PERGAMON International Journal of Non-Linear Mechanics 37 (2002) 1079–1090

INTERNATIONAL JOURNAL OF NON-LINEAR MECHANICS

www.elsevier.com/locate/ijnonlinmec

On the foundations of analytical dynamics

F.E. Udwadia[a,*], R.E. Kalaba[b]

[a] *Aerospace and Mechanical Engineering, Civil Engineering, Mathematics, and Information and Operations Management, 430K Olin Hall, University of Southern California, Los Angeles, CA 90089-1453, USA*

[b] *Biomedical Engineering and Economics, University of Southern California, Los Angeles, CA 90089, USA*

Abstract

In this paper, we present the general structure for the explicit equations of motion for general mechanical systems subjected to holonomic and non-holonomic equality constraints. The constraints considered here *need not* satisfy D'Alembert's principle, and our derivation is *not* based on the principle of virtual work. Therefore, the equations obtained here have general applicability. They show that in the presence of such constraints, the constraint force acting on the system can always be viewed as made up of the sum of two components. The explicit form for each of the two components is provided. The first of these components is the constraint force that *would have existed*, were all the constraints ideal; the second is caused by the non-ideal nature of the constraints, and though it needs specification by the mechanician and depends on the particular situation at hand, this component nonetheless has a specific form. The paper also provides a generalized form of D'Alembert's principle which is then used to obtain the explicit equations of motion for constrained mechanical systems where the constraints may be non-ideal. We show an example where the new general, explicit equations of motion obtained in this paper are used to directly write the equations of motion for describing a non-holonomically constrained system with non-ideal constraints. Lastly, we provide a geometrical description of constrained motion and thereby exhibit the simplicity with which Nature seems to operate. © 2002 Elsevier Science Ltd. All rights reserved.

1. Introduction

Ever since the initial description of constrained motion given by Lagrange [1] the problem of obtaining the equations of motion for constrained mechanical systems has been of considerable interest to both mathematicians and engineers. Various ways have been devised [2–7] for determining these equations under the assumption that the work done by the forces of constraint under virtual displacements is always zero. Such constraints are often referred to as ideal constraints, and the assumption that they do no work under virtual displacements is referred to as D'Alembert's principle [8]. However, the inclusion of general constraint forces (which may do work under virtual displacements) within the framework of Lagrangian mechanics has posed a considerable problem to date. Such forces indeed do exist and are in fact commonplace in Nature, like the force of sliding friction. No general equation of motion within the Lagrangian framework for such systems has so far been developed. As stated by Goldstein (1981), "This [total work done by forces of constraint equal to zero] is no longer true if sliding friction is present, and we must exclude such systems from our [Lagrangian] formulation" [9].

* Corresponding author. Tel.: +1-213-740-0484; fax: +1-213-740-8071.

E-mail address: fudwadia@usc.edu (F.E. Udwadia).

0020-7462/02/$ - see front matter
PII: S0020-7462(01)00033-6

And Pars (1979, p. 14) in his treatise [5] on analytical dynamics writes, "There are in fact systems for which the principle enunciated [D'Alembert's Principle] … does not hold. But such systems will not be considered in this book."

In this paper we obtain the general form of the equation of motion for holonomically and non-holonomically constrained systems that is valid *independent* of whether or not the constraints satisfy D'Alembert's principle. Unlike our previous results [10,11], we derive the general form of the equation of motion *without* considerations of virtual work. We find that the total constraint force can always be thought of as made up of two components. The first component is ever-present and can be uniquely determined from information about the unconstrained system and the kinematic equations of constraint that this unconstrained system is required to satisfy. It is the force of constraint that would have arisen had all the constraints been ideal. The second component is caused by the presence of the non-ideal nature of the constraints, if such non-ideal constraints exist. We show that this second component *must* appear in the equation of motion of a constrained mechanical system in a *specific form*, though. Using this form we then particularize these non-ideal constraint forces to situations where the work that they do under virtual displacements is known and prescribed by the mechanician. This leads us to state a generalization of D'Alembert's principle, and we show that for constrained systems that satisfy this generalized principle one obtains an explicit and unique equation of motion. This equation is then used to further understand the general form of the equation previously obtained. An example (a generalization of Appell's problem [12]) showing the application of our results to a non-holonomically constrained system with non-ideal constraints is provided. The last section of the paper exposes the geometry of constrained motion with non-ideal constraints.

2. Statement of problem

Consider an unconstrained mechanical system described by the Lagrange equations

$$M(q,t)\ddot{q} = Q(q,\dot{q},t), \;\; q(0) = q_0, \;\; \dot{q}(0) = \dot{q}_0, \quad (1)$$

where $q(t)$ is the n-vector (i.e. n by 1 vector) of generalized coordinates, M is an $n \times n$ symmetric, positive-definite matrix, Q is the 'known' n-vector of impressed (also, called 'given') forces, and the dots refer to differentiation with respect to time. By unconstrained, we mean that the components of the n-vector $\dot{q}_0$ can be arbitrarily specified. By 'known', we mean that Q is a known function of its arguments. The acceleration, a, of the unconstrained system at any time t is then given by the relation $a(q,\dot{q},t) = M^{-1}(q,t)Q(q,\dot{q},t)$.

We shall assume that this system is subjected to a set of $m = h + s$ consistent equality constraints of the form

$$\varphi(q,t) = 0 \quad (2)$$

and

$$\psi(q,\dot{q},t) = 0, \quad (3)$$

where φ is an h-vector and ψ an s-vector. Furthermore, we shall assume that the initial conditions q_0 and $\dot{q}_0$ satisfy these constraint equations at time $t = 0$, i.e., $\varphi(q_0, 0) = 0$, and $\psi(q_0, \dot{q}_0, 0) = 0$.

