Pergamon

Int. J. Non-Linear Mechanics, Vol. 31, No. 6, pp. 951–958, 1996

0020-7462/96 $15.00 + 0.00

S0020-7462(96)00116-3

EQUATIONS OF MOTION FOR MECHANICAL SYSTEMS: A UNIFIED APPROACH

Firdaus E. Udwadia

Department of Mechanical Engineering, Civil Engineering and Business Administration, 430K Olin Hall, University of Southern California, Los Angeles, CA 90089-1453, U.S.A.

Abstract—This paper presents a unified framework from which emerge the Lagrange equations, the Gibbs–Appell Equations and the Generalized Inverse Equations for describing the motion of constrained mechanical systems. The unified approach extends the applicability of the first two approaches to systems where the constraints are non-linear functions of the generalized velocities and are not necessarily independent. Furthermore, the approach leads to the Explicit Gibbs–Appell Equations. Copyright © 1996 Published by Elsevier Science Ltd.

Keywords: nonholonomic systems, Gauss's principle, generalized inverse equations of motion, Gibbs–Appell equations of motion, Lagrange's equations of motion, equivalence of different equation sets

The equations of motion for constrained systems initially developed by Lagrange [1] make use of Lagrange multipliers. Gibbs [2] and Appell [3] independently obtained equations of motion for constrained systems by using the concept of quasi-coordinates; these equations are now referred to as the Gibbs–Appell equations. More recently, Udwadia and Kalaba [4] have obtained a third set of equations which utilize the concept of the generalized inverse of a matrix and provide explicit equations of motion for more general constrained systems. Starting with Gauss's principle, in this paper we show that all these three sets of equations can be directly and easily obtained. Furthermore we obtain more general results for the first two sets of equations than have hereto been obtained thereby expanding their applicability. The unified approach presented here shows that the three sets of equations are equivalent.

Consider an unconstrained system of n particles in an inertial Cartesian coordinate frame of reference. Let the masses of the particles be $m_1, m_2, \dots, m_n$, respectively. By unconstrained we mean that the number of coordinates required to specify the configuration of the system equals its degrees of freedom. Let the X-, Y- and Z-components of the position of the jth particle be described by the triplet $\{x_{3j-2}, x_{3j-1}, x_{3j}\}$ so that the configuration of the entire system of n particles, at any time t, is given by the $3n$-vector $\mathbf{x}(t) = [x_1, x_2, x_3, \dots, x_{3n}]^T$. Similarly, let the known impressed forces on the jth particle be denoted by the triplet $\{F_{3j-2}, F_{3j-1}, F_{3j}\}$ so that the forces acting on the system of n particles can be expressed by the $3n$-vector $\mathbf{F}(t) = [F_1, F_2, F_3, \dots, F_{3n}]^T$. By "known" we mean that the forces are known functions of the positions and velocities of the n particles. We shall throughout use Cartesian coordinates, for, as we shall see not much additional fundamental understanding is gained by using generalized coordinates.

Newton's law now yields

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

where along the diagonal of the $3n \times 3n$ matrix $\mathbf{M}$ are the masses of the particles in sets of threes. The matrix $\mathbf{M}$ is positive definite since it is a diagonal matrix with positive entries. Hence the $3n$-vector of acceleration, $\mathbf{a}$, of the unconstrained system at time t is uniquely obtained as

$$\mathbf{a}(t) := \ddot{\mathbf{x}}(t) = \mathbf{M}^{-1}\mathbf{F}(\mathbf{x}, \dot{\mathbf{x}}, t). \tag{2}$$

Now we assume that this system which has so far been unconstrained is further subjected to m smooth constraints of the form

$$\varphi_i(\mathbf{x}, \dot{\mathbf{x}}, t) = 0, \qquad i = 1, 2, \dots, m, \tag{3}$$

which may *not* necessarily be independent. However, we do demand that the set of m equations in (3) be *consistent*, that is, that the satisfaction of any one equation of the set does preclude the satisfaction of any other. These m equations can be appropriately differentiated with respect to time t to yield the set of equations

$$\mathbf{A}(\mathbf{x}, \dot{\mathbf{x}}, t)\ddot{\mathbf{x}} = \mathbf{b}(\mathbf{x}, \dot{\mathbf{x}}, t) \tag{4}$$

where $\mathbf{A}$ is a known $m \times 3n$ matrix and $\mathbf{b}$ is a known $3n$-vector. We note that the constraint set (3) easily accommodates the usual holonomic and non-holonomic constraints met within analytical dynamics, and includes, in addition, constraints which are non-linearly dependent on the velocities.

We shall assume that the vectors $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ are known at some time t, and are compatible with the constraint set (3). Our aim is to obtain the equation of motion describing this constrained system of n particles. Specifically, by this we mean that we want to obtain an expression, which can be evaluated at time t, for the acceleration $\ddot{\mathbf{x}}(t)$ of the constrained system. Because of the presence of constraints described by the constraint set (3), the acceleration $\ddot{\mathbf{x}}(t)$ of the constrained system will, in general, no longer now be the same as $\mathbf{a}(t)$, the acceleration of the unconstrained system.

