dt}\left(\frac{\partial g_k}{\partial \dot{q}_j}\right)-\frac{\partial g_k}{\partial q_j}\right]\quad (j=1,2,\ldots,n), \tag{30}$$

and is zero for $j>n$, Eq. (29) can be rewritten as

$$\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\eta}_j}\right)-\frac{\partial L}{\partial \eta_j}+\dot{\mu}_k\frac{\partial g_k}{\partial \dot{\eta}_j}+\mu_k G_{kj}-g_k\frac{\partial \mu_k}{\partial \eta_j}-Q_j^{NP}\right]\delta\eta_j=0,$$
$$(j=1,2,\ldots,n+c). \tag{31}$$

The first n equations of Eq. (31) provides the equation of state,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)-\frac{\partial L}{\partial q_j}=-\dot{\mu}_k\frac{\partial g_k}{\partial \dot{\eta}_j}-\mu_k G_{kj}+Q_j^{NP}\quad (j=1,2,\ldots,n), \tag{32}$$

as derived from D'Alembert's generalized principle, Eq. (29). The last c equations of Eq. (31) yield the constraint equations (7), as expected. But Eq. (32) reproduces the correct equation (28) of state for the linear constraints in Eq. (25), only when Eq. (30) for linear constraints vanishes, that is, provided

$$G_{kj}^{(L)}=\left[\left(\frac{\partial A_{kj}}{\partial q_i}-\frac{\partial A_{ki}}{\partial q_j}\right)\dot{q}_i+\left(\frac{\partial A_{kj}}{\partial t}-\frac{\partial B_k}{\partial q_j}\right)\right]=0. \tag{33}$$

Because condition Eq. (30) is basic to validity of Eq. (29), the significance of this auxiliary restriction on the linear constraints (25) will now be explored.

In order for Eq. (25) to be a perfect (exact) differential of a function $f_k(\mathbf{q},t)$, we must have

$$A_{ki}(\mathbf{q},t)\dot{q}_i + B_k(\mathbf{q},t) = \frac{d}{dt} f_k = \frac{\partial f_k}{\partial q_i}\dot{q}_i + \frac{\partial f_k}{\partial t}. \tag{34}$$

The correspondence $A_{ki} = \partial f_k/\partial q_i$ and $B_k = \partial f_k/\partial t$ provides the (necessary and sufficient) conditions

$$\frac{\partial A_{ki}}{\partial q_j} = \frac{\partial^2 f_k}{\partial q_j \partial q_i} = \frac{\partial^2 f_k}{\partial q_i \partial q_j} = \frac{\partial A_{kj}}{\partial q_i}, \tag{35}$$

$$\frac{\partial B_k}{\partial q_i} = \frac{\partial^2 f_k}{\partial q_i \partial t} = \frac{\partial^2 f_k}{\partial t \partial q_i} = \frac{\partial A_{ki}}{\partial t}, \tag{36}$$

for the "exactness" of Eq. (25). Provided the linear constraints (25) satisfy conditions (35) and (36), an integrated form f_k therefore exists but may be unknown. Such constraints are termed *semiholonomic* and are denoted by $g_k^{(\mathrm{sh})}(\mathbf{q},\dot{\mathbf{q}},t)=0$. But the conditions (35) and (36) for exactness yield condition Eq. (33), for all $\dot{q}_i$ which satisfy the constraints. Semiholonomic constraints can therefore be correctly treated by D'Alembert's generalized principle, Eq. (29). In addition to exactness, semiholonomic constraints $(G_{kj}^{(L)}=0)$ possess a further important property. The equations of constraint appropriate to the displaced paths $\mathbf{q}+\delta\mathbf{q}$ are

$$g_k(\mathbf{q}+\delta\mathbf{q},\dot{\mathbf{q}}+\delta\dot{\mathbf{q}},t) = g_k(\mathbf{q},\dot{\mathbf{q}},t) + \delta g_k(\mathbf{q},\dot{\mathbf{q}},t). \tag{37}$$

Because $g_k(\mathbf{q},\dot{\mathbf{q}},t)=0$ for the true dynamical path $\mathbf{q}(t)$, the constraint equations for the displaced paths change by

$$\delta g_k = \frac{\partial g_k}{\partial q_j}\delta q_j(t) + \frac{\partial g_k}{\partial \dot{q}_j}\delta \dot{q}_j(t). \tag{38}$$

With the aid of $\delta\dot{q}_j(t) = d[\delta q_j(t)]/dt$, this difference is

$$\delta g_k = \frac{d}{dt}\left[\frac{\partial g_k}{\partial \dot{q}_j}\delta q_j(t)\right] - G_{kj}\delta q_j(t). \tag{39}$$

