

EQUATIONS OF MOTION FOR CONSTRAINED
MECHANICAL SYSTEMS
AND
THE EXTENDED D'ALEMBERT'S PRINCIPLE

BY

FIRDAUS E. UDWADIA (*Dept. of Mechanical Engineering, Civil Engineering
and Decision Systems, Univ. of Southern California, Los Angeles*),

ROBERT E. KALABA (*Dept. of Biomedical Engineering, Electrical
Engineering and Economics, Univ. of Southern California, Los Angeles*),

AND

EUN HEE-CHANG (*Dept. of Civil Engineering, Univ. of Southern California, Los Angeles*)

Abstract. Starting from the principle of virtual work, this paper states and establishes an extended version of D'Alembert's Principle. Using this extended principle and elementary linear algebra, it develops, from first principles, the *explicit* equation of motion for constrained mechanical systems. The results are compared with the authors' previous results. The approach points to new ways of extending these results.

1. Introduction. In the first part of this paper, we state and prove an extended version of the principle of D'Alembert. This principle does not appear in any of the standard treatises on analytical dynamics such as those by Whittaker (1904), Pars (1979), Gantmacher (1964), Neimark and Fufaev (1968), Goldstein (1980), and Rosenberg (1993). Arnold (1980) alludes to it in his third volume on Dynamical Systems but he does not provide a rigorous physically-based argument rooted in the fundamental postulate of the principle of virtual work for its validity, as we do here. However, from a close and careful reading of Gauss (1829) and Gibbs (1879) it appears that both Gauss and Gibbs may possibly have been aware of the extended form of the result we present here; yet, neither mentions the principle explicitly in his writings, nor establishes it.

In the second part of this paper we show that the *explicit* equations of motion for nonholonomically constrained mechanical systems with regular Lagrangians can be obtained in a straightforward manner through the application of this extended principle. The constraints we consider include those that depend explicitly on time and/or non-linearly on the displacements and velocities. The results encompass conservative and nonconservative systems.

Received September 6, 1994.

1991 *Mathematics Subject Classification.* Primary 70F25, 70H35, 90C20.

Work on constrained Hamiltonian systems was first initiated in this century by Dirac. In a series of papers, Dirac (1950, 1958, 1964) developed the Poisson bracket approach where he dealt with Hamiltonian systems with singular Lagrangians. Through a sequential determination of primary and secondary constraints and the enforcement of (weak) consistency conditions, Dirac (1964) provided a *procedure* or algorithm for obtaining the Lagrange multipliers that describe the total Hamiltonian of the constrained system. (See also Hanson, Regge, and Teitelboim, 1976.) The approach presented in this paper is entirely different; unlike Dirac, our approach does not aim at determining the Lagrange multipliers that describe the total Hamiltonian of the system. Our result differs from that of Dirac in that it is applicable to regular Lagrangians, and systems which may be nonconservative and which may be subjected to constraints that depend explicitly on time. The *explicit* equations of motion governing constrained mechanical systems obtained in this paper, though shown to be equivalent to the explicit equations obtained earlier (Udwadia and Kalaba, 1992), shed new light on the fundamental nature of constrained motion and point to new ways of extending previous results.

Consider, for simplicity, an unconstrained system of n particles whose configuration at time t is described by the Cartesian rectangular coordinate $3n$ -vector $\mathbf{x} = [x_1 x_2 \cdots x_{3n}]^T$. The masses m_i , $i = 1, 2, \dots, n$, of the n particles will be taken to be constants. By unconstrained we mean that the $3n$ Cartesian coordinates used to describe the configuration of the system at any time t equals the number of degrees of freedom of the system. The equation of motion for such a system at time t may then be written, using either the Lagrangian or the Newtonian approach, as

$$M\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \dot{\mathbf{x}}, t) \quad (1)$$

where the $3n$ by $3n$ diagonal matrix M is a known positive definite matrix and the vector \mathbf{F} constitutes the $3n$ -vector of the given force. The vector \mathbf{F} is assumed to be a known continuous function of its arguments. From (1), the "scaled" acceleration, $\mathbf{a}_s(t) = M^{1/2}\ddot{\mathbf{x}}(t)$, of the unconstrained system can be expressed as $M^{-1/2}\mathbf{F}$. We note for future reference that were the vectors $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ to be known at a given time t , $\mathbf{a}_s(t)$ would be known.