Assuming that Eqs. (2) and (3) are sufficiently smooth, [1] we differentiate Eq. (2) twice with respect to time, and Eq. (3) once with respect to time, to obtain an equation of the form:

$$A(q,\dot{q},t)\ddot{q} = b(q,\dot{q},t), \quad (4)$$

where the matrix A is $m \times n$, and b is the m-vector that results from carrying out the differentiations. We place no restrictions on the rank of the matrix A.

This set of constraint equations includes among others, the usual holonomic, non-holonomic, scleronomic, rheonomic, catastatic and acatastatic varieties of constraints; combinations of such constraints may also be permitted in Eq. (4). It is important to note that Eq. (4), together with the initial conditions, is equivalent to Eqs. (2) and (3).

Consider now any instant of time t. When the equality constraints (Eqs. (2) and (3)) are imposed

[1] We assume throughout this paper that the presence of constraints does not change the rank of the matrix M. This is almost always true in mechanical systems.

at that instant of time on the unconstrained system, the motion of the unconstrained system is, in general, altered from what it would have been (at that instant of time) in the absence of these constraints. We view this alteration in the motion of the unconstrained system as being caused by an additional set of forces, called the 'forces of constraint', acting on the system at that instant of time. The equation of motion of the constrained system can then be expressed as

$$M(q,t)\ddot{q} = Q(q,\dot{q},t) + Q^{c}(q,\dot{q},t),$$
$$q(0) = q_0, \ \ \dot{q}(0) = \dot{q}_0, \tag{5}$$

where the additional 'constraint force' n-vector, $Q^{c}(q,\dot{q},t)$, arises by virtue of the constraints (2) and (3) imposed on the unconstrained system, which is described by Eq. (1). Our aim is to determine a *general* explicit form for Q^{c} at any time t. In what follows, for brevity, we shall suppress the arguments of the various quantities, unless necessary for purposes of clarification.

Following Gauss [2], another way of viewing constrained motion is as follows. Let us premultiply Eq. (5) by M^{-1}. We then obtain the acceleration equation

$$\ddot{q} = a + \ddot{q}^{c}, \tag{6}$$

where $a = M^{-1}Q$, and $\ddot{q}^{c} = M^{-1}Q^{c}$. Now imagine that the constrained mechanical system has evolved up until some time, say t. We assume that we know the velocity of the constrained system, $\dot{q}(t)$, and its configuration, $q(t)$, at that time. Our aim is to find the *subsequent* motion of the constrained system at the next instant of time. This would naturally be accomplished if we knew the acceleration, $\ddot{q}(t)$, of the constrained system at time t. Now since the acceleration of the unconstrained system, a, is a function of q, $\dot{q}$, and t, it is, by assumption, known at time t. Hence we see that the problem of constrained motion can be viewed as requiring us to find the deviation, $\Delta\ddot{q}$, at each instant of time, of the acceleration of the constrained system from that it would have had, were there no constraints (i.e., from that of the unconstrained system). This deviation can be expressed, using Eq. (6), as

$$\Delta\ddot{q}(t) = \ddot{q} - a = \ddot{q}^{c}. \tag{7}$$

In what follows, instead of the accelerations we shall use, for convenience, the 'scaled' accelerations defined by

$$\ddot{q}_{s} = M^{1/2}\ddot{q}, \tag{8}$$

$$a_{s} = M^{-1/2}Q = M^{1/2}a, \tag{9}$$

and

$$\ddot{q}_{s}^{c} = M^{-1/2}Q^{c} = M^{1/2}\ddot{q}^{c}. \tag{10}$$

Thus 'scaling' consists of simply premultiplication of the acceleration n-vectors at any time t, by the matrix $M^{1/2}$ at that time. Eq. (6) can now be written in terms of the scaled accelerations. These scaled accelerations then satisfy, at each time t, the relation

$$\ddot{q}_{s} = a_{s} + \ddot{q}_{s}^{c}, \tag{11}$$

and, by Eq. (4), also the relation

$$B\ddot{q}_{s} = b, \tag{12}$$

where

$$B = AM^{-1/2}. \tag{13}$$

We note that since, by assumption we know q and $\dot{q}$ at time t, the quantities M, A, B, Q, and b are all known at time t. Eq. (11), therefore, informs us that at time t, the (scaled) acceleration of the constrained system, $\ddot{q}_{s}$, deviates from the known (scaled) acceleration of the unconstrained system, a_{s}, by the quantity $\ddot{q}_{s}^{c}$. (As before, the unconstrained acceleration a_{s} is known at time t, because $a_{s} = M^{-1/2}Q$.) This (scaled) deviation is caused by the presence of the constraints (12) and their nature, and it is the *general form* of this deviation, $\ddot{q}_{s}^{c}$, that we shall now determine. By virtue of Eq. (10) the knowledge of this deviation is tantamount to the knowledge of the constraint force Q^{c}.

3. General form for the equation of motion for constrained mechanical systems

We begin by considering the matrices $T = B^{+}B$ and $N = (I - B^{+}B)$, where the matrix B^{+} is the Moore–Penrose (MP) inverse of the matrix B [13].[2] The matrix T is an orthogonal

[2] Some of the basic properties of the Moore–Penrose inverse used in this paper may be found in Chapter 2 of Reference 8.

projection operator since $(B^+B)^T = B^+B$, and $T^2 = (B^+B)(B^+B) = B^+B = T$. Both these results follow from the definition of the MP inverse of the matrix B. Also, N is an orthogonal projection operator since $(I - B^+B)^T = I - (B^+B)^T = I - B^+B$, and $N^2 = N$.

