Hamiltonian formalism for semiflexible molecules in Cartesian coordinates

G. R. Kneller[a)]
Centre de Biophysique Moléculaire, CNRS, Rue Charles Sadron, 45071 Orléans, France and Synchrotron Soleil, L'Orme de Merisiers, B.P. 48, 91192 Gif-sur-Yuette, France

(Received 14 April 2006; accepted 13 June 2006; published online 21 September 2006)

The article gives a concise description of Hamiltonian dynamics and thermal averages of semiflexible molecules in Cartesian coordinates. Using the concept of constrained inverse matrices introduced by Bott and Duffin [Trans. Am. Math. Soc. **74**, 99 (1953)] explicit expressions are derived for the constrained Hamiltonian, the corresponding equations of motion, and the momentum partition function. In this context Fixman-type corrections of constrained configurational averages are derived for different forms of the constraints. It is shown that the use of mass-weighted coordinates leads to a nonbiased sampling of constrained configurational averages in Cartesian coordinates. The formalism allows moreover to define and to calculate effective masses arising in thermal velocity averages of atoms in semiflexible molecules. These effective masses are identical to the corresponding Sachs-Teller recoil masses, which are here generalized to the case of only partially rigid molecules. © *2006 American Institute of Physics*. [DOI: 10.1063/1.2220037]

I. INTRODUCTION

With the development of computer simulation techniques for simple and complex molecular systems, the field of classical mechanics has regained the interest of many researchers.[1,2] One of the major achievements in the development of classical mechanics was the introduction of variational calculus which allowed to use problem-adapted coordinates.[3] The elegance of this approach, promoted by Lagrange and Hamilton, is very convincing if one looks at the typical examples from astronomy and simple point mechanics. In computer simulations of large molecular systems, the situation is, however, very different. Here thousands of degrees of freedom must be handled and convenient generalized coordinates are not only difficult to find, but they might moreover lead to terms in the equations of motion which are difficult to handle from a numerical point of view. Another disadvantage of generalized coordinates is that they often require problem-specific implementations, which is a major obstacle in the development of all-purpose programs, especially for versions that can be executed on massively parallel computers.

While Cartesian coordinates are the method of choice concerning the algorithmic part in molecular dynamics and Monte Carlo simulations of macromolecular systems, their use is much less straightforward if thermal averages are to be computed in the presence of constraints. Even such a seemingly simple task as the calculation of the mean square velocity of an atom in a rigid molecule poses a serious problem. In contrast to generalized coordinates, which allow to introduce subspaces in the form of subsets of generalized coordinates or momenta, Cartesian coordinates necessitate the consideration of explicit holonomic constraints and corresponding tangential spaces. Formal aspects of Hamiltonian dynamics with constraints have been treated in the works of Andersen and Bergman[4] and Dirac.[5,6] Here the motivation was to set up relativistic quantum theories, treating the required invariance properties with respect to Lorentz transformations and gauge transformations as constraints. Later, de Leeuw *et al.* discussed the problem of constrained Hamiltonian dynamics in the light of molecular dynamics simulations[7] using the method of Lagrange multipliers in order to account for explicit holonomic constraints. The purpose of this paper is to show that a conceptually elegant approach to constrained Hamiltonian dynamics can be developed on the basis of projectors and generalized inverse matrices[9,10] using, in particular, the Bott-Duffin inverse.[11] The focus here is on the derivation of Hamiltonian equations of motion in *Cartesian* coordinates and on the calculation of corresponding phase space averages. A Bott-Duffin inverse can be considered as matrix inverse with respect to a given subspace. Of particular interest are the constrained mass matrix and its Bott-Duffin inverse, which are associated with the mass metric tensor in generalized coordinates and its inverse, respectively. The discriminant of the Bott-Duffin inverse turns out to be very useful for the calculation of thermal averages in momentum space. It is also shown that constrained inverse mass matrices allow to define effective atomic masses for semiflexible molecules. The latter are identical to the Sachs-Teller recoil masses, whose definition is here extended to only partially rigid molecules.

The paper is organized as follows. In Sec. II the Hamiltonian of a constrained dynamical system in Cartesian coordinates is constructed and the corresponding equations of motion are derived. Section III presents examples for the calculation of momentum averages and related effective atomic masses in semiflexible molecules. In Sec. IV the formalism is applied to the correction of thermal averages due to geometrical constraints. The paper is concluded by a final

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

discussion and the appendixes contain the mathematical background in a self-contained form, as well as certain proofs of relations given in the text.

II. INVERSE MASS TENSOR AND CONSTRAINED HAMILTONIAN

A. Constraints and projectors

We consider a system of N atoms which is described by $3N$ Cartesian coordinates. The system may be subjected to s holonomic constraints, which can be specified in terms of a set of conditions for the Cartesian coordinates,

$$h_k(\mathbf{r}) = h_k^{(0)}, \quad k = 1, \ldots, s, \tag{2.1}$$

where $h_k^{(0)}$=const. Here $\mathbf{r}$ is a column vector of length $3N$ which comprises all Cartesian coordinates of the N atoms. In a dynamical system the positions are functions of time, and differentiating (2.1) therefore yields s linear constraints for the velocities. Introducing the $s \times 3N$ matrix $\mathbf{A}$ through

$$A_{ki} = \frac{\partial h_k}{\partial r_i}, \tag{2.2}$$

($i=1,\ldots,3N$, $k=1,\ldots,s$) one has

$$\mathbf{A}\mathbf{v} = \mathbf{0}. \tag{2.3}$$

The above relation indicates that the velocity vector is in the null space of $\mathbf{A}$, which will be denoted as $\mathcal{V}_\parallel$ and represents the tangential space related to the constraints (2.1). A projector onto that tangential space can be written in the form

$$\mathbf{P}_\parallel = \mathbf{1} - \mathbf{A}^+\mathbf{A}, \tag{2.4}$$

where $\mathbf{A}^+$ is the pseudoinverse of $\mathbf{A}$.[9,10,12,13] A short presentation of pseudoinverse matrices is given in Appendix A. If all s rows in $\mathbf{A}$ are linearly independent, $\mathbf{A}^+$ can be expressed as

$$\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}, \tag{2.5}$$

where the superscript T denotes a transposition. By construction, $\mathbf{A}\mathbf{P}_\parallel = 0$ and any velocity vector fulfilling $\mathbf{P}_\parallel\mathbf{v} = \mathbf{v}$ thus verifies the imposed constraints (2.3). We note that

$$\mathbf{P}_\perp = \mathbf{A}^+\mathbf{A} \tag{2.6}$$

is the projector onto the row space of $\mathbf{A}$ and fulfills $\mathbf{P}_\perp\mathbf{P}_\parallel = 0$. The rows of $\mathbf{A}$ thus span the orthogonal complement of $\mathcal{V}_\parallel$, which will be denoted as $\mathcal{V}_\perp$ in the following.

In some cases it may be advantageous to express the Cartesian velocities $\mathbf{v}$ in terms of appropriate generalized velocities $\mathbf{u}=(u_1,\ldots,u_f)^T$,

$$\mathbf{v} = \mathbf{C}\mathbf{u}, \tag{2.7}$$

with f being the number of degrees of freedom and $f+s = 3N$. We recall that s is the number of constraints. In case that the generalized velocities are the time derivatives of f generalized coordinates $q_1,\ldots,q_f$, the $3N \times f$ matrix $\mathbf{C}$ is a Jacobian, with $C_{ik} = \partial r_i / \partial q_k$ ($i=1,\ldots,3N$, $k=1,\ldots,f$). The matrices $\mathbf{A}$ and $\mathbf{C}$ fulfill

$$\mathbf{A}\mathbf{C} = \mathbf{0} \tag{2.8}$$

and the projector $\mathbf{P}_\parallel$ can be written in the alternative form

$$\mathbf{P}_\parallel = \mathbf{C}\mathbf{C}^+. \tag{2.9}$$

Here $\mathbf{C}^+$ is given by

$$\mathbf{C}^+ = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T \tag{2.10}$$

if all columns in $\mathbf{C}$ are linearly independent. If one is dealing with rigid molecules, each composed of N atoms, the generalized velocities are the center-of-mass velocity $\mathbf{v}_{\mathrm{CM}}$ and the angular velocity $\boldsymbol{\omega}$. The latter may refer to a laboratory-fixed or to a co-rotating frame. In the first case the vector of generalized velocities reads

$$\mathbf{u} = \begin{pmatrix} \mathbf{v}_{\mathrm{CM}} \\ \boldsymbol{\omega} \end{pmatrix}, \tag{2.11}$$

and $\mathbf{C}$ has the simple form

$$\mathbf{C} = \begin{pmatrix} \mathbf{1} & -\mathbf{R}_1 \\ \mathbf{1} & -\mathbf{R}_2 \\ \vdots & \vdots \\ \mathbf{1} & -\mathbf{R}_N \end{pmatrix}. \tag{2.12}$$

Here $\mathbf{1}$ is the 3×3 unit matrix, and $\mathbf{R}$ is the skew-symmetric matrix,

$$\mathbf{R} = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}, \tag{2.13}$$

containing the laboratory-fixed coordinates of the position $\mathbf{r}=(x,y,z)^T$ with respect to the center of mass. More complex C matrices for linked rigid bodies have been described in Ref. 14.

