

# COMMENT ON TWO DISTINCT NOTIONS OF FREE ENERGY

CARSTEN HARTMANN\*

**Abstract.** We comment on two different notions of the thermodynamical free energy that are used in Hamiltonian molecular dynamics. Both concepts have different scopes of applications as was pointed out recently in the context of high-friction Langevin dynamics. We show that problems that rely on either definition can be treated in a uniform way using constrained molecular dynamics. Not only proves this useful in designing algorithms that sample the free energy landscape, but it also clarifies the relation between seemingly contradictory results that are present in the literature.

## 1. Introduction

The calculation of free energy profiles along certain prescribed coordinates plays an essential role in physical chemistry and dynamical systems. In particular in molecular dynamics applications there is a variety of phenomena as, for instance, molecular solvation, enzyme catalysis, or conformation dynamics, the analytical understanding of which is directly related to the corresponding free energy landscape.

In recent years progress has been made towards algorithms that efficiently sample free energy profiles, yet the precise mathematical understanding of the notion *free energy as the potential of mean force* was lacking. Even worse there are two incommensurable definitions of free energy in the literature, although they both rely on the same thermodynamical formalism.

From an application viewpoint each of these different definitions makes perfect sense, depending on the problems considered. However from a dynamical or geometrical viewpoint it is, most notably, the standard free energy, which is prevalent in the literature, that remains rather obscure, for it is not possible to fully relate it to the properties of the underlying dynamical system. However we will show that there is a particular definition of the quantity free energy, that is used in transition state theory, which allows for a concise geometric interpretation in the context of Hamiltonian dynamics subject to holonomic constraints.

We briefly review the available concepts of the Helmholtz free energy as presented by E and Vanden-Eijnden [1] for the high-friction Langevin dynamics. Then we show how the apparently contradictory results in the work of Mülders *et al.* [3], or Sprik and Ciccotti [2] can be related to each other within the respective free energy concepts.

## 2. Two definitions

Consider a separable Hamiltonian on the phase space  $T^*\mathbf{R}^n \simeq \mathbf{R}^n \times \mathbf{R}^n$

$$H : T^*\mathbf{R}^n \rightarrow \mathbf{R}, \quad H(q, p) = \frac{1}{2} \langle p, p \rangle + V(q)$$

which is the sum of kinetic and potential energy.<sup>1</sup> Let  $\xi : \mathbf{R}^n \rightarrow \mathbf{R}^s$  be a reaction coordinate, and consider fibres  $\Sigma = \xi^{-1}(x)$  for  $x \in \mathbf{R}^s$  which are smooth and compact submanifolds of codimension  $s$  in  $\mathbf{R}^n$ . To keep things simple we assume that the  $\xi_i, \xi_j$  are mutually orthogonal with respect to the Euclidean metric, that is,

$$\langle \text{grad } \xi_i, \text{grad } \xi_j \rangle = 0 \quad \text{whenever } i \neq j,$$

---

\*Freie Universität Berlin, Institut für Mathematik II, Arnimallee 2-6, 14195 Berlin, Germany, e-mail: chartman@math.fu-berlin.de, supported by the DFG Research Center "Mathematics for Key Technologies" (FZT 86) in Berlin

<sup>1</sup>We have scaled the coordinates according to  $q \mapsto M^{1/2}q$ , and  $p \mapsto M^{-1/2}p$ , where  $M$  is the positive-definite, and symmetric mass matrix. This is clearly a symplectic transform, and it allows us to set  $M = \mathbf{1}$  in what follows.

where  $\langle \cdot, \cdot \rangle : T_q \mathbf{R}^n \times T_q \mathbf{R}^n \rightarrow \mathbf{R}$  is the standard inner product of  $\mathbf{R}^n$ . Let the orbits of  $H(q, p)$  have the invariant distribution  $\rho \propto \exp(-\beta H(q, p))$ , where  $\beta = 1/T$  is the inverse temperature. One possible way to define the *Helmholtz free energy* is by means of the marginal density of the reaction coordinate

$$Q(x) = \int_{\mathbf{R}^n \times \mathbf{R}^n} \exp(-\beta H(q, p)) \delta(\xi(q) - x) dq dp$$

which can be expressed as the equivalent surface integral

$$Q(x) = \int_{\Sigma \times \mathbf{R}^n} \exp(-\beta H(q, p)) (\det J_\xi(q))^{-1} d\sigma(q, p). \quad (2.1)$$