We thus conceptualize the constrained system in *two* steps. First we visualize the unconstrained system described by equation (1), and we then *further* restrict the motion of this unconstrained system by the constraint set (3). The entire problem, at any time t, is thus completely described by the *two* known matrices $\mathbf{M}$ and $\mathbf{A}$, the *two* known vectors $\mathbf{F}$ and $\mathbf{b}$, and the *two* given initial conditions $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ at time t. By "known" we again mean that the quantities are known (or given) functions of $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$ and t.

In what follows we shall use a principle first stated by Gauss in 1829 [5]. Gauss's principle states that, at each instant of time t, of all the acceleration vectors which satisfy the constraint equation (4), Nature "chooses", for the constrained system, the one that minimizes the Gaussian G, defined by

$$G(\ddot{\mathbf{x}}(t)) = \tfrac{1}{2}[\{\ddot{\mathbf{x}}(t) - \mathbf{a}(t)\}^{\mathrm{T}}\,\mathbf{M}\{\ddot{\mathbf{x}}(t) - \mathbf{a}(t)\}]. \tag{5}$$

Note that the vector $\mathbf{a}(t)$ is known since the elements of $\mathbf{M}$ are known constants and those of $\mathbf{F}$ are known functions of $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$ and t, and $\mathbf{x}(t)$ and $\dot{\mathbf{x}}(t)$ are assumed known at time t. We are hence left, at each instant of time t, with a problem of the constrained minimization of a quadratic functional of $\ddot{\mathbf{x}}$ subject to the linear constraint given by equation (4). Dropping the arguments of the various quantities for brevity, the problem of constrained motion can then be simply expressed at each instant of time as:

$$\text{Find } \ddot{\mathbf{x}} \text{ so as to Minimize } [\tfrac{1}{2}(\ddot{\mathbf{x}} - \mathbf{a})^{\mathrm{T}}\,\mathbf{M}(\ddot{\mathbf{x}} - \mathbf{a})] \text{ subject to } \mathbf{A}\ddot{\mathbf{x}} = \mathbf{b}. \tag{6}$$

There are many ways of solving the constrained minimization problem stated in (6). Three different approaches immediately suggest themselves. The first is to directly solve the constrained minimization problem posed in (6). The second is to eliminate the constraint (thereby converting the constrained minimization problem (6) to a new unconstrained minimization problem), and then solve this unconstrained minimization problem. The third approach is to utilize the so-called method of Lagrange multipliers. We shall show below that the first approach leads to the equation of motion obtained in Udwadia and Kalaba [4], the second to the Explicit Gibbs–Appell equations [2, 3] and the third to Lagrange's equation for non-holonomically constrained systems [1].

The direct approach: *the generalized inverse equations*

Using the "scaled" acceleration vectors $\ddot{\mathbf{x}}_s = \mathbf{M}^{1/2}\ddot{\mathbf{x}}$ and $\mathbf{a}_s = \mathbf{M}^{1/2}\mathbf{a}$, problem (6) can be rewritten as

$$\text{Find } \ddot{\mathbf{x}}_s \text{ so as to Minimize } [\tfrac{1}{2}(\ddot{\mathbf{x}}_s - \mathbf{a}_s)^{\mathrm{T}}(\ddot{\mathbf{x}}_s - \mathbf{a}_s)] \text{ subject to } \mathbf{C}\ddot{\mathbf{x}}_s = \mathbf{b}, \tag{7}$$

where the $m \times 3n$ constraint matrix $\mathbf{C} = \mathbf{A}\mathbf{M}^{-1/2}$. Note that since $\mathbf{M}$ is positive definite, the matrices $\mathbf{M}^{1/2}$ and $\mathbf{M}^{-1/2}$ are well defined. Let us now denote $\mathbf{y}_s = (\ddot{\mathbf{x}}_s - \mathbf{a}_s)$. Then problem (7) becomes

$$\text{Find } \mathbf{y}_s \text{ so as to Minimize } [\tfrac{1}{2}\mathbf{y}_s^{\mathrm{T}}\mathbf{y}_s] \text{ subject to } \mathbf{C}\mathbf{y}_s = \mathbf{b} - \mathbf{C}\mathbf{a}_s. \tag{8}$$

But the minimum "length" solution of the consistent linear equation set, $\mathbf{C}\mathbf{y}_s = \mathbf{b} - \mathbf{C}\mathbf{a}_s$ is uniquely given by

$$\mathbf{y}_s = \mathbf{C}^{(1,4)}(\mathbf{b} - \mathbf{C}\mathbf{a}_s) \tag{9}$$

where $\mathbf{C}^{(1,4)}$ denotes *any* of the {1,4}-generalized inverses of matrix $\mathbf{C}$. Though the {1,4}-generalized inverse, $\mathbf{C}^{(1,4)}$, is not uniquely determined for a given matrix $\mathbf{C}$, the right-hand side of equation (9) is unique, and hence the solution $\mathbf{y}_s$ is unique. Noting that $\mathbf{y}_s = (\ddot{\mathbf{x}} - \mathbf{a}_s)$, we obtain

$$\ddot{\mathbf{x}}_s = \mathbf{a}_s + \mathbf{C}^{(1,4)}(\mathbf{b} - \mathbf{C}\mathbf{a}_s) \tag{10}$$

or, noting the definitions of the scaled quantities,

$$\ddot{\mathbf{x}} = \mathbf{a} + \mathbf{M}^{-1/2}(\mathbf{A}\mathbf{M}^{-1/2})^{(1,4)}(\mathbf{b} - \mathbf{A}\mathbf{a}). \tag{11}$$