Furthermore, given an n-vector, w, the vector Nw is such that $B(Nw) = B(I - B^+B)w = 0$; hence Nw belongs to the null space of the matrix B. Also the vector Tw belongs to the range space of B^T [8]. Since $R^n = \mathscr{R}(B^T) \oplus \mathscr{N}(B)$, any n-vector w has a unique orthogonal decomposition $w = B^+Bw + (I - B^+B)w$; and so also our n-vector $\ddot{q}_s$. We can thus write the identity

$$\ddot{q}_s = B^+B\ddot{q}_s + (I - B^+B)\ddot{q}_s, \tag{14}$$

where the right-hand side simplifies to $\ddot{q}_s$. Since the (scaled) acceleration of the constrained system must satisfy Eqs. (11) and (12), we can replace, on the right-hand side of Eq. (14), $\ddot{q}_s$ in the second member by $(a_s + \ddot{q}_s^c)$, and $B\ddot{q}_s$ in the first member by the m-vector b. This yields the relation

$$\begin{aligned}\ddot{q}_s &= B^+b + (I - B^+B)(a_s + \ddot{q}_s^c)\\ &= B^+b + (I - B^+B)a_s + (I - B^+B)\ddot{q}_s^c,\end{aligned} \tag{15}$$

which can also be expressed as

$$\ddot{q}_s = a_s + B^+(b - Ba_s) + (I - B^+B)\ddot{q}_s^c. \tag{16}$$

Comparing the right-hand sides of Eqs. (11) and (16) we see that

$$B^+B\ddot{q}_s^c = B^+(b - Ba_s) \tag{17}$$

which, upon solving for $\ddot{q}_s^c$, gives

$$\begin{aligned}\ddot{q}_s^c &= B^+BB^+(b - Ba_s) + \{I - (B^+B)^+(B^+B)\}z\\ &= B^+(b - Ba_s) + (I - B^+B)z\end{aligned} \tag{18}$$

for *any* n-vector z. We have thus obtained the *most general form* that the (scaled) acceleration, $\ddot{q}_s^c$, of a mechanical system subjected to the constraints (2) and (3) can have. From Eq. (10), we note that the force of constraint $Q^c = M^{1/2}\ddot{q}_s^c$, and Eq. (18) becomes

$$Q^c = M^{1/2}B^+(b - Aa) + M^{1/2}(I - B^+B)z, \tag{19}$$

where we have used the relations (9) and (13) to yield $Ba_s = BM^{1/2}a = Aa$. This then is the *most general form* in which the force of constraint can appear in a mechanical system constrained by Eqs. (2) and (3).

We note that to obtain the unique constraint force acting on a given constrained mechanical system, we need a *further* specification of the constraints (beyond that provided through the knowledge of the matrix A and the vector b); that is, we need a *specification* of the vector z at each instant of time. In the next section we shall consider this in detail.

We now consider the first member on the right-hand side of Eq. (19). We note that were the acceleration, $a = M^{-1}Q$, of the unconstrained system at time t to be inserted into the equation of constraint (4), this equation would not, in general, be satisfied at that time. The extent to which the constraint (Eq. (4)) would not be satisfied by this acceleration, a, of the unconstrained system at time t would then be given by

$$e = b - Aa. \tag{20}$$

Eq. (19) can now be rewritten as

$$Q^c = M^{1/2}B^+e + M^{1/2}(I - B^+B)z \tag{21}$$

and, using Eqs. (18), (7) and (10), the deviation, $\Delta\ddot{q}$, becomes

$$\Delta\ddot{q} = \ddot{q} - a = M^{-1/2}B^+e + M^{-1/2}(I - B^+B)z. \tag{22}$$

Eqs. (21) and (22) encapsulate a new fundamental principle of mechanics which we state in two equivalent statements.

1. A constrained mechanical system evolves in such a way that, at each instant of time, the deviation, $\Delta\ddot{q}$, of its acceleration from what it would have been at that instant had there been no constraints on it, is given by a sum of two components: the first component is proportional to the extent, e, to which the unconstrained acceleration does not satisfy the constraints at that instant of time, the matrix of proportionality being the matrix $M^{-1/2}B^+$; the second is proportional to an n-vector z that needs, in general,

to be specified at each instant of time, the matrix of proportionality being $M^{-1/2}(I - B^+B)$, where $B = AM^{-1/2}$.

2. At each instant of time, the force of constraint acting on a constrained mechanical system is made up of two components: the first component is proportional to the extent, e, to which the unconstrained acceleration of the system does not satisfy the constraints at that instant of time, and the matrix of proportionality is $M^{1/2}B^+$; the second is proportional to an n-vector z that, in general, needs further specification at each instant of time, the matrix of proportionality being $M^{1/2}(I - B^+B)$, where $B = AM^{-1/2}$. This vector z is specific to a given mechanical system and needs to be prescribed by the mechanician who is modeling the system.

We note that these statements are natural extensions of those previously obtained only for the case of ideal constraints [7]. Using Eq. (21) in Eq. (5) we are now ready to state the following fundamental result.

Result 1. The equation of motion of *any* mechanical system subjected to the constraints (2) and (3) has, at each instant of time t, the *explicit general form*

$$\begin{aligned} M\ddot{q} &= Q + M^{1/2}B^+(b - Aa) + M^{1/2}(I - B^+B)z \\ &= Q + Q^c, \end{aligned} \tag{23}$$

where $B = AM^{-1/2}$, and $a = M^{-1}Q$. The n-vector z needs to be specified at each instant of time t, and it depends on the nature of the specific mechanical system under consideration. In general, A, B, a, b, and z are each functions of q, $\dot{q}$ and t; the matrix M is a function of q and t.

We remark that the generality of this result stems from the fact that we have been led to it, somewhat surprisingly, simply through the use of Eqs. (11)–(14). *No appeal to the principle of virtual work was made in obtaining it*, *and therefore the result is valid*, *irrespective of whether the constraint forces are ideal or not*.

We next explore the specification of the vector z, which we shall show reflects the *nature of the non-ideal constraints* in any given mechanical system whose equation of motion we want to obtain.

As we shall soon see, it is useful to express the total force of constraint, Q^c, as being made up of the two additive components:

$$Q_i^c = M^{1/2}B^+(b - Aa) \tag{24}$$

and

$$Q_{ni}^c = M^{1/2}(I - B^+B)z, \tag{25}$$

so that the total constraint force n-vector, Q^c, is given by

$$Q^c = Q_i^c + Q_{ni}^c \tag{26}$$

and Eq. (23) takes the less formidable form

$$M\ddot{q} = Q + Q_i^c + Q_{ni}^c. \tag{27}$$

The reason for using the subscripts 'i' and 'ni' will become clear in the following section.