B. Hamiltonian dynamics of constrained systems

1. Constructing the Hamiltonian

Consider the Lagrangian of a dynamical system consisting of N mass points,

$$L(\mathbf{r},\dot{\mathbf{r}}) = \frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}\dot{\mathbf{r}} - U(\mathbf{r}), \tag{2.14}$$

where $\mathbf{M}$ is the diagonal mass matrix,

$$\mathbf{M} = \begin{pmatrix} m_1\mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & m_2\mathbf{1} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & m_N\mathbf{1} \end{pmatrix}, \tag{2.15}$$

and U is the potential energy of the system. If the system is subjected to geometrical constraints, the velocities are confined to the tangential space $\mathcal{V}_\parallel$, such that $\mathbf{P}_\parallel\mathbf{v} = \mathbf{v}$. In this situation the Lagrangian may be written in the form

$$L_c(\mathbf{r},\dot{\mathbf{r}}) = \frac{1}{2}\dot{\mathbf{r}}^T\mathbf{M}_c\dot{\mathbf{r}} - U(\mathbf{r}), \tag{2.16}$$

where $\mathbf{M}_c$ is the mass matrix projected onto $\mathcal{V}_\parallel$,

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

The momenta are derived in the usual way,

$$\mathbf{p} = \frac{\partial L_c}{\partial \dot{\mathbf{r}}} = \mathbf{M}_c \mathbf{v}, \tag{2.18}$$

and in order to construct the Hamiltonian of the constrained system one must be able to express the velocities through the momenta. For this purpose one uses the relation

$$\mathbf{M}_c \mathbf{M}_e^+ = \mathbf{M}_c^+ \mathbf{M}_c = \mathbf{P}_\parallel, \tag{2.19}$$

where $\mathbf{P}_\parallel$ is at the same time projector on the row and column spaces of the symmetric positive semidefinite matrix $\mathbf{M}_c$. On account of the first Moore-Penrose condition (A1), the above relation allows to write $\mathbf{P}_\parallel \mathbf{p} = \mathbf{M}_c \mathbf{M}_c^+ \mathbf{M}_c \mathbf{v} = \mathbf{M}_c \mathbf{v} = \mathbf{p}$. The relation $\mathbf{M}_c \mathbf{v} = \mathbf{p}$ is thus consistent if $\mathbf{p} \in V_\parallel$. In this case a unique solution for $\mathbf{v} \in \mathcal{V}_\parallel$ exists and reads

$$\mathbf{v} = \mathbf{M}_c^+ \mathbf{p}. \tag{2.20}$$

As shown in Appendix B (Theorems 4 and 5), $\mathbf{M}_c^+$ can be expressed in terms of the matrix $\mathbf{A}$,[8]

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

and equivalently in terms of the matrix $\mathbf{C}$,

$$\mathbf{M}_c^+ = \mathbf{C}(\mathbf{C}^T\mathbf{M}\mathbf{C})^{-1}\mathbf{C}^T. \tag{2.22}$$

We may now construct the Hamiltonian of the constrained system via $H_c = \mathbf{p}^T\dot{\mathbf{r}} - L_c$, expressing the velocities through (2.20) in terms of the momenta,

$$H_c(\mathbf{p},\mathbf{r}) = \frac{1}{2}\mathbf{p}^T\mathbf{M}_c^+\mathbf{p} + U(\mathbf{r}). \tag{2.23}$$

It is important to note that $\mathbf{P}_\parallel \mathbf{M}_c^+ = \mathbf{M}_c^+ \mathbf{P}_\parallel = \mathbf{M}_c^+$, which shows that both $\mathbf{M}_c$ and $\mathbf{M}_c^+$ act only in the subspace $\mathcal{V}_\parallel$, and it must be kept in mind that they both depend on the Cartesian coordinates $\mathbf{r}$.

2. Introducing generalized coordinates

The form (2.23) for the constrained Hamiltonian can be motivated by introducing f generalized coordinates $\mathbf{q} = (q_1, \ldots, q_f)^T$ and the associated momenta $\boldsymbol{\pi}_q = (\pi_{q,1}, \ldots, \pi_{q,f})^T$. As indicated above, one has in this case $\mathbf{v} = \mathbf{C}\mathbf{u}$, where $\mathbf{u} = (\dot{q}_1, \ldots, \dot{q}_f)^T$ and the elements of $\mathbf{C}$ are given by $C_{ik} = \partial r_i / \partial q_k$, i.e., $\mathbf{C}$ is a Jacobian. Introducing the reduced mass metric matrix

$$\mathbf{G}_{qq} = \mathbf{C}^T\mathbf{M}\mathbf{C}, \tag{2.24}$$

the constrained Lagrangian has the form

$$L_c = \frac{1}{2}\dot{\mathbf{q}}^T\mathbf{G}_{qq}\dot{\mathbf{q}} - U(\mathbf{q}) \tag{2.25}$$

and leads to the Hamiltonian

$$H_c = \frac{1}{2}\boldsymbol{\pi}_q^T\mathbf{G}_{qq}^{-1}\boldsymbol{\pi}_q + U(\mathbf{q}), \tag{2.26}$$

where $\boldsymbol{\pi}_q = \partial L_c / \partial \dot{\mathbf{q}} = \mathbf{G}_{qq}\dot{\mathbf{q}}$. Since $\mathbf{P}_\parallel \mathbf{C} = \mathbf{C}$ and $\mathbf{v} = \mathbf{C}\dot{\mathbf{q}}$, it follows from the constrained Cartesian momenta that $\mathbf{C}^T\mathbf{p} = \mathbf{C}^T\mathbf{M}_c\mathbf{v} = \mathbf{C}^T\mathbf{P}_\parallel\mathbf{M}\mathbf{P}_\parallel\mathbf{v} = \mathbf{C}^T\mathbf{M}\mathbf{C}\dot{\mathbf{q}} = \boldsymbol{\pi}_q$. Using the form (2.22) for $\mathbf{M}_c^+$ shows that $\mathbf{p}^T\mathbf{M}_c^+\mathbf{p} = \boldsymbol{\pi}_q^T\mathbf{G}_{qq}^{-1}\boldsymbol{\pi}_q$, proving that H_c as given by (2.23) and (2.26) are equivalent.

3. Equations of motion

The Hamiltonian equations of motion are derived from the variational principle $\delta S = \delta\int_{t_0}^{t_1} d\tau L = 0$, which leads to

$$\delta S = \int_{t_0}^{t_1} d\tau \left\{ \boldsymbol{\delta}\mathbf{p}^T\left(\dot{\mathbf{r}} - \frac{\partial H}{\partial \mathbf{p}}\right) - \boldsymbol{\delta}\mathbf{r}^T\left(\dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{r}}\right)\right\} = 0 \tag{2.27}$$

if coordinates and momenta are considered as independent variables. Since the system under consideration is subject to constraints, the variations $\boldsymbol{\delta}\mathbf{r}$ and $\boldsymbol{\delta}\mathbf{p}$ are *not* arbitrary. It follows from the conditions (2.1) that $\delta h_k(\mathbf{r}) = (\partial h_k(\mathbf{r})/\partial r_i)\delta r_i \equiv 0$, where all variations of the paths must satisfy $\boldsymbol{\delta}\mathbf{r}(t_0) = \boldsymbol{\delta}\mathbf{r}(t_1) = \mathbf{0}$. Here and in the following summation over indices occurring twice is implicitly assumed. Using the definition of the matrix $\mathbf{A}$ given in (2.2) one finds thus $\mathbf{A}\boldsymbol{\delta}\mathbf{r} = \mathbf{0}$, showing that

$$\boldsymbol{\delta}\mathbf{r} \in \mathcal{V}_\parallel. \tag{2.28}$$

The condition for the momenta $\mathbf{p} \in \mathcal{V}_\parallel$, which follows from the constrained Lagrangian, may be cast into the form $\mathbf{A}\mathbf{p} = \mathbf{0}$. Varying this condition with respect to the momenta yields $\mathbf{A}\boldsymbol{\delta}\mathbf{p} = \mathbf{0}$. The momentum variations are thus in the same subspace as the position variations,

$$\boldsymbol{\delta}\mathbf{p} \in \mathcal{V}_\parallel, \tag{2.29}$$

and it thus follows from the variational principle (2.27) that

$$\mathbf{P}_\parallel\left(\dot{\mathbf{r}} - \frac{\partial H}{\partial \mathbf{p}}\right) = \mathbf{0}, \tag{2.30}$$

$$\mathbf{P}_\parallel\left(\dot{\mathbf{p}} + \frac{\partial H}{\partial \mathbf{r}}\right) = \mathbf{0} \tag{2.31}$$

are necessary conditions for the stationarity of the action integral S. By construction, the term contained in the parenthesis of Eq. (2.30) is in $\mathcal{V}_\parallel$ and the corresponding projector can be omitted. In contrast to $\dot{\mathbf{r}}$, the time derivative of the momenta has components in both $\mathcal{V}_\parallel$ and $\mathcal{V}_\perp$. Equation (2.31) thus yields only the component $\dot{\mathbf{p}}_\parallel$. The orthogonal complement can be obtained from the time derivative of the momentum constraint $\mathbf{A}\mathbf{p} = \mathbf{0}$, which yields a linear equation for $\dot{\mathbf{p}}$ of the form $\dot{\mathbf{A}}\mathbf{p} + \mathbf{A}\dot{\mathbf{p}} = \mathbf{0}$. The general solution reads

$$\dot{\mathbf{p}} = \dot{\mathbf{p}}_\parallel \underbrace{- \mathbf{A}^+\dot{\mathbf{A}}\mathbf{p}}_{\dot{\mathbf{p}}_\perp} \tag{2.32}$$

and determines $\dot{\mathbf{p}}_\perp$. Since $\dot{\mathbf{p}}_\parallel$ follows from Eq. (2.31), the Hamiltonian equations of motion can finally be written down,

$$\dot{\mathbf{r}} = \mathbf{M}_c^+\mathbf{p}, \tag{2.33}$$

$$\dot{\mathbf{p}} = \mathbf{P}_\parallel\left\{ - \frac{\partial U}{\partial \mathbf{r}} - \frac{1}{2}\mathbf{p}^T\left(\frac{\partial \mathbf{M}_c^+}{\partial \mathbf{r}}\right)\mathbf{p}\right\} - \mathbf{A}^+\dot{\mathbf{A}}\mathbf{p}. \tag{2.34}$$

It should be noted that the above form of the equations of motion has been derived with out applying the concept of Lagrange multipliers, which is the usual method to take into

account constraints.[7] The equations of motion (2.33) and (2.34) are complete in the sense that all quantities involved can be explicitly expressed in the coordinates and momenta of the particles involved.