The condition for the displaced paths to be all geometrically possible is that $g_k(\mathbf{q}+\delta\mathbf{q},\dot{\mathbf{q}}+\delta\dot{\mathbf{q}},t)=0$, that is $\delta g_k=0$ and the constraints are invariant to displacements. For the linear constraints (25), Eq. (39) reduces to

$$\delta g_k^{(L)} = \frac{d}{dt}(A_{kj}\delta q_j) - G_{kj}^{(L)}\delta q_j. \tag{40}$$

On invoking the basic restriction (26) on the displacements and the exactness condition $G_{kj}^{(L)}=0$, Eq. (40) reduces to $\delta g_k^{(\mathrm{sh})}=0$, which implies geometrically possible paths. D'Alembert's generalized principle (29) with Eq. (25) therefore holds for semiholonomic systems where the displaced paths are all geometrically possible. Semiholonomic systems are, in essence, holonomic, although the integrated holonomic form $f_k=0$ may not be known.

Linear constraints (25) can be integrable and yet violate the exactness condition (33). For example, the constraint,

$$g_1^{(\mathrm{sh})}(\mathbf{q},\dot{\mathbf{q}}) = (3q_1^2+2q_2^2)\dot{q}_1 + 4q_1q_2\dot{q}_2 = 0, \tag{41}$$

is exact because (33) is satisfied and it integrates directly to give $f_1 = q_1^3 + 2q_2^2 q_1 = \text{constant}$. The constraint,

$$g_2^{(I)}(\mathbf{q},\dot{\mathbf{q}}) = (4q_1+3q_2^2)\dot{q}_1 + 2q_1q_2\dot{q}_2 = 0, \tag{42}$$

is not exact but can be integrated via the integrating factor $\Phi_2\ (=q_1^2)$ to give $f_2 = q_1^4 + q_1^3 q_2^2 = \text{constant}$. All exact constraints are therefore integrable, but all integrable constraints are not necessarily exact. The conditions (35) and (36) are too restrictive for integrable constraints $g_k^{(I)}$, which can however be rendered in exact form by multiplying by the integrating factor $\Phi_k(\mathbf{q},t)$. Then $g_k^{(\mathrm{sh})} = \Phi_k g_k^{(I)}$ now satisfies the condition (33) for both exactness and geometrically possible displaced paths. For example, the constraint,

$$g_2^{(\mathrm{sh})}(\mathbf{q},\dot{\mathbf{q}}) = \Phi_2 g_2^{(I)} = (4q_1^3 + 3q_1^2q_2^2)\dot{q}_1 + 2q_1^3q_2\dot{q}_2 = 0, \tag{43}$$

now satisfies condition (33) and is therefore in exact (semiholonomic) form. A known integrating factor Φ_k implies a known integrated holonomic form $f_k=0$, so that the simpler holonomic result Eq. (23) can be used rather than D'Alembert's generalized principle (29).

The linear constraints (25) which do not satisfy the exactness condition (33) are classified as nonholonomic. D'Alembert's generalized principle (29) is therefore not appropriate for nonholonomic constraints (25), as is also confirmed by the fact that Eq. (32) is not the correct equation (28) of state, because $G_{kj}^{(L)} \neq 0$, in general.

D'Alembert's basic principle, Eq. (14), is not amenable to general nonholonomic constraints (7), because there is now no relation such as Eq. (26) which connects the displacements δq_j in a linear form. The fact that Eq. (7) is, in general, not a linear function of $\dot{q}_j$ prohibits writing a linear interrelation between the δq_j's essential for the application of D'Alembert's principle. General nonholonomic constraints (7) are therefore outside the scope of all principles based on virtual displacements.

The key conclusions of Secs. II B and II C are the following:

(1) D'Alembert's basic principle, Eq. (14), is applicable to holonomic and linear nonholonomic constraints, as is already known.
(2) D'Alembert's generalized principle, Eq. (22), applies to holonomic constraints and Eq. (29) applies to semiholonomic systems, because the displaced paths are also geometrically possible paths, an essential criterion for the validity of the underlying multiplier rule. The solution of both sets provides the actual path $\{q_j(t)\}$ and the constraint forces $\{Q_j^c\}$.
(3) The displaced paths $q_j+\delta q_j$ for linear nonholonomic systems are not geometrically possible and therefore do not satisfy the multiplier-rule condition.
(4) It is important to distinguish restrictions imposed on virtual displacements, such as Eq. (26), from the actual equations of constraint, such as Eq. (9), which must only be satisfied within the equations of state that are eventually determined by some variational procedure. The constraint equations $g_k(\mathbf{q},\dot{\mathbf{q}},t)=0$ satisfied by the true dynamical path $\mathbf{q}(t)$ do not necessarily imply that the corresponding equations $g_k(\mathbf{q}+\delta\mathbf{q},\dot{\mathbf{q}}+\delta\dot{\mathbf{q}},t)=0$ are satisfied by the displaced paths.
(5) General nonholonomic constraints (7) are completely outside the scope of even the most fundamental principle

of D'Alembert. The generalization[1–3] of any principle based on Eq. (14) to general nonholonomic constraints is without foundation.

