2012

# Exact relation between potential of mean force and free-energy profile

Kin Yiu Wong
*Hong Kong Baptist University*, wongky@hkbu.edu.hk

Darrin M. York
*The State University of New Jersey*, york@biomaps.rutgers.edu

# Exact Relation between Potential of Mean Force and Free-Energy Profile

*Kin-Yiu Wong*[†‡] *and Darrin M. York*[‡]

[†] Department of Physics
Hong Kong Baptist University
224 Waterloo Road
Kowloon Tong
Hong Kong

[‡] BioMaPS Institute for Quantitative Biology
Department of Chemistry & Chemical Biology
Rutgers, The State University of New Jersey
174 Frelinghuysen Road
Piscataway, NJ 08854
USA

K.-Y. Wong:
Primary Email: wongky@hkbu.edu.hk
Secondary Email: kiniu@alumni.cuhk.net

D. M. York:
Email: york@biomaps.rutgers.edu

Keywords

many-body simulations, curvilinear coordinate, potential of mean force, free-energy profile, Jacobian determinant, differential geometry, covariant and contravariant vector space

Abstract

We apply concepts of covariant and contravariant vector space in differential geometry and general relativity to derive new, general, exact relations between potential of mean force and free-energy profile. These relations are immensely practical in free-energy simulations because a full Jacobian transformation (which is usually unknown) is not required; rather, only knowledge of the (constraint) coordinate of interest is needed. We reveal that in addition to the Jacobian determinant, the Jacobian scale factor and Leibnizian contributions must also be considered, as well a Fixman term with correct mass dependence. Our newly derived relations are verified with new non-trivial benchmark numerical examples for which exact results can be computed, and compared with relations available in the literature that turn out to exhibit significant deviations from the exact values.

Free-energy profiles derived from molecular simulations are widely used in computational physics, biophysics, and chemistry[1-3] to provide valuable insight into biochemical or physical events ranging from folding and conformational changes in proteins [4, 5], lipid-driven aggregation of nanoparticles [6] to chemical reactions [7, 8] that occur on material surfaces [9], in aqueous solution [10], and in the catalytic active sites of RNA [11] or protein [12, 13] enzymes. Two of the most commonly applied approaches for simulating free-energy profiles involve methods based on reweighting biased probability distributions (e.g., umbrella sampling) [14], and constrained mean-force (CMF) samplings (e.g., blue-moon sampling) [15, 16]. The former delivers the free-energy profile directly from the probability distribution along the coordinate of interest, whereas the latter relies on relation between free-energy profile and potential of mean force (PMF), first introduced by Kirkwood over three quarters of century ago [17], which subsequently has become a central underpinning in free-energy simulations (e.g., the two terms, "PMF" and "free-energy profile", are sometimes written as synonyms in the literature).

Advantages of constrained mean-force (CMF) sampling approaches are that they *neither* have inherent assumptions imposed by reweighting algorithms, *nor* binning/overlapping histograms, and also do *not* require biasing forces/potentials [14-16]. Furthermore, it is trivial for the CMF simulations to have *identical* number of samples (sample uniformly) at *any* value of the coordinate of interest (e.g., the number of 'rare' samples at the transition state can easily be as many as the number of 'abundant' samples at the ground state). However, in order to *exactly* equate PMF with free-energy profile, we must consider the Jacobian contribution. Explicitly expressing the Jacobian contribution for *any* curvilinear coordinate in analytical form is *not* straightforward, and in practice, applications are (thus) often limited to use of fairly simple coordinates to describe the events of interest [5, 15, 18].

Substantial theoretical efforts have been made to generalize the original Kirkwood relation, and have led to different, sometimes inconsistent relations between PMF and free-energy profile involving discrepant relation with the integrated mean Lagrange multiplier of the constraint, and incorrect mass-dependence of the Fixman term [15-29]. Further, there is a lack of non-trivial numerical examples, in which *exact* values can be computed, and thus serve as benchmarks to rigorously test different formulations.

In this paper, we first (I) reveal a new free-energy-profile term contributed from Leibniz's rule [30], and then apply the equivalence between the orthogonal covariant and contravariant vector space from differential geometry and general relativity to (II) formulate general equations for the Jacobian contribution to the free-energy profile that requires knowledge only of the (constraint) coordinate of interest (full Jacobian transformation is not needed), (III) present the Fixman term with correct mass dependence, and (IV) disclose a Jacobian scale factor term that is required to exactly relate the integrated mean Lagrange multiplier to the free-energy profile. All new results are compared with those calculated from other relations reported in the literature[15-29] using new rigorous numerical examples that serve as validation benchmarks.

