

# Theory of solutions in the energetic representation. I. Formulation

Nobuyuki Matubayasi<sup>a)</sup> and Masaru Nakahara

*Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan*

(Received 1 May 2000; accepted 21 July 2000)

The energetic representation of the molecular configuration in a dilute solution is introduced to express the solvent distribution around the solute over a one-dimensional coordinate specifying the solute–solvent interaction energy. In this representation, the correspondence is shown to be one-to-one between the set of solute–solvent interaction potentials and the set of solvent distribution functions around the solute. On the basis of the one-to-one correspondence, the Percus–Yevick and hypernetted-chain integral equations are formulated over the energetic coordinate through the method of functional expansion. It is then found that the Percus–Yevick, hypernetted-chain, and superposition approximations in the energetic representation determine the solvent distribution functions correctly to first-order with respect to the solute–solvent interaction potential and to the solvent density. The expressions for the chemical potential of the solute are also presented in closed form under these approximations and are shown to be exact to second-order in the solute–solvent interaction potential and in the solvent density. © 2000 American Institute of Physics.

[S0021-9606(00)51439-1]

## I. INTRODUCTION

In a modern theory of solutions, the solution structure is described by molecular distribution functions. Especially, the two-body distribution functions are the major targets of integral equation theories and are default quantities to compute in molecular simulations.<sup>1,2</sup> The distribution functions are defined over a certain representation of the configurations of the molecules of interest. The full coordinate of a molecule which specifies its configuration completely consists of the position and orientation of the molecule. When the full coordinate is employed to represent the molecular configuration, the distribution functions involve well-behaved mathematical structures and the commonly used Percus–Yevick (PY), hypernetted-chain (HNC), and superposition approximations are exact to first-order in the density of the solution.<sup>1–8</sup> A systematic description of the distribution functions over the full coordinate is possible in principle for a molecule of any symmetry by means of spherical harmonic expansion.<sup>9–11</sup> In the full coordinate representation, however, the multidimensional description cannot be avoided for the distribution functions when molecules of chemical interest are to be treated. Therefore, the full coordinate representation is not desirable in practice to describe the distribution functions of the solution.

The site–site representation is a reduced form of representing the molecular configuration. In this representation, a site–site distribution function is defined over the radial distance between the corresponding interaction sites and the structure of a solution is usually described by a set of site–site radial distribution functions. Since each site–site radial distribution function is represented over a one-dimensional abscissa, the site–site representation is conceptually and computationally convenient in the statistical description of

the intermolecular configuration of a pair of molecules in the solution. Indeed, the method of reference interaction site model (RISM), which implements the PY or HNC approximation in the site–site form, is widely used and achieves reasonable success in high-density molecular fluids.<sup>12–23</sup> A deficiency of the RISM integral equation is that it involves ill-behaved diagrammatical structures. Although this deficiency is removed in the “diagrammatically proper” formalism provided by Chandler *et al.*, the formalism needs to introduce additional correlation functions and loses the simplicity of the diagrammatically ill-behaved RISM approach.<sup>24,25</sup> In addition, the RISM integral equation with the PY or HNC closure does not give the correct zero-density limit and is not useful to describe a low- to medium-density fluid. This point is improved in the RISM-2 integral equations formulated by Chandler and Kojima and Arakawa, which are exact in the limit of zero density.<sup>26–28</sup> In the RISM-2 approximation with the PY or HNC closure, however, the first-order term with respect to the density of the solution is incorrect, whereas the corresponding approximation in the full coordinate representation is exact to first-order in the density. It actually seems difficult, due to the connectivity of the interaction sites contained within a molecule, to systematically devise a simple and improved form of integral equation in the site–site representation.

In this paper, we explore an alternative representation of the molecular configuration in a solution. The system treated is a dilute solution and contains a single solute molecule. The alternative representation, which we call the energetic representation, is introduced by adopting the solute–solvent interaction energy as the coordinate of a solvent molecule around the solute molecule. In the energetic representation, the solvent distribution around the solute is expressed over a one-dimensional abscissa for any type of solute–solvent interaction potential. The density functional theory can then be developed by establishing the one-to-one

<sup>a)</sup> Author to whom correspondence should be addressed.

correspondence between the choice of the solute–solvent interaction potential and the resulting distribution of the solvent in the energetic representation. The integral equations with the PY, HNC, and superposition approximations are further formulated over the coordinate specifying the solute–solvent interaction energy. Since these integral equations treat each of the solute and solvent molecules as one unit and do not separately describe the distinct sites of the molecule, they are exact to first-order in the solvent density of the solution.

The solvent distribution function around the solute molecule provides the chemical potential of the solute when the charging formula is employed and an integration is performed over the intermediate states of the coupling parameter of the solute–solvent interaction potential.<sup>1–3</sup> In fact, the intermediate states are not experimentally realizable and their choice is not unique. Thus, the integration over the coupling parameter not only demands much computational effort, but also inhibits an unambiguous interpretation and a clear understanding of the quantities of interest. In other words, it is desirable, both conceptually and computationally, that the chemical potential of the solute be expressed only in terms of properties at the initial and final states of the coupling parameter. Using such an expression for the chemical potential, which is called a closed form expression, the chemical potential can be evaluated only with the knowledge of the systems of interest.<sup>5,8,29–35</sup> In this paper, we also present the closed form expressions for the chemical potential of a solute under the PY, HNC, and superposition approximations in the energetic representation.

It is a tradition since Kirkwood's days that a theory of solutions is formulated over a coordinate specifying a set of positional variables. Indeed, the specification of the full coordinate of a molecule is equivalent to the specification of the positions of all the points in the molecule, and the site–site representation is implemented by labeling a certain set of points in the molecule and specifying the positions of those points. The energetic representation does not follow the tradition in the sense that the coordinate does not specify the position of any point in the molecule. When the representation is fixed for the molecular configuration, a refined closure relationship leads to an improved description of the solution. We show in this and subsequent papers, on the other hand, that the scheme for representing the molecular configuration may be changed to provide an accurate description of the solution without revising the form of the approximate relationship among correlation functions.

The organization of the present paper is as follows: In Sec. II, the distribution functions are defined over the coordinate specifying the solute–solvent interaction energy and the density functional theory is developed in the energetic representation. In Sec. III, the integral equations are formulated over the energetic coordinate with the PY, HNC, and superposition approximations and the closed form expressions are correspondingly presented for the chemical potential of a solute. In Sec. IV, the paper is concluded with remarks concerning the practical implementations and the comparison to commonly used theories. In a subsequent paper, the approximate procedures developed in the present

paper are employed to study the solvation thermodynamics and solvent-mediated interactions of various types of solutes in water over a wide range of thermodynamic conditions.

## II. DISTRIBUTION FUNCTIONS IN THE ENERGETIC REPRESENTATION

### A. Definitions

The system of our interest is a dilute solution containing a single solute molecule. For the sake of simplicity, it is supposed that the solute and solvent molecules do not involve the intramolecular degrees of freedom. The solute molecule is fixed at the (arbitrarily chosen) origin with an (arbitrarily chosen) fixed orientation. In this case, the solute–solvent interaction can be viewed as an external field for the solvent molecules, and the configuration of a solvent molecule relative to the solute molecule is specified completely by the position and orientation of the solvent molecule. In the present paper, the complete set of the position and orientation is called the full coordinate and is denoted collectively by  $\mathbf{x}$ . When the intramolecular degrees of freedom are present in the solute and/or solvent molecule, the extension of our treatments is actually straightforward and is described in Appendix A.

The full coordinate representation is implemented by expressing the distribution functions over the full coordinate  $\mathbf{x}$ . In the full coordinate representation, the instantaneous distribution  $\hat{\rho}^f$  of the solvent is defined as

$$\hat{\rho}^f(\mathbf{x}) = \sum_i \delta(\mathbf{x} - \mathbf{x}_i), \quad (1)$$

where  $\mathbf{x}_i$  is the full coordinate of the  $i$ th solvent molecule and the sum is taken over all the solvent molecules. The average distribution  $\rho^f$  of the solvent is determined when the intermolecular interaction potentials are given and the thermodynamic state is specified. In the present work, the solvent–solvent interaction and the thermodynamic state are fixed, and  $\rho^f$  is treated as a function of the solute–solvent interaction potential. When the solute–solvent interaction potential is  $u$ , the average distribution  $\rho^f$  of the solvent around the solute is uniquely determined from

$$\rho^f(\mathbf{x}; u) = \langle \hat{\rho}^f(\mathbf{x}) \rangle_u, \quad (2)$$

where  $\langle \cdots \rangle_u$  denotes the ensemble average taken in the solution with the solute–solvent interaction  $u$ . We leave  $u$  as an argument in Eq. (2) to specify the solute–solvent interaction potential.