4. Understanding the two components Q_i^c and Q_{ni}^c of the force of constraint Q^c

To understand the two components of the total constraint force n-vector, Q^c, given by Eqs. (24)–(26) we note firstly that the component, Q_i^c, is explicitly known and is dependent on the description of: the unconstrained system given by Eq. (1); and, the constraints as given by Eq. (4). It thus only depends on the *four* known entities A, M, Q, and b. By 'known' we mean, as usual, known functions of their arguments.

The second component, Q_{ni}^c, of Q^c is dependent on a vector z (which may be, in general, a function of q, $\dot{q}$ and t) that needs to be suitably specified for a given mechanical system, *beyond* the specification of the four quantities just mentioned before. We shall now attempt to understand the physical meaning of the specification of this vector z, that appears in Eq. (23) and remains to be specified so that the appropriate equations of motion relevant to a specific constrained mechanical system may be obtained.

We begin by showing that if the constraints satisfy D'Alemebert's principle, then the n-vector $Q_{ni}^c \equiv 0$ for all time. For, D'Alembert's principle states that at each instant of time t, and for all virtual displacements, v, at that time t, the work done by the force of constraint, Q^c, under these virtual displacements, v, must be zero; that is, $v^T Q^c = 0$ at *each* instant of time t. A virtual displacement [8] at time t is any non-zero n-vector, v, such that at that time, $Av = 0$.

Let us consider a specific instant of time t, and assume that $v^T Q^c = 0$, but *only* at that instant of time t. We shall refer to this assumption as D'Alembert's *prescription*. When D'Alembert's prescription is satisfied at every instant of time, we obviously obtain D'Alembert's Principle, and the constraints are then called ideal. If D'Alembert's Principle is not satisfied, the constraint forces (and, for short, the constraints) will be referred to as non-ideal.

Setting $v = M^{-1/2}\mu$, we note that $Av = 0$ implies $AM^{-1/2}(M^{1/2}v) = B\mu = 0$. D'Alembert's prescription then requires that at the instant of time t,

$$\begin{aligned}\{\mu \mid B\mu = 0, \mu \neq 0\}\} \Rightarrow \mu^T M^{-1/2} Q^c &= \mu^T M^{-1/2} Q^c_i + \mu^T M^{-1/2} Q^c_{ni} \\ &= 0. \end{aligned} \tag{28}$$

The first term on the left-hand side of the last equality in Eq. (28) gives the work done at time t under virtual displacements by the component Q^c_i, and the second term the work done by the component Q^c_{ni}.

But $B\mu = 0$ implies $\mu = (I - B^+B)u$, for any n-vector u at that time [8]. Thus relation (28) implies that for *all* n-vectors u, $u^T(I - B^+B)M^{-1/2}Q^c = 0$ at time t. By Eqs. (24)–(28) we then require that at time t, for all n-vectors u,

$$\begin{aligned} u^T(I - B^+B)M^{-1/2}Q^c &= u^T(I - B^+B)M^{-1/2}Q^c_i \\ &\quad + u^T(I - B^+B)M^{-1/2}Q^c_{ni} \\ &= u^T(I - B^+B)B^+(b - Aa) \\ &\quad + u^T(I - B^+B)(I - B^+B)z \\ &= 0 + u^T(I - B^+B)z = 0. \end{aligned} \tag{29}$$

Notice that we have explicitly pointed out that the first member on the right in the second equation above is always zero, proving that the *work done by the component, Q^c_i, of the total constraint force, Q^c, under virtual displacements is always* (*for all time*) *zero*.

Let us further decompose the n-vector z at time t into its orthogonal components $z = z_n + z_r$, where z_n belongs to the null space of the matrix B, and $z_r = B^T w$ (for some m-vector w) belongs to the range space of B^T. Then the last equality in Eq. (29) requires that at time t, for *all* n-vectors u,

$$\begin{aligned} u^T(I - B^+B)z &= u^T(I - B^+B)(z_n + z_r) \\ &= u^T z_n + u^T(I - B^+B)B^T w \\ &= u^T z_n = 0. \end{aligned} \tag{30}$$

For the last equality in Eq. (30) to be valid for all n-vectors u, we must have $z_n = 0$, and therefore we see that at *each instant of time t* at which the constraint force n-vector Q^c satisfies D'Alembert's prescription, $z = z_r = B^T w$, for some m-vector w. Furthermore, at each such instant of time we have

$$\begin{aligned} Q^c_{ni} &= M^{1/2}(I - B^+B)z \\ &= M^{1/2}(I - B^+B)B^T w \\ &= M^{1/2}\{I - (B^+B)^T\}B^T w \\ &= M^{1/2}\{B^T - B^T(B^T)^+B^T\} = 0. \end{aligned} \tag{31}$$

Hence at those instants of time at which the constraint forces satisfy D'Alembert's prescription, we must have $Q^c = Q^c_i$, and $Q^c_{ni} = 0$! When the constraints are ideal and therefore satisfy D'Alembert's prescription for *every* instant of time, (i.e., they satisfy D'Alembert's Principle), $Q^c = Q^c_i$ is then always true. We have shown the following important result.

Result 2. At each instant of time at which the forces of constraint obey D'Alembert's prescription, the equation of motion is *uniquely* determined *solely* through knowledge of: (a) the unconstrained system as contained in the known matrix M and the known n-vector Q, and, (b) the constraints as contained in Eqs. (2) and (3) (or alternatively by Eq. (4)), and described by the $m \times n$ matrix A and the m-vector b. The contribution, Q^c_{ni}, to the total constraint force n-vector, Q^c, is zero at that instant, and the explicit equation of motion at that instant is then given by

$$M\ddot{q} = Q + Q^c_i = Q + M^{1/2}B^+(b - Aa). \tag{32}$$

When the constraints satisfy D'Alembert's Principle, Eq. (32) is valid for all instants of time. The equation of motion of the constrained system becomes (32). The second member on the right-hand side in Eq. (32) explicitly gives the unique constraint force that the mechanical system is subjected to under the assumption that *all* the constraints are ideal.