III. MOMENTUM AVERAGES AND EFFECTIVE MASSES

A. Mean square momenta in constrained systems

In the following we consider atomic mean square momenta for a system composed of semiflexible molecules, each containing N atoms. The mean square momenta can be computed from the moment generating function,

$$G(\mathbf{k}) = \int d^{3N}p\,\delta(\mathbf{P}_\perp\mathbf{p})w_c(\mathbf{p})\exp(i\mathbf{k}^T\mathbf{p}), \tag{3.1}$$

via the relation

$$\langle p_i p_j\rangle = -\left.\frac{\partial^2 G(\mathbf{k})}{\partial k_i \partial k_j}\right|_{\mathbf{k}=\mathbf{0}}, \tag{3.2}$$

where $w_c(\mathbf{p})$ is the constrained Maxwell distribution,

$$w_c(\mathbf{p}) = \frac{\exp(-(\beta/2)\mathbf{p}^T\mathbf{M}_c^+\mathbf{p})}{\int d^{3N}p\,\delta(\mathbf{P}_\perp\mathbf{p})\exp(-(\beta/2)\mathbf{p}^T\mathbf{M}_c^+\mathbf{p})}. \tag{3.3}$$

Here $\beta=(k_BT)^{-1}$, T is the absolute temperature, and k_B is the Boltzmann constant. The constraint $\mathbf{P}_\parallel\mathbf{p}=\mathbf{p}$, or equivalently $\mathbf{P}_\perp\mathbf{p}=\mathbf{0}$, is accounted for by the Dirac distribution $\delta(\mathbf{P}_\perp\mathbf{p})$. The evaluation of the moment generating function is easily performed in a basis in which $\mathbf{M}_c$ is diagonal. Such a basis is found by singular value decomposition of $\mathbf{M}_c$. Since $\mathbf{M}_c\mathbf{M}_c^+=\mathbf{M}_c^+\mathbf{M}_c=\mathbf{P}_\parallel$, the projectors $\mathbf{P}_\parallel$ and $\mathbf{P}_\perp$ are diagonal in the same basis. The details of the calculation are described in Appendix C, and we give here only the result

$$\langle\mathbf{p}\mathbf{p}^T\rangle = k_BT\mathbf{M}_c. \tag{3.4}$$

The usefulness of the concept of pseudoinverses is illustrated by the calculation of the mean kinetic energy, $E_{\mathrm{kin}}=\langle\frac{1}{2}\mathbf{p}^T\mathbf{M}_c^+\mathbf{p}\rangle=\frac{1}{2}\mathrm{tr}\{\langle\mathbf{p}\mathbf{p}^T\mathbf{M}_c^+\rangle\}=\frac{1}{2}\mathrm{tr}\{\langle\mathbf{p}\mathbf{p}^T\rangle\mathbf{M}_c^+\}$. Using that $\mathbf{M}_c\mathbf{M}_c^+=\mathbf{P}_\parallel$ is the projector on the f-dimensional subspace $\mathcal{V}_\parallel$ and that $\mathrm{tr}\{\mathbf{P}_\parallel\}=f$, one verifies easily that

$$E_{\mathrm{kin}} = \frac{k_BT}{2}\mathrm{tr}\{\mathbf{P}_\parallel\} = \frac{k_BT}{2}f. \tag{3.5}$$

Here "tr" denotes the trace of a matrix.

Relation (3.4) can be readily converted into an expression for the atomic mean square velocities in a semiflexible molecule. In this case one writes $\langle\mathbf{v}\mathbf{v}^T\rangle=\langle\mathbf{M}_c^+\mathbf{p}\mathbf{p}^T\mathbf{M}_c^+\rangle=\mathbf{M}_c^+\langle\mathbf{p}\mathbf{p}^T\rangle\mathbf{M}_c^+$, and with $\mathbf{M}_c^+\mathbf{M}_c\mathbf{M}_c^+=\mathbf{M}_c^+$ one obtains from (3.4)

$$\langle\mathbf{v}\mathbf{v}^T\rangle = k_BT\mathbf{M}_c^+. \tag{3.6}$$

Here $\mathbf{M}_c^+$ can be expressed in the form (2.21) or (2.22). To obtain the mean square velocity of individual atoms, the matrix $\mathbf{M}_c^+$ is partitioned in 3×3 block matrices,

$$\mathbf{M}_c^+ = \begin{pmatrix} \mathbf{m}_{11}^+ & \mathbf{m}_{12}^+ & \cdots & \mathbf{m}_{1N}^+ \\ \mathbf{m}_{21}^+ & \mathbf{m}_{22}^+ & \cdots & \mathbf{m}_{2N}^+ \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{m}_{N1}^+ & \cdots & \cdots & \mathbf{m}_{NN}^+ \end{pmatrix}. \tag{3.7}$$

For atom α one thus has

$$\langle\mathbf{v}_\alpha\mathbf{v}_\alpha^T\rangle = k_BT\mathbf{m}_{\alpha\alpha}^+, \tag{3.8}$$

and it follows that the mean square velocity $\langle\mathbf{v}_\alpha^T\mathbf{v}_\alpha\rangle$ is given by

$$\langle\mathbf{v}_\alpha^T\mathbf{v}_\alpha\rangle = k_BT\,\mathrm{tr}\{\mathbf{m}_{\alpha\alpha}^+\}. \tag{3.9}$$

The effective mass M_α of atom α can be defined through the relation $\langle\mathbf{v}_\alpha^T\mathbf{v}_\alpha\rangle=3k_BT/M_\alpha$ and one obtains

$$M_\alpha^{-1} = \frac{1}{3}\mathrm{tr}\{\mathbf{m}_{\alpha\alpha}^+\}. \tag{3.10}$$

At this point one should note that the definition of effective masses is to some extent arbitrary. One could as well start from the relation $\langle\mathbf{p}_\alpha^T\mathbf{p}_\alpha\rangle=3k_BTM_\alpha$, defining $M_\alpha=\frac{1}{3}\mathrm{tr}\{\mathbf{m}_{\alpha\alpha}\}$, where $\mathbf{m}_{\alpha\beta}$ are 3×3 submatrices of the matrix $\mathbf{M}_c$, and the results are in general not the same. It will be shown below that the effective atomic masses defined according to (3.10) are identical to the well-known Sachs-Teller recoil masses, in case that rigid molecules are considered.

Formula (3.10) will now be applied to a simple model system—a molecule of water which is considered first as rigid and then as semiflexible. In the latter case the H–H distance is not fixed.

(a) *Rigid water*. We consider a rigid water molecule with fixed O–H bond lengths of 1 Å and with a fixed H–O–H angle of 109.47°. The masses are $m_{\mathrm{H}_1}=m_{\mathrm{H}_2}=1.008$ and $m_{\mathrm{O}}=16$ using atomic mass units. In case of rigid water molecules, one can define three distance constraints,

$$h_1(\mathbf{r}) \equiv (\mathbf{r}_{\mathrm{O}}-\mathbf{r}_{\mathrm{H}_1})^2 - 1 = 0,$$

$$h_2(\mathbf{r}) \equiv (\mathbf{r}_{\mathrm{O}}-\mathbf{r}_{\mathrm{H}_2})^2 - 1 = 0,$$

$$h_3(\mathbf{r}) \equiv (\mathbf{r}_{\mathrm{H}_1}-\mathbf{r}_{\mathrm{H}_2})^2 - 8/3 = 0,$$

where all lengths are measured in angstroms. Constructing $\mathbf{A}$ according to (2.2) and $\mathbf{M}$ according to (2.15), the pseudoinverse mass matrix $\mathbf{M}_c^+$ can be obtained from expression (2.21). From the resulting matrix the block matrices $\mathbf{m}_{11}^+$, $\mathbf{m}_{22}^+$, and $\mathbf{m}_{33}^+$ can be extracted, and the effective masses are computed according to definition (3.10). Here 1, 2, and 3 correspond, respectively, to H_1, H_2, and O. The result is

$$m_{\mathrm{H}_1} = m_{\mathrm{H}_2} = 1.896,$$

$$m_{\mathrm{O}} = 17.08.$$

(b) *Semiflexible water*. Here we have one constraint less than in the case of rigid water,

$$h_1(\mathbf{r}) \equiv (\mathbf{r}_O - \mathbf{r}_{H_1})^2 - 1 = 0,$$

$$h_2(\mathbf{r}) \equiv (\mathbf{r}_O - \mathbf{r}_{H_2})^2 - 1 = 0.$$

Following the same steps as for rigid water, one obtains

$$m_{H_1} = m_{H_2} = 1.469,$$

$$m_O = 16.65.$$

Here it has been assumed that the H–O–H angle is the same as in the rigid case. Removing a constraint leads thus to effective masses which are closer to those of the corresponding free atoms.

B. Sachs-Teller problem

We will now show that pseudoinverse mass matrices are also a very elegant tool to compute effective atomic masses which arise in the context of moderation of neutrons by molecular systems. The problem has been studied a long time ago by Sachs and Teller,[15] and we revisit the problem here to show the equivalence with the masses (3.10) arising in thermal velocity averages. At the same time we generalize the Sachs-Teller problem to define effective atomic masses in semiflexible molecules. The Sachs-Teller problem can be formulated as follows. Given a momentum transfer $\mathbf{\Delta p}_\alpha$ transferred to atom α in a rigid molecule, compute the mean velocity change $\mathbf{\Delta v}_\alpha = \mathbf{\Delta p}_\alpha / M_\alpha$ of the same atom, where M_α is the effective Sachs-Teller mass. It is important to note that the atomic mass "seen" by the neutron depends on its energy. Slow neutrons, which cannot excite intermolecular vibrations, "see" effectively mass points in a mechanically rigid molecule.