We now assume that this system is further constrained by the m consistent constraints

$$\varphi_i(\mathbf{x}, \dot{\mathbf{x}}, t) = 0; \quad i = 1, 2, \dots, m \quad (2)$$

of which $k \leq 3n$ are independent. We note in passing that the functions φ_i could involve the \dot{x}'_j 's in a nonlinear manner. Assuming that Eqs. (2) are sufficiently smooth, we differentiate them with respect to t to obtain the linear set of equations

$$A(\mathbf{x}, \dot{\mathbf{x}}, t)\ddot{\mathbf{x}} = \mathbf{b}(\mathbf{x}, \dot{\mathbf{x}}, t) \quad (3)$$

where the matrix A is m by $3n$ and has rank k at time t . Thus A and \mathbf{b} are known functions of their arguments. We assume that they are continuous functions of their arguments; also, that \mathbf{x} , $\dot{\mathbf{x}}$, and $\ddot{\mathbf{x}}$ are smooth functions of time. Given a set of initial conditions $\mathbf{x}(0)$ and $\dot{\mathbf{x}}(0)$ that satisfy the constraint equations (2) we shall determine, from first principles, the equations of motion of the constrained system as defined above.

The presence of the constraint set (2) brings into play forces of constraint, $\mathbf{F}^c(\mathbf{x}, \dot{\mathbf{x}}, t)$, so that the equation of motion at time t of the constrained system can be expressed as

$$M\ddot{\mathbf{x}} = \mathbf{F}(\mathbf{x}, \dot{\mathbf{x}}, t) + \mathbf{F}^c(\mathbf{x}, \dot{\mathbf{x}}, t), \tag{4}$$

where \mathbf{x} , $\dot{\mathbf{x}}$, and $\ddot{\mathbf{x}}$ refer now to the $3n$ -displacement, velocity, and acceleration vectors respectively, at time t of the *constrained* system that has been described above. To *explicitly* determine the constraint force $\mathbf{F}^c(\mathbf{x}, \dot{\mathbf{x}}, t)$, we shall utilize an extension of D'Alembert's Principle. In the next section we establish this result.

Extended version of D'Alembert's principle. We begin with a statement of the extended version of D'Alembert's Principle. Under the assumption that $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$, $\ddot{\mathbf{x}}(t)$, and \mathbf{F} are smooth functions of time, at every instant of time t , the motion of the mechanical system described above proceeds in such a manner that

$$\forall \text{ vectors } \mathbf{v} \text{ such that } A(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)\mathbf{v} = \mathbf{0}, \quad \mathbf{v}^T \mathbf{F}^c(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) = 0, \tag{5}$$

where the m by $3n$ matrix A and the $3n$ -vector \mathbf{F}^c are as defined earlier in Eqs. (3) and (4) respectively.

An alternative geometric statement of this new principle is as follows.

At each instant of time, the force of constraint \mathbf{F}^c must be orthogonal to every vector \mathbf{v} that belongs to the null space of the matrix A at that time.

We note that:

1. this statement differs from the usual statement of D'Alembert's Principle (e.g., Pars, 1979, p. 23) in that the matrix A corresponds to the matrix obtained by differentiating the constraint equations (2) with respect to time, as indicated in equation (3);
2. the matrix A is a function of \mathbf{x} and t , as well as of $\dot{\mathbf{x}}$ since the φ_i 's are, in general, allowed to be nonlinear functions of the \dot{x}_j 's; and
3. when the constraints are linear functions of $\dot{\mathbf{x}}$, the extended Principle coincides with the standard statement of D'Alembert's Principle, as delineated, for example, in Pars (1979).

Starting from the basic postulate of D'Alembert¹ (D'Alembert, 1743; Lagrange, 1788), we establish statement (5) as follows.

We assume that at any given instant of time t , the displacement $\mathbf{x}(t)$ and velocity $\dot{\mathbf{x}}(t)$ are known. Now at the instant of time immediately after t , that is at time $t + dt$, the *actual* displacement of the constrained system can be written as

$$\mathbf{x}(t + dt) = \mathbf{x}(t) + (dt)\dot{\mathbf{x}}(t) + \frac{(dt)^2}{2}\ddot{\mathbf{x}}(t) + O(dt^3), \tag{6}$$

¹D'Alembert stated his principle for a single rigid body around 1743. Lagrange realized its significance and universality of application, and around 1760, stated it in the form that we generally use today. Though, it might be better called the Lagrange-D'Alembert principle, we call it here by its customary name, simply as D'Alembert's principle.

where dt is an infinitesimal quantity. Also, a *possible* displacement, that is, a displacement that satisfies the constraint set (2), at that same instant of time can be expressed as

$$\mathbf{u}(t + dt) = \mathbf{x}(t) + (dt)\dot{\mathbf{x}}(t) + \frac{(dt)^2}{2}\ddot{\mathbf{u}}(t) + O(dt^3). \quad (7)$$

The vector $\ddot{\mathbf{u}}(t)$, though arbitrary, is not entirely so, for it must be such that the constraints are satisfied. We note that since $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ are assumed known, $\mathbf{u}(t) = \mathbf{x}(t)$, and $\dot{\mathbf{u}}(t) = \dot{\mathbf{x}}(t)$ in Eq. (7).