where  $d\sigma(q, p)$  is the surface element of  $\Sigma \times \mathbf{R}^n$  considered as a submanifold of phase space  $T^*\mathbf{R}^n \simeq \mathbf{R}^n \times \mathbf{R}^n$ . The matrix  $J_\xi(q) \in \mathbf{R}^{s \times s}$  is defined as

$$J_\xi(q) = \text{diag}(\|\text{grad} \xi_1(q)\|, \dots, \|\text{grad} \xi_s(q)\|), \quad q \in \Sigma,$$

where  $\|\cdot\|$  is understood as the usual 2-norm associated with  $\langle \cdot, \cdot \rangle$ . The free energy is given by the logarithm of the probability density, that is,

$$F(x) = -\beta^{-1} \ln Q(x). \quad (2.2)$$

A second definition of the Helmholtz free energy utilizes the probability density of the submanifold  $\Sigma \times \mathbf{R}^n \subset \mathbf{R}^n \times \mathbf{R}^n$ , which is given by the expression

$$Z(x) = \int_{\Sigma \times \mathbf{R}^n} \exp(-\beta H(q, p)) d\sigma(q, p) \quad (2.3)$$

Again, this free energy is defined through the thermodynamical relation

$$G(x) = -\beta^{-1} \ln Z(x). \quad (2.4)$$

Although the first definition is certainly the traditional definition of free energy it has the unpleasant property that it depends on the parameterization of  $\Sigma$  as can be readily checked. See the review [1] for details.

Clearly  $F$  is directly available from the marginal density of the reaction coordinate after a very long simulation of  $H(q, p)$ , whereas  $G$  is not. However it is easy to see that one can arbitrarily switch between both definitions, exploiting that

$$\exp(-\beta H) (\det J_\xi)^{-1} = \exp(-\beta H - \ln \det J_\xi). \quad (2.5)$$

Thus comparing (2.1) and (2.3) we can compute either  $F$  or  $G$  with the same algorithm simply by adding or subtracting the so-called *Fixman potential*  $V_F = \beta^{-1} \ln \det J_\xi$ .

### 3. Free energy as the potential of mean force

We shall try to establish a relation between any of the the two free energies and the corresponding mean force transversally to  $\Sigma$  which a particle is exposed to. Once this is done we can use the relation (2.5) in order to switch from  $F$  to  $G$  or vice versa.

Algorithmically we can exploit the dichotomy of the free energy as the potential of mean force, and run constrained simulations on each fibre  $\Sigma = \xi^{-1}(x)$ , sampling the respective mean force. This procedure is known as thermodynamic integration, and goes back to Kirkwood [4]. Formally the constrained Hamilton equations read [5]

$$\dot{q} = \nabla_p H(q, p) \quad (3.1)$$

$$\dot{p} = -\nabla_q H(q, p) - \lambda \nabla_q \xi(q), \quad \xi(q) = x, \quad (3.2)$$

where  $\lambda$  denotes the Lagrange undetermined multiplier. As  $\Sigma \subset \mathbf{R}^n$ , we can take advantage of the natural inclusion  $T\Sigma \subset T\mathbf{R}^n$ , and, similar to the viewpoint of Lagrangian mechanics, construct a constrained Hamiltonian  $H_\Sigma$  by restricting it according to  $H_\Sigma = H|_{T^*\Sigma}$ . The invariant density that is generated by the orbits of the constrained Hamiltonian  $H_\Sigma$  is the function

$$\rho_\Sigma = N_\Sigma^{-1} \exp(-\beta H_\Sigma)$$

with the normalization constant

$$N_\Sigma = \int_{T^*\Sigma} \exp(-\beta H_\Sigma(q, p)) d\zeta(q, p),$$

where  $d\zeta(q, p)$  labels the phase space surface element of the constrained phase space  $T^*\Sigma \subset \mathbf{R}^n \times \mathbf{R}^n$ . We denote averages with respect to the probability density  $\rho_\Sigma$  by

$$\langle \mathcal{O}(q, p) \rangle_\Sigma = \int_{T^*\Sigma} \mathcal{O}(q, p) \rho_\Sigma(q, p) d\zeta(q, p), \quad (3.3)$$

and we shall start with the free energy definition (2.4). By properly extending the surface element  $d\sigma(q, p)$  to the ambient space using a symplectic change of coordinates, we can take the derivative of (2.4) with respect to parameter  $x$ . This yields [6]