The starting point is the velocity constraint $\mathbf{Av}=\mathbf{0}$, which must be fulfilled at any time. Differentiating this constraint with respect to time yields $\mathbf{A}\dot{\mathbf{v}}+\dot{\mathbf{A}}\mathbf{v}=\mathbf{0}$. Writing $\Delta\mathbf{v}=\dot{\mathbf{v}}\Delta t$ and $\mathbf{\Delta A}=\dot{\mathbf{A}}\Delta t$ yields a constraint for the velocity change $\mathbf{\Delta v}$ within the time span Δt, which may be considered as collision time in this context,

$$\mathbf{A\Delta v} = -(\mathbf{\Delta A})\mathbf{v}. \tag{3.11}$$

Provided that $\mathbf{AA}^+(\mathbf{\Delta A})\mathbf{v}=(\mathbf{\Delta A})\mathbf{v}$, the above equation has a set of solutions which reads

$$\mathbf{\Delta v} = \mathbf{\Delta v}_\parallel - \mathbf{A}^+(\mathbf{\Delta A})\mathbf{v}. \tag{3.12}$$

Here $\mathbf{\Delta v}_\parallel$ is an arbitrary vector in $\mathcal{V}_\parallel$, satisfying $\mathbf{A\Delta v}_\parallel=\mathbf{0}$. Consider now a momentum transfer of the form

$$\mathbf{\Delta p} = \begin{pmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \mathbf{\Delta p}_\alpha \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix} \tag{3.13}$$

which describes the situation that atom α is hit by the neutron. The relation between an arbitrary momentum change $\mathbf{\Delta p}$ and the resulting velocity change $\mathbf{\Delta v}$ in the presence of constraints must have the form

$$\mathbf{M\Delta v} + \mathbf{\Delta p}_\perp = \mathbf{\Delta p}, \tag{3.14}$$

where $\mathbf{\Delta v}$ is given by (3.12). Without the term $\mathbf{\Delta p}_\perp$, Eq. (3.14) will in general not have a solution, since $\mathbf{\Delta v}$ is constrained but $\mathbf{\Delta p}$ is not. Imposing $\mathbf{\Delta p}_\perp \in \mathcal{V}_\perp$ makes (3.14) a Bott-Duffin problem[11] (see Appendix B) and ensures that both $\mathbf{\Delta v}_\parallel$ and $\mathbf{\Delta p}_\perp$ can be determined from the same system of linear equations. From a physical point of view, $\mathbf{\Delta p}_\perp$ is proportional to the vector of constraint forces and describes the momentum transfer absorbed by the constraints, i.e., the momentum transfer which will not lead to a change in velocity. Solving the Bott-Duffin problem (3.14) leads to $\mathbf{\Delta v} = \mathbf{M}_c^+\mathbf{\Delta p} + (\mathbf{M}_c^+\mathbf{M}-\mathbf{1})\mathbf{A}^+(\mathbf{\Delta A})\mathbf{v}$, where $\mathbf{M}_c$ is the mass matrix projected onto the subspace $\mathcal{V}_\parallel$, which has been defined in Eq. (2.17). A homogeneous linear relation between $\mathbf{\Delta v}$ and $\mathbf{\Delta p}$,

$$\mathbf{\Delta v} = \mathbf{M}_c^+\mathbf{\Delta p}, \tag{3.15}$$

is thus obtained if one chooses a comoving frame, in which all atomic velocities before the collision are zero,

$$\mathbf{v} = \mathbf{0}. \tag{3.16}$$

In this case the velocity change of atom α reads, in particular,

$$\mathbf{\Delta v}_\alpha = \mathbf{m}_{\alpha\alpha}^+\mathbf{\Delta p}, \tag{3.17}$$

using the partitioning of $\mathbf{M}_c^+$ into 3×3 block matrices $\mathbf{m}_{\alpha\beta}$, as defined in (3.7). The block matrix $\mathbf{m}_{\alpha\alpha}^+$ contains the components of the inverse mass tensor for atom α, and the invariant

$$M_\alpha^{-1} = \frac{1}{3}\mathrm{tr}\{\mathbf{m}_{\alpha\alpha}^+\} \tag{3.18}$$

is the inverse of the corresponding effective mass. Equation (3.10) shows that the latter is identical to the effective mass arising in the calculation of atomic mean square velocities in semiflexible molecules. One verifies that the Sachs-Teller inverse mass tensor, whose elements are given by[15]

$$(\mathbf{m}_{ST}^{-1})_{ii} = \frac{r_j^2}{\Theta_k} + \frac{r_k^2}{\Theta_j} + \frac{1}{M}, \quad (\mathbf{m}_{ST}^{-1})_{ij} = -\frac{r_i r_j}{\Theta_k}, \tag{3.19}$$

yields the identical effective masses if the molecules under consideration are rigid. In (3.19) i, j, and k denote the principal axes of the molecule, r_i are the corresponding coordi-

nates of atom α with respect to the center of mass, and Θ_k are the principal moments of inertia.

IV. PHASE SPACE AVERAGES AND FIXMAN CORRECTION

A. Constrained configurational averages

We now consider phase space averages of position-dependent quantities. In the presence of geometrical constraints, the average of a function $F(\mathbf{r})$ is given by

$$\langle F\rangle = \frac{1}{Z^c}\int d^{3N}r\prod_{i=1}^{s}\delta(h_i(\mathbf{r})-h_i^{(0)})Z_p^c(\mathbf{r})F(\mathbf{r}) \times\exp\left(-\frac{\beta}{2}U(\mathbf{r})\right), \tag{4.1}$$

where Z^c is the partition function of the constrained system,

$$Z^c = \int d^{3N}r\prod_{i=1}^{s}\delta(h_i(\mathbf{r})-h_i^{(0)})Z_p^c(\mathbf{r})\exp\left(-\frac{\beta}{2}U(\mathbf{r})\right), \tag{4.2}$$

and $Z_p^c(\mathbf{r})$ is the partition function in momentum space,

$$Z_p^c(\mathbf{r}) = \int d^{3N}p\,\delta(\mathbf{P}_\perp\mathbf{p})\exp\left(-\frac{\beta}{2}\mathbf{p}^T\mathbf{M}_c^+\mathbf{p}\right). \tag{4.3}$$

The product over the delta functions $\delta(h_i(\mathbf{r})-h_i^{(0)})$ accounts for the constraints in position space and the delta function $\delta(\mathbf{P}_\perp\mathbf{p})$ for the corresponding constraints in momentum space. The latter are again expressed in the form $\mathbf{P}_\perp\mathbf{p}=\mathbf{0}$. As for the calculation of momentum averages, the singular value decomposition of $\mathbf{M}_c$ yields a convenient basis to evaluate Z_p^c. Skipping here the details, which are presented in Appendix D, one finds

$$Z_p^c(\mathbf{r}) = \sqrt{\frac{2\pi}{\beta}}^f\sqrt{\det(\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp)}. \tag{4.4}$$

Here

$$\det(\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp)\equiv D_\parallel \tag{4.5}$$

is the *discriminant* of $\mathbf{M}$ in $\mathcal{V}_\parallel$ (see Appendix B). If no constraints are present, $\mathbf{P}_\parallel$ becomes the unit matrix, $\mathbf{P}_\perp$ vanishes, and one obtains from (4.4) $Z_p=(\sqrt{2\pi/\beta})^{3N}\sqrt{\det(\mathbf{M})}=\text{const}$.

The fact that the momentum partition function is proportional to the square root of the discriminant of a constrained (mass) matrix allows to use the theorems for discriminants which are listed and proven in Appendix B. If the discriminant is to be expressed in terms of the matrix $\mathbf{A}$, it follows from Theorems 3 and 7 that

$$D_\parallel = \det(\mathbf{M})D_\perp, \tag{4.6}$$

where $D_\perp$ is given by

$$D_\perp = \frac{\det(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)}{\det(\mathbf{A}\mathbf{A}^T)}. \tag{4.7}$$

Theorem 6 is useful if the constraints are expressed in terms in the form $\mathbf{v}=\mathbf{C}\mathbf{u}$. In this case one can write

$$D_\parallel = \frac{\det(\mathbf{C}^T\mathbf{M}\mathbf{C})}{\det(\mathbf{C}^T\mathbf{C})}. \tag{4.8}$$

Relations (4.7) and (4.8) show that the physical dimensions of $D_\perp$ and $D_\parallel$ are, respectively, mass^{-s} and mass^f, where s is the number of constraints and f is the number of degrees of freedom. Therefore, Z_p^c has the correct dimension of a partition function of f free linear momenta.

The results derived above can be used for a correction of phase space averages of constrained mechanical systems, which is similar but not identical to the well-known Fixman correction.[16] In contrast to unconstrained systems, where the integral over the momenta in phase space averages will contribute a constant factor, one obtains for constrained systems a position-dependent momentum partition function. The position dependence of the latter has the effect of an additional potential, and one can compensate for this effect by introducing an appropriate bias potential, such that $\exp(-\beta U_{\text{bias}})=\sqrt{\det(\mathbf{M})/D_\parallel}$, neglecting constant terms. If one works with $\mathbf{A}$ matrices, one derives from (4.6) and (4.7) that

$$U_{\text{bias}}(\mathbf{r}) = \frac{k_BT}{2}\ln\left(\frac{\det(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)}{\det(\mathbf{A}\mathbf{A}^T)}\right), \tag{4.9}$$

and in terms of $\mathbf{C}$ matrices one obtains the alternative form

$$U_{\text{bias}}(\mathbf{r}) = \frac{k_BT}{2}\ln\left(\frac{\det(\mathbf{C}^T\mathbf{M}\mathbf{C})}{\det(\mathbf{C}^T\mathbf{C})}\right) \tag{4.10}$$

for the bias potential. It should be noted that the latter vanishes in mass-weighted coordinates, where $\mathbf{M}=\mathbf{1}$. In this case the discriminant (4.5) appearing in momentum partition function (4.4) is, in fact, 1, $\det(\mathbf{P}_\parallel+\mathbf{P}_\perp)=\det(\mathbf{1})=1$.