A virtual displacement \mathbf{w} at time $t + dt$ is any displacement from the actual configuration of the system to a *possible* configuration. We hence have

$$\mathbf{w}(t + dt) = \mathbf{u}(t + dt) - \mathbf{x}(t + dt) = \frac{(dt)^2}{2}[\ddot{\mathbf{u}}(t) - \ddot{\mathbf{x}}(t) + O(dt)]. \quad (8)$$

Since $\mathbf{u}(t) = \mathbf{x}(t)$ and $\dot{\mathbf{u}}(t) = \dot{\mathbf{x}}(t)$, $\mathbf{u}(t)$ and $\dot{\mathbf{u}}(t)$ automatically satisfy the equation set (2) at time t ; it is therefore the *future* positions of the system that must form the subject of our inquiry. What interests us then is finding the condition that equation set (2) is satisfied by (i) every *possible* displacement and, (ii) by the actual displacement at the instant immediately succeeding time t , namely at time $t + dt$. Using a Taylor series expansion of the set (2) about the time t and noting that dt is infinitesimally small, this naturally requires that Eq. (3) be satisfied at time t by every *possible* motion, as well as by the actual motion.

Hence, we require that

$$A(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)\ddot{\mathbf{x}}(t) = \mathbf{b}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t) \quad (9)$$

and

$$A(\mathbf{u}(t), \dot{\mathbf{u}}(t), t)\ddot{\mathbf{u}}(t) = \mathbf{b}(\mathbf{u}(t), \dot{\mathbf{u}}(t), t). \quad (10)$$

Noting that $\mathbf{u}(t) = \mathbf{x}(t)$ and $\dot{\mathbf{u}}(t) = \dot{\mathbf{x}}(t)$, Eq. (10) then leads to

$$A(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)\ddot{\mathbf{u}}(t) = \mathbf{b}(\mathbf{x}(t), \dot{\mathbf{x}}(t), t), \quad (11)$$

and subtracting Eq. (9) from Eq. (11) yields

$$A(\mathbf{x}(t), \dot{\mathbf{x}}(t), t)[\ddot{\mathbf{u}}(t) - \ddot{\mathbf{x}}(t)] = 0. \quad (12)$$

Hence, for $\mathbf{u}(t + dt)$ to represent a *possible* displacement, and therefore for $\mathbf{w}(t + dt)$ to represent a virtual displacement, $\ddot{\mathbf{u}}(t)$, while otherwise arbitrary, must satisfy Eq. (12). Denoting the vector $[\ddot{\mathbf{u}}(t) - \ddot{\mathbf{x}}(t)]$ by $\mathbf{v}(t)$, every virtual displacement $\mathbf{w}(t + dt)$ at time $t + dt$ must be such that Eq. (12) is satisfied, i.e., *such that at time t , $A\mathbf{v} = \mathbf{0}$.*

Next we show that corresponding to any given vector \mathbf{v} at time t for which $A\mathbf{v} = \mathbf{0}$, there exists a virtual displacement vector $\mathbf{w}(t + dt)$. We can write $\ddot{\mathbf{u}}(t) = \ddot{\mathbf{x}}(t) + \mathbf{v}$ and use this value of $\ddot{\mathbf{u}}(t)$ in the right-hand members of expressions (7) and (8) to yield the vectors $\mathbf{u}_1(t + dt)$ and $\mathbf{w}_1(t + dt)$ given by

$$\mathbf{u}_1(t + dt) = \mathbf{x}(t) + dt\dot{\mathbf{x}}(t) + \frac{(dt)^2}{2}(\ddot{\mathbf{x}}(t) + \mathbf{v}) + O(dt^3) \quad (13a)$$

and

$$\mathbf{w}_1(t + dt) = \frac{(dt)^2}{2}[\mathbf{v} + O(dt)]. \tag{13b}$$

Equation (13) yields $\ddot{\mathbf{u}}_1(t) = [\ddot{\mathbf{x}}(t) + \mathbf{v}]$. Since $A\mathbf{v} = \mathbf{0}$, in view of (3), we get $A\ddot{\mathbf{u}}_1(t) = A(\ddot{\mathbf{x}}(t) + \mathbf{v}) = \mathbf{b}$. This shows that $\ddot{\mathbf{u}}_1(t)$ satisfies Eq. (11) and hence that the two vectors $\mathbf{u}_1(t + dt)$ and $\mathbf{w}_1(t + dt)$ constitute a set of *possible* and virtual displacement vectors respectively.