Equation (11)* is a generalization of the equation obtained in Udwadia and Kalaba [4]. Each particular {1,4}-inverse of $\mathbf{A}\mathbf{M}^{-1/2}$ used in equation (11) will yield a particular form of the equation of motion for the constrained system. A different {1,4}-inverse used in equation (11) will yield a different from for the equation of motion; thus several different forms of the equations of motion exist. Yet, despite these different forms, the acceleration $\ddot{\mathbf{x}}$ [i.e. the right-hand side of (11)] when evaluated will of course be uniquely determined. One particular {1,4}-generalized inverse of $\mathbf{C}$ is the {1,2,3,4}-generalized inverse, also called the Moore–Penrose inverse, and hence one particular form of equation (11) is

$$\ddot{\mathbf{x}} = \mathbf{a} + \mathbf{M}^{-1/2}(\mathbf{A}\mathbf{M}^{-1/2})^{+}(\mathbf{b} - \mathbf{A}\mathbf{a}) \tag{12}$$

where the superscript ' + ' on the matrix $(\mathbf{A}\mathbf{M}^{-1/2})$ denotes its Moore–Penrose inverse. The form of the equation given in (12) was the one obtained by Udwadia and Kalaba [4].

Conversion to an unconstrained minimization problem: the Gibbs–Appell Equations

Let the matrix $\mathbf{A}$ have rank r at time t. Without loss of generality, let us say that the rank of the left hand $m \times r$ submatrix of $\mathbf{A}$ is r. If this is not in fact the case, we can easily bring this about by a relabelling of the components of the vector $\ddot{\mathbf{x}}$. We use this relabelled vector in both equations (1) and (4) then. More exactly, we consider a suitable permutation matrix $\mathbf{P}$ such that the first r columns of the matrix $\mathbf{AP}$ are linearly independent. (Clearly, when the first r columns of $\mathbf{A}$ are already of rank r, the matrix $\mathbf{P}$ is the identity matrix.) Then setting $\tilde{\ddot{\mathbf{x}}} = \mathbf{P}^{-1}\ddot{\mathbf{x}}$, the equation $\mathbf{A}(\mathbf{x}, \dot{\mathbf{x}}, t)\ddot{\mathbf{x}} = \mathbf{b}(\mathbf{x}, \dot{\mathbf{x}}, t)$ becomes

$$\mathbf{A}(\mathbf{x}, \dot{\mathbf{x}}, t)\ddot{\mathbf{x}} = \mathbf{A}(\mathbf{P}\tilde{\mathbf{x}}, \mathbf{P}\tilde{\dot{\mathbf{x}}}, t)\mathbf{P}\tilde{\ddot{\mathbf{x}}} = \mathbf{b}(\mathbf{P}\tilde{\mathbf{x}}, \mathbf{P}\tilde{\dot{\mathbf{x}}}, t), \tag{13}$$

or alternately,

$$\tilde{\mathbf{A}}\tilde{\ddot{\mathbf{x}}} = \tilde{\mathbf{b}} \tag{14}$$

where for convenience we have denoted $\mathbf{A}(\mathbf{P}\tilde{\mathbf{x}}, \mathbf{P}\tilde{\dot{\mathbf{x}}}, t)\mathbf{P}$ by $\tilde{\mathbf{A}}$, and $\mathbf{b}(\mathbf{P}\tilde{\mathbf{x}}, \mathbf{P}\tilde{\dot{\mathbf{x}}}, t)$ by $\tilde{\mathbf{b}}$. Thus the first r columns of the matrix $\tilde{\mathbf{A}}$ then, by design, are linearly independent. We thus obtain the equation of constraint $\tilde{\mathbf{A}}\tilde{\ddot{\mathbf{x}}} = \tilde{\mathbf{b}}$, in terms of the relabelled vectors denoted by the tildes over them. In a similar manner, equation (1) transforms, to

$$\tilde{\mathbf{M}}\tilde{\ddot{\mathbf{x}}} = \tilde{\mathbf{F}} \tag{15}$$

where we have denoted $\tilde{\mathbf{M}} = \mathbf{P}^{\mathrm{T}}\mathbf{M}\mathbf{P}$, and $\tilde{\mathbf{F}} = \mathbf{P}^{\mathrm{T}}\mathbf{F}(\mathbf{P}\tilde{\mathbf{x}}, \mathbf{P}\tilde{\dot{\mathbf{x}}}, t)$. The Gaussian now becomes

$$G(\tilde{\ddot{\mathbf{x}}}) = \tfrac{1}{2}(\mathbf{P}\tilde{\ddot{\mathbf{x}}} - \mathbf{P}\tilde{\mathbf{a}})^{\mathrm{T}}\mathbf{M}(\mathbf{P}\tilde{\ddot{\mathbf{x}}} - \mathbf{P}\tilde{\mathbf{a}}) = \tfrac{1}{2}(\tilde{\ddot{\mathbf{x}}} - \tilde{\mathbf{a}})^{\mathrm{T}}\tilde{\mathbf{M}}(\tilde{\ddot{\mathbf{x}}} - \tilde{\mathbf{a}}). \tag{16}$$