*I: Leibnizian contribution (full Jacobian case).* Suppose $q_\xi$ is the generalized coordinate of interest for describing a biochemical or physical process, and is a member of at least one *complete* set of *curvilinear* coordinates, say $\{q\}^N$. Then the (canonical) free-energy profile as a function of $q_\xi = \xi_0$ is:[1-3]

$$G_\xi(\xi_0) = \frac{-1}{\beta}\ln\int\left|\vec{J}\left(\xi_0,\{q\}^{N-1}\right)\right| dq^{N-1}\, e^{-\beta V\left(\xi_0,\{q\}^{N-1}\right)} + C \qquad (1)$$

where $N$ is number of degrees of freedom, $\beta = 1/k_B T$, $k_B$ is Boltzmann's constant, $T$ is absolute temperature, $\left|\vec{J}\right|$ is the determinant of the Jacobian transformation $\vec{J}$ for $\{q\}^N$, $dq^{N-1}$ is a set of

integration variables without $dq_\xi$, and $C$ is a normalization constant from integrating the entire momentum space. Note that in this paper, integrals without explicit limits imply the entire space is integrated. Generalization of Eq. (1) to the isothermal-isobaric ensemble is straightforward [2, 3, 31] [and for the cases in the alchemical free-energy simulations,[5] if the extended degree of freedom, say $q_\lambda$, is completely independent of the original configuration space, such that its Jacobian scale factor is always unity in the extended version of Eq. (1), then there is no Jacobian and Leibnizian contribution in the following Eq. (2)]. Practically, it is rare to determine the constant $C$ because what we often care about is the free-energy differences with various values of $\xi_0$. To exactly equate PMF with the free-energy profile in Eq. (1), we first differentiate it with respect to $\xi_0$:

$$\boxed{\begin{aligned}\frac{dG_\xi(\xi_0)}{d\xi_0} = &\left\langle \left( \frac{\partial V}{\partial q_\xi} \right)_{\{q_{m\neq\zeta}\}^{N-1}} \right\rangle_{\xi_0} - \frac{1}{\beta} \left\langle \frac{1}{\left|\vec{J}\right|} \left( \frac{\partial \left|\vec{J}\right|}{\partial q_\xi} \right)_{\{q_{m\neq\zeta}\}^{N-1}} \right\rangle_{\xi_0} \\ &- \frac{1}{\beta} \left\langle \sum_{m\neq\xi}^{N-1} \left[ \delta\left(q_m - l_{U_m}\right) \frac{dl_{U_m}\left(q_\xi\right)}{dq_\xi} - \delta\left(q_m - l_{L_m}\right) \frac{dl_{L_m}\left(q_\xi\right)}{dq_\xi} \right] \right\rangle_{\xi_0}\end{aligned}} \tag{2}$$

where $\langle \cdots \rangle_{\xi_0}$ is the ensemble average over all configurations with $q_\xi = \xi_0$ (Supp. Info.), $l_{U_m}$ and $l_{L_m}$ are the upper and lower integration limits for $q_m$, respectively. Eq. (2) establishes an important exact relation between free-energy profile and ensemble averages (or mean values) of some physical quantities.

The first term in Eq. (2) is the negative value of the mean force. Integrating the first term over the coordinate of interest provides us with the potential of mean force (PMF). The second term is the contribution due to the Jacobian determinant. These two ensemble-average terms are

sometimes collectively referred to in the literature as the mean value of "generalized force", although we avoid this designation in the present work.

The third term in Eq. (2), which is *not* found in the literature, derives from the change of domains of $\{q_{m\neq\xi}\}^{N-1}$ with respect to $q_\xi$, i.e., from Leibniz's rule [30]. This term arises in the Jacobian transformation between two complete sets of coordinates that have coupling of their integration domains, and is herein referred to as the Leibnizian contribution. In Eq. (2), we restrict ourselves to the cases that if there are domains of $q_{m\neq\xi}$ depending on $q_\xi$, then all are functions of $q_\xi$ *only*, independent of $q_{j\neq m,\xi}$. For all other domains of $q_{j\neq m,\xi}$, even depending on $q_{m\neq\xi}$, as long as they are independent of $q_\xi$, then we treat these domains as constants. A general expression without such restrictions can be found in Ref. [30].

To demonstrate the importance of the Leibnizian contribution in Eq. (2), we consider a three-atom system, initially involving nine degrees of freedom: $\{x_A, y_A, z_A, x_B, y_B, z_B, x_C, y_C, z_C\}$. The motions of the two atoms: *A* and *C* are restricted such that they are always located at the *z*-axis with the values of $R/2$ and $-R/2$, respectively, although their separation, *R*, is a variable. So the degrees of freedom is down to four: $\{R, x_B, y_B, z_B\}$. We would like to express these four degrees of freedom in terms of a set of four elliptic coordinates $\{R, \xi_s, \xi_a, \phi\}$, in which the coordinate of interest $\xi_a \in [-R, R]$ is $\xi_a \equiv \sqrt{x_B^2 + y_B^2 + (z_B - R/2)^2} - \sqrt{x_B^2 + y_B^2 + (z_B + R/2)^2}$. Note that $\xi_a$ is an effective coordinate to describe atom-transfer reaction [3, 10-13, 32] (from atom *A* to *C*), and the domains for $\{R, \xi_s, \xi_a\}$ are related. With the definition of the coordinate of interest $\xi_a$, and with a three-body double-well potential and full Jacobian transformation (details

for the coordinate transformation and the potential are in Supp. Info.), the exact free-energy profile can be directly calculated from Eq. (1).

