Projection formalism for constrained dynamical systems: From Newtonian to Hamiltonian mechanics

Gerald R. Kneller[a)]
Centre de Biophysique Moléculaire, CNRS,[b)] *Rue Charles Sadron, 45071 Orléans, France, University of Orléans, Chateau de la Source-Av. du Parc Floral 45067 Orléans, and Synchrotron Soleil, L'Orme de Merisiers, Boîte Postale 48, 91192 Gif-sur-Yvette, France*

(Received 18 June 2007; accepted 13 August 2007; published online 30 October 2007)

The Hamiltonian of a holonomically constrained dynamical many-particle system in Cartesian coordinates has been recently derived for applications in statistical mechanics [G. R. Kneller, J. Chem. Phys. **125**, 114107 (2006)]. Using the same projector formalism, we show here the equivalence of the corresponding equations of motion with those obtained from a Newtonian and a Lagrangian description. In the case of Newtonian mechanics, the general case of nonholonomic constraints is considered, too. © *2007 American Institute of Physics*. [DOI: 10.1063/1.2779326]

I. INTRODUCTION

In a recent work, the Hamiltonian of a constrained classical mechanical system in Cartesian coordinates and the corresponding Hamiltonian equations of motion have been derived on the basis of a projector formalism, which relies on the concept of generalized inverse (pseudoinverse) matrices, in particular, on the Bott-Duffin inverse.[1] Introducing the projection of the mass matrix to the subspace of linear velocity constraints and its generalized inverse allowed to derive a concise form for the Hamiltonian, which was subsequently used to define effective masses in semiflexible molecules and to revisit the problem of Fixman corrections of constrained phase space averages.[2] The corresponding Hamiltonian equations of motion were derived, too, but the equivalence with more familiar forms in the Newtonian and Lagrangian formulation of mechanics has not been demonstrated. To establish this nontrivial equivalence is the objective of the present article. In this context, the reader is referred to the earlier work on constrained Hamiltonian dynamics by de Leeuw *et al.*[3] and to the classical articles by Dirac[4,5] and by Anderson and Bergman,[6] where constraints are imposed to satisfy invariances in relativistic field theories.

II. EQUATIONS OF MOTION FOR CONSTRAINED SYSTEMS

A. Accelerations in presence of constraints

We consider a system consisting of N pointlike particles, which are located at positions $\mathbf{r}_1,\ldots,\mathbf{r}_N$. In the following, the column vector $\mathbf{r}$ contains the Cartesian components of all position vectors and the dot denotes a derivative with respect to time. We assume that the mechanical system under consideration is subject to arbitrary constraints of the form

$$h_i(\mathbf{r}) = h_i^{(0)}, \quad i = 1 \ldots s_1, \tag{2.1}$$

$$g_j(\mathbf{r},\dot{\mathbf{r}}) = g_j^{(0)}, \quad j = 1 \ldots s_2. \tag{2.2}$$

Differentiating the velocity-dependent constraints once with respect to time and the purely position-dependent constraints twice, one obtains a set of linear constraints for the accelerations,

$$\mathbf{A}\ddot{\mathbf{r}} = \mathbf{b}. \tag{2.3}$$

Here $\mathbf{A}$ is an $s \times 3N$ matrix, with $s = s_1 + s_2$, whose elements are given by

$$A_{ik} = \begin{cases} \partial h_i/\partial r_k, & i = 1,\ldots,s_1 \\ \partial g_{i-s_1}/\partial \dot{r}_k, & i = s_1+1,\ldots,s_1+s_2 \end{cases}, \quad k = 1,\ldots,3N, \tag{2.4}$$

and $\mathbf{b}$ is a vector of length s with elements

$$b_i = \begin{cases} -(\partial^2 h_i/\partial r_k \partial r_l)\dot{r}_k \dot{r}_l, & i = 1,\ldots,s_1 \\ 0, & i = s_1+1,\ldots,s_1+s_2. \end{cases} \tag{2.5}$$

Equation (2.3) may be considered as an incomplete set of linear equations for the components of the acceleration vector, the solution of which is bound to an f-dimensional subspace of $\mathbb{R}^{3N}$, where f is the number of degrees of freedom of the system,

$$f = 3N - \mathrm{rank}(\mathbf{A}). \tag{2.6}$$

To construct the solution for the acceleration vector, we use the generalized inverse of $\mathbf{A}$,[7–10] which is denoted $\mathbf{A}^+$ in the following and which is uniquely defined in terms of the four Moore-Penrose conditions listed in Appendix A. Multiplying Eq. (2.3) by $\mathbf{A}\mathbf{A}^+$ and using relation (A1) shows that the *consistency condition*,

$$\mathbf{A}\mathbf{A}^+\mathbf{b} = \mathbf{b}, \tag{2.7}$$

must be fulfilled. In case that all constraints are independent, the matrix $\mathbf{A}$ has full rank, and its generalized inverse can be expressed in the form

[a)]Electronic mail: kneller@cnrs-orleans.fr
[b)]Affiliated with the University of Orléans.

$$\mathbf{A}^{+}=\mathbf{A}^{T}(\mathbf{A}\mathbf{A}^{T})^{-1}. \tag{2.8}$$

Consequently, $\mathbf{A}\mathbf{A}^{+}=\mathbf{1}_s$, where $\mathbf{1}_s$ is the s-dimensional unit matrix, and consistency condition (2.7) is trivially fulfilled.