The energetic representation is introduced by adopting the solute–solvent interaction energy as the coordinate of the solvent molecule. To formulate the energetic representation, it is necessary to specify the solute–solvent interaction potential  $v$  with respect to which the solvent coordinate is defined. The natural choice of  $v$  is the interaction potential between the solute and solvent in the solution of interest. In our developments,  $v$  is called the defining potential and is fixed at the outset. The coordinate  $\epsilon$  of a solvent molecule in the energetic representation is simply taken to be the value of  $v$ . This coordinate is specified with respect only to the defining potential  $v$  and its information content is reduced

compared to that of the full coordinate  $\mathbf{x}$ . It should be noted that the coordinate  $\epsilon$  is not necessarily continuous. When  $v$  involves a hard core or square well region, for example,  $\epsilon$  is discrete in that region. To develop the density functional theory in the energetic representation in Sec. II B, we treat a set of solute–solvent interaction potentials  $u$  which are constant over equienergy surfaces of the defining potential  $v$  and may be considered to be defined over the energetic coordinate  $\epsilon$ . For a given solute–solvent interaction potential  $u$  contained in this set, the notations  $u^f(\mathbf{x})$  and  $u^e(\epsilon)$  are adopted, respectively, when it is to be emphasized that  $u$  is represented over the full coordinate  $\mathbf{x}$  and over the energetic coordinate  $\epsilon$ . When no emphasis is placed on the representation, the solute–solvent interaction is simply written as  $u$ . Of course, the defining potential  $v$  satisfies  $v^e(\epsilon) = \epsilon$  by definition.

The instantaneous distribution  $\hat{\rho}^e$  of the solvent is defined in the energetic representation as

$$\hat{\rho}^e(\epsilon) = \int d\mathbf{x} \delta(v^f(\mathbf{x}) - \epsilon) \hat{\rho}^f(\mathbf{x}) = \sum_i \delta(v^f(\mathbf{x}_i) - \epsilon). \quad (3)$$

Equation (3) shows that  $\hat{\rho}^e(\epsilon)d\epsilon$  is equal to the (instantaneous) number of solvent molecules whose values of  $v^f(\mathbf{x})$  are between  $\epsilon$  and  $\epsilon + d\epsilon$ . It should be noted that the definition of  $\hat{\rho}^e$  is dependent on the specification of the defining potential  $v$ . The average distribution  $\rho^e$  of the value of the defining potential  $v$  is correspondingly expressed in the presence of a solute–solvent interaction  $u$  as

$$\rho^e(\epsilon; u) = \langle \hat{\rho}^e(\epsilon) \rangle_u = \int d\mathbf{x} \delta(v^f(\mathbf{x}) - \epsilon) \rho^f(\mathbf{x}; u), \quad (4)$$

where  $\langle \dots \rangle_u$  denotes the ensemble average taken in the solution with the solute–solvent interaction  $u$ . Note that the defining potential  $v$  serves to construct the solvent coordinate  $\epsilon$  and does not identify the system in which the average is taken (unless  $u = v$ ). The solute–solvent interaction potential  $u$ , on the other hand, specifies the solution and determines the ensemble in which the averaging is carried out. We leave  $u$  as an argument in Eq. (4) to identify the solute–solvent interaction potential. Equation (4) shows that the solvent distribution  $\rho^e$  in the energetic representation is obtained by integrating the solvent distribution  $\rho^f$  in the full coordinate representation over equienergy surfaces of the defining potential  $v$ . In this process, a multidimensional integration is performed and the resulting  $\rho^e$  is represented over the one-dimensional coordinate  $\epsilon$  for any type of defining potential  $v$ .

$\rho^e(\epsilon; 0)$  is the solvent distribution in the pure solvent system ( $u = 0$ ). When the pure solvent is homogeneous and isotropic,<sup>36</sup>  $\rho^e(\epsilon; 0)$  is simply the product of the solvent density and the density of states for the defining potential  $v$  given by

$$\int d\mathbf{x} \delta(v^f(\mathbf{x}) - \epsilon). \quad (5)$$

This density of states is determined only by  $v$  and is independent of the solvent–solvent interaction and the thermodynamic state. In the full coordinate representation, the indirect

part  $w^f$  of the potential of mean force between the solute and solvent is defined in the presence of a solute–solvent interaction  $u$  as

$$w^f(\mathbf{x}; u) = -k_B T \log \left( \frac{\rho^f(\mathbf{x}; u)}{\rho^f(\mathbf{x}; 0)} \right) - u^f(\mathbf{x}), \quad (6)$$

where  $k_B$  is the Boltzmann constant,  $T$  is the temperature, and  $\rho^f(\mathbf{x}; 0)$  is the solvent distribution in the pure solvent system ( $u = 0$ ). Similarly, the indirect part  $w^e$  of the solute–solvent potential of mean force is defined in the energetic representation as

$$w^e(\epsilon; u) = -k_B T \log \left( \frac{\rho^e(\epsilon; u)}{\rho^e(\epsilon; 0)} \right) - u^e(\epsilon) \quad (7)$$

for a solute–solvent interaction potential  $u$  which is constant over an equienergy surface of the defining potential  $v$ . Both  $w^f(\mathbf{x}; u)$  and  $w^e(\epsilon; u)$  reflect the many-body effects in the solute–solvent correlation and vanish in the limit of zero solvent density.

It is often the case that the solute–solvent interaction potential of interest is essentially of finite range and that its long-range part may be safely neglected to account for the physics of the solution. When the defining potential  $v$  is zero outside a finite region  $\Omega$ , it is useful to employ a reduced form  $\hat{\rho}_\Omega^e$  of instantaneous distribution of the solvent defined as

$$\hat{\rho}_\Omega^e(\epsilon) = \int_\Omega d\mathbf{x} \delta(v^f(\mathbf{x}) - \epsilon) \hat{\rho}^f(\mathbf{x}) = \sum_{i \in \Omega} \delta(v^f(\mathbf{x}_i) - \epsilon), \quad (8)$$

where the integration is performed only over the interaction region  $\Omega$  and the sum is taken over the solvent molecules contained within  $\Omega$ . Note that since the region  $\Omega$  is usually of molecular size, the number of solvent molecules involved in Eq. (8) is microscopic. For a defining potential  $v$  of finite range,  $\hat{\rho}^e$  in Eq. (3) is actually rewritten as

$$\hat{\rho}^e(\epsilon) = \hat{\rho}_\Omega^e(\epsilon) + \delta(\epsilon)(N - \hat{N}_\Omega), \quad (9)$$

where  $N$  is the total number of solvent molecules in the system and  $\hat{N}_\Omega$  is the (instantaneous) number of solvent molecules in  $\Omega$  expressed as

$$\hat{N}_\Omega = \int_\Omega d\mathbf{x} \hat{\rho}^f(\mathbf{x}) = \int d\epsilon \hat{\rho}_\Omega^e(\epsilon). \quad (10)$$

When a solute–solvent interaction potential  $u$  is given, the average form  $\rho_\Omega^e$  is simply determined from

$$\rho_\Omega^e(\epsilon; u) = \langle \hat{\rho}_\Omega^e(\epsilon) \rangle_u. \quad (11)$$

It is then easy to see, in correspondence with Eq. (9), that  $\rho_\Omega^e$  satisfies<sup>37</sup>

$$\rho^e(\epsilon; u) = \rho_\Omega^e(\epsilon; u) + \delta(\epsilon)(N - \langle \hat{N}_\Omega \rangle_u). \quad (12)$$

The definition of the indirect part  $w_\Omega^e$  of the solute–solvent potential of mean force is similar to Eq. (7) and is expressed as

$$w_\Omega^e(\epsilon; u) = -k_B T \log \left( \frac{\rho_\Omega^e(\epsilon; u)}{\rho_\Omega^e(\epsilon; 0)} \right) - u^e(\epsilon). \quad (13)$$

## B. Density functional theory

In the full coordinate representation, Eq. (2) defines a map from a solute–solvent interaction potential  $u^f(\mathbf{x})$  to a distribution function  $\rho^f(\mathbf{x};u)$ . As is well known, the map is one-to-one in the sense that different  $u^f(\mathbf{x})$  give different  $\rho^f(\mathbf{x};u)$ .<sup>2,38–41</sup> This is actually the fundamental theorem of the density functional theory and provides the basis for Percus' method of functional expansion to derive the Percus–Yevick and hypernetted-chain approximations.<sup>2,6</sup>

In order to develop the density functional theory in the energetic representation, it is necessary to restrict the set of solute–solvent interaction potentials. As shown in Appendix B, the suitable set corresponding to the defining potential  $v$  of interest consists of the solute–solvent interaction potentials which are defined over the energetic coordinate  $\epsilon$ . A potential function  $u$  contained in this set is constant over an equienergy surface of the defining potential  $v$  and may be expressed as  $u^e(\epsilon)$ . Equation (4) then defines a map from the set of potential functions  $u^e(\epsilon)$  to a set of distribution functions  $\rho^e(\epsilon;u)$  expressed in the energetic representation. By definition, Eq. (4) maps an element in the set of  $u^e(\epsilon)$  uniquely to an element in the set of  $\rho^e$ . In Appendix B, on the other hand, we show that the converse is also true. Different elements in the set of  $u^e(\epsilon)$  are mapped through Eq. (4) to different elements in the set of  $\rho^e(\epsilon;u)$ . In other words, the correspondence is one-to-one between the set of  $u^e(\epsilon)$  and the set of  $\rho^e$  generated by the map Eq. (4), and it is possible to treat a functional of  $u^e(\epsilon)$  as a functional of  $\rho^e(\epsilon;u)$ . In Sec. III A, this property is exploited to derive approximate integral equations in the energetic representation.

When the defining potential  $v$  is of finite range and vanishes outside a finite region  $\Omega$ , the density functional theory can be developed on the set of solute–solvent interaction potentials which are constant over equienergy surfaces of  $v$  and are zero outside the interaction range  $\Omega$ . As shown in Appendix B, the map is one-to-one from this set to the set of distribution functions  $\rho_{\Omega}^e(\epsilon;u)$  generated by Eq. (11). The present property is useful for a defining potential of finite range because the solvent molecules outside the interaction range do not have to be taken into account in a density-functional treatment.