Fig. 1 shows that integration of Eq. (2) does give back the exact free-energy profile in Eq. (1), but requires consideration of all *three* (mean force, Jacobian and Leibnizian) contributions.[1] In this example, neglect of the Leibnizian contribution disclosed in this work leads to an *overestimate* the free-energy barrier (between minimum and maximum points) by about 25%.

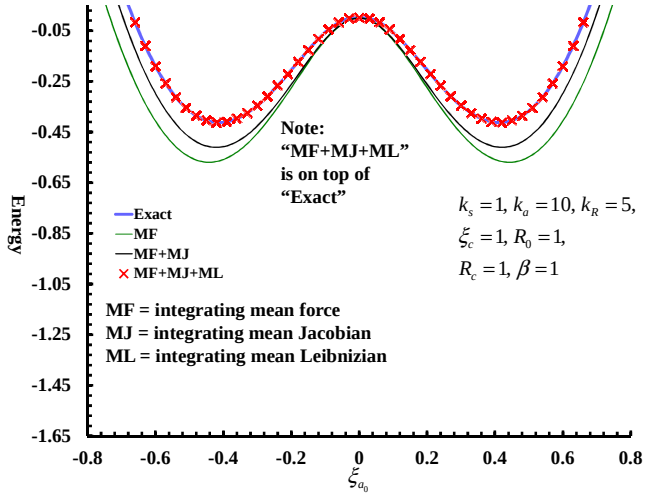


Fig. 1. Exact free-energy profile [Eq. (1)] compared with integrated mean force with and without Jacobian and Leibnizian contributions [Eq. (2)]. All plots are symmetric and anchored at zero value when $\xi_{a_0} = 0$.

*II: Orthogonal contravariant space formulation (no full Jacobian requirement).* To use Eq. (2), we have to express Cartesian coordinates $\{x\}^N$ and then $V$ in terms of $\{q\}^N$ for the instantaneous Jacobian and force contributions, respectively. In the language of differential geometry or general relativity, Eq. (2) is in the *covariant* vector space representation [33]. However, working in the covariant space is often not practical since we usually only know the definition of $q_\xi$ in terms of $\{x\}^N$, rather than having the ability to express all $\{x\}^N$ in terms of $\{q\}^N$.

It would thus be desirable to determine $dG_\xi(\xi_0)/d\xi_0$ in the *contravariant* space, which is the vector space expressing $\{q\}^N$ in terms of $\{x\}^N$. Note that if and only if $q_\xi$ is orthogonal to the

subspace spanned by the rest of coordinates, then the covariant vector space for $q_\xi$ is identical (or trivially related) to the contravariant vector space for $q_\xi$ [33]. We designate these vector spaces as orthogonal covariant and orthogonal contravariant spaces, respectively. Nevertheless, as long as $q_\xi$ can be a member of at least one *complete* set of coordinates, then the orthogonal space always exists, because the rest of the subspace can be simply adjusted to be perpendicular to $q_\xi$ by orthogonalization.

In the orthogonal *contravariant* space, the unit vector for $q_\xi$ is:[1]

$$\boxed{\hat{q}_\xi = \vec{\nabla} q_\xi \Big/ \left|\vec{\nabla} q_\xi\right|} \tag{3}$$

which is identical to the unit vector in the orthogonal *covariant* space (Supp. Info.). Similarly, in the orthogonal *contravariant* space, the Jacobian scale factor for $q_\xi$ is:[1]

$$\boxed{h_{q_\xi} = \left|\vec{\nabla} q_\xi\right|^{-1}} \tag{4}$$

which is identical to the scale factor in the orthogonal *covariant* space (Supp. Info.). Notably, the Jacobian scale factor in Eq. (4) is *misinterpreted* as the *full* Jacobian determinant in Ref. [[20]]. Eq. (3) and (4) are key results for this paper, and in particular the gradient operator in Eq. (3) and (4) that can be calculated in Cartesian coordinates (and in any complete curvilinear coordinates), i.e., Eq. (3) and (4) depend *only* on the definition of $q_\xi$.