D. The δL version of D'Alembert's principle

The Lagrangian for the varied paths is

$$L(\mathbf{q}+\delta\mathbf{q},\dot{\mathbf{q}}+\delta\dot{\mathbf{q}},t)=L(\mathbf{q},\dot{\mathbf{q}},t)+\delta L(\mathbf{q},\dot{\mathbf{q}},t), \tag{44}$$

where the change in L due to the virtual displacement δq_j from the actual path $\mathbf{q}$ is

$$\delta L=\frac{\partial L}{\partial q_j}\delta q_j(t)+\frac{\partial L}{\partial \dot{q}_j}\delta \dot{q}_j(t). \tag{45}$$

With the aid of $\delta\dot{q}_j(t)=d[\delta q_j(t)]/dt$, the change is

$$\delta L=\frac{d}{dt}[p_j\delta q_j(t)]-\left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_j}\right)-\frac{\partial L}{\partial q_j}\right]\delta q_j(t), \tag{46}$$

where the generalized momentum is defined as $p_j=\partial L/\partial\dot{q}_j$. D'Alembert's basic principle (14) can then be recast in δL form as

$$\delta L=\frac{d}{dt}(p_j\delta q_j)-Q_j^{\mathrm{NP}}\delta q_j. \tag{47}$$

The differential version, Eq. (14), and the δL version, Eq. (44), of D'Alembert's principle are equivalent and are fundamental equations of dynamics. When the holonomic constraints (3) are adjoined, rather than embedded, there are c δq_j's in Eq. (46) that are dependent on the remaining $(n-c)$ displacements. Because there is no change, $\delta f_k=0$, to the holonomic equations (3) among the varied paths, we may add $\delta[\lambda_k(t)f_k]=0$ to the left-hand side of Eq. (47). By utilizing the augmented Lagrangian $L^\dagger$ over the extended set of free generalized coordinates $\boldsymbol{\eta}\equiv(\mathbf{q},\boldsymbol{\lambda})$, the generalized version of D'Alembert's principle, Eq. (47), is

$$\delta L^\dagger(\dot{\boldsymbol{\eta}},\boldsymbol{\eta},t)=\delta[L+\lambda_k(t)f_k(\mathbf{q},t)]=\frac{d}{dt}(p_j\delta\eta_j)-Q_j^{\mathrm{NP}}\delta\eta_j. \tag{48}$$

If we use the definition (46) for δL, the generalized version (48) reproduces the correct equations of state, Eq. (21), and provides another example of the multiplier rule.

For semiholonomic systems, the Lagrangian L can also be replaced by $L^{(\mathrm{sh})}=L+\mu_k g_k^{(\mathrm{sh})}$ because the constraints $g_k^{(\mathrm{sh})}(\mathbf{q},\dot{\mathbf{q}},t)=0$ are exact, thereby satisfying the condition $\delta g_k^{(\mathrm{sh})}=0$ for geometrically possible paths. D'Alembert's generalized principle (47) therefore yields the equations of state

$$\frac{d}{dt}\left[\frac{\partial(L+\mu_k g_k^{(\mathrm{sh})})}{\partial\dot{\eta}_j}\right]-\frac{\partial(L+\mu_k g_k^{(\mathrm{sh})})}{\partial\eta_j}=Q_j^{\mathrm{NP}},\quad (j=1,2,\ldots,n+c) \tag{49}$$

for the extended coordinates $(\boldsymbol{\eta}\equiv\mathbf{q},\boldsymbol{\mu})$ for a semiholonomic system. The multiplier rule of replacing L in Eq. (47) by $L^*=L+\mu_k g_k$ is, however, not valid for inexact linear or general nonholonomic constraints, because the displaced paths are not geometrically possible paths, as explained in Sec. II C.

E. Generalization of Hamilton's variational principle

Hamilton's integral principle,

$$\int_{t_1}^{t_2}\delta L\,dt=\delta\int_{t_1}^{t_2}L\,dt=[p_j\delta q_j]_{t_1}^{t_2}-\int_{t_1}^{t_2}[Q_j^{\mathrm{NP}}\delta q_j]dt, \tag{50}$$

is D'Alembert's principle, Eq. (47), integrated between the times t_1 and t_2. The δ operator does not affect the time and was therefore taken outside the integral. The appropriate Eq. (28) for linear nonholonomic constraints is recovered by making the time integration in Eq. (50) redundant. The application of Eq. (50) then reduces simply to an application of D'Alembert's basic principle (14), as in Sec. II C. The main advantage, however, of the integral principle Eq. (50) is that it becomes a variational principle,

$$\delta S=\delta\int_{t_1}^{t_2}L\,dt=0, \tag{51}$$

by admitting only those paths $q_j(t)$ that pass through the fixed end points, $\delta q_j(t_{1,2})=0$, and by considering only potential systems, that is, $Q_j^{\mathrm{NP}}=0$. The virtual variation δ ensures that the transit time $\tau=t_2-t_1$ remains the same for all the varied paths. Equation (51) is Hamilton's principle for the least action $S=\int_{t_1}^{t_2}L\,dt$.