We have therefore shown that for any virtual displacement vector $\mathbf{w}(t + dt)$, there exists a vector $\mathbf{v}(t)$ such that $A\mathbf{v} = \mathbf{0}$ at time t , and vice-versa.

D'Alembert's Principle states that at every instant of time, the work done by the force of constraint \mathbf{F}^c under any virtual displacement \mathbf{w} must equal zero. This implies that

$$\mathbf{w}^T(t + dt)[\mathbf{F}^c]_{t+dt} = 0, \tag{14}$$

where the subscript denotes that the quantity $\mathbf{F}^c(\mathbf{x}, \dot{\mathbf{x}}, t)$ is evaluated at time $t + dt$. Using Eq. (8) and noting that $dt \neq 0$ this simplifies to

$$[\ddot{\mathbf{u}}(t) - \ddot{\mathbf{x}}(t) + O(dt)]^T[\mathbf{F}^c]_{t+dt} = 0. \tag{15}$$

Under our assumptions, the force of constraint \mathbf{F}^c is continuous. Now taking the limit in Eq. (15) as $dt \rightarrow 0$ yields

$$[\ddot{\mathbf{u}}(t) - \ddot{\mathbf{x}}(t)]^T[\mathbf{F}^c]_t = \mathbf{v}(t)^T[\mathbf{F}^c]_t = 0, \tag{16}$$

and statement (5) follows.

Equations of motion for constrained systems. Let us assume that at some given time t , the vectors $\dot{\mathbf{x}}(t)$ and $\mathbf{x}(t)$ are known, and are compatible with the constraints (2). The basic problem of analytical mechanics then reduces to determining the acceleration $\ddot{\mathbf{x}}(t)$ of the constrained system so that it is consistent with equation set (2) and with D'Alembert's Principle. In what follows we shall, for brevity, drop the arguments of the various vectors and matrices, except where their inclusion is essential to our understanding.

We begin by premultiplying equation (4) by $M^{-1/2}$ yielding an equation in the "scaled" acceleration, $M^{1/2}\ddot{\mathbf{x}}(t)$, of the constrained system at time t which we denote by $\ddot{x}_s(t)$. Thus we obtain the set of $3n$ equations

$$\ddot{x}_s(t) - \mathbf{F}_s^c = \mathbf{a}_s(t) \tag{17}$$

where we have denoted the "scaled" force of constraint $M^{-1/2}\mathbf{F}^c$ by \mathbf{F}_s^c , and the "scaled" acceleration, $M^{1/2}\mathbf{F}$, corresponding to the unconstrained system by $\mathbf{a}_s(t)$. Under our assumption that $\dot{\mathbf{x}}(t)$ and $\mathbf{x}(t)$ are known, $\mathbf{a}_s(t)$ is known. Denoting by B the matrix $AM^{-1/2}$ the constraint equation at time t can be expressed as

$$B\ddot{x}_s = \mathbf{b} \tag{18}$$

where the m by $3n$ matrix B and the m -vector \mathbf{b} , being dependent on the known quantities t , $\dot{\mathbf{x}}(t)$, and $\mathbf{x}(t)$, are known.

Since M is a positive definite matrix, the rank of the matrix B at time t is k . Consider the $3n - k$ linearly independent $3n$ -vectors \mathbf{v}_i , $i = 1, 2, \dots, 3n - k$ that span the null space of B . Let the $3n$ by $(3n - k)$ matrix V be denoted by $[\mathbf{v}_1 \mathbf{v}_2 \cdots \mathbf{v}_{3n-k}]$. At time t , relation (5) can then be restated as

$$V^T \mathbf{F}_s^c = \mathbf{0} \quad (19)$$

where V is such that $BV = \mathbf{0}$. Equation (19) yields a set of $(3n - k)$ equations.

In order to determine the equation of motion of the constrained system at time t , we need to determine the $3n$ -vector $\ddot{\mathbf{x}}_s$ and the $3n$ -vector \mathbf{F}_s^c —a total of $6n$ unknowns. We have thus encapsulated the problem of determining, at time t , the equation of motion for a constrained mechanical system in terms of solving the linear equations (17), (18), and (19), along with the condition that the matrix V satisfy the relation $BV = \mathbf{0}$.