Additionally, if we define $q'_\xi \equiv q_\xi$, it is *always* possible that we keep orthogonalizing the subspace $\left\{q_{m\neq\xi}\right\}^{N-1}$ such that we find a *complete* set of coordinates $\left\{q'\right\}^N$ which are all orthogonal to one another. Using Eq. (3), Eq. (4), and the orthogonality of $\left\{q'\right\}^N$, we find:[1]

$$\vec{\nabla}V\cdot\left(h_{q_\xi}\hat{q}_\xi\right)=\left(\frac{\partial V}{\partial q_\xi}\right)_{\left\{q'_{m\neq\xi}\right\}^{N-1}}=\vec{\nabla}V\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right) \tag{5}$$

Eq. (5) is the instantaneous force in the orthogonal contravariant space that depends only on $q_\xi$.

Similarly, using Eq. (3), (4), the Jacobian determinant for $\{q'\}^N$ as $\left|\vec{J}'\right|=\prod_{m=1}^{N}h_{q'_m}$ (due to the orthogonality in $\{q'\}^N$), and $q'_\xi\equiv q_\xi$, now we have the divergence of $h_{q_\xi}\hat{q}_\xi$, i.e., $\vec{\nabla}\cdot\left(h_{q_\xi}\hat{q}_\xi\right)$, as:[1]

$$\vec{\nabla}\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right)=\frac{1}{\prod_{m=1}^{N}h_{q'_m}}\left(\frac{\partial}{\partial q_\xi}\left[h_{q'_\xi}\prod_{m\neq\xi}^{N-1}h_{q'_m}\right]\right)_{\left\{q'_{m\neq\xi}\right\}^{N-1}}=\frac{1}{\left|\vec{J}'\right|}\left(\frac{\partial\left|\vec{J}'\right|}{\partial q_\xi}\right)_{\left\{q'_{m\neq\xi}\right\}^{N-1}} \tag{6}$$

Eq. (6) is the instantaneous Jacobian contribution in the orthogonal contravariant space that depends only on the (constraint) coordinate, $q_\xi$.

Towards this end, since the coordinates in $\{q'\}^N$ are orthogonal to one another, the domains of all coordinates should *not* be related, and should be *constants*. Thus, there should be *no* Leibnizian contribution in Eq. (2) with $\{q'\}^N$. Finally, we can express $dG_\xi(\xi_0)/d\xi_0$ in the orthogonal contravariant space [16, 21] by substituting Eq. (5) and (6) into Eq. (2):[1]

$$\boxed{\frac{dG_\xi(\xi_0)}{d\xi_0}=\left\langle\vec{\nabla}V\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right)\right\rangle_{\xi_0}-\frac{1}{\beta}\left\langle\vec{\nabla}\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right)\right\rangle_{\xi_0}} \tag{7}$$

Unlike Eq. (2) which is in the *covariant* space, Eq. (7) is in the orthogonal *contravariant* space for $q_\xi$, which depends *only* on the definition of $q_\xi$ (without full Jacobian transformation). The desired result, Eq. (7), is correct and general, as long as $q_\xi$ can be a member of a *complete* set of coordinates.[1]

We note that Ref. [21] presents a *similar* (but different) formulation as the above Eq. (7), and it

also indicates the orthogonality requirement that is ignored in Ref. [16]; however, using the formula given in Ref. [21] *cannot* return us the correct *unit* or *dimension* of the mean force when the coordinate of interest is *not* in the dimension of length. This is because, as opposed to our Eq. (7), in which we have a squared power in our denominator, i.e., $\left|\vec{\nabla}q_\xi\right|^2$, the corresponding denominator in Ref. [21] does not have any power index, i.e., $\left|\vec{\nabla}q_\xi\right|$ (instead of $\left|\vec{\nabla}q_\xi\right|^2$).

Figure 2 illustrates that integration of Eq. (7) indeed returns the exact free-energy profile in Eq. (1) without requiring the full Jacobian transformation. Note that *no* Leibnizian contribution to the free energy profile is in the *orthogonal contravariant* space, and that the mean force and the mean Jacobian contributions in the *orthogonal contravariant* space respectively differ from their own counterparts in the *non-orthogonal* covariant space. In the present example, integrating the *non-orthogonal* mean force would *overestimate* the exact free-energy barrier by about 40% (Fig. 1), whereas integrating the *orthogonal* mean force would cause the *overestimate* by about 300% (Fig. 2).

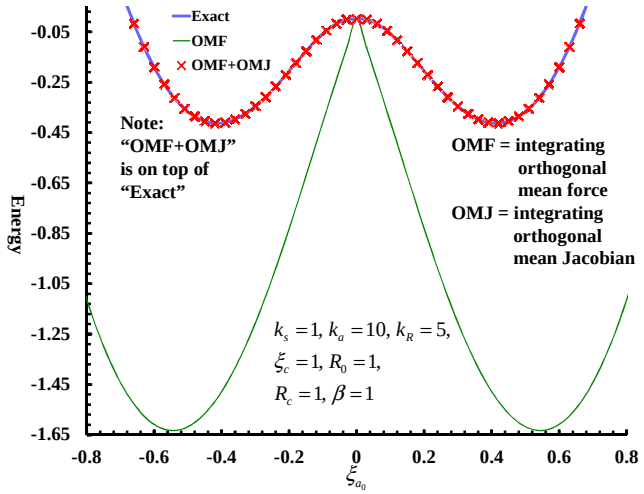


Fig. 2. Exact free-energy profile [Eq. (1)] compared with integrated *orthogonal* mean force with and without Jacobian contributions [Eq. (7)], in which the Leibnizian term vanishes. All curves are symmetric and anchored at zero value when $\xi_{a_0} = 0$.