$$\nabla_x G(x) = Z^{-1}(x) \int_{\Sigma \times \mathbf{R}^n} X(q, p) \exp(-\beta H(q, p)) d\sigma(q, p), \quad (3.4)$$

where the covector field  $X : (T^*\mathbf{R}^n)|_\Sigma \rightarrow T^*\mathbf{R}^s$  is the generalized normal force that is obtained from differentiating the Hamiltonian with respect to  $x$  after the coordinate change. See [6] for details. The normal force can be decomposed according to<sup>2</sup>

$$X(q, p) = -\lambda(q, p_t) + \omega(q, p_t, p_n), \quad q \in \Sigma,$$

where  $p = p_t + p_n$  with  $p_t \in T^*\Sigma$  and  $p_n \in (T^*\Sigma)^\perp$ , and  $\lambda(q, p)$  is the solution of the Lagrange multiplier in the equations of motion (3.1)–(3.2) on condition that the constraint is satisfied, that is, the following conditions are met

$$\xi(q) = x \quad \& \quad \dot{\xi}(q) = 0 \quad \Leftrightarrow \quad q \in \Sigma \quad \& \quad p = p_t.$$

The Lagrange multiplier can easily be computed by differentiating the constraint equation  $\xi(q) = x$  twice with respect to time, and then inserting the result into the equations of motion (3.1)–(3.2). Note that the constraint  $\dot{\xi} = 0$  or  $p = p_t$ , respectively, has *not* been enforced in the derivation of (3.4); otherwise  $\omega = \omega(q, p_t, 0)$  would be zero. However it is a remarkable feature of mechanical systems subject to holonomic constraints that the constraint force, and hence the Lagrange multiplier, is defined solely by points lying in the constrained phase space bundle [8]. The map  $\omega$  is linear in both  $p_t$  and  $p_n$ ; hence it vanishes on average with respect to the conditional average it appears in equation (3.4).<sup>3</sup>

<sup>2</sup>This decomposition is based on the fundamental equations for submanifolds (Gauss formulae and Weingarten equations), which, as a matter of fact, mark a far more general property of vector-fields attached to submanifolds. See [7] for some basic facts about submanifolds.

<sup>3</sup>Strictly speaking,  $\omega$  is the cotangent bundle version of the connection 1-form of the frame associated with the normal bundle over  $\Sigma$ , where due to the mass scaling at the beginning we have  $p = \dot{q}$ . Physically speaking the normal connection couples vectors fields tangential and normal to  $\Sigma$ .

Note that it is only the normal momentum  $p_n$  that distinguishes the constrained average which is computed on  $T^*\Sigma$  from the conditional average on  $(T^*\mathbf{R}^n)|_\Sigma \simeq \Sigma \times \mathbf{R}^n$ . But as the constraint force does not depend on the normal momenta [5], nor does the Lagrange multiplier, we can replace the conditional average by the constrained average  $\langle \cdot \rangle_\Sigma$ , and we recover indeed the result of Mülders and his co-workers [3]

$$\nabla_x G(x) = -\langle \lambda(q, p) \rangle_\Sigma. \quad (3.5)$$

In point of fact this result was derived by Mülders *et al.* imposing the constraint from the outset. Anyway this does not make a difference, as should be clear now.

#### 4. Algorithmic realization

By definition the mean force (3.4) is an integrable vector field on  $S \subset \mathbf{R}^s$ , and we can recover  $G(x)$  by numerical integration over the components of the reaction coordinate. The last equation is very useful in order to compute the mean force during the course of the numerical integration of the constrained Hamiltonian system: the mean force is simply obtained by a time average of the Lagrange multiplier, provided the dynamics is ergodic with respect to the constrained invariant density  $\rho_\Sigma$ ; no further function evaluations are required. We can even use the same simple formula to access  $F(x)$ , too. Taking into account (2.5) enables us to compute the corresponding mean force  $-\nabla_x F$  in a straightforward manner [9]. We find

$$\nabla_x F(x) = -\langle \mu(q, p) \rangle_{\Sigma, \text{aug}}. \quad (4.1)$$

where the former Hamiltonian has been augmented by the Fixman potential  $V_F$  of the reaction coordinate  $H \mapsto H + \beta^{-1} \ln \det(J_\xi)$ . Accordingly  $\mu$  is the Lagrange undetermined multiplier that corresponds to the modified Hamiltonian, and  $\langle \cdot \rangle_{\Sigma, \text{aug}}$  denotes the respective constrained average. It is easy to check that

