Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER Applied Mathematics and Computation 164 (2005) 313–320
APPLIED MATHEMATICS AND COMPUTATION
www.elsevier.com/locate/amc

On constrained motion

Firdaus E. Udwadia

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

In honor of Bob Kalaba…friend, and colleague

Abstract

The general explicit equations of motion for constrained discrete dynamical systems are obtained. These new equations lead to a simple and new fundamental view of constrained motion where the forces of constraint may be ideal and/or non-ideal.

The general problem of obtaining the equations of motion of a constrained discrete mechanical system is one of the central issues in analytical dynamics. While it was formulated at least as far back as Lagrange, the determination of the explicit equations of motion, even within the restricted compass of lagrangian dynamics, has been a major hurdle. The Lagrange multiplier method relies on problem specific approaches to the determination of the multipliers which are often difficult to obtain for systems with a large number of degrees of freedom and many non-integrable constraints. Formulations offered by Gibbs, Volterra, Appell, Boltzmann, and Poincare require a felicitous choice of problem specific quasi-coordinates and suffer from similar problems in dealing with systems with large numbers of degrees of freedom and many

E-mail address: fudwadia@usc.edu

0096-3003/$ - see front matter
doi:10.1016/j.amc.2004.06.039

non-integrable constraints. Dirac offers a formulation for hamiltonian systems with singular lagrangians where the constraints do not explicitly depend on time.

The explicit equations of motion obtained by Udwadia and Kalaba [4] provide a new and different perspective on constrained motion. They introduce the notion of generalized inverses in the description of such motion and, through their use, obtain a simple and general explicit equation of motion for constrained mechanical systems without the use of, or any need for, the notion of Lagrange multipliers. These equations are applicable to general mechanical systems and include situations where the constraints may be: (1) nonlinear functions of the velocities, (2) explicitly dependent on time, and, (3) functionally dependent. However, they deal only with systems where the constraints are ideal and satisfy D'Alembert's principle. This principle says that the motion of a constrained mechanical system occurs in such a way that at every instant of time the sum total of the work done under virtual displacements by the forces of constraint is zero.

In this paper, we extend these results along two directions. First, we extend D'Alembert's Principle to include constraints that may be, in general, non-ideal so that the forces of constraint may therefore do positive, negative, or zero work under virtual displacements at any given instant of time during the motion of the constrained system. We thus expand lagrangian mechanics to include non-ideal constraint forces within its compass. We then obtain the general, explicit equations of motion for such systems. Second, we use a different kind of generalized inverse that makes the explicit equations of motion much simpler and leads to deeper insights into the way Nature seems to work. With the help of these equations we provide a new fundamental principle governing the motion of constrained mechanical systems.

Consider first an unconstrained, discrete dynamical system whose configuration is described by the n generalized coordinates $q = [q_1, q_2, \ldots, q_n]^{\mathrm{T}}$. By 'unconstrained' we mean that the components, $\dot{q}_i$, of the velocity of the system can be independently assigned at any given initial time, say, $t = t_0$. Its equation of motion can be obtained, using newtonian or lagrangian mechanics, by the relation

$$M(q,t)\ddot{q} = Q(q,\dot{q},t), \tag{1}$$

where the n by n matrix M is symmetric and positive definite. The matrix $M(q,t)$ and the generalized force n-vector (n by 1 matrix), $Q(q,\dot{q},t)$, are known. In this paper, by 'known' we shall mean known functions of their arguments. The generalized acceleration of the unconstrained system, which we denote by the n-vector a, is then given by

$$\ddot{q} = M^{-1}Q = a(q,\dot{q},t). \tag{2}$$

We next suppose that the system is subjected to h holonomic constraints of the form

$$\varphi_i(q,t) = 0 \quad i = 1,2,\ldots,h \tag{3}$$

and $m-h$ nonholonomic constraints of the form

$$\varphi_i(q,\dot{q},t) = 0, \quad i = h+1, h+2,\ldots,m. \tag{4}$$

The initial conditions $q_0 = q(t = t_0)$ and $\dot{q}_0 = \dot{q}(t = t_0)$ are assumed to satisfy these constraints so that $\varphi_i(q_0,t_0) = 0$, $i = 1,2,\ldots,h$, and $\varphi_i(q_0,\dot{q}_0,t_0) = 0$, $i = h + 1,h + 2,\ldots,m$. We note that the constraints may be explicit functions of time, and the nonholonomic constraints may be nonlinear in the velocity components $\dot{q}_i$. Under the assumption of sufficient smoothness, we can differentiate equations (3) twice with respect to time and Eq. (4) once with respect to time to obtain the consistent equation set