We point out that Eq. (32) agrees with the explicit equation of motion given by Udwadia and Kalaba [7,8] for constrained systems that satisfy D'Alembert's principle. Comparing Eq. (32) with the general form (23) stated in Result 1, we obtain the following fundamental result.

Result 3. In *any* constrained motion of a mechanical system subjected to the constraints (2) and (3), whether or not the constraints are ideal, the total force of constraint, $Q^c = Q^c_i + Q^c_{ni}$, is made up of two additive components. The component $Q^c_i = M^{1/2}B^+(b - Aa)$ is *always* present in general, and it is the force of constraint that *would have been generated had all the constraints been ideal*. The work done by this component Q^c_i of the constraint force under virtual displacements is always *zero*. The second component Q^c_{ni} of the total constraint force is caused by the non-ideal nature of the constraints.

The work done by the constraint force, Q^c, under any virtual displacement v at any time t is, therefore, always given by

$$v^T Q^c = v^T Q^c_i + v^T Q^c_{ni} = v^T Q^c_{ni}, \tag{33}$$

because, as pointed out in the comment following Eq. (29), at each instant of time $v^T Q^c_i = 0$ for all virtual displacements, v, at that instant. We then have the next result.

Result 4. Only the component, Q^c_{ni}, of the total constraint force, Q^c, can do work under virtual displacements. The amount of work done by this component of the total constraint force at any instant of time t under a virtual displacement, v, at that time, always has the explicit form

$$v^T Q^c = v^T Q^c_{ni} = v^T M^{1/2}(I - B^+B)z \tag{34}$$

and, in general, at each instant of time t, the work done may be *positive*, *negative or zero*. (At instants of time when this work done is zero, the constraints satisfy D'Alembert's prescription, and hence $Q^c_{ni} = 0$, at those instants, by our previous result.)

The component Q^c_{ni} depends on the specification of the vector z for any given mechanical system. Using Eq. (34), for a given constrained mechanical system, the vector z (which, in general, may be a function of q, $\dot{q}$ and t) may be viewed as providing the needed *additional* specification related to the constraints in the system (beyond that contained in the matrix A and the vector b) that informs us of the extent of work done by the constraint forces at each instant of time t. During such a specification, it should be remembered that *no* work is done under virtual displacements by the component z_r of z that lies in the range space of the matrix B^T.

We now provide further insight into the physical meaning of the vector z. Consider again the instant of time t, and a given virtual displacement v of the constrained system at that instant. Furthermore, let us say that we could prescribe the work done, $W^c(t)$, by the total constraint force n-vector, Q^c, under any given infinitesimal virtual displacement v at the instant t, by the quantity

$$W^c(t) = v^T(t)C(q, \dot{q}, t) \tag{35}$$

which may be *positive*, *negative or zero* at that instant of time. The n-vector C is, in general, time-dependent.

For any given mechanical system at hand, C comes from a proper examination of the specific system at hand; it depends on our understanding of the nature of the constraint forces specific to the given system, as best discernable to the mechanician. In particular, it would be different for a body undergoing sliding friction (subjected to a set of impressed forces) on a rough surface as opposed to a smooth surface (and to the same set of impressed forces). The n-vector C has the dimensions of force and can be thought of as a kind of 'generalized force' for describing the non-ideal nature of the constraints in any given, specific, mechanical system. In general, as shown below it does *not* equal Q^c_{ni}.

Eq. (35), which prescribes, for each instant of time t, the work done under any infinitesimal virtual displacement v at time t by the forces of constraint acting at that time, constitutes a natural generalization of D'Alembert's prescription, and reduces to D'Alembert's Principle when $C \equiv 0$ for all instants of time. We shall refer to Eq. (35) as the *Generalized D'Alembert's Principle*.

We now show that the prescription (at each instant of time) of this vector C, allows us to *uniquely* determine the contribution, Q^{c}_{ni}, which is caused by the presence of non-ideal constraints, to the total constraint force, Q^{c}. In short, such a prescription in effect, informs us of the relevant nature of z.

Our generalized D'Alembert's Principle then states that $v^{T}Q^{c} = v^{T}Q^{c}_{ni} = W^{c}$ at each instant of time t, where $W^{c}(t)$ is specified by the n-vector C. We again set $v = M^{-1/2}\mu = M^{-1/2}(I - B^{+}B)u$, as we did before in Eqs. (28) and (29). This would require that at each instant of time, for *all* n-vectors u,

$$u^{T}(I - B^{+}B)z = u^{T}(I - B^{+}B)M^{-1/2}C, \tag{36}$$

from which it follows that

$$Q^{c}_{ni} = M^{1/2}(I - B^{+}B)z = M^{1/2}(I - B^{+}B)M^{-1/2}C, \tag{37}$$

and Q^{c}_{ni} is, therefore, *uniquely* determined. Using this in Result 1, we obtain our next result.

Result 5. If for a given constrained mechanical system, the work done, W^{c}, by the total force of constraint, Q^{c}, under virtual displacements is prescribed by the mechanician through a specification of the vector $C(q, \dot{q}, t)$ (see Eq. (35)), then the *unique* equation describing the constrained motion of the system is given by

$$M\ddot{q} = Q + M^{1/2}B^{+}(b - Aa) + M^{1/2}(I - B^{+}B)M^{-1/2}C. \tag{38}$$

It is indeed satisfying that the generalized D'Alembert's principle through a specification of the 'generalized' constraint force, C, yields just enough information to uniquely determine (the contribution of the non-ideal *nature* of the constraints to) the constraint force component, Q^{c}_{ni}, that arises in the equation of motion of the constrained mechanical system. It should be noted that this n-vector C that describes the non-ideal constraints *does not itself appear directly* in the equation of motion (38). What appears instead is the projection of $M^{-1/2}C$ on the null space of the matrix B. As seen from Eq. (38), *only* when $M^{-1/2}C$ belongs to the null space of B does $Q^{c}_{ni} = C$. Also, if C belongs to the range space of B^{T}, $Q^{c}_{ni} = 0$.