When attempting to generalize Hamilton's variational principle, Eq. (51), the conditions for generalization of the more fundamental differential and δL versions, Eqs. (14) and (47) of D'Alembert's principle by the multiplier rule, are still in effect. Equation (51) can be directly applied to holonomic systems with the embedded constraints in Eq. (3) to recover the correct equations of state (16) with $Q_j^{\mathrm{NP}}=0$. When holonomic constraints are adjoined in order to determine the constraint forces, then L in Eq. (51) can be replaced by $L^\dagger=L+\lambda_k(t)f_k(\mathbf{q},t)$, because $\delta f_k=0$, to give Hamilton's generalized principle

$$\delta S^\dagger=\delta\int_{t_1}^{t_2}L^\dagger(\dot{\boldsymbol{\eta}},\boldsymbol{\eta},t)dt=\delta\int_{t_1}^{t_2}[L+\lambda_k(t)f_k(\mathbf{q},t)]dt=0, \tag{52}$$

where the $\delta\eta_j$'s involved are free and independent. For semiholonomic constraints, Hamilton's principle is generalized to

$$\delta S^{(\mathrm{sh})}=\delta\int_{t_1}^{t_2}L^{(\mathrm{sh})}(\dot{\boldsymbol{\eta}},\boldsymbol{\eta},t)dt=\delta\int_{t_1}^{t_2}[L+\mu_k(t)g_k^{(\mathrm{sh})}(\mathbf{q},\dot{\mathbf{q}},t)]dt=0. \tag{53}$$

The essential reason for the validity of (52) and (53) is that the paths $\mathbf{q}+\delta\mathbf{q}$ admitted into the variational procedures are all geometrically possible, that is $\delta f_k=0$ and $\delta g_k^{(\mathrm{sh})}=0$ and that the δ and $\int$ operations commute. The correct equations of state (19) and (49) with $Q_j^{\mathrm{NP}}=0$ are recovered from (52) and (53), respectively. Because $g_k^{(\mathrm{sh})}$ is, by definition, the perfect differential df_k/dt, then provided that f_k is known, Eq. (53) reduces to

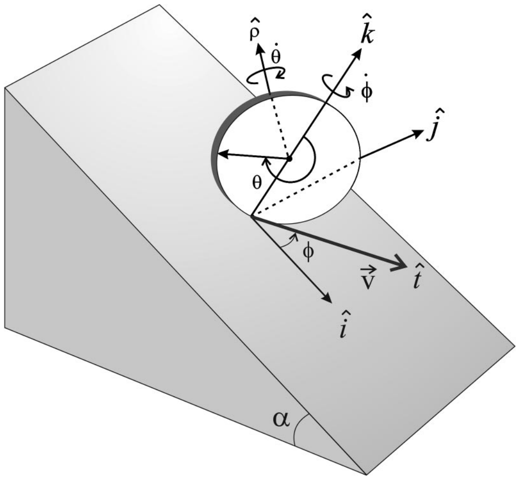

Fig. 1. An upright coin rolls and spins down an inclined plane of angle α. Directions of space-fixed axes are $\hat{i}$, $\hat{j}$, and $\hat{k}$, as indicated. Coin rolls with angular velocity $\vec{\omega}_{\rm Rot} = \dot{\theta}\hat{\rho}$ about axis $\hat{\rho}$ which in turn spins with angular velocity $\vec{\omega}_S = \dot{\phi}\hat{k}$ about fixed axis $\hat{k}$. The center of mass has velocity $\vec{v} = R\dot{\theta}\hat{t}$.

$$\delta S^{(\rm sh)} = \delta \int_{t_1}^{t_2} \left[L + \frac{d}{dt}(\mu_k f_k) - \dot{\mu}_k f_k \right] dt$$

$$= \delta \int_{t_1}^{t_2} [L - \dot{\mu}_k(t) f_k(\mathbf{q},t)] dt = 0, \qquad (54)$$

the holonomic form (52), as expected. The relationship between the multipliers is $\lambda_k = -\dot{\mu}_k$, as also shown in Sec. II C.