*III: Fixman term (constrained MD).* In MD simulations, once we put a constraint to eliminate a degree of freedom in configuration space, e.g., $q_\xi = \xi_0$, then the degree of freedom of its velocity must vanish , i.e., $\dot{q}_\xi = 0$. Thus, the constant nature of *C* in Eq. (1) does not exist because we no longer integrate the *entire* momentum space with $\dot{q}_\xi = 0$. To figure out the consequence of this, we re-express Eq. (1) in the phase space *explicitly* consisting of $q_\xi$ and $\dot{q}_\xi$. First, we transform the Cartesian coordinates and momenta to the mass-scaled coordinates $\{s\}^N$, $s_i = \sqrt{M_i}\, x_i$, and the mass-scaled velocities $\{\dot{s}\}^N$, $\dot{s}_i = p_i \big/ \sqrt{M_i} = \sqrt{M_i}\, \dot{x}_i$, respectively. We denote $\vec{\vec{J}}_s$ as the Jacobian transformation for mass-scaled $ds_j$ in terms of a *complete* set of *orthogonal* coordinates $\{w\}^N$, in which $w_\xi \equiv q_\xi$. $\vec{\vec{J}}_s$ is a function of masses and $\{w\}^N$ only. The same $\vec{\vec{J}}_s$ is also the transformation for the kinetic energy and for mass-scaled $d\dot{s}_j$ in terms of $\{\dot{w}\}^N$, in which $\dot{w}_\xi \equiv \dot{q}_\xi$. So Eq. (1) in the phase space of $\{w\}^N$ and $\{\dot{w}\}^N$ is (Supp. Info.):

$$\boxed{G_\xi(\xi_0) = \frac{-1}{\beta} \ln \iint \frac{\left|\vec{\vec{J}}_s\right| dw^{N-1} \left|\vec{\vec{J}}_s\right| d\dot{w}^N}{h^N} e^{\frac{-\beta}{2}\left[\sum_{j=1}^{N} h_{w_j}^2 \dot{w}_j^2 + V\left(\xi_0, \{w\}^{N-1}\right)\right]}} \tag{8}$$

where *h* is Planck's constant, $h_{w_j}$ is a Jacobian scale factor with $\left|\vec{\vec{J}}_s\right| = \prod_{m=1}^{N} h_{w_m}$. By putting one more constraint: $\dot{w}_\xi \equiv \dot{q}_\xi = 0$ in Eq. (8), then integrating out all velocities, the *modified* free-energy profile $G_{\xi\xi\xi}$ is:

$$G_{\xi\xi\xi}(\xi_0) = \frac{-1}{\beta} \ln \int h_{w_\xi} \left|\vec{\vec{J}}_s\right| dw^{N-1} e^{-\beta V\left(\xi_0, \{w\}^{N-1}\right)} \frac{1}{h} \left(\frac{k_B T}{2\pi\hbar^2}\right)^{\frac{N-1}{2}} \tag{9}$$

By Zwanzig's free-energy perturbation theory [34], Eq. (8) and (9) are related (Supp. Info.):

$$\boxed{G_\xi(\xi_0) = G_{\xi\xi\xi}(\xi_0) - k_B T \ln \left\langle h_{w_\xi}^{-1} \right\rangle_{\xi_0 \dot{q}_\xi} - k_B T \ln \sqrt{2\pi k_B T}} \tag{10}$$

The second term on the RHS of Eq. (10) is called the Fixman term [28].

Now if we do the perturbation on ensemble average, then Eq. (7) can be obtained from the constrained MD (Supp. Info.):

$$\boxed{\frac{dG_\xi(\xi_0)}{d\xi_0} = \frac{1}{\left\langle h_{w_\xi}^{-1}\right\rangle_{\xi_0\dot{q}_\xi}}\left(\left\langle \frac{1}{h_{w_\xi}}\vec{\nabla}V\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right)\right\rangle_{\xi_0\dot{q}_\xi} - \frac{1}{\beta}\left\langle \frac{1}{h_{w_\xi}}\vec{\nabla}\cdot\left(\frac{\vec{\nabla}q_\xi}{\left|\vec{\nabla}q_\xi\right|^2}\right)\right\rangle_{\xi_0\dot{q}_\xi}\right)} \tag{11}$$