The last result and the remarks preparatory to it have been aimed at understanding how one might prescribe the vector z that arises in Eq. (23) for any given practical situation. We have found that the use of the generalized D'Alembert's Principle (which requires a specification of the vector C at each instant of time) yields a unique characterization of Q^{c}_{ni}, and therefore of Q^{c}. We have thus obtained, within the framework of Lagrangian dynamics, the explicit equation of motion for systems with non-ideal constraints.

But were we to simply set $z = M^{-1/2}C$ (and this we can always do, since M is positive definite) then Eqs. (23) and (38) would be identical! This points to the generality of Eq. (38), and gives us our last result.

Result 6. The equation

$$\begin{aligned} M\ddot{q} &= Q + M^{1/2}B^{+}(b - Aa) \\ &\quad + M^{1/2}(I - B^{+}B)M^{-1/2}C \\ &= Q + Q^{c}_{i} + Q^{c}_{ni} \end{aligned} \tag{39}$$

is the general equation of motion for any mechanical system that is holonomically and/or non-holonomically constrained. The entities A, B, Q, b, and C are, in general, functions of q, $\dot{q}$, and t.

The n-vector $C(q, \dot{q}, t)$ needs to be specified at each instant of time t, and it depends on the nature of the non-ideal constraints in the specific mechanical system under consideration. This n-vector can *always* be viewed as that vector that specifies the work done by the force of constraint, Q^{c}, at time t under virtual displacements v (at that instant of time) via the relation $W^{c}(t) \equiv v^{T}Q^{c} = v^{T}C$.

Eq. (39) informs us that to model a given specific mechanical system and obtain its equation of motion, once the coordinates, q_{i}, are chosen, and the constraint equation (4) determined by the mechanician, (s)he needs to further prescribe, in general, the vector $C(t)$. Furthermore, no matter how (s)he arrives at (perhaps by inspection, or otherwise) the requisite vector $C(t)$ (or, equivalently $z(t)$) that characterizes the nature of the non-ideal constraints in the system, this vector can *always be viewed* in terms of the virtual work done by the forces of constraint. Also, when $C \equiv 0$, and the constraints are ideal then Eq. (39) reduces to the equation given in Ref. [7].

5. Illustrative example

To illustrate the simplicity with which we can write out the equations of motion for non-holonomic systems with non-ideal constraints we consider here a generalization of a well-known problem that was first introduced by Appell [12].

Consider a particle of unit mass moving in a Cartesian inertial frame subjected to the known impressed (given) forces $F_x(x,y,z,t), F_y(x,y,z,t)$, and $F_z(x,y,z,t)$ acting in the X-, Y- and Z-directions. Let the particle be subjected to the non-holonomic constraint $\dot{x}^2+\dot{y}^2-\dot{z}^2=2\alpha g(x,y,z,t)$, where α is a given scalar constant and g is a given, known function of its arguments. (Appell [12] took $\alpha=0$.) Let us say that the mechanician (who has supposedly examined the physical mechanism which is being modeled here) ascertains that this constraint subjects the particle to a force proportional to the square of its velocity and opposing its motion, so that the virtual work done by the force of constraint on the particle is prescribed (by the mechanician) as

$$W^{\mathrm{c}}(t) = -a_0 v^{\mathrm{T}}(t)\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\end{bmatrix}\frac{u^2}{|u|}, \tag{39′}$$

where $u(t)$ is the speed of the particle. Thus the constraint is non-holonomic and non-ideal. We point out that this force is caused *solely because of the presence of the constraint*. In the absence of the non-holonomic constraint, it would disappear.

We shall obtain the equations of motion of this system.

On differentiating the constraint equation with respect to time, we obtain

$$[\dot{x} \quad \dot{y} \quad -\dot{z}]\begin{bmatrix}\ddot{x}\\ \ddot{y}\\ \ddot{z}\end{bmatrix} = \alpha\dot{g}, \tag{40}$$

where

$$\dot{g} = \frac{\partial g}{\partial x}\dot{x} + \frac{\partial g}{\partial y}\dot{y} + \frac{\partial g}{\partial z}\dot{z} + \frac{\partial g}{\partial t}.$$

Thus we have $A=[\dot{x} \;\; \dot{y} \;\; -\dot{z}]$, and the scalar $b=\alpha\dot{g}$. Since $M=I_3$, $B=A$. Hence $B^{+}=1/u^2[\dot{x} \;\; \dot{y} \;\; -\dot{z}]^{\mathrm{T}}$, and the vector $C=-(a_0u^2/|u|)[\dot{x} \;\; \dot{y} \;\; \dot{z}]^{\mathrm{T}}$.