B. Fixman correction revisited

It is illustrative to compare the correction of phase space averages discussed above to the classical Fixman correction. For this purpose we introduce $3N$ generalized coordinates x_i which are partitioned into f "soft" variables, $\mathbf{q}=(q_1,\ldots,q_f)^T$, and s "hard" variables, $\mathbf{h}=(h_1,\ldots,h_s)^T$, such that that $\mathbf{r}=\mathbf{r}(\mathbf{q},\mathbf{h})$. The associated generalized momenta are denoted by $\boldsymbol{\pi}_q$ and $\boldsymbol{\pi}_h$, respectively. We suppose that all generalized hard and soft coordinates are independent and that the relation $\mathbf{r}=\mathbf{r}(\mathbf{q},\mathbf{h})$ can be inverted. Using the definition of the full Jacobian matrix $\mathbf{J}$, whose elements are given by $J_{ik}=\partial r_i/\partial x_k$, we thus require that $\mathbf{J}^{-1}$ exists. With the Jacobian we define the mass metric tensor of the unconstrained system through $\mathbf{G}=\mathbf{J}^T\mathbf{M}\mathbf{J}$. Using the partitioning into soft and hard variables, the latter may be written in the form

$$\mathbf{G} = \begin{pmatrix}\mathbf{G}_{qq} & \mathbf{G}_{qh}\\ \mathbf{G}_{hq} & \mathbf{G}_{hh}\end{pmatrix}. \tag{4.11}$$

To obtain the Fixman correction for constrained configurational averages of a variable $F(\mathbf{q})$, we write

$$\langle F\rangle = \frac{1}{Z^c}\int d^fqZ_p^c(\mathbf{q})F(\mathbf{q})\exp\left(-\frac{\beta}{2}U(\mathbf{q})\right), \tag{4.12}$$

where Z^c is the partition function of the constrained system,

$$Z^c = \int d^f q Z_p^c(\mathbf{q}) \exp\left(-\frac{\beta}{2} U(\mathbf{q})\right), \tag{4.13}$$

and $Z_p^c(\mathbf{q})$ is the partition function in the space of *generalized* momenta,

$$Z_p^c(\mathbf{q}) = \int d^f \pi_q \exp\left(-\frac{\beta}{2} \boldsymbol{\pi}^T \mathbf{G}_{qq}^{-1} \boldsymbol{\pi}\right) = \left(\sqrt{\frac{2\pi}{\beta}}\right)^f \sqrt{\det(\mathbf{G}_{qq})}. \tag{4.14}$$

Introducing the projectors

$$\mathbf{P}_q = \begin{pmatrix} \mathbf{1}_f & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \mathbf{P}_h = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix}, \tag{4.15}$$

where $\mathbf{1}_f$ and $\mathbf{1}_s$ are unit matrices of dimension f and s, respectively, the momentum partition function can be written in the same form as for Cartesian coordinates,

$$Z_p^c(\mathbf{q}) = \left(\sqrt{\frac{2\pi}{\beta}}\right)^f \sqrt{\det(\mathbf{G}\mathbf{P}_q + \mathbf{P}_h)}, \tag{4.16}$$

replacing $\mathbf{M}$ by $\mathbf{G}$ and the projectors $\mathbf{P}_\parallel$ and $\mathbf{P}_\perp$ by the *diagonal* projectors $\mathbf{P}_q$ and $\mathbf{P}_h$, respectively. Formally Z_p^c is thus again expressed by a discriminant, and one may use that $\det(\mathbf{G}\mathbf{P}_q + \mathbf{P}_h) = \det(\mathbf{G})\det(\mathbf{G}^{-1}\mathbf{P}_h + \mathbf{P}_q)$. Due to the simple form of the projectors $\mathbf{P}_q$ and $\mathbf{P}_h$, it follows that

$$\mathbf{G}\mathbf{P}_q + \mathbf{P}_h = \begin{pmatrix} \mathbf{G}_{qq} & \mathbf{0} \\ \mathbf{G}_{hq} & \mathbf{1}_s \end{pmatrix} \tag{4.17}$$

and

$$\mathbf{G}^{-1}\mathbf{P}_h + \mathbf{P}_q = \begin{pmatrix} \mathbf{1}_f & \mathbf{G}_{qh}^{-1} \\ \mathbf{0} & \mathbf{G}_{hh}^{-1} \end{pmatrix}. \tag{4.18}$$

It is now easy to see that $\det(\mathbf{G}^{-1}\mathbf{P}_h + \mathbf{P}_q) = \det(\mathbf{G}_{hh}^{-1})$, where

$$(\mathbf{G}_{hh}^{-1})_{ij} = \left(M_{kl}^{-1} \frac{\partial h_i}{\partial r_k} \frac{\partial h_j}{\partial r_l}\right) = (\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)_{ij}, \tag{4.19}$$

such that

$$\det(\mathbf{G}_{qq}) = \det(\mathbf{G})\det(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T). \tag{4.20}$$

This is precisely Fixman's theorem[16] and we construct here the bias potential U_F according to $\exp(-\beta U_F) = \sqrt{\det(\mathbf{G})/\det(\mathbf{G}_{qq})}$,

$$U_F(\mathbf{r}) = \frac{k_B T}{2} \ln(\det(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)), \tag{4.21}$$

which is *not* identical to the one given by Eq. (4.9). The reason is that the Fixman potential corrects for configurational bias induced by constrained generalized momenta, whereas (4.9) corrects for configurational bias induced by constrained Cartesian momenta. Although the total partition function is invariant with respect to any coordinate transformation, the same is not true for partial integrals, such as the momentum partition function. The latter depends clearly on the choice of the generalized coordinates—in the unconstrained as well as in the constrained case. Therefore Fixman-type corrections depend of the choice of coordinates as well.

In this context it is worthwhile to consider again a subject of debate since the early papers of Kirkwood,[17] Erpenbeck and Kirkwood,[18] and Kramers,[19] on polymer solutions, mentioned by Fixman,[16] and which concerns the way constraints are treated in generalized coordinates and momenta. One may either set the constrained velocities to zero,

$$\dot{\mathbf{h}} = \mathbf{0}, \tag{4.22}$$

as we did in this paper, or require that the corresponding associated *momenta* vanish,

$$\boldsymbol{\pi}_h = \mathbf{0}, \tag{4.23}$$

as, for example, considered by Fixman. Applying condition (4.22) one finds from the relation between the generalized momenta and velocities,

$$\begin{pmatrix} \boldsymbol{\pi}_q \\ \boldsymbol{\pi}_h \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{qq} & \mathbf{G}_{qh} \\ \mathbf{G}_{hq} & \mathbf{G}_{hh} \end{pmatrix} \begin{pmatrix} \dot{\mathbf{q}} \\ \mathbf{0} \end{pmatrix}, \tag{4.24}$$

that the momenta associated with the constrained variables, $\boldsymbol{\pi}_h$, can be expressed through those of the free variables,

$$\boldsymbol{\pi}_h = \mathbf{G}_{hq}\mathbf{G}_{qq}^{-1}\boldsymbol{\pi}_q. \tag{4.25}$$

Here condition (4.23) is thus not fulfilled, but this does nevertheless not influence the momentum partition function. Expressing the latter in *all* momenta, we have

$$Z_p^c = \int d^f \pi_q \int d^s \pi_h \delta(\boldsymbol{\pi}_h - \mathbf{G}_{hq}\mathbf{G}_{qq}^{-1}\boldsymbol{\pi}_q) \times \exp\left(-\frac{\beta}{2} \boldsymbol{\pi}_q^T \mathbf{G}_{qq}^{-1} \boldsymbol{\pi}_q\right), \tag{4.26}$$

where the integration over the momenta $\boldsymbol{\pi}_h$ drops out from the calculation since the Boltzmann factor does depend only on the free momenta $\boldsymbol{\pi}_q$ and since

$$\int d^s \pi_h \delta(\boldsymbol{\pi}_h - \mathbf{G}_{hq}\mathbf{G}_{qq}^{-1}\boldsymbol{\pi}_q) = 1.$$

Therefore one is again left with the momentum partition function used by Fixman given by Eq. (4.14), and one can exclude that the differences in the bias potentials found in the paper and by Fixman are due to a different treatment of the constraints in generalized coordinates.

V. CONCLUSION AND DISCUSSION

It has been shown that projectors and pseudoinverse matrices, in particular, the Bott-Duffin inverse, are useful tools for a Hamiltonian description of constrained dynamical systems in Cartesian coordinates. The key object in this context is the constrained inverse of the mass matrix with respect to the velocity space defined by the imposed geometrical constraints. It allows first of all to construct the constrained Hamiltonian in Cartesian coordinates and to relate it to the well-known representation in generalized coordinates. The constrained inverse mass matrix allows, in particular, to define effective atomic masses, as they appear in thermal velocity averages and in the Sachs-Teller problem. The influ-

ence of constraints on the effective atomic masses has been demonstrated with the simple example of rigid and semiflexible water, where the H–H distance is flexible and only the O–H bonds are fixed. The Sachs-Teller tensors and the corresponding masses are retrieved in case of rigid molecules. The discriminants related to the projected mass matrix and its inverse were shown to be particularly useful for the calculation of momentum partition functions of constrained systems in Cartesian coordinates. In order to remove the bias from sampling of constrained configurational averages, explicit Fixman-type potentials have been derived for different forms of the constraints. It has been highlighted that the classical Fixman correction removes the configurational bias induced by constrained generalized momenta, whereas the corrections presented in this paper remove configurational bias due to constrained Cartesian momenta. In this context it is important to note that *no* configurational bias is generated if mass-weighted coordinates are used.