$$A(q,\dot{q},t)\ddot{q} = b(q,\dot{q},t), \tag{5}$$

where the constraint matrix, A, is a known m by n matrix and b is a known m-vector. It is important to note that for a given set of initial conditions, equation set (5) is equivalent to Eqs. (3) and (4), which can be obtained by appropriately integrating the set (5).

The presence of the constraints (5) imposes additional forces of constraint on the system that alter its acceleration so that the explicit equation of motion of the constrained system becomes

$$M\ddot{q} = Q(q,\dot{q},t) + Q^c(q,\dot{q},t). \tag{6}$$

The additional term, Q^c, on the right-hand side arises by virtue of the imposed constraints prescribed by Eq. (5).

We begin by generalizing D'Alembert's Principle to include forces of constraint that may do positive, negative, or zero work under virtual displacements.

We assume that for any virtual displacement vector, $v(t)$, the total work done, $W = v^{\mathrm{T}}Q^c$, by the forces of constraint at each instant of time t, is prescribed (for the given, specific dynamical system under consideration) through the specification of a known n-vector $C(q,\dot{q},t)$ such that

$$W = v^{\mathrm{T}}C. \tag{7}$$

Eq. (7) reduces to the usual D'Alembert's Principle when $C(t) \equiv 0$, for then the total work done under virtual displacements is prescribed to be zero, and the constraints are then said to be ideal. In general, the prescription of C is the task of the mechanician who is modeling the specific constrained system whose equation of motion is to be found. It may be determined for the specific system at hand through experimentation, analogy with other systems, or otherwise. We include the situation here when the constraints may be ideal over certain intervals of time and non-ideal over other intervals. Also, W at any given instant of time may be negative, positive, or zero, allowing us to include mechanical systems where energy may be extracted from, or fed into, them through the presence of the constraints. We shall denote the acceleration of the unconstrained system subjected to this prescribed force C by $c(t) = M^{-1}C$.

We begin by stating our result for the constrained system described above. For convenience we state it in two equivalent forms.

1. The explicit equation of motion that governs the evolution of the constrained system is

$$M\ddot{q} = Q + Q_{\mathrm{i}}^{c} + Q_{\mathrm{ni}}^{c} = Q + MA_{\mathrm{M}}^{+}(b - AM^{-1}Q) + M(I - A_{\mathrm{M}}^{+}A)M^{-1}C \tag{8}$$

or

$$\ddot{q} = a + A_{\mathrm{M}}^{+}(b - Aa) + (I - A_{\mathrm{M}}^{+}A)c \tag{9}$$

or

$$\varDelta = \ddot{q} - a = A_{\mathrm{M}}^{+}(b - Aa) + (I - A_{\mathrm{M}}^{+}A)c, \tag{10}$$

where, A_{M}^{+} denotes the generalized Moore–Penrose M-inverse [7] of the constraint matrix A. In Eq. (10) we have denoted by $\varDelta(t)$ the deviation of the acceleration of the constrained system, $\ddot{q}$, at time t from its unconstrained value at that time, $a(t)$. The quantity $e(t) := (b - Aa)$ in Eqs. (9) and (10) represents the extent to which the acceleration a, at the time t, corresponding to the unconstrained motion does not satisfy the constraint equation (5).