## C. Ornstein–Zernike equation

In commonly used approaches to a solution system, the Ornstein–Zernike equation or its site–site form is employed to introduce the direct correlation function and a closure relationship is adopted to give a self-consistent integral equation for pair correlation functions.<sup>2,4–7,10–28</sup> Since the direct correlation function is useful, at least notationally, to develop an approximate integral equation, it is desirable to define the direct correlation function in the energetic representation. In this section, we introduce the solute–solvent direct correlation function in the energetic representation and formulate the Ornstein–Zernike equation for a dilute solution over the energetic coordinate  $\epsilon$ .

To formulate the Ornstein–Zernike equation for a dilute solution of our interest, it is necessary to define the two-body

correlation function of the solvent.<sup>42</sup> In the full coordinate representation, the two-body correlation function  $\chi^f$  under the presence of a solute–solvent interaction potential  $u$  is expressed as

$$\chi^f(\mathbf{x}, \mathbf{y}; u) = \frac{\delta \rho^f(\mathbf{x}; u)}{\delta(-\beta u^f(\mathbf{y}))} = \langle \hat{\rho}^f(\mathbf{x}) \hat{\rho}^f(\mathbf{y}) \rangle_u - \langle \hat{\rho}^f(\mathbf{x}) \rangle_u \langle \hat{\rho}^f(\mathbf{y}) \rangle_u, \quad (14)$$

where  $\beta$  is the inverse of the thermal energy ( $k_B T$ ) and  $u$  is left as an argument to specify the solute–solvent interaction. The two-body correlation function  $\chi^e$  in the energetic representation is similarly defined as

$$\chi^e(\epsilon, \eta; u) = \frac{\delta \rho^e(\epsilon; u)}{\delta(-\beta u^e(\eta))} = \langle \hat{\rho}^e(\epsilon) \hat{\rho}^e(\eta) \rangle_u - \langle \hat{\rho}^e(\epsilon) \rangle_u \langle \hat{\rho}^e(\eta) \rangle_u. \quad (15)$$

It is obvious that both of  $\chi^f$  and  $\chi^e$  are positive definite and invertible and are symmetric with respect to the two arguments of the solvent coordinate. In the full coordinate representation, the direct correlation function  $c^f$  between the solute and solvent in a dilute solution satisfies

$$\rho^f(\mathbf{x}; u) - \rho^f(\mathbf{x}; 0) = \int d\mathbf{y} c^f(\mathbf{y}; u) \chi^f(\mathbf{y}, \mathbf{x}; 0), \quad (16)$$

where  $u$  appears as an argument to identify the solute–solvent interaction and  $\chi^f(\mathbf{x}, \mathbf{y}; 0)$  is the two-body correlation function in the pure solvent system ( $u=0$ ). The direct correlation function  $c^e$  in the energetic representation can then be defined by writing an expression similar in form to Eq. (16) as

$$\rho^e(\epsilon; u) - \rho^e(\epsilon; 0) = \int d\eta c^e(\eta; u) \chi^e(\eta, \epsilon; 0). \quad (17)$$

Since the two-body correlation function  $\chi^e(\epsilon, \eta; 0)$  in the pure solvent is invertible, Eq. (17) defines  $c^e$  over the coordinate  $\epsilon$  in terms of the average distributions  $\rho^e$  and  $\chi^e$ . The structure of Eq. (17) is similar to that of the Ornstein–Zernike equation for a dilute solution in the full coordinate representation given by Eq. (16). Thus, we call Eq. (17) the Ornstein–Zernike equation for a dilute solution in the energetic representation.

In the full coordinate representation, it is a common practice to introduce the total correlation function  $h^f$  of the solvent by<sup>2</sup>

$$\chi^f(\mathbf{x}, \mathbf{y}; u) = \rho^f(\mathbf{x}; u) \delta(\mathbf{x} - \mathbf{y}) + \rho^f(\mathbf{x}; u) \rho^f(\mathbf{y}; u) h^f(\mathbf{x}, \mathbf{y}; u). \quad (18)$$

The Ornstein–Zernike relation Eq. (16) is then rewritten in a familiar form as

$$\rho^f(\mathbf{x}; u) = \rho^f(\mathbf{x}; 0) \left[ 1 + c^f(\mathbf{x}; u) + \int d\mathbf{y} c^f(\mathbf{y}; u) \rho^f(\mathbf{y}; 0) h^f(\mathbf{y}, \mathbf{x}; 0) \right]. \quad (19)$$

In the energetic representation, it is also useful to define the total correlation function  $h^e$  of the solvent. The definition of  $h^e$  is similar in form to Eq. (18) and is expressed as

$$\chi^e(\epsilon, \eta; u) = \rho^e(\epsilon; u) \delta(\epsilon - \eta) + \rho^e(\epsilon; u) \rho^e(\eta; u) h^e(\epsilon, \eta; u). \quad (20)$$

In this case, the Ornstein–Zernike relation Eq. (17) is equivalent to

$$\rho^e(\epsilon; u) = \rho^e(\epsilon; 0) \left[ 1 + c^e(\epsilon; u) + \int d\eta c^e(\eta; u) \rho^e(\eta; 0) h^e(\eta, \epsilon; 0) \right]. \quad (21)$$

Thus, the introduction of the direct correlation function and the formulation of the Ornstein–Zernike equation in the energetic representation can be made in parallel to those in the full coordinate representation.

When the defining potential  $v$  is zero outside a finite region  $\Omega$ , it is useful to employ a reduced form  $\chi_\Omega^e$  of two-body correlation function of the solvent defined as

$$\chi_\Omega^e(\epsilon, \eta; u) = \langle \hat{\rho}_\Omega^e(\epsilon) \hat{\rho}_\Omega^e(\eta) \rangle_u - \langle \hat{\rho}_\Omega^e(\epsilon) \rangle_u \langle \hat{\rho}_\Omega^e(\eta) \rangle_u. \quad (22)$$

Using  $\chi_\Omega^e$  and the reduced form  $\rho_\Omega^e$  of average distribution given by Eq. (11), the corresponding direct correlation function  $c_\Omega^e$  is introduced by

$$\rho_\Omega^e(\epsilon; u) - \rho_\Omega^e(\epsilon; 0) = \int d\eta c_\Omega^e(\eta; u) \chi_\Omega^e(\eta, \epsilon; 0). \quad (23)$$

This form is useful for a defining potential  $v$  of finite range since it does not involve the  $\delta(\epsilon)$  singularity illustrated in Eq. (12) for  $\rho^e$ . Equation (23) is actually the Ornstein–Zernike equation and is rewritten as

$$\rho_\Omega^e(\epsilon; u) = \rho_\Omega^e(\epsilon; 0) \left[ 1 + c_\Omega^e(\epsilon; u) + \int d\eta c_\Omega^e(\eta; u) \rho_\Omega^e(\eta; 0) h_\Omega^e(\eta, \epsilon; 0) \right], \quad (24)$$

where the total correlation function  $h_\Omega^e$  is defined as

$$\chi_\Omega^e(\epsilon, \eta; u) = \rho_\Omega^e(\epsilon; u) \delta(\epsilon - \eta) + \rho_\Omega^e(\epsilon; u) \rho_\Omega^e(\eta; u) h_\Omega^e(\epsilon, \eta; u). \quad (25)$$

### III. APPROXIMATIONS

#### A. Integral equation

In an exact formulation of a solution, a correlation function of interest is related to the higher-order correlation functions through a hierarchical set of integral equations.<sup>2</sup> Since this exact set of equations simply connects the correlation function in question to another unknown correlation functions, an approximation needs to be introduced to close the set of equations and obtain a self-consistent equation for the correlation function of interest. The method of integral equation provides, commonly at the two-body level, an approximate and solvable equation for a correlation function of interest under a given set of intermolecular interaction

potentials. In this section, we formulate the method of integral equation in the energetic representation with the Percus–Yevick (PY), hypernetted-chain (HNC), and superposition approximations.