Since $\mathbf{M}$ is diagonal, the matrix $\tilde{\mathbf{M}}$ is also diagonal and positive definite. We note that equations (4), (1) and (5) are of the same form as equations (14), (15) and (16), respectively,

* Given any real matrix $\mathbf{X}$, the real matrix $\mathbf{Y}$ is called its Moore–Penrose in verse if: (1) $\mathbf{XYX} = \mathbf{X}$, (2) $\mathbf{YXY} = \mathbf{Y}$, (3) $\mathbf{XY}$ is a symmmetric matrix and, (4) $\mathbf{YX}$ is a symmetric matrix. *Any* matrix which satisfies the first and the fourth of these four conditions is called the {1,4}-inverse of $\mathbf{X}$. The matrix which satisfies all four of these conditions can also be written as the {1,2,3,4}-inverse of $\mathbf{X}$.

with the tildes, substantiating our statement that, from a theoretical standpoint, no loss of generality occurs by assuming that the first r columns of $\mathbf{A}$ are linearly independent. From here on, for ease of notation, we shall omit the tildes on the various quantities except when their presence is necessary to facilitate a better understanding.

We now express the constraint equation (14) at time t in terms of the partitioned acceleration vector $\ddot{\mathbf{x}}^T = [\ddot{\mathbf{x}}_e^T \; \ddot{\mathbf{x}}_I^T]^T$ as

$$[\mathbf{A}_e \; \mathbf{A}_I]\begin{bmatrix}\ddot{\mathbf{x}}_e \\ \ddot{\mathbf{x}}_I\end{bmatrix} = \mathbf{b} \tag{17}$$

$$\ddot{\mathbf{x}}_e = \mathbf{A}_e^+ \mathbf{b} - \mathbf{R}\ddot{\mathbf{x}}_I, \tag{18}$$

where we have denoted $\mathbf{A}_e^+ \mathbf{A}_I$ by $\mathbf{R}$. We next express the Gaussian (16) in terms of the partitioned subvectors $\ddot{\mathbf{x}}_e$ and $\ddot{\mathbf{x}}_I$ as

$$G(\ddot{\mathbf{x}}_e, \ddot{\mathbf{x}}_I) = \frac{1}{2}(\ddot{\mathbf{x}} - \mathbf{a})^T\mathbf{M}(\ddot{\mathbf{x}} - \mathbf{a}) = \frac{1}{2}\begin{bmatrix}\ddot{\mathbf{x}}_e - \mathbf{a}_e \\ \ddot{\mathbf{x}}_I - \mathbf{a}_I\end{bmatrix}^T \begin{bmatrix}\mathbf{M}_e & 0 \\ 0 & \mathbf{M}_I\end{bmatrix}\begin{bmatrix}\ddot{\mathbf{x}}_e - \mathbf{a}_e \\ \ddot{\mathbf{x}}_I - \mathbf{a}_I\end{bmatrix} \tag{19}$$

where $\mathbf{M}_e$ and $\mathbf{M}_I$ are diagonal matrices with $\mathbf{M}_e$ being $r \times r$. Eliminating the subvector $\ddot{\mathbf{x}}_e$ from the expression (19) by using equation (18), we obtain

$$G(\ddot{\mathbf{x}}_I) = \tfrac{1}{2}\{\ddot{\mathbf{x}}_e^T \mathbf{M}_e \ddot{\mathbf{x}}_e + \ddot{\mathbf{x}}_I^T \mathbf{M}_I \ddot{\mathbf{x}}_I\} - \mathbf{a}_e^T \mathbf{M}_e \mathbf{A}_e^+(\mathbf{b} - \mathbf{A}_I \ddot{\mathbf{x}}_I) - \mathbf{a}_I^T \mathbf{M}_I \ddot{\mathbf{x}}_I + \tfrac{1}{2}\mathbf{a}^T\mathbf{M}\mathbf{a}. \tag{20}$$

We notice that the first member on the right-hand side in the brackets is simply the "kinetic energy of acceleration" of the system; this term can be expressed entirely in terms of the subvector $\ddot{\mathbf{x}}_I$ as

$$S(\ddot{\mathbf{x}}_I) = \tfrac{1}{2}\{(\mathbf{b} - \mathbf{A}_I \ddot{\mathbf{x}}_I)^T \mathbf{A}_e^{+T} \mathbf{M}_e \mathbf{A}_e^+(\mathbf{b} - \mathbf{A}_I \ddot{\mathbf{x}}_I) + \ddot{\mathbf{x}}_I^T \mathbf{M}_I \ddot{\mathbf{x}}_I\}, \tag{21}$$

so that

$$G(\ddot{\mathbf{x}}_I) = S(\ddot{\mathbf{x}}_I) - \mathbf{a}_e^T \mathbf{M}_e \mathbf{A}_e^+(\mathbf{b} - \mathbf{A}_I \ddot{\mathbf{x}}_I) - \mathbf{a}_I^T \mathbf{M}_I \ddot{\mathbf{x}}_I + \tfrac{1}{2}\mathbf{a}^T\mathbf{M}\mathbf{a}. \tag{22}$$