2. At each instant of time t, the total force of constraint, Q^{c}, is made up of two additive parts. The first part, Q_{i}^{c}, is the force of constraint that would have been generated were the constraints ideal at the time t; the second part, Q_{ni}^{c}, is created by the non-ideal nature of the constraints at the time t. These two contributions to the total constraint force are explicitly given by

$$Q_{\mathrm{i}}^{c} = MA_{\mathrm{M}}^{+}(b - Aa) \tag{11}$$

and

$$Q_{\mathrm{ni}}^{c} = (I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}C = M(I - A_{\mathrm{M}}^{+}A)M^{-1}C, \tag{12}$$

where $Q^{c} = Q_{\mathrm{i}}^{c} + Q_{\mathrm{ni}}^{c}$. The subscripts 'i' and 'ni' refer to ideal and non-ideal, respectively. When $C(t) \equiv 0$, the constraints are all ideal and then $Q^{c} = Q_{\mathrm{i}}^{c}$.

Eq. (10) leads to the following new fundamental principle of motion of constrained mechanical systems:

The motion of a discrete dynamical system subjected to constraints evolves, at each instant in time, in such a way that the deviation in its acceleration from what it would have at that instant if there were no constraints on it, is the sum of two M-orthogonal components; the first component is directly proportional to the *extent, e, to which the accelerations corresponding to its unconstrained motion, at that instant, do not satisfy the constraints, the matrix of proportionality being* A_M^+*; and, the second component is proportional to the given n-vector c, the matrix of proportionality being* $(I - A_M^+ A)$.

We define two n-vectors u and w to be M-orthogonal if $u^T M w = 0$. Since the generalized Moore–Penrose M-inverse of a matrix, A_M^+, is not as well known as the (regular) generalized Moore–Penrose inverse, A^+, we provide here some of its properties, which we shall shortly use. Given an m by n matrix A, and an n by n positive definite matrix M, the n by m matrix A_M^+ is a unique matrix that satisfies the following four relations [2]:

$$(1)\ AA_M^+A = A, \tag{13a}$$

$$(2)\ A_M^+AA_M^+ = A_M^+, \tag{13b}$$

$$(3)\ (AA_M^+)^T = AA_M^+ \quad \text{and} \tag{13c}$$

$$(4)\ (A_M^+A)^T = MA_M^+AM^{-1}. \tag{13d}$$

We note that A_M^+ differs from the regular Moore–Penrose (MP) generalized inverse, A^+, in the fourth property stated in Eq. (13d). When $M = \mu I$, relation (13d) reduces to the standard so-called fourth MP condition, and then $A_M^+ = A^+$.

As stated in our fundamental principle above, the two components of acceleration engendered by the presence of the constraints are explicitly given by the last two members of Eq. (9). Their M-orthogonality follows from the relations $(I - A_M^+A)^T M(A_M^+) = [I^T - M(A_M^+A)M^{-1}]MA_M^+ = M(I - A_M^+A)A_M^+ = 0$, where we have used relation (13d) in the first equality and Eq. (13b) in the last.

The derivation of our result is as follows. The acceleration, $\ddot{q}$, of the constrained system must satisfy two requirements. It must be such that:

(1) at each instant of time it must satisfy the constraints given by Eq. (5), and,
(2) the work W done under any virtual displacement by the force of constraint, Q^c, must, at each instant of time t, be as prescribed by relation (7).

Since we require the acceleration of the constrained system to satisfy the consistent set of equations $A\ddot{q} = A(\Delta + a) = b$, we have, from the theory of generalized inverses

$$\Delta = A_M^+(b - Aa) + (I - A_M^+A)z \tag{14}$$

where z is any arbitrary n-vector, and A_M^+ is the generalized Moore–Penrose M-inverse (of the constraint matrix A) whose properties are described in Eqs. (13a)–(13d). From Eq. (14) we then have

$$M\ddot{q} = M(a + \Delta) = Q + MA_{\mathrm{M}}^{+}(b - Aa) + M(I - A_{\mathrm{M}}^{+}A)z = Q + Q^{c}, \tag{15}$$

so that

$$Q^{c} = MA_{\mathrm{M}}^{+}(b - Aa) + M(I - A_{\mathrm{M}}^{+}A)z. \tag{16}$$

To explicitly find Q^c, we next determine the second member on the right in Eq. (16) in such a way as to ensure that the second of the above-mentioned requirements is satisfied. A virtual displacement at time t is any displacement that satisfies the relation $Av = 0$ at that time [6]. But the explicit solution of this homogeneous set of equations is simply