Supposing that consistency condition (2.7) is verified, the solution of Eq. (2.3) is given by

$$\ddot{\mathbf{r}}=\mathbf{A}^{+}\mathbf{b}+\ddot{\mathbf{r}}_{\parallel}, \tag{2.9}$$

where $\ddot{\mathbf{r}}_{\parallel}$ is a yet undetermined vector satisfying $\mathbf{A}\ddot{\mathbf{r}}_{\parallel}=\mathbf{0}$.

The Moore-Penrose conditions show that $\mathbf{A}^{+}\mathbf{A}$ and $\mathbf{A}\mathbf{A}^{+}$ are the respective projectors on the row and column spaces of $\mathbf{A}$ (the spaces spanned by the rows and columns, respectively), and in the following we consider, in particular, the projector on the row space and its orthogonal complement,

$$\mathbf{P}_{\perp}=\mathbf{A}^{+}\mathbf{A}, \tag{2.10}$$

$$\mathbf{P}_{\parallel}=\mathbf{1}-\mathbf{P}_{\perp}. \tag{2.11}$$

The subspaces of $\mathbb{R}^{3N}$ onto which $\mathbf{P}_{\perp}$ and $\mathbf{P}_{\parallel}$ project are denoted by $\mathcal{V}_{\perp}$ and $\mathcal{V}_{\parallel}$, respectively. The corresponding dimensions are $\dim(\mathcal{V}_{\perp})=\mathrm{rank}(\mathbf{A})$ and $\dim(\mathcal{V}_{\parallel})=f$. One sees easily that

$$\ddot{\mathbf{r}}_{\perp}=\mathbf{A}^{+}\mathbf{b}\in\mathcal{V}_{\perp}. \tag{2.12}$$

The general form [Eq. (2.9)] of the acceleration vector in the presence of constraints thus represents a decomposition into two mutually orthogonal components.

B. Newtonian dynamics and the Bott-Duffin problem

We consider now the situation where N particles move according to Newton's laws of motion in the presence of constraints. In the following, $\mathbf{f}$ denotes a column vector containing the Cartesian components of all external forces acting upon the particles and $\mathbf{z}$ contains the components of the constraint forces. With these definitions, Newton's equation of motion have the form

$$\mathbf{M}\ddot{\mathbf{r}}=\mathbf{f}+\mathbf{z}. \tag{2.13}$$

The matrix $\mathbf{M}$ is supposed to be diagonal with only positive entries representing the masses of the particles in the system ($\mathbf{1}$ is here the 3×3 unit matrix),

$$\mathbf{M}=\begin{pmatrix} m_1\mathbf{1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & m_2\mathbf{1} & \cdots & \mathbf{0}\\ \vdots & \vdots & \vdots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & m_N\mathbf{1}\end{pmatrix}. \tag{2.14}$$

We may now write $\ddot{\mathbf{r}}=\ddot{\mathbf{r}}_{\parallel}+\ddot{\mathbf{r}}_{\perp}$, where $\ddot{\mathbf{r}}_{\perp}$ is determined by the coordinates and velocities according to Eq. (2.12) and $\ddot{\mathbf{r}}_{\parallel}$ is unknown. Newton's equations of motion thus represent a system of linear equations of motion in which $\ddot{\mathbf{r}}_{\parallel}$ and $\mathbf{z}$ are unknown. Obviously, a solution can exist only if $\mathbf{z}\perp\ddot{\mathbf{r}}_{\parallel}$, i.e., if

$$\mathbf{z}\in\mathcal{V}_{\perp}. \tag{2.15}$$

The solution of Eq. (2.13) is a classical Bott-Duffin problem, the solution of which is described in Ref. 11. Here we will take a slightly different route. Projecting Newton's equation of motion onto $\mathcal{V}_{\parallel}$ will eliminate the constraint forces and using explicit form (2.9) for the acceleration vector yields a linear equation for $\ddot{\mathbf{r}}_{\parallel}$ only,

$$\mathbf{M}_c\ddot{\mathbf{r}}_{\parallel}=\mathbf{P}_{\parallel}(\mathbf{f}-\mathbf{M}\ddot{\mathbf{r}}_{\perp}). \tag{2.16}$$