We begin by stating our result.

(a) The acceleration at time t of the constrained system described above is unique and given by

$$\ddot{\mathbf{x}}(t) = M^{-1/2} J [B^T \mathbf{b} + E \mathbf{a}_s]. \quad (20a)$$

Alternatively stated, the equation of motion of the constrained system at time t is explicitly given by

$$M \ddot{\mathbf{x}}(t) = M^{1/2} J [B^T \mathbf{b} + E \mathbf{a}_s]. \quad (20b)$$

(b) The $3n$ -vector $\mathbf{F}^c(t)$ is unique and is given by

$$\mathbf{F}^c(t) = M^{1/2} J [B^T \mathbf{b} - D \mathbf{a}_s]. \quad (21)$$

In equations (20) and (21),

$$D = B^T B, \quad E = V V^T, \quad \text{and the symmetric matrix } J = (D + E)^{-1}. \quad (22)$$

Equation (20) is the explicit general equation of motion for the constrained system; the constraint force is explicitly provided, at each instant of time t , by Eq. (21).

At time t , Eqs. (17), (18), and (19) can be expressed more compactly as

$$L \mathbf{r} = \mathbf{s} \quad (23)$$

where the $(6n + m - k)$ by $3n$ partitioned matrix L is given by the relation

$$L = \begin{bmatrix} [I]_{3n \times 3n} & -[I]_{3n \times 3n} \\ [B]_{m \times 3n} & [0]_{m \times 3n} \\ [0]_{(3n-k) \times 3n} & [V^T]_{(3n-k) \times 3n} \end{bmatrix}, \quad (24)$$

the partitioned, unknown $6n$ -vector \mathbf{r} is given by $\mathbf{r} = \{\ddot{\mathbf{x}}_s^T \quad [\mathbf{F}_s^c]^T\}^T$, and the partitioned, known $(6n + m - k)$ -vector \mathbf{s} is given by $\mathbf{s} = \{\mathbf{a}_s^T \quad \mathbf{b}^T \quad \mathbf{0}^T\}^T$.

The matrix B has k independent rows. Since $BV = \mathbf{0}$, every row of the matrix B is orthogonal to every row of the matrix V^T . Since the number of independent rows of V^T

equals $(3n - k)$, it follows then that the number of independent rows of L , and therefore its rank, is $6n$. A detailed proof of this statement may be found in Appendix 1.

We now premultiply both sides of Eq. (23) by L^T to obtain the equation

$$G\mathbf{r} = \begin{bmatrix} [I + D]_{3n \times 3n} & -[I]_{3n \times 3n} \\ -[I]_{3n \times 3n} & [I + E]_{3n \times 3n} \end{bmatrix} \mathbf{r} = \begin{bmatrix} \mathbf{a}_s + B^T \mathbf{b} \\ -\mathbf{a}_s \end{bmatrix} \quad (25)$$

where we have denoted the $6n$ by $6n$ symmetric matrix $L^T L$ by G . Since the rank of L is $6n$, the rank of G is also $6n$. Hence its inverse exists and the vector \mathbf{r} can be uniquely determined.

We have therefore shown that a unique acceleration vector, $\ddot{\mathbf{x}}$, and a unique constraint force vector, \mathbf{F}^c , exist at time t such that Eqs. (17), (18), and (19) are satisfied along with the condition $BV = \mathbf{0}$.

In fact, we can denote the inverse of the matrix G by the partitioned matrix

$$G^{-1} = \begin{bmatrix} P & J \\ J & S \end{bmatrix} \quad (26)$$

where each of the matrices P , J , and R is $3n$ by $3n$. The condition $GG^{-1} = I$ then yields the relations $P = (I + E)J$, $S = (I + D)J$, and $J = (D + E + DE)^{-1}$. Because $BV = \mathbf{0}$, $DE = B^T B V V^T = B^T (BV) V^T = \mathbf{0}$, and the last relation simplifies to $J = (D + E)^{-1}$. That the matrix G is nonsingular can also be seen by noting that the $3n$ by $3n$ symmetric matrix $D + E$ can be expressed as

$$D + E = [B^T B + V V^T] = [B^T \quad V] \begin{bmatrix} B \\ V^T \end{bmatrix}. \quad (27)$$