We have thus obtained the Gaussian in terms of the independent subvector $\ddot{\mathbf{x}}_I$ and hence converted Gauss's constrained minimization problem to an unconstrained one. The extremization of $G(\ddot{\mathbf{x}}_I)$ over all possible subvectors $\ddot{\mathbf{x}}_I$ is then obtained by simply setting $dG(\ddot{\mathbf{x}}_I)/d\ddot{\mathbf{x}}_I = \mathbf{0}$, which then yields

$$\frac{\partial S(\ddot{\mathbf{x}}_I)}{\partial \ddot{\mathbf{x}}_I} = \mathbf{M}_I \mathbf{a}_I - \mathbf{R}^T \mathbf{M}_e \mathbf{a}_e = \mathbf{F}_I - \mathbf{R}^T \mathbf{F}_e. \tag{23}$$

In the second equality on the right we have used equation (2) after expressing the impressed force vector $\mathbf{F}$ in its partitioned form as $\mathbf{F}^T = [\mathbf{F}_e^T \; \mathbf{F}_I^T]$ so that $\mathbf{F}_e$ is an r-vector and $\mathbf{F}_I$ is a $(3n - r)$-vector.

Equation (23) *is the Gibbs–Appell Equation.* The left-hand side is directly recognized as the derivative with respect to $\ddot{\mathbf{x}}_I$ of the "kinetic energy of acceleration" expressed in terms of the independent acceleration subvector $\ddot{\mathbf{x}}_I$. Taking the inner product of the right-hand side with the generalized virtual displacement vector, yields the virtual work done by the impressed forces acting on the system of particles. For, a virtual displacement can be expressed as any vector $\mathbf{v}^T = [\mathbf{v}_e^T \; \mathbf{v}_I^T]^T$ which satisfies the relation

$$[\mathbf{A}_e \; \mathbf{A}_I]\begin{bmatrix}\mathbf{v}_e \\ \mathbf{v}_I\end{bmatrix} = 0, \tag{24}$$

which then yields the relation $\mathbf{v}_e = -\mathbf{R}\mathbf{v}_I$. The virtual work, W, done by the impressed force $\mathbf{F}$ under virtual displacements which are consistent with the constraints is then simply given by $W = \mathbf{v}_I^T \mathbf{F}_I + \mathbf{v}_e^T \mathbf{F}_e = \mathbf{v}_I^T(\mathbf{F}_I - \mathbf{R}^T \mathbf{F}_e)$. The term in brackets in the last expression is now recognized as the one appearing on the right-hand side of equation (23).

Furthermore, since $S(\ddot{\mathbf{x}}_I)$ is explicitly given in equation (21), we can obtain the Explicit Gibbs–Appell equation as

$$(\mathbf{R}^T \mathbf{M}_e \mathbf{R} + \mathbf{M}_I)\ddot{\mathbf{x}}_I = \mathbf{F}_I - \mathbf{R}^T \mathbf{F}_e + \mathbf{R}^T \mathbf{M}_e \mathbf{A}_e^+ \mathbf{b}. \tag{25}$$

Note that all the quantities on the right-hand side of equation (25) are known functions of $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$ and t. Thus relation (25) provides the acceleration $\ddot{\mathbf{x}}_I$ of the constrained system at each instant of time t, given that at that instant the vectors $\mathbf{x}$ and $\dot{\mathbf{x}}$ are known. Knowing the vector $\ddot{\mathbf{x}}_I(t)$, the vector $\ddot{\mathbf{x}}_e(t)$ can be obtained by using equation (18), and hence the vector $\ddot{\mathbf{x}}(t)$ describing the acceleration of the constrained system.

However, viewed upon as a system of differential equations in $\ddot{\mathbf{x}}_I(t)$, we observe that the right-hand side of equation (25) may be a function of quantities such as $\mathbf{x}_e$ and therefore, in general, to complete the system of differential equations we would need to add equation (17). More precisely, it is sufficient to add the equation

$$[\hat{\mathbf{A}}_e\ \hat{\mathbf{A}}_I]\begin{bmatrix}\ddot{\mathbf{x}}_e\\ \ddot{\mathbf{x}}_I\end{bmatrix}=\hat{\mathbf{b}} \tag{26}$$

where $\hat{\mathbf{A}}_e$ is the $r\times r$ non-singular matrix comprising any r linearly independent rows of $\tilde{\mathbf{A}}_e$, and $\hat{\mathbf{A}}_I$ is the $r\times(3n-r)$ matrix which comprises the corresponding r rows of $\tilde{\mathbf{A}}_I$. The equations of motion describing the constrained system then become

$$\begin{bmatrix}\hat{\mathbf{A}}_e & \hat{\mathbf{A}}_I\\ 0 & \mathbf{R}^T\mathbf{M}_e\mathbf{R}+\mathbf{M}_I\end{bmatrix}\begin{bmatrix}\ddot{\mathbf{x}}_e\\ \ddot{\mathbf{x}}_I\end{bmatrix}=\begin{bmatrix}\hat{\mathbf{b}}\\ \mathbf{F}_I-\mathbf{R}^T\mathbf{F}_e+\mathbf{R}^T\mathbf{M}_e\mathbf{A}_e^+\end{bmatrix}. \tag{27}$$

We may call these the Explicit Gibbs–Appell Equations. As derived here, they are applicable to more general constraints than those usually circumscribed in the literature, because the constraint equations are: (1) allowed to be dependent, and (2) allowed to be non-linear functions of the velocities.