Here $\mathbf{M}_c$ is the projected mass matrix,

$$\mathbf{M}_c=\mathbf{P}_{\parallel}\mathbf{M}\mathbf{P}_{\parallel}, \tag{2.17}$$

and $\ddot{\mathbf{r}}_{\perp}$ is given by relation (2.12). To solve Eq. (2.16) for $\ddot{\mathbf{r}}_{\parallel}$, we use relations (A9) and (A10), which have been proven in Ref. 1. Multiplying Eq. (2.16) from the left by the generalized inverse of $\mathbf{M}_c$ yields

$$\ddot{\mathbf{r}}_{\parallel}=\mathbf{M}_c^{+}(\mathbf{f}-\mathbf{M}\ddot{\mathbf{r}}_{\perp}). \tag{2.18}$$

The above solution can be formally obtained from the Gauss principle of least constraint,[12] which we write here in the form

$$Q=\|\mathbf{M}^{1/2}\ddot{\mathbf{r}}-\mathbf{M}^{-1/2}\mathbf{f}\|^2=\mathrm{Min}\{\ddot{\mathbf{r}}\}. \tag{2.19}$$

The minimum has to be taken with respect to the accelerations, $\ddot{\mathbf{r}}=\ddot{\mathbf{r}}_{\parallel}+\ddot{\mathbf{r}}_{\perp}$, where $\ddot{\mathbf{r}}_{\perp}$ is given by Eq. (2.12) and $\ddot{\mathbf{r}}_{\parallel}\in\mathcal{V}_{\parallel}$ is to be determined. Writing $\ddot{\mathbf{r}}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{x}$, where $\mathbf{x}$ is an *arbitrary* vector in $\mathcal{R}^{3N}$, the minimum principle [Eq. (2.19)] takes the form

$$Q=\|\mathbf{M}^{1/2}[\mathbf{P}_{\parallel}\mathbf{x}+\ddot{\mathbf{r}}_{\perp}]-\mathbf{M}^{-1/2}\mathbf{f}\|^2=\mathrm{Min}\{\mathbf{x}\}.$$

The necessary condition $\partial Q/\partial\mathbf{x}=\mathbf{0}$ leads then to $\mathbf{P}_{\parallel}(\mathbf{M}[\mathbf{P}_{\parallel}\mathbf{x}+\ddot{\mathbf{r}}_{\perp}]-\mathbf{f})=\mathbf{0}$. This equation is equivalent to Eq. (2.16), replacing $\mathbf{P}_{\parallel}\mathbf{x}$ by $\ddot{\mathbf{r}}_{\parallel}$.

The equation for the total acceleration, $\ddot{\mathbf{r}}=\ddot{\mathbf{r}}_{\parallel}+\ddot{\mathbf{r}}_{\perp}$, is finally given by

$$\ddot{\mathbf{r}}=\mathbf{M}_c^{+}\mathbf{f}+(\mathbf{1}-\mathbf{M}_c^{+}\mathbf{M})\ddot{\mathbf{r}}_{\perp}, \tag{2.20}$$

where $\mathbf{M}_c^{+}$ can be expressed in the form

$$\mathbf{M}_c^{+}=\mathbf{M}^{-1}-\mathbf{M}^{-1}\mathbf{A}^{T}(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^{T})^{-1}\mathbf{A}\mathbf{M}^{-1}, \tag{2.21}$$

if $\mathbf{A}$ has full rank.[1] For completeness, we state also the expression for the vector of constraint forces,

$$\mathbf{z}=(\mathbf{M}\mathbf{M}_c^{+}-\mathbf{1})\mathbf{f}+\mathbf{M}(\mathbf{1}-\mathbf{M}_c^{+}\mathbf{M})\ddot{\mathbf{r}}_{\perp}. \tag{2.22}$$

Using that $\mathbf{M}_c\mathbf{M}_c^{+}=\mathbf{P}_{\parallel}$, one finds easily that $\mathbf{P}_{\parallel}\mathbf{z}=\mathbf{0}$.