Hamilton's variational principle (51) cannot be generalized to inexact linear or more general nonholonomic constraints, Eq. (9) or (7), by replacing L by $L + \mu_k g_k$ in Eq. (51), as has been suggested.[1–3] The fact that $\delta g_k \neq 0$ for these cases implies that the varied paths are not geometrically possible. We have shown that generalization of Hamilton's and D'Alembert's principles rests on the multiplier rule which demands that the varied paths be geometrically possible, a property reserved only for holonomic and semiholonomic systems.

F. Validity of generalized principles and multiplier rule

The generalized principles of D'Alembert and Hamilton are effected by the multiplier rule (see the Appendix). The theorem (rule) applies only when all varied paths $(\mathbf{q} + \delta\mathbf{q})$ preserve the side conditions $g_k(\mathbf{q} + \delta\mathbf{q}, \dot{\mathbf{q}} + \delta\dot{\mathbf{q}}, t) = 0$, that is the $\delta\mathbf{q}$ variation causes no change $\delta g_k = 0$ to g_k. The displaced paths are then geometrically possible in that they satisfy the same equations of constraint. It is only for holonomic and semiholonomic constraints that the appropriate criteria, $\delta f_k = 0$ and $\delta g_k^{(\rm sh)} = 0$, are satisfied. For all nonholonomic constraints, the conditions $g_k = 0$ cannot be satisfied by the displaced paths and are therefore not good constant side conditions, as the multiplier rule demands. The invariance of the constraint equations to displacements is the key condition for application of the multiplier rule. The application of Eq. (6) to nonholonomic constraints is therefore without justification.

III. A TEST CASE

Some of these key points may be tested by the physical system depicted in Fig. 1. The solution of this spinning–rolling problem does not appear to have been provided in any standard textbook, although the limiting cases of rolling without spinning down a plane[1] and rolling–spinning on a horizontal plane[8,10] have been analyzed. Let $\mathbf{r}_{\rm c.m.} = x\hat{i} + y\hat{j} + z\hat{k}$ be the Cartesian coordinate of the center of mass (c.m.) of the coin of mass M and radius R, where the origin O is at the top of plane and where the directions $\hat{i}$, $\hat{j}$, and $\hat{k}$ form a Cartesian (X,Y,Z) fixed set of axes, with $\hat{i}$ pointing directly downward along the plane. Let θ and ϕ be the angles associated with the rolling and spinning motions about the symmetry axis (which is perpendicular to the coin) and the axis pointing along $\hat{k}$, the fixed outward normal to the plane. The Lagrangian is

$$L = \tfrac{1}{2} M(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2} I_S \dot{\theta}^2 + \tfrac{1}{2} I_D \dot{\phi}^2 + Mgx \sin\alpha, \qquad (55)$$

where $I_S = \beta M R^2$ and I_D are the moments of inertia of the body about the symmetry axis and the fixed Z-figure axis, respectively. Cases involving a solid sphere, coin, solid cylinder, spherical shell, hoop, or cylindrical shell, can be treated by taking $\beta = 2/5$, 1/2, 1/2, 2/3, 1, and 1, respectively.

Rolling without spinning: $\dot{y} = 0$, $\dot{\phi} = 0$. This example is a simple test of our proof that semiholonomic (exact linear) constraints $g_k^{(\rm sh)}(\mathbf{q}, \dot{\mathbf{q}}, t) = 0$ are covered by D'Alembert's and Hamilton's generalized principles, Eq. (49) or Eq. (53), respectively. The rolling constraint $g = \dot{x} - R\dot{\theta} = 0$ is exact so that the generalized principles should work. If we apply either Eq. (49) or (53) to the augmented Lagrangian,

$$L^{(\rm sh)}(\boldsymbol{\eta}, \dot{\boldsymbol{\eta}}) = \tfrac{1}{2} M\dot{x}^2 + \tfrac{1}{2} I_S \dot{\theta}^2 + Mgx \sin\alpha + \mu(\dot{x} - R\dot{\theta}), \qquad (56)$$

for the extended set $\boldsymbol{\eta} = (x, \theta, \mu)$ of free coordinates, we obtain the equations of state, $M\ddot{x} = Mg \sin\alpha - \dot{\mu}$, $I_S \ddot{\theta} = \dot{\mu} R$, and $\dot{x} = R\dot{\theta}$. When decoupled, these equations yield the acceleration $\ddot{x} = g \sin\alpha/(1+\beta)$ and the frictional constraint force $\dot{\mu}$ which produces the torque needed for rolling motion, $[(\beta/(1+\beta)]Mg \sin\alpha$, in agreement with standard results[1,8–10,13] obtained from holonomic theory, Eq. (22).