The equation of motion for this constrained system can now be written down directly by using Eq. (38). It is given by

$$\begin{bmatrix}\ddot{x}\\ \ddot{y}\\ \ddot{z}\end{bmatrix} = \begin{bmatrix}F_x\\ F_y\\ F_z\end{bmatrix} + \frac{(\alpha\dot{g}-\dot{x}F_x-\dot{y}F_y+\dot{z}F_z)}{u^2}\begin{bmatrix}\dot{x}\\ \dot{y}\\ -\dot{z}\end{bmatrix} - \frac{a_0}{|u|}\begin{bmatrix}(\dot{y}^2+\dot{z}^2) & -\dot{x}\dot{y} & \dot{x}\dot{z}\\ -\dot{y}\dot{x} & (\dot{x}^2+\dot{z}^2) & \dot{y}\dot{z}\\ \dot{x}\dot{z} & \dot{y}\dot{z} & (\dot{x}^2+\dot{y}^2)\end{bmatrix}\begin{bmatrix}\dot{x}\\ \dot{y}\\ \dot{z}\end{bmatrix}, \tag{41}$$

which simplifies to

$$\begin{bmatrix}\ddot{x}\\ \ddot{y}\\ \ddot{z}\end{bmatrix} = \begin{bmatrix}F_x\\ F_y\\ F_z\end{bmatrix} + \frac{(\alpha\dot{g}-\dot{x}F_x-\dot{y}F_y+\dot{z}F_z)}{u^2}\begin{bmatrix}\dot{x}\\ \dot{y}\\ -\dot{z}\end{bmatrix} - \frac{a_0}{|u|}\begin{bmatrix}2\dot{x}\dot{z}^2\\ 2\dot{y}\dot{z}^2\\ 2\dot{z}(\dot{x}^2+\dot{y}^2)\end{bmatrix}. \tag{42}$$

The first term on the right-hand side is the impressed force. The second term on the right-hand side is the constraint force $Q_{\mathrm{i}}^{\mathrm{c}}$ that would prevail were all the constraints ideal so that they did no work under virtual displacements. The third term on the right-hand side of (42) is the contribution, $Q_{\mathrm{ni}}^{\mathrm{c}}$, to the total constraint force generated by virtue of the fact that the constraint force is not ideal, and its nature in the given physical situation is specified by the vector C, which gives the work done by this constraint force under virtual displacements. Note that $Q_{\mathrm{ni}}^{\mathrm{c}} \neq C$.

We observe that it is because we do not eliminate any of the q's or the $\dot{q}$'s (as is customarily done in the development of the equations of motion for constrained systems) that we can explicitly assess the effect of the 'given' force, and of the components $Q_{\mathrm{i}}^{\mathrm{c}}$ and $Q_{\mathrm{ni}}^{\mathrm{c}}$ on the motion of the constrained system.

Problems that arise with sliding friction can be handled in a similar manner [10]. We note that holonomic constraints that *do* work could at times be handled by Newtonian mechanics. However, to date we know of no general formulations of mechanics that provide the explicit equations of motion for non-holonomically constrained mechanical systems where the non-holonomic constraints *do* work.

6. The geometry of constrained motion with non-ideal constraints

We present here a geometrical description of constrained motion. Using Eq. (18) in Eq. (1) the scaled acceleration of the constrained system can be written as

$$\ddot{q}_{\mathrm{s}} = N(a_{\mathrm{s}} + c_{\mathrm{s}}) + B^{+}b, \tag{43}$$

where we have denoted here the n-vector z by c_{s}. Recalling that z can be expressed as $z = M^{-1/2}C$, the acceleration c_{s} can be viewed as the 'scaled' acceleration $[c_{\mathrm{s}} = M^{1/2}(M^{-1}C)]$ created by the 'generalized' force C that specifies the virtual work done by the non-ideal constraint forces. As shown in Fig. 1, Eq. (43) thus informs us that the scaled acceleration of the constrained system is simply the sum of two *orthogonal* vectors, one belonging to the null space of B—denoted $\mathcal{N}(B)$, and the other belonging to the range space of B^{T}—denoted $\mathcal{R}(B^{\mathrm{T}})$.

Our geometrical understanding is also enhanced by considering the *deviation* of the (scaled) acceleration, $\Delta\ddot{q}_{\mathrm{s}}(t)$, of the constrained system from what it would have been, had there been no constraints on it at time t, given by (see Eqs. (11) and (18))

$$\Delta\ddot{q}_{\mathrm{s}} = \ddot{q}_{\mathrm{s}} - a_{\mathrm{s}} = B^{+}(b - Ba_{\mathrm{s}}) + (I - B^{+}B)c_{\mathrm{s}}. \tag{44}$$

Recalling Eq. (20), Eq. (44) becomes

$$\Delta\ddot{q}_{\mathrm{s}} = \ddot{q}_{\mathrm{s}} - a_{\mathrm{s}} = B^{+}b - Ta_{\mathrm{s}} + Nc_{\mathrm{s}} = B^{+}e + Nc_{\mathrm{s}}. \tag{45}$$

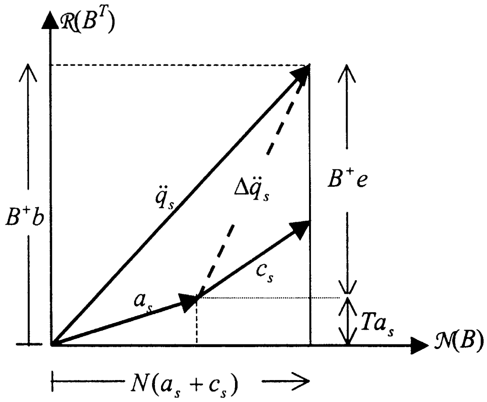

Fig. 1. The geometry of constrained motion is depicted using projections on $\mathcal{N}(B)$ and $\mathcal{R}(B^{\mathrm{T}})$. The projection of $\ddot{q}_{\mathrm{s}}$ on $\mathcal{N}(B)$ is the same as that of $(a_{\mathrm{s}}+c_{\mathrm{s}})$ because $N\ddot{q}_{\mathrm{s}} = N(a_{\mathrm{s}}+c_{\mathrm{s}})$. The vector $B^{+}b$ is orthogonal to this projection.

Using the last equality, $B^{+}b$, the component of the (scaled) acceleration of the constrained system that lies in $\mathcal{R}(B^{\mathrm{T}})$ is given by (see Fig. 1)

$$B^{+}b = B^{+}e + Ta_{\mathrm{s}}. \tag{46}$$

Fig. 1 depicts the relations (43)–(46) pictorially, and reveals the geometrical elegance with which Nature appears to operate. The n-vectors $B^{+}e, Ta_{\mathrm{s}}$ $B^{+}b$ are seen to belong to $\mathcal{R}(B^{\mathrm{T}})$ which is orthogonal to $\mathcal{N}(B)$ to which the n-vectors Nc_{s} (or Nz) and Na_{s} belong.