Because $BV = \mathbf{0}$, the rank of each of the matrices (in the brackets) in the last expression on the right equals $3n$; hence $D + E$ is nonsingular and its inverse exists. Furthermore, the condition $GG^{-1} = I = G^{-1}G$ yields the following four relations:

$$DP = PD, \quad (28)$$

$$EJ = JE, \quad (29)$$

$$DJ = JD = I - ES, \quad (30)$$

$$ES = SE. \quad (31)$$

Solving Eq. (25) for the partitioned vector \mathbf{r} after using the relations obtained for P , J , and S in G^{-1} , we obtain the following relations:

$$\ddot{\mathbf{x}}_s = (I + E)J(\mathbf{a}_s + B^T \mathbf{b}) - J\mathbf{a}_s \quad (32)$$

and

$$\mathbf{F}_s^c = J(\mathbf{a}_s + B^T \mathbf{b}) - (I + D)J\mathbf{a}_s. \quad (33)$$

By Eq. (29), $(I + E)J = J(I + E)$; also $EB^T = V V^T B^T = V(BV)^T = \mathbf{0}$. Hence Eq. (32) reduces to

$$\ddot{\mathbf{x}}_s = J(I + E)(\mathbf{a}_s + B^T \mathbf{b}) - J\mathbf{a}_s = JB^T \mathbf{b} + JE\mathbf{a}_s, \quad (34)$$

from which Eq. (20) follows, because $\ddot{\mathbf{x}} = M^{-1/2}\ddot{\mathbf{x}}_s$. Also, simplifying Eq. (33) and using (30), Eq. (21) follows.

We observe that Eqs. (20) and (21), which describe the motion of constrained mechanical systems, are obtained using (a) the extended Principle of D'Alembert established earlier, and (b) elementary linear algebra. Their derivation is found to require no use or knowledge of concepts such as the Moore-Penrose generalized inverse—concepts in terms of which our earlier statements of these equations were obtained.

These equations can be further simplified by adding and subtracting the quantity $JB^T B\mathbf{a}_s$ on the right-hand side of Eq. (34). Noting the definition of J in (22), the “scaled” acceleration, $\ddot{\mathbf{x}}_s$, can now be expressed as

$$\ddot{\mathbf{x}}_s = \mathbf{a}_s + JB^T(\mathbf{b} - B\mathbf{a}_s). \quad (35)$$

Thus the constrained, scaled acceleration equals the unconstrained, scaled acceleration plus an additional term that is brought into play by virtue of the constraint set (2). Furthermore, since $D = B^T B$, Eq. (21) can be expressed as

$$\mathbf{F}^c(t) = M^{1/2}JB^T[\mathbf{b} - B\mathbf{a}_s]. \quad (36)$$

Comparison of Eqs. (35) and (36) with the corresponding equations obtained by Udwadia Kalaba (1992) indicates that the matrix JB^T must be the Moore-Penrose inverse, B^+ , of the matrix B . That this is indeed the case is shown in Appendix 2.

Our description of constrained motion in terms of the three sets of Eqs. (17), (18), and (19), and the determination of the equations of motion by solving these sets of linear algebraic equations has the following advantage. We observe that we could have solved the linear system of Eq. (23) pertinent to the time t , by premultiplying both sides of the equation by the matrix $L^T W$, where $W(t)$ is, say, any positive definite matrix whose elements are continuous functions of t , instead of simply premultiplying by the matrix L^T , as we did in Eq. (25). This would lead to the least-squares system of equations

$$G_W \mathbf{r} = (L^T W L)\mathbf{r} = (L^T W)\mathbf{s} \quad (37)$$

whose solution is

$$\mathbf{r} = (L^T W L)^{-1}(L^T W)\mathbf{s} \quad (38)$$

where the square matrix $(L^T W L)^{-1}$ exists. Since the equation sets (17), (18), and (19) are consistent, their solution is unique and hence the weighting matrix W has obviously no influence on the solution vector \mathbf{r} , and hence on the acceleration, $\ddot{\mathbf{x}}(t)$, of the constrained system, or on the force of constraint \mathbf{F}^c .

However, were the constraint equations (3) and/or the equations of motion of the unconstrained system (1) inferred from a set of measurements, then these two sets of equations could conceivably be *inconsistent* and in that case the elements of the weighting matrix could serve as indicators of the accuracy of these measurements. The solution vector \mathbf{r} given by Eq. (38) would then indeed depend on the matrix W , yielding a least-squares solution for the now inconsistent set of equations given by (17), (18), and (19). Thus the general approach presented here, although straightforward, is powerful enough that it can be extended to such situations. While these considerations may be important in certain practical applications, a detailed development of Eq. (38) will lead us far afield from the original intent of this paper, and will therefore not be pursued any further.