Rolling and spinning in two dimensions. We now test to see if linear conditions exist between the displacements δq_j needed for D'Alembert's basic principle (14) and then see if the constraints imply geometrically possible displaced paths, as needed for the generalized principles. The constraint for rolling is now

$$g_1'(\dot{x}, \dot{y}, \dot{\theta}) = [\dot{x}^2 + \dot{y}^2]^{1/2} - (R\dot{\theta})^2 = 0, \qquad (57)$$

which is nonintegrable and quadratic in the generalized velocities. There is no velocity component perpendicular to $\hat{\mathbf{v}}$ so that a second constraint is

$$g_2(\dot{x}, \dot{y}) = \dot{x} \sin\phi - \dot{y} \cos\phi = 0, \qquad (58)$$

which is also nonintegrable, but linear in the generalized velocities. That the coin remains upright implies that the center of mass coordinates (x,y) are also those for the point of contact of the coin with the plane and that $z = R$, a holonomic constraint which can be embedded from the outset unless the normal reaction (constraint) of the plane on the coin is sought. From Eqs. (57) and (58), the virtual displacements satisfy

$$(\delta x)^2+(\delta y)^2-R^2(\delta\theta)^2=0, \tag{59}$$

$$\delta x \sin\phi-\delta y \cos\phi=0. \tag{60}$$

The relation (60) is linear in δq_j and therefore amenable to being absorbed into D'Alembert's principle, Eq. (14). The quadratic relation (59) cannot be directly absorbed. Fortunately, for this case, the offending quadratic constraint (57) can be replaced by the combination $g_1^2=g_1'^2-g_2^2$ of g_1' and g_2 to give

$$g_1(\dot{x},\dot{y},\dot{\theta})=\dot{x}\cos\phi+\dot{y}\sin\phi-R\dot{\theta}=0, \tag{61}$$

which leads to the linear form,

$$\delta x\cos\phi+\delta y\sin\phi-R\,\delta\theta=0, \tag{62}$$

which is now suitable for application of D'Alembert's principle. The displaced paths $q_j+\delta q_j$ cause the changes,

$$\delta g_1=\delta\dot{x}\cos\phi+\delta\dot{y}\sin\phi-R\,\delta\dot{\theta}$$
$$-(\dot{x}\sin\phi-\dot{y}\cos\phi)\,\delta\phi, \tag{63a}$$

$$\delta g_2=\delta\dot{x}\sin\phi-\delta\dot{y}\cos\phi+(\dot{x}\cos\phi+\dot{y}\sin\phi)\,\delta\phi, \tag{63b}$$

in the constraint conditions (58) and (61). Because $\delta\dot{q}_j=d(\delta q_j)/dt$, then, on using the time derivatives of Eqs. (60) and (62) together with the relations (58)–(62), δg_1 and δg_2 reduce to 0 and $R(\dot{\theta}\delta\phi-\dot{\phi}\delta\theta)$, respectively. Therefore, the constraint (61) is semiholonomic. Integration yields the holonomic form $x^2+y^2-R^2\theta^2=0$. Because the sum $\delta(\lambda_k g_k)=\delta(\lambda_2 g_2)\neq 0$, we cannot use D'Alembert's or Hamilton's generalized principles, Eqs. (29) and (53), respectively, as predicted.

Because the conditions (60) and (62) on the displacements are now all linear, the problem can be solved by D'Alembert's basic principle (14), or by its time-integrated version, Hamilton's integral principle (50). The solution is straightforward and reduces to the standard results[8,10] for horizontal motion ($\alpha=0$).

IV. SUMMARY AND CONCLUSIONS

This paper has presented the basic reason why Hamilton's variational principle and the more basic principle of D'Alembert cannot be generalized by substituting the augmented Lagrangian Eq. (8) in either Eq. (2) or Eq. (14) to cover general nonholonomic constraints, as the multiplier rule[5–7] in the calculus of variations might suggest.[1–3] The multiplier rule requires that the side conditions $g_k=0$ be satisfied by all varied paths, which must therefore be geometrically possible. The displacements δq_j in nonholonomic systems violate this rule because they cause nonzero changes $\delta g_k\neq 0$ in the constraint conditions and the displaced paths are not geometrically possible. The constraint $g_k=0$ is satisfied only by the actual physical path $\mathbf{q}(t)$ in configuration space and not by the individual members of the family of varied paths for nonholonomic systems. The multiplier rule cannot therefore be used to generalize Hamilton's or D'Alembert's principles to cover nonholonomic constraints. It can however be applied to all holonomic and semiholonomic (exact linear) constraints which have the property that all displaced paths are geometrically possible in accord with the multiplier rule.

We have traced the development of various generalized principles from D'Alembert's basic principle in such a way as to render transparent their scope of application. It is useful to keep the following conclusions in mind.