At any time t, the component, $B^{+}b$, of the (scaled) acceleration of the constrained system only depends on the three entities A, M, and b (at time t) that occur in Eqs. (1) and (4); it is *not* affected by whether or not the constraints are ideal. The presence of constraint forces that do *not* satisfy D'Alembert's principle (what we have referred to in this paper as non-ideal constraint forces) *only* affects the component of the (scaled) acceleration that lies in $\mathcal{N}(B)$ (see Fig. 1). Such forces engender the n-vector c_{s}, which then *needs to be suitably specified* in order to obtain the relevant equation of motion for a given constrained mechanical system. Their only effect is on the component of the acceleration that lies in $\mathcal{N}(B)$, and is seen to be Nc_{s}. At those instants of time at which D'Alembert's prescription is satisfied, $c_{\mathrm{s}} = 0$, and Fig. 1 reverts to the one that was obtained earlier, but only for constrained mechanical systems that obeyed D'Alembert's Principle [14].

7. Conclusions

In this paper we present a comprehensive description of constrained motion where the forces of constraint need not satisfy D'Alembert's principle. We summarize the main results of this paper as follows:

1. We obtain the general form of the explicit equations of motion for constrained mechanical systems. *This general form is applicable to all holonomically and non-holonomically constrained systems irrespective of whether they satisfy D'Alembert's Principle or not*. It is obtained *without* appealing to the principle of virtual work. The general form shows that the total force of constraint, Q^c, is made up of the sum of two contributions. The first contribution, Q^c_i, is what would have been caused were all the constraints ideal and had the unconstrained system been required to satisfy the prescribed constraint equations during its motion. This contribution is *uniquely* determined once M, and Q (which describe the unconstrained system), and A and b (which characterize the kinematics of the constraints), are known. The second contribution, Q^c_{ni}, to the total constraint force deals with the non-ideal nature of the constraints. It *needs to be specified by the mechanician* after an inspection of the *given* mechanical system at hand through specification of the vector $z(q, \dot{q}, t)$. This specification depends, in general, on an understanding of the underlying physics that is involved in generating the forces of constraint; one needs to rely here on the judgement and discernment of the mechanician. Yet, it is shown that *the contribution Q^c_{ni} takes a specific form in the equation of motion* that governs the constrained system, as shown in Eq. (25).
2. We provide two fundamental principles that govern the motion of general constrained mechanical systems.
3. In an effort to understand the nature of the n-vector z, we are led to generalize D'Alembert's Principle to include situations in which the constraints are not ideal, and the forces of constraint may do positive, negative, or zero work under virtual displacements. This generalized principle reduces to the usual D'Alembert's Principle when the constraints are ideal.
4. The framework of Lagrangian mechanics is used to show that this generalized D'Alembert's Principle provides a deeper insight into the nature of the vector z. In addition, this principle is shown to provide just the right extent of information to uniquely yield the accelerations of the constrained system (subjected to non-ideal constraints), as demanded by practical observation. In the situation that the constraints are ideal, these accelerations agree with those determined using formalisms (like those of Gibbs, Appell, and Gauss) that have been developed earlier [8], and that are *only* applicable to the case of ideal constraints.
5. The general equations that describe constrained motion (Eqs. (23) and (39)) obtained here appear to be the simplest and most comprehensive so far discovered.
6. The paper shows how we can include general constraint forces that *do* work within the scope of the Lagrangian formulation of analytical dynamics. Specifically, we provide the explicit general equations of motion for *non-holonomically* constrained systems where the constraints *do* work. So far, this has been beyond the reach of Lagrangian formulations of mechanics (see Refs. [5,9]). We thus surmount one of the long-standing difficulties in mechanics.
7. The geometry of constrained motion described here reveals the simplicity and elegance with which Nature seems to operate.

References

[1] J.L. Lagrange, Mecanique Analytique, Mme Ve Courcier, Paris, 1787.
[2] C.F. Gauss, Uber Ein Neues Allgemeines Grundgesetz der Mechanik, J. Reine Angew. Math. 4 (1829) 232–235.
[3] J.W. Gibbs, On the fundamental formulae of dynamics, Amer. J. Math. 2 (1879) 49–64.

[4] P. Appell, Sur une forme generale des equations de la dynamique, C. R. Acad. Sci. Paris 129 (1899) 459–460.
[5] L.A. Pars, A Treatise on Analytical Dynamics, Oxbow Press, Woodridge, CT, 1979.
[6] P.A.M. Dirac, Lectures in Quantum Mechanics, Yeshiva Univ., New York, NY, 1964.
[7] F.E. Udwadia, R.E. Kalaba, A new perspective on constrained motion, Proc. Roy. Soc. London 439 (1992) 407–410.
[8] F.E. Udwadia, R.E. Kalaba, Analytical Dynamics: A New Approach, Cambridge University Press, Cambridge, England, 1996.
[9] H. Goldstein, Classical Mechanics, Addison-Wesley, Reading, MA, 1981.
[10] F.E. Udwadia, R.E. Kalaba, Non-ideal constraints and lagrangian dynamics, J. Aerosp. Eng. 13 (2000) 17–22.
[11] F.E. Udwadia, R.E. Kalaba, Explicit equations of motion for systems with non-ideal constraints, ASME J. Appl. Mech., to appear.
[12] P. Appell, Example de mouvement d'un point assujeti a une liason exprimee par une relation non lineaire entre les composantes de la vitesse, Comptes Rendus (1911) 48–50.
[13] R. Penrose, A generalized inverse of matrices, Proc. Cambridge Philos. Soc. 51 (1955) 406–413.
[14] F.E. Udwadia, R.E. Kalaba, The geometry of constrained motion, Z. Angew. Math. Mech. 75 (8) (1995) 637–640.