It is noteworthy that, in essence, the Gibbs–Appell equations are simply *the necessary conditions for the Gaussian G to be an extremum.* Indeed, it can be shown that the $\ddot{\mathbf{x}}_I(t)$ which extremizes the Gaussian G, actually minimizes it, and furthermore that this $\ddot{\mathbf{x}}_I(t)$ is uniquely determined. Though of fundamental importance to analytical dynamics, we provide a proof of this result in Appendix A so as not to interrupt our flow of thought here. Thus, at each instant of time t, the unique solution to the minimization problem posed in (6), is given by $\ddot{\mathbf{x}}$ as described by equation (27).

The method of Lagrange multipliers: the Lagrange equations

We consider, at each instant of time t, the minimization of the function

$$G_L(\ddot{\mathbf{x}})=G(\ddot{\mathbf{x}})+\lambda^T(\mathbf{A}\ddot{\mathbf{x}}-\mathbf{b}) \tag{28}$$

with respect to the acceleration vector $\ddot{\mathbf{x}}$. Setting $\mathrm{d}G_L(\ddot{\mathbf{x}})/\mathrm{d}\ddot{\mathbf{x}}=\mathbf{0}$, this yields the equation

$$\mathbf{M}(\ddot{\mathbf{x}}-\mathbf{a})=\mathbf{A}^T\lambda \tag{29}$$

or, by (1), the equation

$$\mathbf{M}\ddot{\mathbf{x}}=\mathbf{F}+\mathbf{A}^T\lambda \tag{30}$$

which must be solved along with the constraint equation

$$\mathbf{A}\ddot{\mathbf{x}}=\mathbf{b}. \tag{31}$$

Equations (30) and (31) are indeed the Lagrange equations of the first kind. They appear to fall out quite naturally from Gauss's principle. In fact, we have shown that these equations are valid beyond their usual compass, even when the constraints are (1) non-linear functions of the velocities, and (2) not necessarily independent. Now, we need only to show that the $\ddot{\mathbf{x}}$ obtained from equations (30) and (31) at each instant of time is unique and furthermore minimizes G_L. We relegate the proof of this to Appendix B.

We have therefore shown that the equation sets represented by equations (11) and (27), and the combination of equations (30) and (31) are all *equivalent* to each other, each set yielding the unique acceleration vector $\ddot{\mathbf{x}}$ which minimizes the constrained Gaussian as stated in (6).

We next present the three alternative sets of equations for a problem of constrained motion which was the subject of considerable investigation by Appell in 1911 [6].

Illustrative example

Consider a particle of mass M moving in three-dimensional space whose coordinates relative to an inertial Cartesian coordinate frame of reference are x, y, and z. The particle is subjected to the given impressed forces $MF_x(x, y, z)$, $MF_y(x, y, z)$, and $MF_z(x, y, z)$ in the X-, Y- and Z-directions. The particle is constrained by the two constraints [6]

$$\dot{z}^2 = \dot{x}^2 + \dot{y}^2 \tag{32}$$

and,

$$2\dot{z}^2 = 2\dot{x}^2 + 2\dot{y}^2. \tag{33}$$

Clearly, both constraints are not independent, the second being simply a restatement of the first. We deliberately take these constraint equations to be dependent, in order to show the influence of taking constraints which are not dependent on the three different sets of equations. In this trivial example, the dependence is indeed quite obvious; however, in more realistic problems where a system may be subjected to several tens of nonintegrable non-linear constraints, determining which of the constraint equations are independent may not be such a simple matter, both analytically and computationally.

The Generalized Inverse Equations

These equations of constraint can be expressed, upon differentiation with respect to time, in the form $\mathbf{A}\ddot{\mathbf{x}} = \mathbf{b}$, as

$$\begin{bmatrix} \dot{x} & \dot{y} & -\dot{z} \\ 2\dot{x} & 2\dot{y} & -2\dot{z} \end{bmatrix} \begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \mathbf{0}, \tag{34}$$

whence,

$$\mathbf{A}^+ = \frac{1}{5(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)} \begin{bmatrix} \dot{x} & 2\dot{x} \\ \dot{y} & 2\dot{y} \\ -\dot{z} & -2\dot{z} \end{bmatrix}, \tag{35}$$

and the Generalized Inverse Equations of motion are then obtained, using (11), in one step as

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} - \frac{(\dot{x}F_x + \dot{y}F_y - \dot{z}F_z)}{(\dot{x}^2 + \dot{y}^2 + \dot{z}^2)} \begin{bmatrix} \dot{x} \\ \dot{y} \\ -\dot{z} \end{bmatrix}. \tag{36}$$

It should be noted that we could have opted to use *any* {1,4}-inverse in this example in equation (11); here we have chosen to use the {1,2,3,4}-inverse, commonly known as the Moore–Penrose inverse.