C. Lagrangian dynamics with holonomic constraints

We consider now the variational approach to classical mechanics, which leads to the celebrated Euler-Lagrange equations of classical mechanics. Starting from the Lagrangian of a dynamical system,

$$L(\mathbf{r},\dot{\mathbf{r}})=\tfrac{1}{2}\dot{\mathbf{r}}^{T}\mathbf{M}\dot{\mathbf{r}}-U(\mathbf{r}), \tag{2.23}$$

where U is the potential energy of the system, the equations of motion are derived from the variational principle

$$S:=\int_{t_0}^{t_1}dtL(\mathbf{r},\dot{\mathbf{r}})=\mathrm{Min}. \tag{2.24}$$

In the following, we consider only *holonomic* constraints, i.e., constraints of form (2.1), which yield linear velocity constraints of the form $\mathbf{A}\dot{\mathbf{r}}=\mathbf{0}$. Using the projector $\mathbf{P}_{\parallel}$ introduced above, one may write

$$\dot{\mathbf{r}} = \mathbf{P}_{\|}\dot{\mathbf{r}}, \tag{2.25}$$

$$\ddot{\mathbf{r}}_{\perp} = \dot{\mathbf{P}}_{\|}\dot{\mathbf{r}}, \tag{2.26}$$

where the second equation follows from the first by differentiation and the identity $\mathbf{P}_{\|}\dot{\mathbf{P}}_{\|}\mathbf{P}_{\|}=\mathbf{0}$, which holds for any time-dependent projector.

Requiring that the action integral S has a minimum leads to the condition

$$\delta S = \int_{t_0}^{t_1} dt\,\delta\mathbf{r}^T\left(\frac{d}{dt}\frac{\partial L}{\partial\dot{\mathbf{r}}} - \frac{\partial L}{\partial\mathbf{r}}\right) = 0. \tag{2.27}$$

Due to the imposed holonomic constraints, the paths are not arbitrary. We require that any admissible path $\mathbf{r}(t)$ must fulfill the holonomic constraints at any time. Calling the optimal path $\mathbf{r}_0(t)$, an arbitrary path in the neighborhood can be decomposed as $\mathbf{r}_0(t)+\delta\mathbf{r}(t)$, where $\delta\mathbf{r}(t_0)=\delta\mathbf{r}(t_1)=\mathbf{0}$, and we have

$$h_i(\mathbf{r}_0(t) + \delta\mathbf{r}(t)) \equiv 0, \quad i = 1, \ldots, s.$$

Developing the above expression up to first order in the variation yields the condition

$$\mathbf{A}\,\delta\mathbf{r} = \mathbf{0}. \tag{2.28}$$

Therefore, $\delta\mathbf{r} \in \mathcal{V}_{\|}$ and condition (2.27) leads to Euler-Lagrange equations of the form

$$\mathbf{P}_{\|}\left\{\frac{d}{dt}\frac{\partial L}{\partial\dot{\mathbf{r}}} - \frac{\partial L}{\partial\mathbf{r}}\right\} = \mathbf{0}. \tag{2.29}$$

This means simply that the vector between the curly brackets is an element of $\mathcal{V}_{\perp}$. If we define $\mathbf{z} := \{\cdots\}$ and use that the force vector is given by

$$\mathbf{f} = -\,\partial U/\partial\mathbf{r}, \tag{2.30}$$

we obtain Newton's equations of motion [Eq. (2.13)], where the constraint forces may be expressed by Eq. (2.22). Equivalently, one can write

$$\ddot{\mathbf{r}} = \mathbf{M}_c^{+}\mathbf{f} + (\mathbf{1} - \mathbf{M}_c^{+}\mathbf{M})\ddot{\mathbf{r}}_{\perp}, \tag{2.31}$$

which is identical with Eq. (2.20).

D. Redundant velocity constraints in Lagrangian dynamics

Instead from the Lagrangian [Eq. (2.23)], one could also start from the *constrained* Langrangian

$$L_c(\mathbf{r},\dot{\mathbf{r}}) = \tfrac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}} - U(\mathbf{r}), \tag{2.32}$$

to derive the equations of motion. Here one uses explicitly that $\dot{\mathbf{r}}=\mathbf{P}_{\|}\dot{\mathbf{r}}$, and on the constraint surface defined by Eq. (2.1), we have obviously

$$L = L_c. \tag{2.33}$$

Replacing L by L_c in the Euler-Langrange equations [Eq. (2.29)], we obtain by straightforward calculation

$$\mathbf{M}_c\ddot{\mathbf{r}} = \mathbf{P}_{\|}\left\{\mathbf{f} - \dot{\mathbf{M}}_c\dot{\mathbf{r}} + \frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}}\right)\right\}.$$

The component of the acceleration in $\mathcal{V}_{\|}$ is thus given by

$$\ddot{\mathbf{r}}_{\|} = \mathbf{M}_c^{+}\left\{\mathbf{f} - \mathbf{P}_{\|}\mathbf{M}\dot{\mathbf{P}}_{\|}\dot{\mathbf{r}} - \dot{\mathbf{P}}_{\|}\mathbf{M}\mathbf{P}_{\|}\dot{\mathbf{r}} + \frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}}\right)\right\},$$

expanding $\mathbf{M}_c=\mathbf{P}_{\|}\mathbf{M}\mathbf{P}_{\|}$ before differentiation with respect to t. Using Eq. (2.26), the resulting equation of motion becomes

$$\ddot{\mathbf{r}} = \mathbf{M}_c^{+}\mathbf{f} + (\mathbf{1} - \mathbf{M}_c^{+}\mathbf{M})\ddot{\mathbf{r}}_{\perp} + \mathbf{M}_c^{+}\left\{\frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}}\right) - \dot{\mathbf{P}}_{\|}\mathbf{M}\mathbf{P}_{\|}\dot{\mathbf{r}}\right\}.$$