$$v = (I - A_{\mathrm{M}}^{+}A)y, \tag{17}$$

where y is any arbitrary n-vector. And so from relation (7) we require that

$$W = v^{\mathrm{T}}Q^{c} = v^{\mathrm{T}}[MA_{\mathrm{M}}^{+}(b - Aa) + M(I - A_{\mathrm{M}}^{+}A)z] = v^{\mathrm{T}}C, \tag{18}$$

where, at each instant of time, C is specified by the mechanician who is modeling the specific mechanical system. Using Eq. (17) in the last equality in (18) we get

$$y^{\mathrm{T}}(I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}[MA_{\mathrm{M}}^{+}(b - Aa) + M(I - A_{\mathrm{M}}^{+}A)z] = y^{\mathrm{T}}(I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}C, \tag{19}$$

which, because y is arbitrary, yields

$$M(I - A_{\mathrm{M}}^{+}A)z = (I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}C = M(I - A_{\mathrm{M}}^{+}A)M^{-1}C. \tag{20}$$

Relation (20) follows from (19) through the use of relations (13d) and (13b) because $(I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}MA_{\mathrm{M}}^{+} = [I - (A_{\mathrm{M}}^{+}A)^{\mathrm{T}}]MA_{\mathrm{M}}^{+} = [I - MA_{\mathrm{M}}^{+}AM^{-1}]MA_{\mathrm{M}}^{+} = M(I - A_{\mathrm{M}}^{+}A)A_{\mathrm{M}}^{+} = 0$, and $(I - A_{\mathrm{M}}^{+}A)^{\mathrm{T}}M(I - A_{\mathrm{M}}^{+}A) = M(I - A_{\mathrm{M}}^{+}A\ )(I - A_{\mathrm{M}}^{+}A) = M(I - A_{\mathrm{M}}^{+}A)$.

Using (20) in the second member on the right of Eq. (16) we obtain Q^c, and the result follows from Eq. (15).

We observe that Eq. (9) explicitly gives the acceleration of the constrained system in terms of the two matrices A_{M}^{+} and A, and the three column vectors a, b, and c. Of these, the matrix A and the vectors a, b, and c are known functions of their arguments. What remains to be found on the right-hand side of Eq. (9) is the matrix A_{M}^{+}. Here we explain the connection of A_{M}^{+} to the singular value decomposition of the m by n matrix A for which there are several fast and robust numerical algorithms available. One way of obtaining such a decomposition is to determine the positive eigenvalues, λ_i^2, of the semi-definite m by m matrix $AM^{-1}A^{\mathrm{T}}$ and the corresponding orthonormal eigenvectors w_i, $1, 2, \ldots, r$, where r is the rank of A. Then the singular value decomposition of the matrix A can be expressed as $A = W\Lambda V^{\mathrm{T}}$ where, the m by r matrix $W = [w_1\ w_2\ \ldots\ w_r]$, the n by r matrix V is such that $V^{\mathrm{T}}M^{-1}V = I_r$, and the diagonal matrix Λ has as its nonzero elements the r positive singular values λ_i. The matrix A_{M}^{+} is then simply $A_{\mathrm{M}}^{+} = M^{-1}V\Lambda^{-1}W^{\mathrm{T}}$. In most computing environments, such as MATLAB, the computations needed to determine A_M^+ can be carried out with considerable ease, speed, and reliably; the quantities MA_M^+ and $M(I - A_M^+A)M^{-1}$ that appear in Eqs. (11) and (12) are then simply $V\Lambda^{-1}W^T$ and $(I - VV^TM^{-1})$, respectively. Furthermore, we point out that A^+ and A_M^+ are related, because $(AM^{-1/2})^+ = M^{-1/2}V\Lambda^{-1}W^T = M^{1/2}A_M^+$.

We see from Eqs. (8)–(12), however, that in the dynamics of constrained motion a fundamental role is more directly played by the generalized Moore–Penrose M-inverse, A_M^+, rather than the (regular) generalized Moore–Penrose inverse, A^+, though the two are related. The matrix A_M^+ is in fact the minimum length least-squares solution to the equation $Ax = b$ when the length of the vector x is defined as $|x|_M = +\sqrt{x^TMx}$. The length defined in this way pays heed to the fact that for a dynamical system of discrete particles the kinematical line element in configuration space has a 'length', ds, defined by $ds^2 = 2T dt^2 = \dot{q}^T M \dot{q} dt^2 = dq^T M dq$ where T is the kinetic energy of the system. Hence the matrix M arises quite naturally in defining the metric in the configuration space of the system, and the elements of M form the components of the metric tensor [3].