The Explicit Gibbs–Appell Equations. To obtain the Gibbs–Appell equations we first need to obtain the submatrices $\mathbf{A}_e$ and $\mathbf{A}_I$. The rank $\mathbf{A}$ is unity, and so the matrix $\mathbf{A}_e$ must have one column. (It should be noted that determining the rank of the matrix, the rank of $\mathbf{A}_e$ and the rank of the matrix $\mathbf{A}$ may not be a simple matter both analytically and computationally when there are a large number of constraint equations, especially in large systems with several non-holonomic constraints.) We may take these matrices to be (assuming $\dot{x} \neq 0$).

$$\mathbf{A}_e = \begin{bmatrix} \dot{x} \\ 2\dot{x} \end{bmatrix} \text{ and } \mathbf{A}_I = \begin{bmatrix} \dot{y} & -\dot{z} \\ 2\dot{y} & -2\dot{z} \end{bmatrix}, \tag{37}$$

so that

$$\mathbf{R} = \mathbf{A}_e^+ \mathbf{A} = \begin{bmatrix} \frac{\dot{y}}{\dot{x}} & \frac{-\dot{z}}{\dot{x}} \end{bmatrix} \tag{38}$$

and the explicit equation (25), which forms part of the Explicit Gibbs–Appell Equation, yields

$$\begin{bmatrix} 1+\frac{\dot{y}^2}{\dot{x}^2} & -\frac{\dot{y}\dot{z}}{\dot{x}^2} \\ -\frac{\dot{y}\dot{z}}{\dot{x}^2} & 1+\frac{\dot{z}^2}{\dot{x}^2} \end{bmatrix}\begin{bmatrix} \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} F_y(x,y,z) - \frac{F_x(x,y,z)\dot{y}}{\dot{x}} \\ F_z(x,y,z) + \frac{F_x(x,y,z)\dot{z}}{\dot{x}} \end{bmatrix}. \tag{39}$$

Were we to be given x, y, z, $\dot{x}$, $\dot{y}$, and $\dot{z}$ at time t, then, since all the elements of the vector on the right-hand side of (39) and the elements of the matrix on the left-hand side are known functions of these given quantities, we can determine from (39), the acceleration components $\ddot{y}$ and $\ddot{z}$ at time t. Knowing these two components of the acceleration, the components $\ddot{x}(t)$ can be obtained from any one of the two equations of constraint in the set (34).

However, when viewed as a set of differential equations in the variables y and z, these equations do *not* form a complete set, for the quantities F_x, F_y and F_z on the right-hand side of equation (39) are functions of x also. Thus to complete the system, we augment it with any one of the equations in the set (34). As our Explicit Gibbs–Appell Equation (27) shows we can then explicitly write the equation of motion as

$$\begin{bmatrix} \dot{x} & \dot{y} & -\dot{z} \\ 0 & 1+\frac{\dot{y}^2}{\dot{x}^2} & -\frac{\dot{y}\dot{z}}{\dot{x}^2} \\ 0 & -\frac{\dot{y}\dot{z}}{\dot{x}^2} & 1+\frac{\dot{z}^2}{\dot{x}^2} \end{bmatrix}\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} 0 \\ F_y - \frac{F_x\dot{y}}{\dot{x}} \\ F_z + \frac{F_x\dot{z}}{\dot{x}} \end{bmatrix}. \tag{40}$$

We note that the forms of equations (36) and (40) are different, though they are equivalent. That equations (36) and (40) are equivalent, can be easily shown by substituting for the acceleration vector in (40) from (36) (and vice versa); such a substitution will render (40) an identity.

The Lagrange Multiplier Equations. Using equation (3), the Lagrange Multiplier Equations yield for this constrained system.

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = \begin{bmatrix} F_x \\ F_y \\ F_z \end{bmatrix} + \begin{bmatrix} \dot{x} & 2\dot{x} \\ \dot{y} & 2\dot{y} \\ -\dot{z} & -2\dot{z} \end{bmatrix}\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix}. \tag{41}$$

This equations along with the constraint equation (31) is a complete characterization of the motion of the constrained system. Clearly, from equation (41) it is evident that λ_1 and λ_2 cannot be determined uniquely, though $(\lambda_1 + 2\lambda_2)$ is uniquely determined. In fact, comparing equation (36) with (41) we can explicitly obtain the 2-vector λ as

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \frac{(-F_x\dot{x} - F_y\dot{y} + F_z\dot{z})}{5(\dot{x}^2+\dot{y}^2+\dot{z}^2)}\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} \frac{4}{5} & -\frac{2}{5} \\ -\frac{2}{5} & \frac{1}{5} \end{bmatrix}\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}, \tag{42}$$

where w_1 and w_2 are *arbitrary*. We note thus that the Lagrange multipliers are now not unique. Yet, the equation of motion (41) uniquely determines the acceleration vector at time t, since the arbitrary vector comprising the second member on the right-hand side of equation (42) will always lie in the null space of the matrix $\mathbf{A}^T$ so that

$$\mathbf{A}^T\lambda = -\frac{(\mathbf{F}_x\dot{x} + \mathbf{F}_y\dot{\mathbf{y}} - \mathbf{F}_z\dot{z})}{(\dot{x}^2+\dot{y}^2+\dot{z}^2)}\begin{bmatrix} \dot{x} \\ \dot{y} \\ -\dot{z} \end{bmatrix}, \tag{43}$$

a result in obvious conformity with equation (36).