Before presenting the approximate integral equations in the energetic representation, we briefly review the PY and HNC integral equations in the full coordinate representation using the notations introduced in Sec. II. In the full coordinate representation, a systematic route to formulating the PY and HNC integral equations is Percus' method of functional expansion.<sup>2,6</sup> This method is based on the property that the correspondence is one-to-one between the solute–solvent interaction potential  $u^f(\mathbf{x})$  and the distribution function  $\rho^f(\mathbf{x}; u)$  generated by Eq. (2). For a given solute–solvent interaction potential  $v$  of interest, the PY and HNC integral equations are obtained by expanding  $\rho^f(\mathbf{x}; v) \exp(\beta v^f(\mathbf{x}))$  and  $(\log \rho^f(\mathbf{x}; v) + \beta v^f(\mathbf{x}))$  to first-order, respectively, in terms of  $(\rho^f(\mathbf{x}; v) - \rho^f(\mathbf{x}; 0))$ , where  $\rho^f(\mathbf{x}; 0)$  is the distribution function in the pure solvent system. By using the direct correlation function  $c^f$  introduced by the Ornstein–Zernike relation Eq. (16) and the indirect part  $w^f$  of the solute–solvent potential of mean force defined by Eq. (6), the PY and HNC approximations in the full coordinate representation are expressed, respectively, as

$$c^f(\mathbf{x}; v) = (1 - \exp(\beta v^f(\mathbf{x}))) \frac{\rho^f(\mathbf{x}; v)}{\rho^f(\mathbf{x}; 0)} \quad (26)$$

and

$$c^f(\mathbf{x}; v) = \frac{\rho^f(\mathbf{x}; v)}{\rho^f(\mathbf{x}; 0)} - 1 + \beta w^f(\mathbf{x}; v). \quad (27)$$

When the solute–solvent interaction  $v$  is small, Eq. (16) reduces in combination with Eq. (26) or (27) to the standard result of the first-order perturbation theory written as

$$\rho^f(\mathbf{x}; v) = \exp(-\beta v^f(\mathbf{x})) \rho^f(\mathbf{x}; 0) \times \left[ 1 + \int d\mathbf{y} f(\mathbf{y}) \rho^f(\mathbf{y}; 0) h^f(\mathbf{y}, \mathbf{x}; 0) \right] + O(f^2), \quad (28)$$

where  $f$  is the Mayer function for the solute–solvent interaction defined as

$$f = \exp(-\beta v) - 1 \quad (29)$$

and  $O(f^2)$  denotes a second- or higher-order term with respect to  $f$ . Equation (28) shows that the PY and HNC approximations are exact to first-order in the solute–solvent interaction  $v$ . Furthermore, the solute-induced modification of the solution structure expressed by  $w^f$  in Eq. (6) is given correctly to first-order with respect to the solvent density in the PY and HNC approximations.

In order to formulate the PY and HNC approximations in the energetic representation, we also employ the method of functional expansion. For a given solute–solvent interaction potential  $v$  of interest, the energetic coordinate  $\epsilon$  can be constructed by taking  $v$  as the defining potential. It is then stated in Sec. II B that the correspondence is one-to-one between the set of potential functions  $u^e(\epsilon)$  defined over the coordinate  $\epsilon$  and the set of distribution functions  $\rho^e(\epsilon; u)$  generated by the map Eq. (4). In this case, it is possible to

convert an expansion in terms of  $u^e(\epsilon)$  into an expansion in terms of  $(\rho^e(\epsilon;u) - \rho^e(\epsilon;0))$ , where  $\rho^e(\epsilon;0)$  is the distribution function in the pure solvent system ( $u=0$ ). The PY approximation in the energetic representation can be formulated by expanding  $\rho^e(\epsilon;u)\exp(\beta u^e(\epsilon))$  to first-order in terms of  $(\rho^e(\epsilon;u) - \rho^e(\epsilon;0))$  and setting  $u=v$ . Similarly, the HNC approximation is provided by expanding  $(\log \rho^e(\epsilon;u) + \beta u^e(\epsilon))$  and setting  $u=v$ . When the direct correlation function  $c^e$  introduced by the Ornstein–Zernike relation Eq. (17) and the indirect part  $w^e$  of the solute–solvent potential of mean force defined by Eq. (7) are used, the PY and HNC approximations in the energetic representation are written, respectively, as

$$c^e(\epsilon;v) = (1 - \exp(\beta\epsilon)) \frac{\rho^e(\epsilon;v)}{\rho^e(\epsilon;0)} \quad (30)$$

and

$$c^e(\epsilon;v) = \frac{\rho^e(\epsilon;v)}{\rho^e(\epsilon;0)} - 1 + \beta w^e(\epsilon;v), \quad (31)$$

where the property is used that the solute–solvent interaction  $v$  of interest is the defining potential and satisfies  $v^e(\epsilon) = \epsilon$  over the energetic coordinate. In combination with the Ornstein–Zernike relation Eq. (17), Eq. (30) or (31) constitutes a self-consistent integral equation for the solvent distribution around the solute in the energetic representation. It should be noted that although the integral equations in the energetic representation are similar in form to those in the full coordinate representation, they are formulated over a one-dimensional coordinate specifying the solute–solvent interaction energy. When the solute–solvent interaction  $v$  is small, both the PY and HNC integral equations in the energetic representation reduce to

$$\begin{aligned} \rho^e(\epsilon;v) = & \exp(-\beta\epsilon)\rho^e(\epsilon;0) \\ & \times \left[ 1 + \int d\eta(\exp(-\beta\eta) - 1) \rho^e(\eta;0) h^e(\eta, \epsilon;0) \right] \\ & + O(v^2). \end{aligned} \quad (32)$$

This equation can actually be obtained by integrating Eq. (28) over the equienergy surface given by  $\epsilon = v^f(\mathbf{x})$ . Thus, the PY and HNC approximations in the energetic representation are exact to first-order in the solute–solvent interaction  $v$ . In addition, the solute-induced effect expressed by  $w^e$  in Eq. (7) is provided correctly to first-order with respect to the solvent density in the PY and HNC approximations in the energetic representation.

In order to formulate the superposition approximation in the energetic representation, we adopt the approach presented in Ref. 8. In this approach, no derivative of the solute–solvent interaction potential is involved and the integral equation will be easier to handle than the Born–Green equation. Using the total correlation function  $h^e$  of the solvent defined by Eq. (20), the superposition approximation is expressed as

$$h^e(\epsilon, \eta;u) = h^e(\epsilon, \eta;0) \quad (33)$$

when a solute–solvent interaction  $u$  is given. In other words, the superposition approximation Eq. (33) states that the total correlation function under the presence of a solute–solvent

interaction is the same as that in the pure solvent system ( $u=0$ ). The derivation of the integral equation in the energetic representation is parallel to that in the full coordinate representation described in Ref. 8. The integral equation is then expressed in terms of the indirect part  $w^e$  of the solute–solvent potential of mean force as

$$\begin{aligned} w^e(\epsilon;v) = & -k_B T \int d\eta(\exp(-\beta\eta) - 1) \\ & \times \frac{\beta w^e(\eta;v)}{\exp(\beta w^e(\eta;v)) - 1} \rho^e(\eta;0) h^e(\eta, \epsilon;0). \end{aligned} \quad (34)$$

When the solute–solvent interaction  $v$  is small, Eq. (34) reduces to Eq. (32). Therefore, the superposition approximation in the energetic representation provides the exact  $w^e$  to first-order with respect to the solute–solvent interaction  $v$  and to the solvent density.

## B. Chemical potential

It is a principle of statistical thermodynamics that the stability of a solute in solution is determined by its chemical potential. Indeed, the equilibrium and rate constants of a chemical reaction in a solution are governed by the chemical potentials of the solute species involved in the reaction. In this section, we relate the chemical potential of a solute in closed form to the distribution functions of the solvent in the energetic representation by employing the PY, HNC, and superposition approximations.

In order to treat the thermodynamics of solvation on a definite basis, it is necessary to specify the solution process precisely. We consider the insertion process of the solute at the (arbitrarily chosen) fixed origin with an (arbitrarily chosen) fixed orientation. The free energy change for this process is the excess chemical potential  $\Delta\mu$  of the solute.<sup>2</sup>  $\Delta\mu$  involves only the contribution from the potential energy, and the ideal (translational) contribution is excluded at the outset. In the insertion process, the solute–solvent interaction is gradually turned on and its extent is described by the coupling parameter  $\lambda$  of the solute–solvent interaction. When  $\lambda=0$ , there is no interaction between the solute and solvent and the system is the pure solvent. When  $\lambda=1$ , the solute interacts with the solvent at full coupling under the solute–solvent interaction potential  $v$  of interest. When  $0 < \lambda < 1$ , the system is at an intermediate state between the pure solvent and the solution with the fully coupled solute–solvent interaction.

As done in Sec. III A, the energetic representation is implemented for the solute–solvent interaction potential  $v$  of interest by adopting  $v$  as the defining potential. It is then possible to introduce a family of solute–solvent interaction potentials  $u_\lambda$  defined over the coupling parameter  $\lambda$  and the energetic coordinate  $\epsilon$ . Of course, although  $u_\lambda$  may be chosen arbitrarily at the intermediate states  $0 < \lambda < 1$ , it needs to be imposed that

$$\begin{aligned} u_0^e(\epsilon) &= 0, \\ u_1^e(\epsilon) &= v^e(\epsilon) = \epsilon. \end{aligned} \quad (35)$$

In this case, the average distribution  $\rho^e(\epsilon; u_\lambda)$  of the solvent at the coupling parameter  $\lambda$  is given by Eq. (4) and is related to the excess chemical potential  $\Delta\mu$  of the solute of interest through

$$\Delta\mu = \int_0^1 d\lambda \int d\epsilon \frac{\partial u_\lambda^e(\epsilon)}{\partial \lambda} \rho^e(\epsilon; u_\lambda). \quad (36)$$

Equation (36) is the charging formula in the energetic representation, and the value of  $\Delta\mu$  determined from Eq. (36) does not depend on the choice of the intermediate states in an exact treatment.

When the analytical integration over the coupling parameter is not attainable in Eq. (36), the determination of the excess chemical potential  $\Delta\mu$  requires a procedure to explicitly treat the intermediate states of the coupling parameter  $\lambda$ . It should be noted, however, that the intermediate states are not of physical interest. Thus, an explicit treatment of the intermediate states is not only computationally expensive, but is also undesirable for the physical understanding of the chemical potential. When the  $\lambda$  integration in Eq. (36) is performed,  $\Delta\mu$  is expressed only in terms of distribution functions in the pure solvent ( $\lambda=0$ ) and in the solution with the fully coupled solute–solvent interaction ( $\lambda=1$ ).