$$\langle \mathcal{O}(q, p) \rangle_{\Sigma, \text{aug}} = \frac{\langle (\det J_\xi(q))^{-1} \mathcal{O}(q, p) \rangle_\Sigma}{\langle (\det J_\xi(q))^{-1} \rangle_\Sigma}$$

which is the famous blue moon relation of Carter *et al.* [10], but note that nowhere throughout the calculation we have imposed the constraint  $\xi = 0$ . Moreover we have

$$\mu_i(q, p) = \lambda_i(q, p) - \beta^{-1} \frac{1}{\|\text{grad } \xi_i(q)\|^2} \langle \text{grad } \xi_i(q), \text{grad } V_F \rangle$$

which is in perfect agreement with the formulae that have been derived in various works [2, 11, 12, 13] for the case when  $\Sigma$  is a hypersurface of codimension one in  $\mathbf{R}^n$ . In this case the unaveraged force field  $X(q, p)$  is even identical to the Lagrange multiplier  $\lambda(q, p_t)$ , for, as a general property of embedded submanifolds of codimension one, the coefficient of the normal connection identically vanishes, and so does  $\omega(q, p_t, p_n)$  in the expression below (3.4). Moreover it has been shown recently [14] that the codimension-one-formula for the mean force also applies for each of the components of the mean force if the reaction coordinates are orthogonal.

However likely important, our simple considerations reveal the origin of the terms that appear additionally to the averaged Lagrange multiplier  $\lambda(q, p_t)$  when calculating  $F(x)$  by means of thermodynamic integration methods. These terms are neither geometrically nor dynamically relevant or meaningful. In fact it is purely a matter of the particular definition starting either from the marginal distribution (2.1) or the conditional density (2.3). This links the seemingly contradictory results derived in the papers by Sprik and Ciccotti [2], or Mülders *et al.* [3], each of which is consistent within the respective concept of free energy.

**Acknowledgement.** The author thanks Tony Lelièvre for many fruitful discussions on the present topic.

#### REFERENCES

- [1] W. E and E. Vanden-Eijnden. Metastability, conformation dynamics, and transition pathways in complex systems. In S. Attinger and P. Koumoutsakos, editors, *Multiscale, Modelling, and Simulation*, pages 35–68. Springer, Berlin, 2004.
- [2] M. Sprik and G. Ciccotti. Free energy from constrained molecular dynamics. *J. Chem. Phys.*, 109(18):7737–7744, 1998.
- [3] T. Mülders, P. Krüger, W. Swegat, and L. Schlitter. Free energy as the potential of mean force. *J. Chem. Phys.*, 104(12):4869–4870, 1996.
- [4] J.G. Kirkwood. Statistical mechanics of fluid mixtures. *J. Chem. Phys.*, 3:300–313, 1935.
- [5] J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry*. Springer, New York, 1999.
- [6] C. Hartmann and C. Schütte. A constrained hybrid monte-carlo algorithm and the problem of calculating the free energy in several variables. *ZAMM*, 85(10):700–710, 2005.
- [7] M. Spivak. *Differential Geometry*, volume 4. Publish or Perish, Boston, 1975.
- [8] P.A.M. Dirac. Generalized Hamiltonian dynamics. *Can. J. Math.*, 2:129–148, 1950.
- [9] M. Fixman. Classical statistical mechanics of constraints: a theorem and applications to polymers. *PNAS*, 71:3050–3053, 1974.
- [10] E.A. Carter, G. Ciccotti, J.T. Hynes, and R. Kapral. Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Lett.*, 156(5):472–477, 1989.
- [11] W.K. den Otter. Thermodynamic integration of the free energy along a reaction coordinate in Cartesian coordinates. *J. Chem. Phys.*, 112(17):7283–7286, 2000.
- [12] E. Darve, M.A. Wilson, and A. Pohorille. Calculating free energies using a scaled-force molecular dynamics algorithm. *Mol. Sim.*, 28(1-2):113–144, 2002.
- [13] C. Hartmann and C. Schütte. A geometric approach to free energy calculations. *Comm. Math. Sci.*, 3(1):1–20, 2005.
- [14] G. Ciccotti, R. Kapral, and E. Vanden-Eijnden. Blue moon sampling, vectorial reaction coordinates, and unbiased constrained dynamics. *ChemPhysChem*, 6(9):1809–1814, 2005.