In Appendix B, it is shown that

$$\mathbf{P}_{\|}\left\{\frac{\partial}{\partial\mathbf{r}}\left(\frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}}\right) - \dot{\mathbf{P}}_{\|}\mathbf{M}\mathbf{P}_{\|}\dot{\mathbf{r}}\right\} = \mathbf{0}. \tag{2.34}$$

Knowing that $\mathbf{M}_c^{+}$ projects implicitly onto $\mathcal{V}_{\|}$, one obtains again explicit form (2.20) for the acceleration vector. Identity (2.34) shows that the use of L_c instead of L does not change the equations of motion, which means that the explicit velocity constraints in L_c are *redundant*. It should be mentioned that the redundancy of the velocity constraints is much easier to see in mass-weighted coordinates, where it follows simply from the relation $\mathbf{P}_{\|}\mathbf{P}_{\|}'\mathbf{P}_{\|}=\mathbf{0}$, which was already mentioned and which holds for any projector. The prime indicates here either a differentiation of $\mathbf{P}_{\|}$ with respect to time or with respect to the positions.

E. Hamiltonian dynamics with holonomic constraints

As described in Ref. 1, one starts from the *constrained* Lagrangian L_c to construct the corresponding Hamilton function. Using the momenta

$$\mathbf{p} := \frac{\partial L_c}{\partial\dot{\mathbf{r}}} = \mathbf{M}_c\dot{\mathbf{r}}, \tag{2.35}$$

one performs the Legendre transform $dL_c(\mathbf{r},\dot{\mathbf{r}}) \rightarrow dH_c(\mathbf{r},\mathbf{p})$, where $H_c=\mathbf{p}^T\dot{\mathbf{r}}-L_c$. Here the velocities are to be eliminated in favor of the momenta, i.e., Eq. (2.35) must be inverted. This is possible since the constrained Lagrangian L_c generates by construction momenta, which are in the same subspace as the velocities, $\mathbf{p} \in \mathcal{V}_{\|}$, and one can write

$$\dot{\mathbf{r}} = \mathbf{M}_c^{+}\mathbf{p}. \tag{2.36}$$

The Hamiltonian thus has the form

$$H_c(\mathbf{r},\mathbf{p}) = \tfrac{1}{2}\mathbf{p}^T\mathbf{M}_c^{+}\mathbf{p} + U(\mathbf{r}). \tag{2.37}$$

The Hamiltonian equations of motion are derived in the usual way from variational condition (2.27), expressing the Langrangian as $L_c=\mathbf{p}^T\dot{\mathbf{r}}-H_c$ and considering the positions and momenta as independent dynamical variables,

$$\delta S = \int_{t_0}^{t_1} dt \left\{ \boldsymbol{\delta}\mathbf{p}^T \left(\dot{\mathbf{r}} - \frac{\partial H_c}{\partial \mathbf{p}} \right) - \boldsymbol{\delta}\mathbf{r}^T \left(\dot{\mathbf{p}} + \frac{\partial H_c}{\partial \mathbf{r}} \right) \right\} = 0. \tag{2.38}$$

We require that any admissible momentum path $\mathbf{p}(t)$ in the variational principle [Eq. (2.38)] fulfills at any time $\mathbf{p}(t) \in \mathcal{V}_{\parallel}$. Taking the optimal path $\mathbf{p}_0(t)$ and a variation $\mathbf{p}(t) = \mathbf{p}_0(t) + \delta\mathbf{p}(t)$, with $\delta\mathbf{p}(t_0) = \delta\mathbf{p}(t_1) = 0$, we thus have (the time argument is omitted) $\mathbf{A}\mathbf{p}_0 = \mathbf{0}$ and $\mathbf{A}(\mathbf{p}_0 + \delta\mathbf{p}) = \mathbf{0}$. In addition to condition (2.28) for the variations of the coordinate paths, we also have

$$\mathbf{A}\,\delta\mathbf{p} = \mathbf{0}, \tag{2.39}$$

for the variations of the momentum paths. We thus obtain from Eq. (2.38) the necessary conditions

$$\mathbf{P}_{\parallel}\left(\dot{\mathbf{r}} - \frac{\partial H}{\partial \mathbf{p}} \right) = \mathbf{0}, \tag{2.40}$$