In the PY, HNC, and superposition approximations, the integration over the coupling parameter may be analytically performed in the energetic representation and  $\Delta\mu$  can be expressed in terms of  $\rho^e$ ,  $w^e$ , and  $c^e$  defined by Eqs. (4), (7), and (17), respectively. As shown in Appendix C,  $\Delta\mu$  is given in the PY and HNC approximations, respectively, by

$$\Delta\mu = -k_B T \int d\epsilon \rho^e(\epsilon; 0) (\exp(-\beta\epsilon) - 1) \times \frac{\beta w^e(\epsilon; v)}{\exp(\beta w^e(\epsilon; v)) - 1} \quad (37)$$

and

$$\Delta\mu = -k_B T \int d\epsilon \rho^e(\epsilon; 0) [c^e(\epsilon; v) + \frac{1}{2}\beta w^e(\epsilon; v) \times (c^e(\epsilon; v) - \beta w^e(\epsilon; v))]. \quad (38)$$

In the superposition approximation, the derivation of  $\Delta\mu$  is parallel to that in the full coordinate representation described in Ref. 8 and leads explicitly to Eq. (37).<sup>43</sup>

When the solute–solvent interaction potential  $v$  of interest is small, both Eqs. (37) and (38) reduce to a result of the second-order perturbation theory written as

$$\Delta\mu = -k_B T \left[ \int d\epsilon \rho^e(\epsilon; 0) (\exp(-\beta\epsilon) - 1) + \frac{1}{2} \int d\epsilon d\eta \rho^e(\epsilon; 0) \rho^e(\eta; 0) (\exp(-\beta\epsilon) - 1) \times (\exp(-\beta\eta) - 1) h^e(\epsilon, \eta; 0) \right] + O(f^3), \quad (39)$$

where  $f$  is the Mayer function defined by Eq. (29). Equation (39) shows for the PY, HNC, and superposition approximations in the energetic representation that the excess chemical

potential  $\Delta\mu$  is given exactly to second-order in the Mayer function for the solute–solvent interaction  $v$  of interest. This is actually due to the property shown in Eq. (32) that  $w^e$  obtained from the integral equation in the PY, HNC, or superposition approximation involves the correct first-order term with respect to  $f$ . Furthermore, the excess chemical potential  $\Delta\mu$  is exact to second-order in the solvent density when the PY, HNC, or superposition approximation is employed in the energetic representation. Therefore, the solvation thermodynamics of a solute is expected to be reproduced accurately in a low- to medium-density fluid when the solution is studied with the integral equation in the energetic representation.

When the solute–solvent interaction potential  $v$  of interest is zero outside a finite region  $\Omega$ , it is useful to employ the reduced form of functions  $\rho_\Omega^e$ ,  $w_\Omega^e$ ,  $c_\Omega^e$ , and  $h_\Omega^e$  defined by Eqs. (11), (13), (23), and (25), respectively. In this case, the PY, HNC, and superposition approximations are obtained simply by replacing  $\rho^e$ ,  $w^e$ ,  $c^e$ , and  $h^e$  with the corresponding variables involving the subscript  $\Omega$  in Eqs. (30), (31), (34), (37), and (38). In other words, the Ornstein–Zernike equation is Eq. (23) and the PY and HNC approximations are written, respectively, as

$$c_\Omega^e(\epsilon; v) = (1 - \exp(\beta\epsilon)) \frac{\rho_\Omega^e(\epsilon; v)}{\rho_\Omega^e(\epsilon; 0)} \quad (40)$$

and

$$c_\Omega^e(\epsilon; v) = \frac{\rho_\Omega^e(\epsilon; v)}{\rho_\Omega^e(\epsilon; 0)} - 1 + \beta w_\Omega^e(\epsilon; v). \quad (41)$$

The expressions for the excess chemical potential  $\Delta\mu$  in the PY and HNC closures are further given, respectively, by

$$\Delta\mu = -k_B T \int d\epsilon \rho_\Omega^e(\epsilon; 0) (\exp(-\beta\epsilon) - 1) \times \frac{\beta w_\Omega^e(\epsilon; v)}{\exp(\beta w_\Omega^e(\epsilon; v)) - 1} \quad (42)$$

and

$$\Delta\mu = -k_B T \int d\epsilon \rho_\Omega^e(\epsilon; 0) [c_\Omega^e(\epsilon; v) + \frac{1}{2}\beta w_\Omega^e(\epsilon; v) (c_\Omega^e(\epsilon; v) - \beta w_\Omega^e(\epsilon; v))]. \quad (43)$$

On the other hand, the integral equation in the superposition approximation takes the form that

$$w_\Omega^e(\epsilon; v) = -k_B T \int d\eta (\exp(-\beta\eta) - 1) \times \frac{\beta w_\Omega^e(\eta; v)}{\exp(\beta w_\Omega^e(\eta; v)) - 1} \rho_\Omega^e(\eta; 0) h_\Omega^e(\eta, \epsilon; 0), \quad (44)$$

and the corresponding expression for  $\Delta\mu$  is Eq. (42).

#### IV. DISCUSSION

This paper has presented a method of distribution function for describing a dilute solution over a one-dimensional

coordinate specifying the solute–solvent interaction energy. The one-to-one correspondence is then established between the set of solute–solvent interaction potentials defined over the energetic coordinate and the set of solvent distribution functions around the solute in the energetic representation. On the basis of the one-to-one correspondence, the integral equations with the Percus–Yevick (PY) and hypernetted-chain (HNC) approximations can be formulated over the energetic coordinate through the method of functional expansion. The integral equation may also be formulated in the superposition approximation in a manner similar to that presented in Ref. 8. The PY, HNC, and superposition approximations in the energetic representation determine the solvent distribution functions around the solute exactly to first-order with respect to the solute–solvent interaction potential and to the solvent density. Correspondingly, the excess chemical potential of the solute is given in these approximations correctly to second-order in the solute–solvent interaction potential and in the solvent density.

In order to solve the integral equation for a dilute solution with the PY, HNC, or superposition approximation in the energetic representation, it is necessary to treat the two-body correlation function  $\chi^e(\epsilon, \eta; 0)$  in the pure solvent system as a known input. Actually,  $\chi^e(\epsilon, \eta; 0)$  can be readily obtained by performing a computer simulation of the pure solvent system. Indeed, the instantaneous distribution  $\hat{\rho}^e(\epsilon)$  defined by Eq. (3) is to be evaluated by placing the solute molecule as a test particle and is to be averaged through Eq. (15) over the configurations generated. Therefore, as far as the pure solvent system is easily simulated, the approximate procedures developed in the present paper are useful routes to assessing the behavior of the solute at infinite dilution.

When the solute–solvent interaction potential depends only on the position of a single site in the solvent molecule relative to the solute molecule, the radial distribution function of that site in the pure solvent may be utilized to construct the two-body correlation function  $\chi^e(\epsilon, \eta; 0)$ . In Appendix D, we treat this specific case and provide an explicit expression for  $\chi^e(\epsilon, \eta; 0)$  in terms of the radial distribution function. Furthermore, when the solute–solvent interaction potential is constant over its interaction range, the PY, HNC, and superposition approximations are particularly simple in the energetic representation. In Appendix E, we investigate this simple case and present closed form expressions for the excess chemical potential.

The integral equations presented in this paper do not directly provide the potential of mean force between two solute molecules. The evaluation of the potential of mean force at a given configuration is readily possible, however, by noting that it is the difference between the free energy changes of the system upon insertion of the two solute molecules at infinite separation and upon insertion at the configuration of interest. The free energy change upon insertion at infinite separation is simply the sum of the individual excess chemical potentials of the two solute molecules. The free energy change upon insertion at the configuration of interest may be evaluated, on the other hand, by treating the two solute molecules at that configuration as one unit and determining the “excess chemical potential” of the unit. In

this procedure, since the excess chemical potentials are given exactly to second-order in the solute–solvent interaction potential and in the solvent density, the potential of mean force is also exact to second-order.

The extension of the procedures in the present paper to a mixed solvent system is straightforward. In this case, the energetic coordinate  $\epsilon_\alpha$  is constructed for the solvent species  $\alpha$  from the interaction potential between the solute and the solvent species  $\alpha$ . The integral equations and the expressions for the excess chemical potential may then be reformulated simply by adding the subscripts representing the solvent species to the energetic coordinates and supplementing the integrals over the energetic coordinates with the sums over the solvent species.

An accurate and efficient route to the excess chemical potential of a hard core solute is provided by the information-theoretic approach of Hummer *et al.*<sup>44,45</sup> In this approach, the distribution of the number of solvent molecules in the hard core region is determined from its first and second cumulants in the pure solvent through the maximization of a properly defined information entropy, and the probability that the number is equal to zero gives the excess chemical potential of the hard core solute. The information-theoretic approach may be viewed as a method of evaluating the excess chemical potential of a hard core solute in the energetic representation since the hard core region is an equienergy surface with the infinite solute–solvent interaction energy. When the solute–solvent interaction of interest is soft and the corresponding coordinate in the energetic representation is continuous over its whole range, the method of entropy maximization leads to an energetic version of the Gaussian field model elaborated by Chandler.<sup>46</sup> As pointed out by Hummer *et al.*, however, the Gaussian field model does not satisfy the condition that the (instantaneous) solvent distribution is non-negative.<sup>44</sup> Moreover, when the method of entropy maximization is modified by imposing the condition of non-negativity, it is not tractable in practice to evaluate the excess chemical potential of a solute. This is because the evaluation requires an integration over the (fluctuating) solvent distribution within the condition of non-negativity and the solvent distribution involves a continuous coordinate as an argument. Therefore, the method of entropy maximization has not been adopted in the present work to derive an approximate expression for the excess chemical potential of a solute in the energetic representation.