APPENDIX A: PSEUDOINVERSE MATRICES

(1) *Definition*. Let $\mathbf{A}$ be an arbitrary $m\times n$ matrix. A $n\times m$ matrix $\mathbf{X}$ is called the pseudoinverse of $\mathbf{A}$ if it fulfills the relations

$$\mathbf{AXA}=\mathbf{A}, \tag{A1}$$

$$\mathbf{XAX}=\mathbf{X}, \tag{A2}$$

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

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

Relations (A1)–(A4) are the *Moore-Penrose conditions*[9,10,12,13] which define the generalized inverse of a matrix uniquely. The pseudoinverse is also called Moore-Penrose inverse or generalized inverse and is usually denoted by

$$\mathbf{A}^+\equiv\mathbf{X}. \tag{A5}$$

It is easy to see that $\mathbf{P}_r=\mathbf{A}^+\mathbf{A}$ is the projector on the row space of $\mathbf{A}$ and the projector on the column space of $\mathbf{A}^+$, whereas $\mathbf{P}_c=\mathbf{AA}^+$ projects on the column space of $\mathbf{A}$ and on the row space of $\mathbf{A}^+$. On account of the Moore-Penrose conditions, one has $\mathbf{AP}_r=\mathbf{A}$, $\mathbf{P}_r\mathbf{A}^+=\mathbf{A}^+$, $\mathbf{P}_c\mathbf{A}=\mathbf{A}$, and $\mathbf{A}^+\mathbf{P}_c=\mathbf{A}^+$.

(2) *Full-rank factorization*. Suppose that $\mathbf{A}$ is an $m\times n$ matrix which can be factorized as

$$\mathbf{A}=\mathbf{FG}, \tag{A6}$$

where $\mathbf{F}$ is an $m\times r$ matrix of full column rank r and $\mathbf{G}$ a $r\times n$ matrix with full row rank r. In this case the pseudoinverse of $\mathbf{B}$ can be given explicitly,[9]

$$\mathbf{A}^+=\mathbf{G}(\mathbf{GG}^T)^{-1}(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T. \tag{A7}$$

(3) *Singular value decomposition*. Another representation of pseudoinverse matrices is based on the singular value decomposition (SVD) of $\mathbf{A}$,

$$\mathbf{A}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T, \tag{A8}$$

where $\mathbf{U}$ is an orthogonal $m\times m$ matrix, $\mathbf{V}$ is an orthogonal $n\times n$ matrix, and $\boldsymbol{\Sigma}$ is the diagonal $m\times n$ matrix

$$\boldsymbol{\Sigma}=\left.\begin{pmatrix} \overbrace{\begin{matrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_f\end{matrix}}^{f} & \overbrace{\begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix}}^{n-f} \\ \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix}\end{pmatrix}\right\}m \tag{A9}$$

containing the singular values of $\mathbf{A}$, which fulfill $\sigma_i\geqslant 0$. Using SVD, the pseudoinverse of $\mathbf{A}$ can be expressed in the form

$$\mathbf{A}^+=\mathbf{V}\boldsymbol{\Sigma}^+\mathbf{U}^T, \tag{A10}$$

where $\boldsymbol{\Sigma}^+$ is the $n\times m$ matrix

$$\boldsymbol{\Sigma}^+=\left.\begin{pmatrix} \overbrace{\begin{matrix}\sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_f^{-1}\end{matrix}}^{f} & \overbrace{\begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix}}^{m-f} \\ \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix}\end{pmatrix}\right\}n, \tag{A11}$$

and the projectors on the row and column space of $\mathbf{A}$, respectively, can be cast into the form

$$\mathbf{P}_r=\mathbf{A}^+\mathbf{A}=\mathbf{V}\tilde{\mathbf{P}}_r\mathbf{V}^T, \tag{A12}$$

$$\mathbf{P}_c=\mathbf{AA}^+=\mathbf{U}\tilde{\mathbf{P}}_c\mathbf{U}^T. \tag{A13}$$

Here

$$\tilde{\mathbf{P}}_r=\boldsymbol{\Sigma}^+\boldsymbol{\Sigma}=\left.\begin{pmatrix}\mathbf{1}_f & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\right\}n \tag{A14}$$

and

$$\tilde{\mathbf{P}}_c=\boldsymbol{\Sigma}\boldsymbol{\Sigma}^+=\left.\begin{pmatrix}\mathbf{1}_f & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix}\right\}m \tag{A15}$$

are quadratic matrices of dimensions $n\times n$ and $m\times m$, respectively, and $\mathbf{1}_f$ is a unit matrix of dimension f.

APPENDIX B: BOTT-DUFFIN INVERSES

In the following a short and self-contained description of Bott-Duffin inverses and their relation to pseudoinverses are given. Most of the theorems listed in the following and their proofs can be found in the original article by Bott and Duffin[11] and in textbooks,[9,10] but sometimes without proof.

Bott-Duffin inverses are special forms of pseudoinverse matrices. In the original work by Bott and Duffin on the theory of electrical networks, one considers the following

problem. Given an arbitrary quadratic $n\times n$ matrix $\mathbf{G}$, one searches the solution of the system of linear equations

$$\mathbf{G}\mathbf{v}_{\parallel}+\mathbf{v}_{\perp}=\mathbf{h}, \tag{B1}$$

where $\mathbf{h}$ is an arbitrary column vector of length n and $\mathbf{v}_{\parallel}\in\mathcal{V}_{\parallel}$ and $\mathbf{v}_{\perp}\in\mathcal{V}_{\perp}$ are mutually orthogonal column vectors. The projectors on $\mathcal{V}_{\parallel}$ and $\mathcal{V}_{\perp}$ are represented by the $n\times n$ matrices $\mathbf{P}_{\parallel}$ and $\mathbf{P}_{\perp}$, respectively. Introducing the column vector $\mathbf{y}$, such that $\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{y}$ and $\mathbf{v}_{\perp}=\mathbf{P}_{\perp}\mathbf{y}$, Eq. (B1) may be written in the form $(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})\mathbf{y}=\mathbf{h}$. Under the condition that

$$D_{\parallel}\equiv\det(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})\neq 0, \tag{B2}$$

where $D_{\parallel}$ is called the *discriminant* of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$, the solutions for $\mathbf{v}_{\parallel}$ and $\mathbf{v}_{\perp}$ thus read $\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})^{-1}\mathbf{h}$ and $\mathbf{v}_{\perp}=\mathbf{P}_{\perp}(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})^{-1}\mathbf{h}$. The matrix

$$\mathbf{T}_{\parallel}=\mathbf{P}_{\parallel}(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})^{-1} \tag{B3}$$

is called the *constrained inverse* of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$.

Theorem 1. The constrained inverse of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$ is given by $\mathbf{T}_{\parallel}=(\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel})^{+}$ and fulfills $\mathbf{P}_{\parallel}\mathbf{T}_{\parallel}=\mathbf{T}_{\parallel}\mathbf{P}_{\parallel}=\mathbf{T}_{\parallel}$.

Proof. Applying the projector $\mathbf{P}_{\parallel}$ from the left to the system of equations (B1) and using $\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{v}_{\parallel}$ yield $\mathbf{G}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{h}$, where $\mathbf{G}_{\parallel}:=\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}$. With $\mathbf{P}_r=\mathbf{G}_{\parallel}^{+}\mathbf{G}_{\parallel}$ and $\mathbf{P}_c=\mathbf{G}_{\parallel}\mathbf{G}_{\parallel}^{+}$, one thus has from the Moore-Penrose conditions $\mathbf{P}_c\mathbf{G}_{\parallel}=\mathbf{P}_c\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}$ and $\mathbf{G}_{\parallel}\mathbf{P}_r=\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}\mathbf{P}_r=\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}$. Since the above relations must be fulfilled for any $\mathbf{G}$, in particular for $\mathbf{G}=1$, it thus follows that $\mathbf{P}_c=\mathbf{P}_r=\mathbf{P}_{\parallel}$, and from the Moore-Penrose conditions one obtains $\mathbf{P}_{\parallel}\mathbf{G}_{\parallel}^{+}=\mathbf{G}_{\parallel}^{+}\mathbf{P}_{\parallel}$.

Since $\mathbf{G}_{\parallel}\mathbf{G}_{\parallel}^{+}\mathbf{G}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{G}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{G}_{\parallel}\mathbf{G}_{\parallel}^{+}\mathbf{P}_{\parallel}\mathbf{h}=\mathbf{P}_{\parallel}\mathbf{h}$, the system of equations $\mathbf{G}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{h}$ is consistent and the solution for $\mathbf{v}_{\parallel}$ is found by applying $\mathbf{G}_{\parallel}^{+}$ from the left. One obtains $\mathbf{G}_{\parallel}^{+}\mathbf{G}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{P}_{\parallel}\mathbf{v}_{\parallel}=\mathbf{v}_{\parallel}=\mathbf{G}_{\parallel}^{+}\mathbf{P}_{\parallel}\mathbf{h}$. It follows now from the Moore-Penrose conditions that $\mathbf{P}_{\parallel}\mathbf{G}_{\parallel}^{+}=\mathbf{G}_{\parallel}^{+}\mathbf{P}_{\parallel}=\mathbf{G}_{\parallel}^{+}$ and therefore $\mathbf{v}_{\parallel}=\mathbf{G}_{\parallel}^{+}\mathbf{h}$. Identifying $\mathbf{G}_{\parallel}^{+}\equiv\mathbf{T}_{\parallel}$ the theorem is thus proven.

Theorem 2. If $\mathbf{G}^{-1}$ exists, the solution of $\mathbf{G}\mathbf{v}_{\parallel}+\mathbf{v}_{\perp}=\mathbf{h}$ for $\mathbf{v}_{\perp}$ is given by $\mathbf{v}_{\perp}=\mathbf{T}_{\perp}\mathbf{G}^{-1}\mathbf{h}$, where $\mathbf{T}_{\perp}=\mathbf{P}_{\perp}(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})^{-1}$ and $\mathbf{T}_{\perp}$ fulfills $\mathbf{T}_{\perp}=(\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{P}_{\perp})^{+}$ and $\mathbf{P}_{\perp}\mathbf{T}_{\perp}=\mathbf{T}_{\perp}\mathbf{P}_{\perp}=\mathbf{T}_{\perp}$.