$$\mathbf{P}_{\parallel}\left(\dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{r}} \right) = \mathbf{0}, \tag{2.41}$$

for the stationarity of the action integral S. Since $\dot{\mathbf{r}} \in \mathcal{V}_{\parallel}$ and since $\partial H_c / \partial \mathbf{p} = \mathbf{M}_c^{+}\mathbf{p} \in \mathcal{V}_{\parallel}$, the projector in Eq. (2.40) can be omitted. The time derivative of $\mathbf{p}$ has, in contrast, components in $\mathcal{V}_{\parallel}$ and in its orthogonal complement, $\mathcal{V}_{\perp}$. As for the time derivative of the velocity vector $\dot{\mathbf{r}}$, one derives

$$\dot{\mathbf{p}}_{\perp} = \dot{\mathbf{P}}_{\parallel}\mathbf{p}, \tag{2.42}$$

and the equations of motion may be written in the form

$$\dot{\mathbf{r}} = \mathbf{M}_c^{+}\mathbf{p}, \tag{2.43}$$

$$\dot{\mathbf{p}} = \mathbf{P}_{\parallel}\left\{ -\frac{\partial U}{\partial \mathbf{r}} - \frac{\partial}{\partial \mathbf{r}}\left(\frac{1}{2}\mathbf{p}^T \mathbf{M}_c^{+} \mathbf{p} \right) \right\} + \dot{\mathbf{P}}_{\parallel}\mathbf{p}. \tag{2.44}$$

To establish the equivalence to the Langrangian equations of motion, we reintroduce the velocities as dynamical variables, writing $\mathbf{p} = \mathbf{M}_c \dot{\mathbf{r}}$. Using identity (A11), which is proven in Appendix A, it follows that

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{r}}\left(\frac{1}{2}\mathbf{p}^T \mathbf{M}_c^{+} \mathbf{p} \right) &= \frac{1}{2}\mathbf{p}^T \left(\frac{\partial \mathbf{M}_c^{+}}{\partial \mathbf{r}} \right) \mathbf{p} \\
&= -\frac{1}{2}\dot{\mathbf{r}}^T \mathbf{M}_c \left(\mathbf{M}_c^{+} \frac{\partial \mathbf{M}_c}{\partial \mathbf{r}} \mathbf{M}_c^{+} \right) \mathbf{M}_c \dot{\mathbf{r}} \\
&= -\frac{1}{2}\dot{\mathbf{r}}^T \mathbf{P}_{\parallel} \left(\frac{\partial \mathbf{M}_c}{\partial \mathbf{r}} \right) \mathbf{P}_{\parallel} \dot{\mathbf{r}} \\
&= -\frac{1}{2}\dot{\mathbf{r}}^T \left(\frac{\partial \mathbf{M}_c}{\partial \mathbf{r}} \right) \dot{\mathbf{r}} \\
&= -\frac{\partial}{\partial \mathbf{r}}\left(\frac{1}{2}\dot{\mathbf{r}}^T \mathbf{M}_c \dot{\mathbf{r}} \right).
\end{aligned}$$

With $\mathbf{p} = \partial L_c / \partial \dot{\mathbf{r}}$, Eq. (2.44) may be cast into the form

$$\frac{d}{dt}\frac{\partial L_c}{\partial \dot{\mathbf{r}}} = \mathbf{P}_{\parallel}\frac{\partial L_c}{\partial \mathbf{r}} + \dot{\mathbf{p}}_{\perp},$$

and projecting the above equation onto $\mathcal{V}_{\parallel}$ yields again the Euler-Langrange equation [Eq. (2.29)]. The equivalence of the Hamiltonian and Langrangian equations of motion—and consequently also of the Hamiltonian and Newtonian equations of motion—is therefore established.

III. CONCLUSION

It has been proven that the Hamiltonian equations of motion of a constrained dynamical system derived in Ref. 1 are equivalent to the corresponding Newtonian and Lagrangian equations of motion. Splitting the acceleration vector into a component tangential to the surface defined by the imposed constraints and an orthogonal component, the solution of the Newtonian equations of motion can be phrased as a Bott-Duffin problem, which is known from linear algebra. The equivalence of the latter with the minimization of a quadratic form in the presence of linear constraints establishes the equivalence with Gauss' principle of least constraint. The projector formalism makes transparent that the equivalence of the Newtonian and the Lagrangian formulation of the equations of motion is effectively due to the redundancy of the associated velocity constraints. This redundancy allows to introduce a constrained Lagrangian by projecting the velocities explicitly on the tangential space defined by the constraints and yields a one-to-one correspondence between the velocities and the momenta, which is the prerequisite to construct the Hamiltonian for the constrained system. A useful identity for the differentiation of generalized inverse matrices, which is proven in the Appendix, allows to demonstrate the equivalence of the Hamiltonian equations of motion derived in Ref. 1 and the Lagrangian equations of motion. The equivalence of the Newtonian, Lagrangian, and Hamiltonian formulation of the dynamics of constrained systems is thus proven.