Conclusion and remarks. Starting from the principle of virtual work, we have presented here a precise statement of an extended version of the Principle of D’Alembert for systems subjected to the types of constraints described by equation set (2). These constraints include those that may be nonlinear in the velocities (see Appell, (1911), for examples of such constraints). It appears that in some instances this extended version may have a more sharpened applicability than D’Alembert’s original postulate. We have applied this new principle to a long-standing problem in classical mechanics—the determination of the *explicit*, general equation of motion for nonholonomically constrained mechanical systems.

Looking over the long, three hundred year history of analytical dynamics, one is initially inclined to believe that a simple, *explicit* equation of motion for constrained mechanical systems was so late in arriving (Kalaba and Udawadia, 1992; Udawadia and Kalaba, 1993) because it relied on developments in the field of linear algebra, specifically, the development of the concept of the Moore-Penrose generalised inverse—a relatively recent invention. That this line of reasoning may indeed be specious, is shown in this paper.

Rather than new concepts related to the generalized inverses of matrices, what appears to have been lacking is perhaps a deeper understanding of one of the primal postulates of mechanics—the Principle of D’Alembert¹, which itself dates almost as far back as the very statement, by Lagrange, of the problem of constrained motion. For, by using this principle, the *explicit* equation of motion for general nonholonomically constrained mechanical systems is shown to be determined in a straightforward manner, obviating the need for notions like quasicordinates (Appell, 1911; Poincaré, 1901), extended Poisson brackets (Dirac, 1964), and generalized inverses of matrices (Kalaba and Udawadia, 1992)—notions which have hereto been developed for grappling with the problem of constrained motion. It appears that the tools for obtaining the *explicit* equation of motion for constrained mechanical systems have all along been with us, practically from the time the problem was first formulated by Lagrange.

Appendix 1. We shall prove that the matrix L in Eq. (24) has $6n$ independent rows by considering the linearly independent rows provided by the three different blocks of the matrix L .

We begin by denoting each of the first $3n$ rows of the matrix L by the $6n$ component vector $\hat{\mathbf{e}}_i^T$, $1 \leq i \leq 3n$; note that each row vector $\hat{\mathbf{e}}_i^T$ has all its entries zero except for its i th entry which is 1 and its $(3n + i)$ th entry which is -1 . Since the rank of the matrix B is k , we clearly need consider only those rows of L that contain these k independent rows of B . Let us denote these rows of L by $\hat{\mathbf{b}}_i^T$, $1 \leq i \leq k$; note that the first $3n$ elements of each row vector $\hat{\mathbf{b}}_i^T$ are identical to those of each of the k linearly independent rows of B , the remaining $3n$ elements being all zero. Referring to the $6n$ vector $[\mathbf{0} \ \mathbf{v}_i^T]$ (where \mathbf{v}_i are the columns of the matrix V), by $\hat{\mathbf{v}}_i^T$, $1 \leq i \leq (3n - k)$, we need to prove that the only solution to the system of $6n$ equations given by

$$\sum_{i=1}^{3n} \alpha_i \hat{\mathbf{e}}_i + \sum_{i=1}^k \beta_i \hat{\mathbf{b}}_i + \sum_{i=1}^{3n-k} \gamma_i \hat{\mathbf{v}}_i = \mathbf{0} \tag{A1}$$

is $\alpha = [\alpha_1 \alpha_2 \dots \alpha_{3n}]^T = \mathbf{0}$, $\beta = [\beta_1 \beta_2 \dots \beta_k]^T = \mathbf{0}$, and $\gamma = [\gamma_1 \gamma_2 \dots \gamma_{3n-k}]^T = \mathbf{0}$.

We express Eq. (A1) in matrix notation as

$$\begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -I \end{bmatrix} \begin{Bmatrix} \alpha \\ \alpha \end{Bmatrix} + \begin{bmatrix} B_1^T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \beta \\ \beta \end{Bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ V & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \gamma \\ \gamma \end{Bmatrix} = \mathbf{0} \tag{A2}$$

where the k by $3n$ matrix B_1 contains the k linearly independent rows of the matrix B . Solving the lower $3n$ block of equations in the set (A2) yields

$$\alpha = V\gamma. \tag{A3}$$

Using this value of α in the first $3n$ block of equations of (A2) now gives

$$V\gamma + B_1^T\beta = \mathbf{0}. \tag{A4}$$

Premultiplying Eq. (A4) by V^T and noting that $B_1V = \mathbf{0}$ yields

$$V^TV\gamma = \mathbf{0}. \tag{A5}$$

Since the matrix V has rank $(3n - k)$, the square matrix V^TV has full rank, and it follows from (A5) that $\gamma = \mathbf{0}$. Then by Eq. (A3), $\alpha = \mathbf{0}$, and by (A2) we must have $\beta = \mathbf{0}$ since the matrix B_1^T has linearly independent columns. This proves our assertion. \square