Conclusion

We have shown that the Generalized Inverse Equation of motion, the Gibbs–Appell Equation and the Lagrange Multiplier Equation of motion constitute three equations which are equivalent to each other; each equation gives the unique solution to the constrained minimization problem posed by virtue of Gauss's Principle. It is interesting to note that though these three equations were developed at intervals of about a century between them, thus indicating a gradual development in the field of analytical mechanics over about the last 200 years, they all flow very naturally from Gauss's principle—indeed a principle both aesthetic and of enormous applicability.

REFERENCES

1. J. L. Lagrange, *Mechanique Analytique*, Vols 1 and 2. Gauthier-Villars, Paris (1788).
2. W. Gibbs, On the fundamental formulae of mechanics. *Am. J. Math.* **2**, 49 (1879).
3. P. Appell, *Traite de Mecanique Rationelle*, Tome II (2nd edition). Paris (1904).
4. F. E. Udwadia and R. E. Kalaba, A new perspective on constrained motion. *Proc. Roy. Soc. London* **A 439**, 407 (1992).
5. C. F. Gauss, Uber ein Neues Allgemeine Grundgesatz der Mechanik. *J. Reine Angewandte Mathematik* **4**, 232 (1829).
6. P. Appell, Exemple de movement d'un point assujeti a une liason exprime par une relation non lineaire entre les composantes de la vitesse. *Comptes Rendus* p. 48 (1911).

APPENDIX A

Uniqueness

Let us assume that there exist two sets of accelerations $\ddot{\mathbf{x}}_I$ and $\hat{\ddot{\mathbf{x}}}_I$ which both satisfy equation (25) at time t. Since the right-hand side of equation (25) is only a function of the given values of $\mathbf{x}(t)$, $\dot{\mathbf{x}}(t)$ and t, the difference $(\ddot{\mathbf{x}} - \hat{\ddot{\mathbf{x}}})$ satisfies the equation

$$(\mathbf{R}^T\mathbf{M}_e\mathbf{R} + \mathbf{M}_I)(\ddot{\mathbf{x}}_I - \hat{\ddot{\mathbf{x}}}_I) = \mathbf{0}. \tag{A1}$$

But the matrix $(\mathbf{R}^T\mathbf{M}_e\mathbf{R} + \mathbf{M}_I)$ is positive definite, hence $\ddot{\mathbf{x}}_I = \hat{\ddot{\mathbf{x}}}_I$, and uniqueness follows.

Minimization

Since the Hessian matrix

$$\frac{\partial^2 G(\ddot{\mathbf{x}}_I)}{\partial \ddot{\mathbf{x}}_I^2} = (\mathbf{R}^T\mathbf{M}_e\mathbf{R} + \mathbf{M}_I), \tag{A2}$$

is positive definite, we have a minimum.

Hence the vector $\ddot{\mathbf{x}}_I$ which extremizes $G(\ddot{\mathbf{x}}_I)$ also minimizes it, and it is unique. □

APPENDIX B

Uniqueness

Let there be two vectors $\ddot{\mathbf{x}}$ and $\hat{\ddot{\mathbf{x}}}$ (and two corresponding vectors $\boldsymbol{\lambda}$ and $\hat{\boldsymbol{\lambda}}$) each of which satisfy equation (30) at time t. We then have, from equation (30), that

$$\mathbf{M}(\ddot{\mathbf{x}} - \hat{\ddot{\mathbf{x}}}) = \mathbf{A}^T(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}). \tag{B1}$$

Also, equation (31) yields

$$\mathbf{A}(\ddot{\mathbf{x}} - \hat{\ddot{\mathbf{x}}}) = \mathbf{0}. \tag{B2}$$

Premultiplying equation (B1) by $\mathbf{M}^{-1}$ we then get

$$(\ddot{\mathbf{x}} - \hat{\ddot{\mathbf{x}}}) = \mathbf{M}^{-1}\mathbf{A}^T(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}) \tag{B3}$$

which in view of (B2) yields,

$$\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}) = \mathbf{0}. \tag{B4}$$

But this implies that

$$\mathbf{A}^T(\boldsymbol{\lambda} - \hat{\boldsymbol{\lambda}}) = \mathbf{0}. \tag{B5}$$

Hence, by equation (B1) $(\ddot{\mathbf{x}} - \hat{\ddot{\mathbf{x}}}) = \mathbf{0}$ since $\mathbf{M}$ is positive definite, and uniqueness is proved.

Minimization

Since the Hessian matrix of $G_L(\ddot{\mathbf{x}})$ is positive definite, we have a minimum.

Thus the $\ddot{\mathbf{x}}$ which extremizes $G_L(\ddot{\mathbf{x}})$, also minimizes it, and it is unique.

Incidentally, we have also shown, by equation (B5), that though the acceleration vector $\ddot{\mathbf{x}}$ is uniquely determined, the Lagrange multiplier vector $\boldsymbol{\lambda}$ is not, in general, unique; in fact it is determined only up to an arbitrary vector which belongs to the null space of $\mathbf{A}^T$. □