When the intermolecular interaction is of the site–site form, the method of reference interaction site model (RISM) is a convenient route to the solution structure and is particularly useful in a high-density molecular fluid.<sup>12–23</sup> The RISM integral equations with the PY and HNC closures, however, exhibit unphysical dependence on “auxiliary” sites which simply label points in a molecule and make no contribution to the intermolecular interaction.<sup>15</sup> Since the integral equations in the energetic representation do not separately treat the distinct sites of a molecule, in contrast, they are not affected by the presence of “auxiliary” sites.<sup>47</sup> In addition, the RISM integral equations are not exact in the limit of zero solvent density and are not useful to evaluate the excess chemical potential of a solute in a low-density



fluid. The RISM-2 integral equations provide improved descriptions in the low-density regime, but determine the excess chemical potential correctly only to first-order in the density.<sup>26–28</sup> Therefore, since the PY, HNC, and superposition approximations in the energetic representation give the excess chemical potential of a solute exactly to second-order in the solvent density, they will be suitable for describing a low- to medium-density fluid. In subsequent work, the approximate procedures developed in this paper are applied to the solvation thermodynamics and solvent-mediated interactions of various types of solutes in water over a wide range of thermodynamic conditions.

## ACKNOWLEDGMENTS

N.M. is grateful to the Research Grant-in-Aid from the Ministry of Education, Science, and Culture (No. 11740322) and to the Supercomputer Laboratory of Institute for Chemical Research, Kyoto University, for generous allocation of computation time. In addition, this work is supported by the Research Grant-in-Aid from the Ministry of Education, Science, and Culture (No. 10304047) and by CREST (Core Research for Evolutional Science and Technology) of Japan Science and Technology Corporation (JST).

## APPENDIX A

In Sec. II A, it has been supposed for the sake of simplicity that the solute and solvent molecules do not involve the intramolecular degrees of freedom such as the molecular vibrations and intramolecular polarization. This supposition can actually be removed in the method of distribution function in the energetic representation by changing the content of the full coordinate  $\mathbf{x}$ . Let  $\psi$  and  $\gamma$  be the complete sets of variables specifying the configurations of the solute and solvent molecules, respectively.  $\psi$  and  $\gamma$  contain the intramolecular degrees of freedom of the solute and solvent molecules, respectively. The developments in the present paper is then valid simply by redefining the full coordinate  $\mathbf{x}$  as a collective set of  $\psi$  and  $\gamma$  and using an appropriate expression for the average of a given quantity  $Q$ .

The solute–solvent interaction potential  $u$  is a function of  $\psi$  and  $\gamma$  and may be expressed as  $u^f(\psi, \gamma)$  in the full coordinate representation. When the intramolecular energy of the solute is  $\phi(\psi)$  and the solvent–solvent interaction energy is  $U$ , the average  $\langle Q \rangle_u$  of a quantity  $Q$  is expressed in the presence of the solute–solvent interaction  $u$  as

$$\langle Q \rangle_u = \frac{\int d\psi d\Gamma Q \exp(-\beta\{\phi(\psi) + \sum_i u^f(\psi, \gamma_i) + U\})}{\int d\psi d\Gamma \exp(-\beta\{\phi(\psi) + \sum_i u^f(\psi, \gamma_i) + U\})}, \quad (\text{A1})$$

where  $\gamma_i$  is the variable to specify the configuration of the  $i$ th solvent molecule and  $\Gamma$  represents the solvent configuration collectively. The corresponding expression for the excess chemical potential  $\Delta\mu$  of the solute is given by

$$\exp(-\beta\Delta\mu) = \frac{\int d\psi d\Gamma \exp(-\beta\{\phi(\psi) + \sum_i v^f(\psi, \gamma_i) + U\})}{\int d\psi d\Gamma \exp(-\beta\{\phi(\psi) + U\})} \quad (\text{A2})$$

when the solute–solvent interaction is  $v^f(\psi, \gamma)$  in the full coordinate representation. It is then possible to show, by adopting Eqs. (A1) and (A2), that the method of distribution function in the energetic representation developed in this paper is valid by taking  $\mathbf{x}=(\psi, \gamma)$ .

In Secs. II and III, we have fixed the solute molecule at the (arbitrarily chosen) origin with an (arbitrarily chosen) fixed orientation. This is also unnecessary when the position and orientation of the solute molecule are incorporated into the coordinate  $\psi$  and Eqs. (A1) and (A2) are employed.

## APPENDIX B

The method of functional expansion exploited in Sec. III A to develop the PY and HNC approximations is based on the property that the correspondence is one-to-one between the set of solute–solvent interaction potentials and the set of distribution functions of the solvent. The purpose of this Appendix is to prove the validity of the one-to-one correspondence in the energetic representation when the set of solute–solvent interaction potentials is properly chosen. Our procedure of the proof is parallel to that described by Hansen and McDonald in their textbook.<sup>2</sup> We present the proof only in the canonical ensemble since the extension to the other ensembles is straightforward.

When the solute–solvent interaction  $u$  is given, we define a functional  $Z[P; u]$  of the probability distribution function  $P$  of the solvent in the configuration space as

$$Z[P; u] = \int d\Gamma \left[ \sum_i u^f(\mathbf{x}_i) + U + k_B T \log P(\Gamma) \right] P(\Gamma), \quad (\text{B1})$$

where  $\Gamma$  represents the solvent configuration collectively and  $U$  denotes the solvent–solvent interaction energy. It is then possible to show that

$$Z[P; u] \geq Z[P_u^0; u], \quad (\text{B2})$$

where  $P_u^0$  is the equilibrium distribution function given by

$$P_u^0(\Gamma) = \frac{\exp(-\beta\{\sum_i u^f(\mathbf{x}_i) + U\})}{\int d\Gamma \exp(-\beta\{\sum_i u^f(\mathbf{x}_i) + U\})}. \quad (\text{B3})$$

The equality holds in Eq. (B2) only when  $P = P_u^0$ .

Suppose that the average distributions given by Eq. (2) are identical between two distinct solute–solvent interactions  $u$  and  $w$  and that

$$\rho^f(\mathbf{x}; u) = \rho^f(\mathbf{x}; w) \quad (\text{B4})$$

is valid. It then follows from Eqs. (B1) and (B2) that

$$Z[P_w^0; w] < Z[P_u^0; u] + \int d\mathbf{x} \rho^f(\mathbf{x}; u) (w^f(\mathbf{x}) - u^f(\mathbf{x})), \quad (\text{B5})$$

$$Z[P_u^0; u] < Z[P_w^0; w] + \int d\mathbf{x} \rho^f(\mathbf{x}; u) (u^f(\mathbf{x}) - w^f(\mathbf{x})),$$

where the functional  $Z[P; w]$  and the equilibrium distribution function  $P_w^0$  are given by expressions similar to Eqs. (B1) and (B3), respectively. Actually, since the two inequalities in Eq. (B5) are not compatible with each other, the supposition Eq. (B4) needs to be negated. Thus, the one-to-one

correspondence is established between the solute–solvent interaction potential and the average distribution of the solvent in the full coordinate representation.

We now introduce the energetic representation by adopting the solute–solvent interaction potential  $v$  of interest as the defining potential and constructing the energetic coordinate  $\epsilon$  with respect to  $v$ . We restrict the set of solute–solvent interaction potentials and consider only the potential functions defined over the coordinate  $\epsilon$ . A solute–solvent interaction potential under consideration is then constant over an equienergy surface of  $v$  and may be expressed in the form involving only one argument  $\epsilon$ . Suppose again that the average distributions given by Eq. (4) are identical between two distinct solute–solvent interactions  $u$  and  $w$  and that

$$\rho^e(\epsilon; u) = \rho^e(\epsilon; w) \quad (\text{B6})$$

holds. In that case, it is possible to show by virtue of Eqs. (4) and (B2) that

$$\begin{aligned} Z[P_w^0; w] < Z[P_u^0; u] + \int d\epsilon \rho^e(\epsilon; u)(w^e(\epsilon) - u^e(\epsilon)), \\ Z[P_u^0; u] < Z[P_w^0; w] + \int d\epsilon \rho^e(\epsilon; u)(u^e(\epsilon) - w^e(\epsilon)). \end{aligned} \quad (\text{B7})$$

Since the two inequalities in Eq. (B7) are not consistent with each other, the supposition Eq. (B6) cannot be true. In other words, different solute–solvent interactions are mapped through Eq. (4) to different distributions of the solvent. Therefore, for a given defining potential  $v$ , the correspondence is one-to-one between the set of  $u^e(\epsilon)$  defined over the energetic coordinate  $\epsilon$  and the set of  $\rho^e(\epsilon; u)$  expressed in the energetic representation.