The explicit equations of motion obtained herein, like those obtained earlier for ideal constraints [5], are completely innocent of the notion of Lagrange multipliers. Over the last 200 years, Lagrange multipliers have been so widely used in the development of the equations of motion of constrained mechanical systems that it is sometimes tempting to mistakenly believe that they have an instrinsic presence in the description of constrained motion. This is not true. As shown in this paper, neither in the formulation of the physical problem of the motion of constrained mechanical systems nor in the equations governing their motion are any Lagrange multipliers involved. The use of Lagrange multipliers (a mathematical tool invented by Lagrange [1]) constitutes just one of the several intermediary mathematical devices invented for handling constraints. And, in fact, the direct use of this device appears difficult when the constraints are functionally dependent. Lagrange multipliers do not appear in the physical description of constrained motion, and therefore cannot, and do not, ultimately appear in the equations governing such motion.

The simplicity of the general explicit equation of motion obtained herein relies on the interplay of four central observations:

(1) No transformation of coordinates, or their elimination, is undertaken when constraints are present; the coordinates in which the unconstrained system is described are the same as those used to describe the constrained system. This, at first, appears to be counter-intuitive and indeed goes against a 200 year-old, well-accepted current of practice in analytical dynamics and theoretical physics that was first initiated by Lagrange. Though such transformations and eliminations are often useful in handling problems of mathematical physics, it is the fact that we do not use them that appears

to be ultimately responsible for the simplicity of the explicit equation obtained herein, and the fundamental insights about the nature of constrained motion provided by it.

(2) The constraints are described in their differentiated form by Eq. (5); this a consequence of the realization that, at any instant of time t, the 'state' of the system $(q(t), \dot{q}(t))$ is assumed known, and it is the state immediately following this instant that must then be the focus of our inquiry.

(3) For a physical system where the constraint forces do work the equations of motion *cannot* be obtained solely through knowledge of the kinematical nature of the constraints as described by Eqs. (3) and (4); one needs to have an additional dynamical characterization of the constraints given by the extension of D'Alembert's principle (or some equivalent of it), as stated in Eq. (7). Such a characterization yields a unique equation of motion, as expected from, and consistent with, practical observation.

(4) The generalized M-inverse of the constraint matrix A plays a quintessential role in describing constrained motion in Nature. The constraints in our equations can be nonlinear in the $\dot{q}_i$'s, functionally dependent, and/or explicitly depend on time. Knowledge of the rank of the matrix A is not required to obtain the explicit form of the equation of motion governing the constrained dynamical system; it is this fact that allows deeper insights into the nature of constrained motion.

In this paper we have extended the Lagrangian formulation of mechanics to include constraints that may be ideal and/or non-ideal, and the equations of motion presented in this paper are applicable to mechanical systems that include such constraints. They appear to be the simplest and most comprehensive equations of motion so far discovered for such systems.

References

[1] J.L. Lagrange, Mechanique Analytique, Mme Ve Coureier, Paris, 1811.

[2] C. Rao, S. Mitra, Generalized Inverse of Matrices and its Applications, John Wiley and Sons, New York, 1971.

[3] J.L. Synge, On the geometry of dynamics, Philos. Trans. Royal Soc. London Ser. A 226 (1927) 31–106.

[4] F.E. Udwadia, R.E. Kalaba, A new perspective on constrained motion, Proc. Royal Soc. London Ser. A 439 (1992) 407–410.

[5] F.E. Udwadia, R.E. Kalaba, Analytical Dynamics: A New Approach, Cambridge University Press, England, 1996.

[6] F.E. Udwadia, R.E. Kalaba, E. Hee-Chang, Equations of motion for constrained mechanical systems and the extended D'Alembert principle, Quart. Appl. Math. 56 (1997) 321–331.