The $V$ in Eq. (11) is the original potential energy. Following from Eq. (4), it is now straightforward to yield the Jacobian scale factor $h_{w_\xi}$ for $w_\xi \equiv q_\xi$ in the *mass-scaled* orthogonal *contravariant* space:

$$\boxed{\frac{1}{h_{w_\xi}} = \left|\vec{\nabla}_{\{s\}}w_\xi\right| = \left|\sum_{j=1}^{N}\left(\frac{\partial w_\xi}{\partial s_j}\right)_{\{s_{m\neq j}\}^{N-1}}\hat{s}_j\right| = \left|\sum_{j=1}^{N}\left(\frac{\partial q_\xi}{\partial x_j}\right)_{\{x_{m\neq j}\}^{N-1}}\frac{\hat{x}_j}{\sqrt{M_j}}\right|} \tag{12}$$

Eq. (12) *correctly* indicates the velocity contribution to the partition function in Eq. (10) and. (11) *decreases* with mass, as opposed to momentum (see Supp. Info. for a simple one-dimensional one-body example).[1] Notably, the *entire* correction term in Eq. (10), i.e., $h_{w_\xi}^{-1}$, is the *inverse* of the entire term reported in other literature [15, 16, 18, 22-27, 29].

To verify the mass-dependence of the Fixman term shown in Eq. (10) and (12) is correct, we consider a two-degree-of-freedom system: $x_1$ and $x_2$ with mass $m_1$ and $m_2$, respectively. We would like to express these two degrees of freedom in terms of a set of two standard elliptic coordinates $\{\mu,\nu\}$, in which the coordinate of interest $\nu\in[0,2\pi]$ is: $2a\cos\nu \equiv \sqrt{(x_1+a)^2+x_2^2}-\sqrt{(x_1-a)^2+x_2^2}$. Fig. 3, which is plotted from Eq. (1), (9), and (10) (details for the two-body double-well potential and full Jacobian transformation are in the Supp. Info.) assures the mass-dependence of the Fixman term derived in Eq. (10), (11), and (12) is

correct and exact (i.e., FEP:VC+FT*), while using the Fixman term found in the literature[15, 16, 18, 22-27, 29] (i.e., FEP:VC+FT) *does not* return the exact results and *underestimates* the exact free-energy barrier by about 35%.

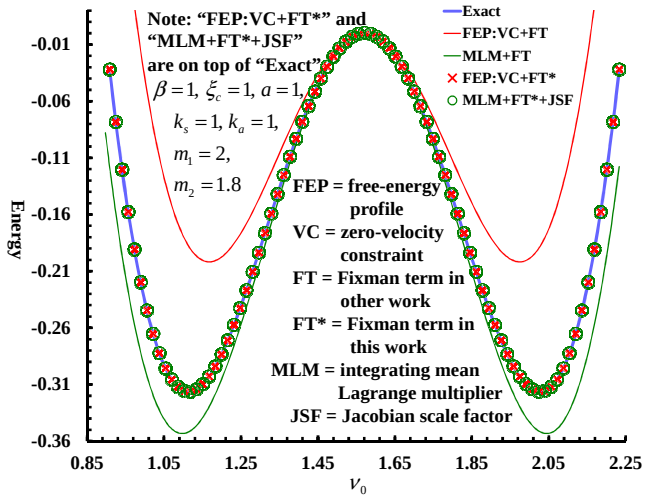


Fig. 3. Comparison of exact free-energy profile [Eq. (1)], profile with zero-velocity constraint [Eq. (9)], Fixman term in this work: "FT*" [Eq. (12)], Fixman term in the literature: "FT" [i.e., inverse of Eq. (12) [15, 16, 18, 22-27, 29]], and integrated mean Lagrange multiplier with and without Jacobian scale factor derived in this work [i.e., Eq. (14) and (13), respectively]. All curves are symmetric and anchored at zero value when $\nu_0 = \pi/2$.

*IV: Integrated mean Lagrange multiplier (constrained MD).* Certain literature [15, 16, 22-27] suggests that integration of the mean of the Lagrange multiplier $\lambda_\nu$ associated with the constraint force would return $G_{\nu\dot{\nu}}(\nu_0)$ [Eq. (9)]. If generally true, then this would be a convenient approach because the explicit knowledge of the Laplacian of the coordinate of interest in Eq. (11) is not required. In their derivations [23, 26, 27] that equate $\lambda_\nu$ with the exact free-energy profile, the Jacobian contribution is not discussed (or else is assumed to vanish) and the Fixman term is the inverse of the Fixman term shown in Eq. (10) and (12). As a result, the following quantity has been claimed to be equal to the original free-energy profile $G_\nu(\nu_0)$ [Eq. (1)] [15, 16, 22-27]:

$$\int^{\nu_0} \langle \lambda_\nu \rangle_{\nu'\dot{\nu}} \, d\nu' - k_B T \ln \langle h_\nu \rangle_{\nu_0 \dot{\nu}} \tag{13}$$