Proof. If $\mathbf{G}^{-1}$ exists it follows from $\mathbf{G}\mathbf{v}_{\parallel}+\mathbf{v}_{\perp}=\mathbf{h}$ that $\mathbf{G}^{-1}\mathbf{v}_{\perp}+\mathbf{v}_{\parallel}=\mathbf{G}^{-1}\mathbf{h}$. Therefore $\mathbf{v}_{\perp}=\mathbf{P}_{\perp}(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})^{-1}\mathbf{G}^{-1}\mathbf{h}\equiv\mathbf{T}_{\perp}\mathbf{G}^{-1}\mathbf{h}$. Comparing $\mathbf{T}_{\perp}=\mathbf{P}_{\perp}(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})^{-1}$ with $\mathbf{T}_{\parallel}=\mathbf{P}_{\parallel}(\mathbf{G}\mathbf{P}_{\perp}+\mathbf{P}_{\perp})^{-1}$ shows that in the two Bott-Duffin inverses $\mathbf{G}$ is interchanged with $\mathbf{G}^{-1}$ and $\mathbf{P}_{\parallel}$ with $\mathbf{P}_{\perp}$. Using the properties of $\mathbf{T}_{\parallel}$ one thus obtains $\mathbf{T}_{\perp}=(\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{P}_{\perp})^{+}$ and $\mathbf{P}_{\perp}\mathbf{T}_{\perp}=\mathbf{T}_{\perp}\mathbf{P}_{\perp}=\mathbf{T}_{\perp}$.

Theorem 3. The discriminant of $\mathbf{T}_{\perp}$, given by $D_{\perp}\equiv\det(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})$, and the discriminant of $\mathbf{T}_{\parallel}$ are related through $D_{\parallel}=D_{\perp}\det(\mathbf{G})$.

Proof. Write $D_{\parallel}=\det(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})=\det(\mathbf{G}[\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel}])=\det(\mathbf{G})\det(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})=\det(\mathbf{G})D_{\perp}$.

Theorem 4. If the columns of $\mathbf{C}$ are a basis of $\mathcal{V}_{\parallel}$ and if $\det(\mathbf{C}^T\mathbf{G}\mathbf{C})\neq 0$, the constrained inverse of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$ can be written in the form $\mathbf{T}_{\parallel}=\mathbf{C}(\mathbf{C}^T\mathbf{G}\mathbf{C})^{-1}\mathbf{C}^T$.

Proof. By definition $\mathbf{C}$ is a matrix whose column vectors span $\mathcal{V}_{\parallel}$. The columns in $\mathbf{C}$ are thus linearly independent, such that the projector on $\mathcal{V}_{\parallel}$ can be expressed as $\mathbf{P}_{\parallel}=\mathbf{C}\mathbf{C}^{+}=\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T$ and $\mathbf{v}=\mathbf{C}\mathbf{u}$, where $\mathbf{u}$ is a column vector of length f and f is the dimension of $\mathcal{V}_{\parallel}$. The relation $\mathbf{G}_{\parallel}\mathbf{v}=\mathbf{P}_{\parallel}\mathbf{G}\mathbf{P}_{\parallel}\mathbf{v}=\mathbf{P}_{\parallel}\mathbf{h}$ takes the form $\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{G}\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{C}\mathbf{u}=\mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{h}$. Multiplication from the left by $\mathbf{C}^T$ thus yields $\mathbf{C}^T\mathbf{G}\mathbf{C}\mathbf{u}=\mathbf{C}^T\mathbf{h}$. Solving for $\mathbf{u}$ and inserting the result into the relation $\mathbf{v}=\mathbf{C}\mathbf{u}$ yield $\mathbf{v}=\mathbf{C}(\mathbf{C}^T\mathbf{G}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{h}$, from which one can conclude $\mathbf{T}_{\parallel}=\mathbf{C}(\mathbf{C}^T\mathbf{G}\mathbf{C})^{-1}\mathbf{C}^T$.

Theorem 5. If the rows of $\mathbf{A}$ are a basis of $\mathcal{V}_{\perp}$ and if $\det(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)\neq 0$, the constrained inverse of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$ can be written in the form $\mathbf{T}_{\parallel}=\mathbf{G}_{\parallel}^{+}=\mathbf{G}^{-1}-\mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}$.

Proof. One starts with the relation $\mathbf{G}\mathbf{v}_{\parallel}+\mathbf{v}_{\perp}=\mathbf{h}$, which is equivalent to $\mathbf{G}^{-1}\mathbf{v}_{\perp}+\mathbf{v}_{\parallel}=\mathbf{G}^{-1}\mathbf{h}$ if $\mathbf{G}^{-1}$ exists. A subsequent projection with $\mathbf{P}_{\perp}$ eliminates $\mathbf{v}\in\mathcal{V}_{\parallel}$ and yields $\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{v}_{\perp}=\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{h}$. Since $\mathbf{P}_{\perp}\mathbf{v}_{\perp}=\mathbf{v}_{\perp}$ one may also write $\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{P}_{\perp}\mathbf{v}_{\perp}=\mathbf{P}_{\perp}\mathbf{G}^{-1}\mathbf{h}$. Here one may now insert $\mathbf{P}_{\perp}=\mathbf{A}^{+}\mathbf{A}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}$; since the rows of $\mathbf{A}$ are by definition linearly independent, this yields $\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{v}_{\perp}=\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}\mathbf{h}$. Multiplication from the left with $\mathbf{A}$ and using that $\mathbf{v}_{\perp}$ can be represented as a linear combination of the rows of $\mathbf{A}$, $\mathbf{v}_{\perp}=\mathbf{A}^T\boldsymbol{\lambda}$, leads to $\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T\boldsymbol{\lambda}=\mathbf{A}\mathbf{G}^{-1}\mathbf{h}$. The solution for $\boldsymbol{\lambda}$ may then be used to write $\mathbf{v}_{\perp}=\mathbf{A}^T(\mathbf{A}\mathbf{M}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}\mathbf{h}$. Since $\mathbf{v}_{\parallel}=\mathbf{G}^{-1}(\mathbf{h}-\mathbf{v}_{\perp})$ one-obtains finally $\mathbf{v}_{\parallel}=\{\mathbf{G}^{-1}-\mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}\}\mathbf{h}$. Consequently $\mathbf{T}_{\parallel}=\mathbf{G}^{-1}-\mathbf{G}^{-1}\mathbf{A}^T(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{G}^{-1}$.

Theorem 6. Let $\mathbf{C}$ be a matrix whose columns span $\mathcal{V}_{\parallel}$. The discriminant of $\mathbf{G}$ in $\mathcal{V}_{\parallel}$ can be written in the form $D_{\parallel}=\det(\mathbf{C}^T\mathbf{G}\mathbf{C})/\det(\mathbf{C}^T\mathbf{C})$.[11]

Proof. Consider the matrix $\mathbf{U}=(\mathbf{C}|\mathbf{A}^T)$ whose first $n<m$ columns form a basis of $\mathcal{V}_{\parallel}$ and whose last $m-n$ columns form a basis of $\mathcal{V}_{\perp}$. The respective bases are chosen to be columns of the matrix $\mathbf{C}$ and the rows of the matrix $\mathbf{A}$, which have both been introduced earlier. With $\mathbf{U}$ we construct the matrix

$$\mathbf{Z}=\mathbf{U}^T(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})\mathbf{U}=\begin{pmatrix}\mathbf{C}^T\\ \mathbf{A}\end{pmatrix}(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})(\mathbf{C}|\mathbf{A}^T)$$

$$=\begin{pmatrix}\mathbf{C}^T\\ \mathbf{A}\end{pmatrix}(\mathbf{G}\mathbf{C}|\mathbf{A}^T)=\begin{pmatrix}\mathbf{C}^T\mathbf{G}\mathbf{C} & \mathbf{0}\\ \mathbf{A}\mathbf{G}\mathbf{C} & \mathbf{A}\mathbf{A}^T\end{pmatrix}.$$

Here $\mathbf{P}_{\parallel}\mathbf{C}=\mathbf{C}$, $\mathbf{P}_{\perp}\mathbf{C}=\mathbf{0}$, and $\mathbf{A}\mathbf{C}=\mathbf{0}$ have been used. From the above form of $\mathbf{Z}$ one sees that $\det(\mathbf{Z})=\det(\mathbf{C}^T\mathbf{G}\mathbf{C})\det(\mathbf{A}\mathbf{A}^T)$. On the other hand $\det(\mathbf{Z})=\det(\mathbf{G}\mathbf{P}_{\parallel}+\mathbf{P}_{\perp})\det(\mathbf{U}\mathbf{U}^T)=D_{\parallel}\det(\mathbf{U}\mathbf{U}^T)=D_{\parallel}\det(\mathbf{U}^T\mathbf{U})$, where

$$\mathbf{U}^T\mathbf{U}=\begin{pmatrix}\mathbf{C}^T\\ \mathbf{A}\end{pmatrix}(\mathbf{C}|\mathbf{A}^T)=\begin{pmatrix}\mathbf{C}^T\mathbf{C} & \mathbf{0}\\ \mathbf{0} & \mathbf{A}\mathbf{A}^T\end{pmatrix}.$$

Here $\mathbf{A}\mathbf{C}=\mathbf{0}$ has again been used. From the above block matrix form for $\mathbf{U}^T\mathbf{U}$, it follows that $\det(\mathbf{U}^T\mathbf{U})=\det(\mathbf{C}^T\mathbf{C})\det(\mathbf{A}\mathbf{A}^T)$, such that $D_{\parallel}=\det(\mathbf{C}^T\mathbf{G}\mathbf{C})/\det(\mathbf{C}^T\mathbf{C})$.

Theorem 7. Let $\mathbf{A}$ be a matrix whose rows span $\mathcal{V}_{\perp}$ and let $\mathbf{G}$ be a quadratic matrix whose inverse exists. It follows that $D_{\perp}=\det(\mathbf{G}^{-1}\mathbf{P}_{\perp}+\mathbf{P}_{\parallel})=\det(\mathbf{A}\mathbf{G}^{-1}\mathbf{A}^T)/\det(\mathbf{A}\mathbf{A}^T)$.