(1) D'Alembert's basic principle, Eq. (14), is the most fundamental of all the principles considered here.
(2) D'Alembert's basic principle, Eq. (14), and Hamilton's variational principle, Eq. (2), are well designed for holonomic systems. Equation (16) is the equation of state.
(3) When constraint forces in holonomic systems are sought, D'Alembert's generalized principle, Eq. (22), and Hamilton's generalized principle, Eq. (5), are appropriate, because the varied paths under holonomic constraints are all geometrically possible and the underlying multiplier rule is then valid. Equation (6) is the equation of state.
(4) The correct equations of state (28) for general linear nonholonomic constraints are furnished only by D'Alembert's basic principle, Eq. (14), or its time-integrated version, Hamilton's integral principle, Eq. (50).
(5) As shown here, the generalized principles, Eqs. (29) and (53), are valid for semiholonomic systems. In these generalized principles, the constraints are automatically included and the displacements $\delta\eta_j$ are all free. Equation (49) is the equation of state for semiholomic systems, that is, those which satisfy conditions for exactness and therefore geometrically possible displaced paths.
(6) Generalized principles are inappropriate for linear nonholonomic constraints, because the constraint equations $g_k=0$ are not exact and change from varied path to varied path. The underlying multiplier rule then loses validity.
(7) The theory for nonholonomic constraints with a general velocity dependence remains outside the scope of the most fundamental principle, Eq. (14) of D'Alembert. It is impossible to extract from the equations $g_k=0$ of general nonholonomic constraints the linear relation between the δq_j's required for the application of D'Alembert's principle unless the constraints are either linear in velocity or holonomic. Nonholonomic constraints are therefore outside the scope of any of the principles based on D'Alembert's principle.

The above conclusions reflect the intrinsic merit of reconstructing the variational principle, Eq. (2), from the more fundamental D'Alembert principle, Eq. (14) via Eq. (47), so that the validity of the various stages involved becomes directly exposed. Pitfalls[1–3] can easily occur by arbitrarily invoking the multiplier rule to assert generalized principles such as Eqs. (29) and (53), without first ascertaining the critical but hidden condition that the varied paths must be geometrically possible. We have shown here that the condition is satisfied only for holonomic and semiholonomic systems.

General nonholonomic constraints (7) can be analyzed by other principles[13] that involve, for example, the virtual velocity (Jourdain) displacements, constructed by maintaining both the configuration $\mathbf{q}$ and time t fixed, in contrast to virtual displacements δq_j which keep only t fixed. The Jourdain variational principle is the subject of a separate paper.[14]

APPENDIX: THE MULTIPLIER THEOREM

We will determine the paths $q_i(t)$; $i=1,2,\ldots,n$ that provide an extremum to the functional

$$J=\int_{t_1}^{t_2} F(\mathbf{q},\dot{\mathbf{q}},t)\,dt, \tag{A1}$$

subject to the $c<n$-finite auxiliary (side) conditions

$$g_k(\mathbf{q},\dot{\mathbf{q}},t)=0. \quad (k=1,2,\ldots,c). \tag{A2}$$

A basic theorem[5–7] in the calculus of variations can be invoked, provided we admit to the variational competition only those curves $\mathbf{q}(t)$ that satisfy fixed end-point conditions $\delta\mathbf{q}(t_{1,2})=0$ and c-finite fixed side conditions as in Eq. (A2). The varied curves must all be geometrically possible by satisfying $g_k(\mathbf{q}+\delta\mathbf{q},\dot{\mathbf{q}}+\delta\dot{\mathbf{q}},t)=g_k(\mathbf{q},\dot{\mathbf{q}},t)+\delta g_k(\mathbf{q},\dot{\mathbf{q}},t)=0$, so that $\delta g_k=0$. The physical path $\mathbf{q}(t)$ is then determined by the extremum determined by the free variation of the modified functional,

$$J^\dagger=\int_{t_1}^{t_2} F^\dagger(\mathbf{q},\dot{\mathbf{q}},t)\,dt\equiv\int_{t_1}^{t_2}[F(\mathbf{q},\dot{\mathbf{q}},t)+\lambda_k(t)g_k(\mathbf{q},\dot{\mathbf{q}},t)]\,dt, \tag{A3}$$

without any side conditions imposed. The physical path $\mathbf{q}(t)$ then satisfies the Euler–Lagrange system of equations,

$$\frac{d}{dt}\left(\frac{\partial F^\dagger}{\partial\dot{\eta}_j}\right)-\frac{\partial F^\dagger}{\partial\eta_j}=0, \quad (j=1,2,\ldots,n+c) \tag{A4}$$

for the extended set $\boldsymbol{\eta}\equiv\{q_1,q_2,\ldots,q_n,\lambda_1,\lambda_2,\ldots,\lambda_c\}$ of $(n+c)$ variables. Because $F^\dagger$ does not depend on $\dot{\lambda}_k(t)$, the last c members of the set of equations (A4) reproduce the side conditions (A2). The validity of the multiplier theorem, Eqs. (A3) and (A4), rests on the fact that conditions (A2) must be satisfied by *all* the varied paths therein, that is, $\delta g_k=0$. This condition is satisfied for holonomic and semiholonomic constraints. It is not satisfied for nonholonomic constraints because $\delta g_k\neq 0$ for this case; the condition $g_k=0$ is satisfied only by the physical path to be eventually determined. The theorem is therefore irrelevant to nonholonomic systems.