APPENDIX A: SOME USEFUL RELATIONS FOR GENERALIZED INVERSE MATRICES

Let $\mathbf{A}$ be an $m \times n$ matrix. The $n \times m$ matrix $\mathbf{A}^{+}$ is called the generalized inverse (pseudoinverse) of $\mathbf{A}$ if it fulfills the following four relations, which are known as Moore-Penrose conditions,[7–10]

$$\mathbf{A}\mathbf{A}^{+}\mathbf{A} = \mathbf{A}, \tag{A1}$$

$$\mathbf{A}^{+}\mathbf{A}\mathbf{A}^{+} = \mathbf{A}^{+}, \tag{A2}$$

$$(\mathbf{A}\mathbf{A}^{+})^T = \mathbf{A}\mathbf{A}^{+}, \tag{A3}$$

$$(\mathbf{A}^{+}\mathbf{A})^T = \mathbf{A}^{+}\mathbf{A}. \tag{A4}$$

If $m \geqslant n$ and all *columns* of $\mathbf{A}$ are linearly independent, $\mathbf{A}^{+}$ may be expressed as

$$\mathbf{A}^{+} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T. \tag{A5}$$

If, in contrast, $m \leqslant n$ and all *rows* of $\mathbf{A}$ are linearly independent, we have

$$\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}. \tag{A6}$$

A special class of generalized inverse matrices are the Bott-Duffin inverses.[9–11] Let $\mathbf{M}$ be an $n\times n$ matrix, $\mathbf{P}$ be an $n\times n$ projector matrix with $\mathbf{P}^2=\mathbf{P}$ and $\mathbf{P}^T=\mathbf{P}$, and $\mathbf{Q}$ its orthogonal complement. If $\det(\mathbf{MP}+\mathbf{Q})\neq 0$, the Bott-Duffin inverse exists and is given by

$$\mathbf{M}_{\mathrm{BD}}^{(-1)} = \mathbf{P}(\mathbf{MP}+\mathbf{Q})^{-1}. \tag{A7}$$

If we define the projected matrix $\mathbf{M}_c=\mathbf{PMP}$, the Bott-Duffin inverse can be expressed as

$$\mathbf{M}_{\mathrm{BD}}^{(-1)} = \mathbf{M}_c^+ \tag{A8}$$

and is thus the inverse of a quadratic matrix with respect to a subspace of $\mathbb{R}^n$. The above relation is mentioned in different references and a simple proof can be found in Ref. 1. We note that $\mathbf{M}_c^+$ verifies the relations

$$\mathbf{M}_c^+\mathbf{M}_c = \mathbf{M}_c\mathbf{M}_c^+ = \mathbf{P}, \tag{A9}$$

$$\mathbf{P}\mathbf{M}_c^+ = \mathbf{M}_c^+\mathbf{P} = \mathbf{M}_c^+. \tag{A10}$$

Consider now the case that the matrix $\mathbf{M}_c$ depends on a parameter λ, $\mathbf{M}_c\equiv\mathbf{M}_c(\lambda)$. With the Moore-Penrose relation [Eq. (A2)], we thus have

$$\begin{aligned}\frac{\partial}{\partial\lambda}\mathbf{M}_c^+ &= \frac{\partial}{\partial\lambda}(\mathbf{M}_c^+\mathbf{M}_c\mathbf{M}_c^+)\\ &= \left(\frac{\partial\mathbf{M}_c^+}{\partial\lambda}\right)\mathbf{M}_c\mathbf{M}_c^+ + \mathbf{M}_c^+\left(\frac{\partial\mathbf{M}_c}{\partial\lambda}\right)\mathbf{M}_c^+\\ &\quad + \mathbf{M}_c^+\mathbf{M}_c\left(\frac{\partial\mathbf{M}_c^+}{\partial\lambda}\right)\\ &= \left(\frac{\partial\mathbf{M}_c^+}{\partial\lambda}\right)\mathbf{P} + \mathbf{M}_c^+\left(\frac{\partial\mathbf{M}_c}{\partial\lambda}\right)\mathbf{M}_c^+ + \mathbf{P}\left(\frac{\partial\mathbf{M}_c^+}{\partial\lambda}\right).\end{aligned}$$

Multiplying this equation from the left and from the right with the projector matrix $\mathbf{P}$ and using Eq. (A10) show that

$$\mathbf{P}\left(\frac{\partial\mathbf{M}_c^+}{\partial\lambda}\right)\mathbf{P} = -\mathbf{M}_c^+\left(\frac{\partial\mathbf{M}_c}{\partial\lambda}\right)\mathbf{M}_c^+. \tag{A11}$$

The corresponding relation for nonsingular matrices is retrieved by setting $\mathbf{P}=\mathbf{1}$, $\mathbf{M}_c=\mathbf{M}$, and $\mathbf{M}_c^+=\mathbf{M}^{-1}$.