From Fig. 3, in contrast to our exact relation (i.e., FEP:VC+FT*), it is clear that Eq. (13) (i.e., MLM+FT) does *not* return us the exact $G_\nu(\nu_0)$, and in fact (not shown in Fig. 3) integrating $\langle \lambda_\nu \rangle_{\nu\dot{\nu}}$ (i.e., MLM) also does not return $G_{\nu\dot{\nu}}(\nu_0)$ (i.e., FEP:VC). The use of Eq. (13), for this example, would lead to an *overestimate* of the exact free-energy barrier by about 10%. In order to correct Eq. (13) to obtain the exact free-energy profile, not only do we first need to use the Fixman term derived in this work [i.e., Eq. (10) and (12)], but also we need to add a new contribution from mass-scaled Jacobian scale factor as follows:

$$G_\nu(\nu_0) = \int^{\nu_0} \langle \lambda_{\nu'} \rangle_{\nu'\dot{\nu}} \, d\nu' - k_B T \ln \left\langle h_\nu^{-1} \right\rangle_{\nu_0 \dot{\nu}} - 2k_B T \int^{\nu_0} \left\langle \bar{\nabla}_{\{s\}} h_\nu \cdot \left( \frac{\bar{\nabla}_{\{s\}} \nu}{\left| \bar{\nabla}_{\{s\}} \nu \right|} \right) \right\rangle_{\nu'\dot{\nu}} d\nu' + \text{constant} \tag{14}$$

This exact relation is illustrated in Fig. 3 (MLM+FT*+JSF). The new third term on the RHS of Eq. (14) is associated with the differentiation of the mass-scaled Jacobian scale factor of the coordinate of interest, and arises from the inherent $\dot{\nu} = 0$ condition in constrained MD simulation where the corresponding coordinate is held fixed. With the zero-velocity constraint, this differentiation information is unavailable in the Lagrange multiplier and requires explicit knowledge of the Laplacian of the coordinate of interest to recover it fully [18, 29] (Supp. Info.).

*V: Conclusion*. In this work, new exact relations between PMF and free-energy profile (with or without full Jacobian transformation) are presented. Specifically, we apply mathematical and physical concepts of covariant and contravariant vector space [e.g., Eq. (3), (4), (12)] from differential geometry and general relativity to derive explicit, practical expressions that equate the PMF obtained from (constrained) MD simulations with the free-energy profile in terms of the Cartesian coordinates [i.e., Eq. (7) and (11)]. These general expressions require knowledge only

of the (constraint) coordinate of interest, and thus are immensely practical in molecular simulations that may require complex coordinate constraints used as basic variables in the free-energy profile. Further, we demonstrate that, in general, in addition to the Jacobian contribution, the Leibnizian contribution derived here still needs to be considered when there is an interdependence of the integration domains [Eq. (2)] (even if a full Jacobian transformation is available). Moreover, we illustrate that the individual contribution from the mean force, and from the Jacobian can vary significantly (e.g., from ~140% to ~400% in terms of the free-energy barriers considered in Figs. 1 and 2) with different complete sets of coordinates for which the (constraint) coordinate of interest is a simultaneous member. A sufficient condition to make the contributions be invariant is to be in the orthogonal contravariant space [Eq. (5) and (6)], in which the Leibnizian contribution vanishes. Next, we present a definition of the Fixman term with correct dependence on mass in constrained MD simulations [Eq. (10), (11), (12)]. Finally, we divulge that an additional term arising from the Jacobian scale factor contribution that requires explicit knowledge of the Laplacian of the coordinate of interest [Eq. (14)] must be included in order to exactly equate the integrated mean Lagrange multiplier with the free-energy profile. All formulations presented here are verified and illustrated by new non-trivial benchmark numerical results, and are compared with those reported in other literature, which, for the examples studied here, result in discrepancies from the exact values by about 10% to 35% in terms of free-energy barriers (Figs. 1-3). Future work will involve application of the present formulations to simulations of more complex, real-world molecular processes, and comparison with other methods to determine the free-energy profiles.

Acknowledgement

The authors are grateful for financial support provided by NIH (GM084149 to D.Y.).