Proof. Using Theorem 6, the proof is trivial. Simply replace $\mathbf{G}\rightarrow\mathbf{G}^{-1}$ and exchange $\mathbf{C}\leftrightarrow\mathbf{A}^T$.

APPENDIX C: PROOF OF RELATION (3.4)

The moment generating function $G(\mathbf{k})$ defined trough (3.1) can be evaluated in a basis in which $\mathbf{M}_c$ is diagonal. For this purpose one uses the singular value decomposition $\mathbf{M}_c=\mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^T$, where $\boldsymbol{\Sigma}$ contains the singular values, $\boldsymbol{\Sigma}=\mathrm{diag}(\sigma_1,\ldots,\sigma_f,0,\ldots,0)$. Here $\sigma_k>0$ and $\mathbf{U}$ is a $3N\times 3N$

orthogonal matrix. With $\mathbf{M}_c^+=\mathbf{U}\mathbf{\Sigma}^+\mathbf{U}^T$ and $\mathbf{\Sigma}^+=\mathrm{diag}(\sigma_1^{-1},\ldots,\sigma_f^{-1},0,\ldots,0)$, it follows that $\mathbf{P}_\parallel=\mathbf{M}_c^+\mathbf{M}_c=\mathbf{M}_c\mathbf{M}_c^+=\mathbf{U}\mathbf{\Sigma}^+\mathbf{\Sigma}\mathbf{U}^T$. Since

$$\mathbf{\Sigma}^+\mathbf{\Sigma}=\mathrm{diag}(\underbrace{1,\ldots,1}_{f\text{ times}},0,\ldots,0),$$

the projector can thus be written as

$$\mathbf{P}_\parallel=\sum_{k=1}^{f}\mathbf{u}_k\mathbf{u}_k^T.$$

The above expression shows that $\mathcal{V}_\parallel$ is spanned by the first f columns of $\mathbf{U}$. Correspondingly, one has

$$\mathbf{P}_\perp=\sum_{k=f+1}^{3N}\mathbf{u}_k\mathbf{u}_k^T.$$

Introducing now new momenta via $\tilde{\mathbf{p}}=\mathbf{U}^T\mathbf{p}$, it follows that $\tilde{p}_k=\mathbf{u}_k^T\mathbf{p}$, and consequently

$$\delta(\mathbf{P}_\perp\mathbf{p})=\prod_{k=f+1}^{3N}\delta(\mathbf{u}_k^T\mathbf{p})=\prod_{k=f+1}^{3N}\delta(\tilde{p}_k).$$

Since the Jacobian determinant is 1, $J=|(\partial\tilde{\mathbf{p}})/(\partial\mathbf{p})|=|\mathbf{U}|=1$, it follows that

$$G(\mathbf{k})=\int d\tilde{p}_1\cdots d\tilde{p}_f\tilde{w}(\tilde{\mathbf{p}})\exp(i[\mathbf{U}^T\mathbf{k}]^T\tilde{\mathbf{p}}),$$

where

$$\tilde{w}_c(\tilde{\mathbf{p}})=\frac{\exp(-(\beta/2)\tilde{\mathbf{p}}^T\mathbf{\Sigma}^+\tilde{\mathbf{p}})}{\int d\tilde{p}_1\cdots d\tilde{p}_f\exp(-(\beta/2)\tilde{\mathbf{p}}^T\mathbf{\Sigma}^+\tilde{\mathbf{p}})}.$$

Defining $\tilde{\mathbf{k}}=\mathbf{U}^T\mathbf{k}$, the Gaussian integrals over $\tilde{w}_c(\tilde{\mathbf{p}})$ can be evaluated straightforwardly, and one obtains

$$G(\tilde{\mathbf{k}})=\exp\left(-\frac{1}{2\beta}[\sigma_1\tilde{k}_1^2+\cdots+\sigma_f\tilde{k}_f^2]\right).$$

Therefore

$$\langle(\mathbf{p}\mathbf{p}^T)_{ij}\rangle=-U_{im}U_{jn}\left.\frac{\partial^2G(\tilde{\mathbf{k}})}{\partial\tilde{k}_m\partial\tilde{k}_n}\right|_{\mathbf{k}=0},$$

where summation over pairwiselike indices is implied. Using the simple form of $G(\tilde{\mathbf{k}})$, we find

$$\left.\frac{\partial^2G(\tilde{\mathbf{k}})}{\partial\tilde{k}_i\partial\tilde{k}_j}\right|_{\mathbf{k}=0}=-\beta^{-1}(\mathbf{\Sigma})_{ij},$$

and with $\mathbf{M}_c=\mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$ it follows that

$$\langle\mathbf{p}\mathbf{p}^T\rangle=k_BT\mathbf{M}_c,$$

replacing β^{-1} by k_BT.

APPENDIX D: PROOF OF RELATION (4.4)

To evaluate the momentum partition function Z_p^c (the position dependence is omitted), one proceeds as for the calculation of thermal averages for the velocities, working in particular, in a basis in which $\mathbf{M}_c$ is diagonal, $\mathbf{M}=\mathbf{U}\mathbf{\Sigma}\mathbf{U}^T$. Using the new momenta $\tilde{\mathbf{p}}=\mathbf{U}^T\mathbf{p}$ one finds

$$Z_p^c=\int d\tilde{p}_1\cdots d\tilde{p}_f\exp\left(-\frac{\beta}{2}\tilde{\mathbf{p}}^T\mathbf{\Sigma}^+\tilde{\mathbf{p}}\right)=\left(\sqrt{\frac{2\pi}{\beta}}\right)^f\sqrt{\sigma_1\cdots\sigma_f}.$$

To prove relation (4.4) one writes

$$\begin{aligned}\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp&=\mathbf{P}_\parallel\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp\\&=\mathbf{U}\mathbf{\Sigma}\mathbf{U}^T+\mathbf{U}\tilde{\mathbf{P}}_\perp\underbrace{\mathbf{U}^T\mathbf{M}\mathbf{U}}_{\tilde{\mathbf{M}}}\tilde{\mathbf{P}}_\parallel\mathbf{U}^T+\mathbf{U}\tilde{\mathbf{P}}_\perp\mathbf{U}^T\\&=\mathbf{U}(\mathbf{\Sigma}+\tilde{\mathbf{P}}_\perp\tilde{\mathbf{M}}\tilde{\mathbf{P}}_\parallel+\tilde{\mathbf{P}}_\perp)\mathbf{U}^T.\end{aligned}$$

Using that $\mathbf{U}^T\mathbf{U}=\mathbf{1}$ one thus obtains $\det(\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp)=\det(\mathbf{\Sigma}+\tilde{\mathbf{P}}_\perp\tilde{\mathbf{M}}\tilde{\mathbf{P}}_\parallel+\tilde{\mathbf{P}}_\perp)$, where

$$\begin{aligned}\tilde{\mathbf{P}}_\parallel&=\mathrm{diag}(\underbrace{1,\ldots,1}_{f},\underbrace{0,\ldots,0}_{s})\text{ and }\tilde{\mathbf{P}}_\perp\\&=\mathrm{diag}(\underbrace{0,\ldots,0}_{f},\underbrace{1,\ldots,1}_{s}),\end{aligned}$$

such that (the indices indicate the dimensions of the block matrices)

$$\mathbf{\Sigma}+\tilde{\mathbf{P}}_\perp\tilde{\mathbf{M}}\tilde{\mathbf{P}}_\parallel+\tilde{\mathbf{P}}_\perp=\left(\begin{array}{ccc|c}\sigma_1&&&\\&\ddots&&\mathbf{0}_{fs}\\&&\sigma_f&\\\hline&(\tilde{\mathbf{P}}_\perp\tilde{\mathbf{M}}\tilde{\mathbf{P}}_\parallel)_{sf}&&\mathbf{1}_{ss}\end{array}\right).$$

Consequently $\det(\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp)=\det(\mathbf{\Sigma}+\tilde{\mathbf{P}}_\perp\tilde{\mathbf{M}}\tilde{\mathbf{P}}_\parallel+\tilde{\mathbf{P}}_\perp)=\sigma_1\ldots\sigma_f$, which thus proves that

$$Z_p^c=\left(\sqrt{\frac{2\pi}{\beta}}\right)^f\sqrt{\det(\mathbf{M}\mathbf{P}_\parallel+\mathbf{P}_\perp)}.$$

[1] W. G. Hoover, *Molecular Dynamics*, Lecture Notes in Physics Vol. 258 (Springer, Berlin, 1986).
[2] W. G. Hoover, *Time Reversibility, Computer Simulation, and Chaos*, Advanced Series in Nonlinear Dynamics Vol. 13 (World Scientific, Singapore, 1999).
[3] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1971).
[4] J. L. Anderson and P. G. Bergmann, Phys. Rev. **83**, 1018 (1951).
[5] P. A. M. Dirac, Can. J. Math. **2**, 129 (1950).
[6] P. A. M. Dirac, Proc. R. Soc. London, Ser. A **246**, 326 (1958).
[7] S. W. de Leeuw, J. W. Perram, and H. G. Petersen, J. Stat. Phys. **61**, 1203 (1990).
[8] See also Ref. 7, Eq. (7.17).
[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] E. H. Moore, Bull. Am. Math. Soc. **26**, 394 (1920).
[13] R. Penrose, Proc. Cambridge Philos. Soc. **51**, 406 (1955).
[14] G. R. Kneller and K. Hinsen, Phys. Rev. E **50**, 1559 (1994).
[15] R. G. Sachs and E. Teller, Phys. Rev. **60**, 18 (1941).
[16] M. Fixman, Proc. Natl. Acad. Sci. U.S.A. **8**, 3050 (1974).
[17] J. G. Kirkwood, J. Chem. Phys. **14**, 180 (1946).
[18] J. Erpenbeck and J. G. Kirkwood, J. Chem. Phys. **29**, 909 (1958).
[19] H. A. Kramers, J. Chem. Phys. **14**, 415 (1946).