When the defining potential  $v$  vanishes outside a finite region  $\Omega$ , it is possible to repeat the similar arguments for the reduced form  $\rho_\Omega^e$  of average distribution determined from Eq. (11). In this case, the map to the set of distribution functions  $\rho_\Omega^e(\epsilon; u)$  is one-to-one from a set of solute–solvent interaction potentials  $u$  which are constant over equienergy surfaces of  $v$  and are zero outside the interaction range  $\Omega$ .

## APPENDIX C

When the PY approximation is adopted in the full coordinate representation for a solute–solvent interaction potential  $v$  of interest, it was shown by Lee that the excess chemical potential  $\Delta\mu$  of the solute is given by<sup>29–31</sup>

$$\Delta\mu = -k_B T \int d\mathbf{x} \rho^f(\mathbf{x}; 0) \log(1 + \zeta(\mathbf{x})) \frac{c^f(\mathbf{x}; v)}{\zeta(\mathbf{x})}, \quad (\text{C1})$$

where an auxiliary function  $\zeta$  is defined as

$$\zeta(\mathbf{x}) = \frac{\rho^f(\mathbf{x}; v)}{\rho^f(\mathbf{x}; 0)} - c^f(\mathbf{x}; v) - 1. \quad (\text{C2})$$

Equation (C1) then reduces, by virtue of the PY closure Eq. (26), to

$$\Delta\mu = -k_B T \int d\mathbf{x} \rho^f(\mathbf{x}; 0) f(\mathbf{x}) \frac{\beta w^f(\mathbf{x}; v)}{\exp(\beta w^f(\mathbf{x}; v)) - 1}, \quad (\text{C3})$$

where  $f$  is the Mayer function for the solute–solvent interaction  $v$  of interest given by Eq. (29). It should be noted that Eq. (C3) is identical to Eq. (2.11) of Ref. 8.<sup>43</sup> The derivation of the  $\Delta\mu$  expression under the PY approximation in the energetic representation is parallel to that presented by Lee and Kjellander and Sarman in the full coordinate representation,<sup>29–31</sup> and Eq. (37) is obtained.

In the HNC approximation, the method elaborated by Lee and Kjellander and Sarman in the full coordinate representation<sup>29–31</sup> may be readily developed in the energetic representation to give Eq. (38).

## APPENDIX D

A specific form of the solute–solvent interaction potential is often adopted which depends only on the position of a single site in the solvent molecule relative to the solute molecule. For example, the interaction potential between water and a hydrophobic solute is commonly formulated as a function only of the position of the oxygen site relative to the hydrophobic solute.<sup>48</sup> In such a specific case, it is possible to construct the two-body correlation function  $\chi^e(\epsilon, \eta; 0)$  in the pure solvent from the knowledge of the radial distribution function of the site determining the solute–solvent interaction.

When the solute molecule is fixed at the origin with a fixed orientation and the solute–solvent interaction potential  $v$  of interest depends only on the position  $\mathbf{r}$  of the site  $\tau$  of the solvent molecule,  $v$  may be expressed as  $v_\tau(\mathbf{r})$ . In this case, the correlation function relevant for constructing  $\chi^e(\epsilon, \eta; 0)$  is the radial distribution function  $g_{\tau\tau}$  of the site  $\tau$  in the pure solvent, provided that the pure solvent system is homogeneous and isotropic.<sup>36</sup> It is then easy to see that Eq. (15) reduces in the pure solvent to

$$\begin{aligned} \chi^e(\epsilon, \eta; 0) &= \rho \int d\mathbf{r} \delta(v_\tau(\mathbf{r}) - \epsilon) \delta(\epsilon - \eta) \\ &+ \rho^2 \int d\mathbf{r} d\mathbf{R} \delta(v_\tau(\mathbf{r}) - \epsilon) \delta(v_\tau(\mathbf{R}) - \eta) \\ &\times (g_{\tau\tau}(\mathbf{r} - \mathbf{R}) - 1). \end{aligned} \quad (\text{D1})$$

This expression indeed relates  $\chi^e(\epsilon, \eta; 0)$  to the radial distribution function  $g_{\tau\tau}$ .

## APPENDIX E

A simplest model of the solute–solvent interaction is such that the potential is constant over its interaction range  $\Omega$ . In this case, the solute–solvent interaction potential  $v$  of interest is written as

$$v^f(\mathbf{x}) = \begin{cases} \xi & \text{when } \mathbf{x} \in \Omega \\ 0 & \text{otherwise} \end{cases} \quad (\text{E1})$$

and  $\xi$  is the only parameter for the solute–solvent interaction. Note that a hard core solute is a special case of Eq. (E1) for which  $\beta\xi \gg 1$ . When the specific form Eq. (E1) of the solute–solvent interaction is adopted as the defining potential, the reduced instantaneous distribution  $\hat{\rho}_\Omega^e$  defined by Eq. (8) is expressed as

$$\hat{\rho}_\Omega^e(\epsilon) = \hat{N}_\Omega \delta(\epsilon - \xi), \quad (\text{E2})$$

where  $\hat{N}_\Omega$  is the (instantaneous) number of solvent molecules in the interaction range  $\Omega$ . The average distributions  $\rho_\Omega^e$  and  $\chi_\Omega^e$ , which are given in the presence of a solute–solvent interaction  $u$  by Eqs. (11) and (22), respectively, then reduce to

$$\begin{aligned}\rho_\Omega^e(\epsilon; u) &= \langle \hat{N}_\Omega \rangle_u \delta(\epsilon - \xi), \\ \chi_\Omega^e(\epsilon, \eta; u) &= (\langle \hat{N}_\Omega^2 \rangle_u - \langle \hat{N}_\Omega \rangle_u^2) \delta(\epsilon - \xi) \delta(\eta - \xi),\end{aligned}\quad (\text{E3})$$

where  $\langle \cdots \rangle_u$  denotes the ensemble average taken in the solution with the solute–solvent interaction  $u$ . Equation (E3) shows that  $\rho_\Omega^e$  and  $\chi_\Omega^e$  correspond to the first and second cumulants of  $\hat{N}_\Omega$ , respectively, when the defining potential involves a specific form Eq. (E1). The purpose of this Appendix is to show for the PY, HNC, and superposition approximations in the energetic representation that the excess chemical potential of the solute with the solute–solvent interaction of the form Eq. (E1) is determined from the first and second cumulants of  $\hat{N}_\Omega$  in the pure solvent system.

Under the specific form Eq. (E1) of the solute–solvent interaction, the PY approximation consisting of Eqs. (23) and (40) is actually identical to the superposition approximation given by Eq. (44). In these approximations, the indirect part  $w_\Omega^e$  of the solute–solvent potential of mean force defined by Eq. (13) is determined to be

$$\begin{aligned}\exp(\beta w_\Omega^e(\xi; v)) &= 1 - (\exp(-\beta\xi) - 1) \\ &\quad \times \frac{\langle \hat{N}_\Omega(\hat{N}_\Omega - 1) \rangle_0 - \langle \hat{N}_\Omega \rangle_0^2}{\langle \hat{N}_\Omega \rangle_0},\end{aligned}\quad (\text{E4})$$

where  $\langle \cdots \rangle_0$  denotes the ensemble average taken in the pure solvent system [ $\xi=0$  in Eq. (E1)]. The corresponding expression for the excess chemical potential  $\Delta\mu$  is obtained from Eq. (42) and is given by

$$\begin{aligned}\Delta\mu &= -k_B T \langle \hat{N}_\Omega \rangle_0 (\exp(-\beta\xi) - 1) \\ &\quad \times \frac{\beta w_\Omega^e(\xi; v)}{\exp(\beta w_\Omega^e(\xi; v)) - 1}.\end{aligned}\quad (\text{E5})$$

In the HNC approximation consisting of Eqs. (23) and (41), on the other hand,  $w_\Omega^e$  is the solution of

$$\begin{aligned}-\beta w_\Omega^e(\xi; v) &= [\exp(-\beta\{\xi + w_\Omega^e(\xi; v)\}) + \beta w_\Omega^e(\xi; v) - 1] \\ &\quad \times \frac{\langle \hat{N}_\Omega(\hat{N}_\Omega - 1) \rangle_0 - \langle \hat{N}_\Omega \rangle_0^2}{\langle \hat{N}_\Omega \rangle_0}\end{aligned}\quad (\text{E6})$$

and the excess chemical potential  $\Delta\mu$  is expressed, by virtue of Eq. (43), as

$$\begin{aligned}\Delta\mu &= -k_B T \langle \hat{N}_\Omega \rangle_0 \left[ \left( \frac{\beta w_\Omega^e(\xi; v)}{2} + 1 \right) \right. \\ &\quad \left. \times \exp(-\beta\{\xi + w_\Omega^e(\xi; v)\}) + \frac{\beta w_\Omega^e(\xi; v)}{2} - 1 \right].\end{aligned}\quad (\text{E7})$$

Equations (E4), (E5), (E6), and (E7) show that the excess chemical potential of the solute can be expressed in terms of

the first and second cumulants of  $\hat{N}_\Omega$  in the pure solvent system when the solute–solvent interaction potential takes the simple form Eq. (E1).