However, the multiplier theorem is directly relevant to the related principle[1,8,11,13]

$$\Delta S_A=\Delta\int_{t_1}^{t_2} p_i\dot{q}_i\,dt=[p_i\Delta q_i]_{t_1}^{t_2}=0 \tag{A5}$$

of least abbreviated action S_A, valid for varied curves, all chosen to obey the same constant Hamiltonian H and to pass through the end points, that is, $\Delta q_i(t_{1,2})=0$. It is similar in form to Eqs. (50) and (51). The Δ operator causes nonsimultaneous variations $\Delta q_i=\delta q_i+\dot{q}_i\Delta t$, which also involve displacements Δt in time, in addition to the usual virtual displacements δq_i. When the kinetic energy T reduces to a homogeneous quadratic function $T_2=\frac{1}{2}M_{ij}(\mathbf{q})\dot{q}_i\dot{q}_j$ of the generalized velocities $\dot{q}_i$, then $p_i\dot{q}_i=2T$ and the least action principle, Eq. (A5), reduces to the Euler–Lagrange–Maupertuis principle,[1,8,11,13]

$$\Delta\int_{t_1}^{t_2} 2T\,dt=0, \tag{A6}$$

of least action. The multiplier theorem, Eqs. (A3) and (A4), can now be applied to extract Lagrange's equations from Eq. (A6). The condition for the variation (A6) is that the Hamiltonian H does not depend on time and remains fixed at the same value for all the paths considered. In the sense that $(t,-H)$ are conjugate variables, the principle (A6), which admits paths with the same constant H, is complementary to Hamilton's variational principle, Eq. (2), which admits only those paths with the same transit times $\tau=t_1-t_2$ into the variation. For $T=T_2$, H equals the total energy $E=T+V$, so that Eq. (A6) becomes modified, under the fixed constraint $g=(T+V)-E=0$ for all varied paths, to finding a stationary value of

$$\Delta\int_{t_1}^{t_2}[2T(\mathbf{q},\dot{\mathbf{q}})+\lambda(t)\{T(\mathbf{q},\dot{\mathbf{q}})+V(\mathbf{q})-E\}]\,dt=0. \tag{A7}$$

The application[8,11,13] of this modified version (A7) of the Euler–Lagrange–Maupertuis principle leads directly to the standard Lagrange's equations (16), with $Q_j^{\mathrm{NP}}=0$ for potential systems.

[a]Electronic mail: ray.flannery@physics.gatech.edu

[1]H. Goldstein, C. Poole, and J. Safko, *Classical Dynamics* (Addison–Wesley, New York, 2002), 3rd ed., pp. 46–48. This generalized action principle is retracted without explanation at ⟨http://astro.physics.sc.edu/Goldstein/⟩, pp. 356–360.

[2]G. H. Goedecke, "Undetermined multiplier treatments of the Lagrange problem," Am. J. Phys. **34**, 571–574 (1966).

[3]J. R. Ray, "Nonholonomic constraints," Am. J. Phys. **34**, 406–408 (1966).

[4]J. R. Ray, "Erratum: Nonholonomic constraints," Am. J. Phys. **34**, 1202–1203 (1966).

[5]D. A. Pierre, *Optimization Theory with Applications* (Dover, New York, 1986), p. 92.

[6]I. M. Gelfand and S. Fomin, *Calculus of Variations* (Dover, New York, 2000), p. 48.

[7]L. Elsgolc, *Calculus of Variations* (Addison–Wesley, Reading, MA, 1962), p. 136.

[8]R. M. Rosenberg, *Analytical Dynamics of Discrete Systems* (Plenum, New York, 1977), pp. 176–178, 220, 232–235, 265–268.

[9]D. T. Greenwood, *Classical Dynamics* (Dover, New York, 1997), pp. 159–162.

[10]J. V. José and E. J. Saletan, *Classical Dynamics* (Cambridge UP, Cambridge, 1998), pp. 114–116.

[11]L. A. Pars, *An Introduction to the Calculus of Variations* (Wiley, New York, 1962), pp. 253, 249.

[12]I. Gatland, "Nonholonomic constraints: A test case," Am. J. Phys. **72**, 941–942 (2004).

[13]L. A. Pars, *A Treatise on Analytical Dynamics* (Ox Bow, Woodbridge, CT, 1979), pp. 543–545.

[14]M. R. Flannery, "Theory for general nonholonomic constraints," (unpublished).