APPENDIX B: PROOF OF RELATION (2.34)

We start from the variation $\delta\mathbf{r}(t)$ of the admissible paths in the variation problem [Eq. (2.24)] and consider the variation of the corresponding velocity. Using that for holonomic constraints $\mathbf{A}\dot{\mathbf{r}}=\mathbf{0}$, such that $\dot{\mathbf{r}}=\mathbf{P}_\parallel\dot{\mathbf{r}}$, we thus have

$$\delta\dot{\mathbf{r}} = \delta(\mathbf{P}_\parallel\dot{\mathbf{r}}) = (\delta\mathbf{P}_\parallel)\dot{\mathbf{r}} + \mathbf{P}_\parallel\delta\dot{\mathbf{r}}.$$

Since $\delta\dot{\mathbf{r}}=d/dt\,\delta\mathbf{r}$ and since $\delta\mathbf{r}=\mathbf{P}_\parallel\delta\mathbf{r}$, it follows that

$$\delta\dot{\mathbf{r}} = \frac{d}{dt}\delta\mathbf{r} = \frac{d}{dt}(\mathbf{P}_\parallel\delta\mathbf{r}) = \dot{\mathbf{P}}_\parallel\delta\mathbf{r} + \mathbf{P}_\parallel\delta\dot{\mathbf{r}}.$$

Equating the above two expressions for $\delta\dot{\mathbf{r}}$ shows that

$$(\delta\mathbf{P}_\parallel)\dot{\mathbf{r}} = \dot{\mathbf{P}}_\parallel\delta\mathbf{r}.$$

With $P_{ij}^\parallel$ being the components of $\mathbf{P}_\parallel$, it follows from the above relation that

$$\frac{\partial P_{ij}^\parallel}{\partial r_k}\delta r_k\dot{r}_j = \frac{\partial P_{ij}^\parallel}{\partial r_k}\delta r_j\dot{r}_k.$$

Interchanging the indices j and k on the right-hand side thus leads to

$$\left(\frac{\partial P_{ij}^\parallel}{\partial r_k} - \frac{\partial P_{ik}^\parallel}{\partial r_j}\right)\delta r_k\dot{r}_j = 0,$$

which may be written in the alternative form

$$\left(\frac{\partial}{\partial r_k}(P_{ij}^\parallel\dot{r}_j) - \dot{P}_{ik}^\parallel\right)\delta r_k = 0.$$

Since $\delta\mathbf{r}$ is an arbitrary vector in $\mathcal{V}_\parallel$, the above relation is equivalent to

$$\left(\frac{\partial}{\partial r_k}(P_{ij}^\parallel\dot{r}_j) - \dot{P}_{ik}^\parallel\right)P_{kl}^\parallel = 0. \tag{B1}$$

Consider now identity (2.34), which may be cast into the form

$$\mathbf{P}_\parallel\left\{\left(\frac{\partial}{\partial\mathbf{r}}\mathbf{P}_\parallel\dot{\mathbf{r}}\right)^T - \dot{\mathbf{P}}_\parallel\right\}\mathbf{M}\mathbf{P}_\parallel\dot{\mathbf{r}} = \mathbf{0},$$

due to the symmetry of $\mathbf{M}_c$. Since $\mathbf{P}_\parallel\{\cdots\}=0$ on account of Eq. (B1), the above relation is indeed fulfilled and identity (2.34) is proven.

[1] G. R. Kneller, J. Chem. Phys. **125**, 114107 (2006).
[2] M. Fixman, Proc. Natl. Acad. Sci. U.S.A. **8**, 3050 (1974).
[3] S. W. de Leeuw, J. W. Perram, and H. G. Petersen, J. Stat. Phys. **61**, 1203 (1990).
[4] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).
[5] P. A. M. Dirac, Proc. R. Soc. London, Ser. A **246**, 326 (1958).
[6] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1018 (1951).
[7] E. H. Moore, Bull. Am. Math. Soc. **26**, 394 (1920).
[8] R. Penrose, Proc. Cambridge Philos. Soc. **51**, 406 (1955).
[9] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications* (Wiley, New York, 1974).
[10] S. L. Campbell and C. D. Meyer, Jr., *Generalized Inverses of Linear Transformations* (Pitman, New York, 1979).
[11] R. Bott and R. J. Duffin, Trans. Am. Math. Soc. **74**, 99 (1953).
[12] C. F. Gauss, J. Reine Angew. Math. **4**, 232 (1829).