## Supporting Information

This information is available free of charge via the Internet at http://pubs.acs.org

## References


1. Wong, K.-Y., Developing a systematic approach for *ab initio* path-integral simulations. In *Molecular Dynamics / Book 1 - Theoretical Developments and Applications in Nanotechnology and Energy*, Wang, L., Ed. InTech: 2012; pp 107-132; Open access from: http://www.intechopen.com/books/molecular-dynamics-theoretical-developments-and-applications-in-nanotechnology-and-energy/developing-a-systematic-approach-for-ab-initio-path-integral-simulations.
2. Chipot, C.; Pohorille, A., *Free Energy Calculations: Theory and Applications in Chemistry and Biology*. Springer Berlin Heidelberg: 2007; Vol. 86, p 535.
3. Tuckerman, M. E., *Statistical mechanics : theory and molecular simulation*. Oxford University Press: Oxford ; New York, 2010; p xv, 696 p.
4. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y.; Wriggers, W., *Science* **2010,** *330* (6002), 341-346.
5. Zheng, L.; Chen, M.; Yang, W., *Proc. Natl. Acad. Sci. U. S. A.* **2008,** *105* (51), 20227-20232.
6. Šarić, A.; Cacciuto, A., *Phys. Rev. Lett.* **2012,** *108* (11), 118101.
7. Tuckerman, M. E.; Marx, D., *Phys. Rev. Lett.* **2001,** *86* (21), 4946-4949.
8. Tuckerman, M. E.; Marx, D.; Parrinello, M., *Nature (London)* **2002,** *417* (6892), 925-929.
9. Minary, P.; Tuckerman, M. E., *J. Am. Chem. Soc.* **2004,** *126* (43), 13920-13921.
10. Wong, K.-Y.; Gu, H.; Zhang, S.; Piccirilli, J. A.; Harris, M. E.; York, D. M., *Angew. Chem. Int. Ed.* **2012,** *51* (3), 647-651.
11. Wong, K.-Y.; Lee, T.-S.; York, D. M., *J. Chem. Theory Comput.* **2011,** *7* (1), 1-3.
12. Wong, K.-Y.; Gao, J., *FEBS Journal* **2011,** *278* (14), 2579-2595.
13. Wong, K.-Y.; Gao, J., *Biochemistry* **2007,** *46* (46), 13352-13369.
14. Kästner, J., *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011,** *1* (6), 932-942.
15. Darve, E., Thermodynamic Integration Using Constrained and Unconstrained Dynamics. In *Free Energy Calculations*, Chipot, C.; Pohorille, A., Eds. Springer Berlin Heidelberg: 2007; Vol. 86, pp 119-170.
16. den Otter, W. K., *J. Chem. Phys.* **2000,** *112* (17), 7283-7292.
17. Kirkwood, J. G., *J. Chem. Phys.* **1935,** *3* (5), 300-313.
18. Doltsinis, N., Free Energy and Rare Events in Molecular Dynamics. In *Computational Nanoscience: Do It Yourself! NIC Series*, Grotendorst, J.; Blügel, S.; Marx, D., Eds. John von Neumann Institute for Computing, Jülich: 2006; Vol. 31, pp 375-387.
19. Hénin, J. r.; Fiorin, G.; Chipot, C.; Klein, M. L., *J. Chem. Theory Comput.* **2010,** *6* (1), 35-47.
20. Ruiz-Montero, M. J.; Frenkel, D.; Brey, J. J., *Mol. Phys.* **1997,** *90* (6), 925-942.
21. Ciccotti, G.; Kapral, R.; Vanden-Eijnden, E., *ChemPhysChem* **2005,** *6*, 1809-1814.
22. den Otter, W. K., *J. Chem. Phys.* **2009,** *131*, 205101.
23. Li, W.; Rudack, T.; Gerwert, K.; Gräter, F.; Schlitter, J., *J. Chem. Theory Comput.* **2012**, Article ASAP.
24. Schlitter, J., *Eur. Phys. J. Special Topics* **2011,** *200*, 91-105.
25. Burisch, C.; Markwick, P. R. L.; Doltsinis, N. L.; Schlitter, J., *J. Chem. Theory Comput.* **2007,** *4* (1), 164-172.
26. Schlitter, J.; Klahn, M., *J. Chem. Phys.* **2003,** *118* (5), 2057-2060.
27. Mülders, T.; Kruger, P.; Swegat, W.; Schlitter, J., *J. Chem. Phys.* **1996,** *104* (12), 4869-4870.
28. Fixman, M., *Proc. Natl. Acad. Sci. U. S. A.* **1974,** *71* (8), 3050-3053.
29. Sprik, M.; Ciccotti, G., *J. Chem. Phys.* **1998,** *109*, 7737-7744.
30. Flanders, H., *The American Mathematical Monthly* **1973,** *80* (6), 615-627.
31. McQuarrie, D. A., *Statistical mechanics*. University Science Books: Sausalito, Calif., 2000; p xii, 641 p.
32. Cohen-Tannoudji, C.; Diu, B.; Laloë, F., *Quantum mechanics*. Wiley: New York, 1977; Vol. 1-2, p xv, 1524 p.
33. Arfken, G. B.; Weber, H.-J., *Mathematical methods for physicists*. 5th ed.; Academic Press: San Diego, 2001; p xiv, 1112 p.
34. Zwanzig, R. W., *J. Chem. Phys.* **1954,** *22* (8), 1420-1426.