When the solute–solvent interaction is given by Eq. (E1), the excess chemical potential  $\Delta\mu$  is closely connected to the probability distribution of the number  $\hat{N}_\Omega$  of solvent molecules contained in the interaction range  $\Omega$ . Indeed,  $\Delta\mu$  can be expressed as

$$\exp(-\beta\Delta\mu) = \sum_n \exp(-n\beta\xi) p_n, \quad (\text{E8})$$

where  $p_n$  is the probability that  $\hat{N}_\Omega = n$  in the pure solvent system. When  $p_n$  is given, Eq. (E8) determines  $\Delta\mu$  for any interaction parameter  $\xi$  in Eq. (E1). Conversely, when  $\Delta\mu$  is provided as a function of  $\xi$ , the probability  $p_n$  can be obtained, for example, by expanding  $\exp(-\beta\Delta\mu)$  with respect to  $\exp(-\beta\xi)$ . In the information-theoretic approach by Hummer *et al.*,  $p_n$  is evaluated from the first and second cumulants of  $\hat{N}_\Omega$  in the pure solvent through the maximization of a properly defined information entropy.<sup>44,45</sup> It is of interest to note, on the other hand, that Eqs. (E5) and (E7) provide routes, by virtue of Eq. (E8), to determining  $p_n$  from the first and second cumulants of  $\hat{N}_\Omega$  under the PY, HNC, and superposition approximations in the energetic representation.

When the solute molecule interacts with the solvent molecule through Eq. (E1) at an interaction parameter  $\xi$ , the average number  $N_\Omega(\xi)$  of solvent molecules in the interaction range  $\Omega$  is uniquely determined from

$$N_\Omega(\xi) = \langle \hat{N}_\Omega \rangle_\xi, \quad (\text{E9})$$

where  $\langle \cdots \rangle_\xi$  denotes the ensemble average taken in the solution with the solute–solvent interaction given by Eq. (E1). Equation (E9) defines a map from  $\xi$  to  $N_\Omega(\xi)$ , and the map is one-to-one as seen from the arguments in Appendix B. In Ben-Naim's treatments of water and hydrophobic effects, the degree of water structure was discussed by assuming the hydrogen bond interaction of the form Eq. (E1).<sup>49–51</sup> The one-to-one correspondence between  $\xi$  and  $N_\Omega(\xi)$  was then utilized to adopt  $N_\Omega(\xi)$  as a measure of the degree of water structure. In Ben-Naim's treatments, however, the quantitative argument was restricted to the linear regime of the  $\xi$  variation and the nonlinear regime was explored only qualitatively. The approximate method described in this Appendix can examine the nonlinear regime quantitatively. Furthermore, the method of integral equation developed in the present paper is useful to treat an interaction which is more general in form than Eq. (E1) and is not characterized by a single parameter.

<sup>1</sup>T. L. Hill, *Statistical Mechanics* (McGraw–Hill, New York, 1956).

<sup>2</sup>J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic Press, London, 1986).

<sup>3</sup>J. G. Kirkwood, *J. Chem. Phys.* **3**, 300 (1935).

<sup>4</sup>J. K. Percus and G. J. Yevick, *Phys. Rev.* **110**, 1 (1958).

<sup>5</sup>T. Morita, *Prog. Theor. Phys.* **23**, 829 (1960).

<sup>6</sup>J. K. Percus, *Phys. Rev. Lett.* **8**, 462 (1962).

<sup>7</sup>G. Stell, *Physica (Amsterdam)* **29**, 517 (1963).

<sup>8</sup>N. Matubayasi, *J. Chem. Phys.* **104**, 5265 (1996).

<sup>9</sup>W. A. Steele, *J. Chem. Phys.* **39**, 3197 (1963).

<sup>10</sup>L. Blum and A. J. Torruella, *J. Chem. Phys.* **56**, 303 (1972).

<sup>11</sup>L. Blum, *J. Chem. Phys.* **57**, 1862 (1972).

- <sup>12</sup>D. Chandler and H. C. Andersen, *J. Chem. Phys.* **57**, 1930 (1972).
- <sup>13</sup>G. Stell, G. N. Patey, and J. S. Høye, *Adv. Chem. Phys.* **48**, 183 (1981).
- <sup>14</sup>L. J. Lowden and D. Chandler, *J. Chem. Phys.* **59**, 6587 (1973).
- <sup>15</sup>C. S. Hsu, D. Chandler, and L. J. Lowden, *Chem. Phys.* **14**, 213 (1976).
- <sup>16</sup>F. Hirata and P. J. Rossky, *Chem. Phys. Lett.* **83**, 329 (1981).
- <sup>17</sup>P. J. Rossky, B. M. Pettitt, and G. Stell, *Mol. Phys.* **50**, 1263 (1983).
- <sup>18</sup>S. Ten-no, F. Hirata, and S. Kato, *J. Chem. Phys.* **100**, 7443 (1994).
- <sup>19</sup>M. Ikeguchi and J. Doi, *J. Chem. Phys.* **103**, 5011 (1995).
- <sup>20</sup>D. Beglov and B. Roux, *J. Phys. Chem. B* **101**, 7821 (1997).
- <sup>21</sup>C. M. Cortis, P. J. Rossky, and R. A. Friesner, *J. Chem. Phys.* **107**, 6400 (1997).
- <sup>22</sup>A. Kovalenko and F. Hirata, *J. Chem. Phys.* **110**, 10095 (1999).
- <sup>23</sup>Q. Du, D. Beglov, and B. Roux, *J. Phys. Chem. B* **104**, 796 (2000).
- <sup>24</sup>D. Chandler, R. Silbey, and B. M. Ladanyi, *Mol. Phys.* **46**, 1335 (1982).
- <sup>25</sup>P. J. Rossky and R. A. Chiles, *Mol. Phys.* **51**, 661 (1984).
- <sup>26</sup>D. Chandler, *J. Chem. Phys.* **59**, 2742 (1973).
- <sup>27</sup>K. Kojima and K. Arakawa, *Bull. Chem. Soc. Jpn.* **53**, 1795 (1980).
- <sup>28</sup>Y. Takebayashi, Y. Kimura, N. Hirota, and M. Ohba, *J. Chem. Phys.* **105**, 10084 (1996).
- <sup>29</sup>L. L. Lee, *J. Chem. Phys.* **60**, 1197 (1974).
- <sup>30</sup>R. Kjellander and S. Sarman, *J. Chem. Phys.* **90**, 2768 (1989).
- <sup>31</sup>L. L. Lee, *J. Chem. Phys.* **97**, 8606 (1992).
- <sup>32</sup>S. J. Singer and D. Chandler, *Mol. Phys.* **55**, 621 (1985).
- <sup>33</sup>P. Attard, *J. Chem. Phys.* **94**, 2370 (1991).
- <sup>34</sup>X. S. Chen, F. Forstmann, and M. Kasch, *J. Chem. Phys.* **95**, 2832 (1991).
- <sup>35</sup>L. Lue and D. Blankschtein, *J. Chem. Phys.* **100**, 3002 (1994).
- <sup>36</sup>In our general treatments, it is not necessary to assume that the pure solvent system is homogeneous and isotropic.
- <sup>37</sup>When the grand canonical ensemble is adopted, the total number  $N$  of solvent molecules in the system fluctuates, and  $N$  in Eq. (12) needs to be replaced by the average  $\langle N \rangle_u$ .
- <sup>38</sup>J. L. Lebowitz and J. K. Percus, *J. Math. Phys.* **4**, 116 (1963).
- <sup>39</sup>R. Evans, *Adv. Phys.* **28**, 143 (1979).
- <sup>40</sup>D. Chandler, J. D. McCoy, and S. J. Singer, *J. Chem. Phys.* **85**, 5971 (1986); **85**, 5977 (1986).
- <sup>41</sup>R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
- <sup>42</sup>Since the solute is fixed at the origin with a fixed orientation, the two-body correlation functions of the solvent defined by Eqs. (14), (15), and (22) actually describe the three-body correlation among one solute molecule and two solvent molecules. We call them the two-body correlation functions since the solute-solvent interaction can be viewed as an external field for the solvent molecules.
- <sup>43</sup>The coincidence between the  $\Delta\mu$  expressions in the PY and superposition approximations is probably accidental.
- <sup>44</sup>G. Hummer, S. Garde, A. E. García, A. Pohorille, and L. R. Pratt, *Proc. Natl. Acad. Sci. USA* **93**, 8951 (1996).
- <sup>45</sup>G. E. Crooks and D. Chandler, *Phys. Rev. E* **56**, 4217 (1997).
- <sup>46</sup>D. Chandler, *Phys. Rev. E* **48**, 2898 (1993).
- <sup>47</sup>Although the RISM-2 integral equations are not affected by the presence of “auxiliary” sites, they do not satisfy a moment condition which is related to the indistinguishability of sites at large wavelengths (Refs. 26–28). This type of difficulty is always absent in the full coordinate representation and the energetic representation since each of the solute and solvent molecules is treated as a single unit.
- <sup>48</sup>W. L. Jorgensen, J. Gao, and C. Ravimohan, *J. Phys. Chem.* **89**, 3470 (1985).
- <sup>49</sup>A. Ben-Naim, *J. Phys. Chem.* **79**, 1268 (1975).
- <sup>50</sup>A. Ben-Naim, *Hydrophobic Interactions* (Plenum, New York, 1980).
- <sup>51</sup>In Ref. 50, Ben-Naim mentioned a more general form of  $v^f(\mathbf{x}) = \xi G(\mathbf{x})$  for the hydrogen bond interaction. The arguments are valid even in this case when  $N_\Omega(\xi)$  is replaced with the corresponding average  $\langle \sum_i G(\mathbf{x}_i) \rangle_\xi$ .