Appendix 2. We shall show that the matrix JB^T satisfies the following four Moore-Penrose (MP) conditions:

$$\begin{aligned} (1) \quad & B(JB^T)B = B, & (2) \quad & (JB^T)B(JB^T) = JB^T, \\ (3) \quad & (BJB^T)^T = BJB^T, & (4) \quad & (JB^TB)^T = JB^TB \end{aligned}$$

and therefore equals B^+ , the Moore-Penrose inverse of B .

(1) $B(JB^T)B = BJD = B(I - ES) = B - BES$, where we have used Eq. (30). Since $BV = \mathbf{0}$, $BE = BVV^T = (BV)V^T = \mathbf{0}$, and the first MP condition is satisfied.

(2) $(JB^T)B(JB^T) = J(B^TB)JB^T = JDJB^T = J(I - ES)B^T = JB^T - JESB^T$ where we have again used Eq. (30). But by Eq. (31) $JESB^T = JSEB^T$, and $EB^T = VV^TB^T = V(BV)^T = \mathbf{0}$; therefore, the second MP condition is satisfied.

(3) $(BJB^T)^T = BJ^TB^T = BJB^T$ since J is symmetric.

(4) $(JB^TB)^T = (B^TB)J^T = DJ = JD = JB^TB$, since by Eq. (30) $DJ = JD$.

REFERENCES

- [1] J. le R. D'Alembert. *Traite de Dynamique*, Paris, 1743
- [2] P. Appell, *Traite de Mecanique Rationnelle*, Third Edition, Paris, 1911
- [3] V. I. Arnold, *Dynamical Systems III*, Springer-Verlag, 1980, p. 17
- [4] P. A. M. Dirac, *On generalized Hamiltonian dynamics*, *Canad. J. Math.* **2**, 129-148 (1950)
- [5] P. A. M. Dirac, *Generalized Hamiltonian dynamics*, *Proc. Roy. Soc. London Ser. A* **246**, 326-332 (1958)
- [6] P. A. M. Dirac, *Lectures on Quantum Mechanics*, Belfer Graduate School of Science, Yeshiva University, New York, 1964
- [7] F. Gantmacher, *Lectures in Analytical Mechanics*, Mir Publishers, Moscow, 1970

- [8] C. F. Gauss, *Über ein Neues Allgemeines Grundgesetz der Mechanik*, J. Reine Angew. Math. **4**, 232–235 (1829)
- [9] W. Gibbs, *On the fundamental formulae of dynamics*, Amer. J. Math. **2**, 49–64 (1879)
- [10] H. Goldstein, *Classical Mechanics*, Second Edition, Addison-Wesley, 1980
- [11] A. Hanson, T. Regge, and C. Teitelboim, *Constrained Hamiltonian Systems*, Accademia Nazionale Dei Lincei, Roma, 1976
- [12] J. L. Lagrange, *Mécanique Analytique*, Mme. De Courcier, Paris, 1787
- [13] E. H. Moore, *On the reciprocal of the general algebraic matrix*, Abstract, Bull. Amer. Math. Soc. **26**, 394–395 (1920)
- [14] Yu. I. Neimark and N. A. Fufaev, *Dynamics of Nonholonomic Systems*, 1968, English translation by Amer. Math. Soc., Vol. 33, Providence, Rhode Island, 1972
- [15] L. A. Pars, *A Treatise on Analytical Dynamics*, Second Printing, Ox Bow Press, Connecticut, 1979 (originally published in 1965)
- [16] R. Penrose, *A generalized inverse of matrices*, Proc. Cambridge Philos. Soc. **51**, 406–413, 1955
- [17] H. Poincaré, *Les Méthodes Nouvelles de la Mécanique Celeste*, Gauthier-Villars, Paris, 1899
- [18] R. M. Rosenberg, *Analytical Mechanics of Particles*, Third Printing, Plenum Press, 1993
- [19] F. E. Udwardia and R. E. Kalaba, *A new perspective on constrained motion*, Proc. Roy. Soc. London Ser. A **439**, 407–410 (1992)
- [20] F. E. Udwardia and R. E. Kalaba, *On Motion*, Journal of the Franklin Institute **330**, 571–577 (1993)
- [21] E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Fourth Edition, Cambridge Univ. Press, 1989 (originally published 1904)