THE JOURNAL OF CHEMICAL PHYSICS **128**, 244511 (2008)

Kirkwood–Buff theory of four and higher component mixtures

Myungshim Kang and Paul E. Smith[a)]
Department of Chemistry, 111 Willard Hall, Kansas State University, Manhattan, Kansas 66506-3701, USA

(Received 17 April 2008; accepted 21 May 2008; published online 24 June 2008)

Explicit expressions are developed for the chemical potential derivatives, partial molar volumes, and isothermal compressibility of solution mixtures involving four components at finite concentrations using the Kirkwood–Buff theory of solutions. In addition, a general recursion relationship is provided which can be used to generate the chemical potential derivatives for higher component solutions. © *2008 American Institute of Physics*. [DOI: 10.1063/1.2943318]

I. INTRODUCTION

Kirkwood–Buff (KB) theory is an exact theory of solutions that relates properties of a solution mixture to radial distribution functions (RDFs) between the different components of the solution.[1,2] KB theory has been widely used to understand the basic properties of solutions,[2–4] the effects of additives on the solubility of solutes (from small hydrocarbons to proteins)[5–11] and biomolecular equilibria,[12–16] to investigate the local composition of solutions in the context of preferential solvation,[17,18] to study the effects of additives on the surface tension of liquids,[19,20] to interpret computer simulation data,[13,21,22] and to develop models for many of the above effects.[23]

The central focus of KB theory are the KB integrals (G_{ij}) between the different species i and j in the solution mixture,[1]

$$G_{ij} = G_{ji} = 4\pi\int_0^\infty [g_{ij}^{\mu VT}(r) - 1]r^2 dr, \tag{1}$$

where g_{ij} is the corresponding RDF and r is the intermolecular separation. The above RDFs are defined in a grand canonical (μVT) ensemble open to all species. Chemical potential derivatives for closed or semiopen systems in terms of the KB integrals and number densities ($\rho_i = n_i/V$) are then obtained after suitable thermodynamic transformations.[2,24] The KB integrals, together with the corresponding excess coordination numbers, have provided a simple physical picture of changes in the local solution composition around each species.[4]

Unfortunately, as the number of solution components (n) increases and/or as one moves from open to semiopen to closed ensembles, the resulting expressions become more cumbersome and involve significant algebraic manipulation.[4] Expressions for two component solutions were provided in the original Kirkwood and Buff[1] paper. Subsequently, O'Connell[25] presented a general matrix formulation of KB theory, and Ben-Naim[12] developed a method to simplify the matrices involved for a general n component mixture. O'Connell[25] also developed expressions based on the direct correlation function, as defined by the Ornstein–Zernike equation, instead of the total correlation function. This also makes KB theory highly compatible with integral equation theories. However, the physical interpretation of integrals involving the direct correlation function is more complicated than that of the standard KB integral at normal solution densities. Furthermore, the direct correlation function can only be obtained from computer simulations after Fourier transforming the original total correlation function.[26]

Of course, one could always use the general matrix formulation of KB theory and solve numerically using values for the RDFs or KB integrals provided by some other approach (theory or simulation). However, this tends to obscure the contribution from the different KB integrals and hinder our understanding of specific effects. Therefore, it is often desirable to use explicit expressions that involve combinations of KB integrals and number densities. Explicit expressions for three component solutions have been provided by Ruckenstein and Shulgin (using an algebraic software),[27] Ben-Naim,[4] and Smith.[24] Ben-Naim[12] also developed expressions for some properties of four component systems, where several of the components appear at infinite dilution. To our knowledge, explicit expressions for chemical potential derivatives in four or higher component systems have not appeared for the case where all components are present at finite concentrations. Here, we use some of the relationships provided previously by Hall in an alternative derivation of the KB theory[28] to generate expressions for four component solutions. A general recursion relationship is then developed for higher component mixtures.

II. THEORY

A. General approach

Hall rederived the KB theory using a different approach to Kirkwood and Buff.[28] In doing so Hall produced two primary equations from which many of the expressions required here can be generated. However, his approach was still somewhat involved. Here we present a simpler derivation of the Hall equations. The first focuses on changes in the molar concentration of any component. If we consider the species

---

[a)] Author to whom correspondence should be addressed. Tel.: 785-532-5109. FAX: 785-532-6666. Electronic mail: pesmith@ksu.edu.

number densities (or molarities) in the grand canonical ensemble to be functions of T and all the chemical potentials (μ), then we can write

$$d \ln \rho_i = \sum_{j=1}^{n} \left(\frac{\partial \ln \rho_i}{\partial \mu_j} \right)_{T,\mu_{k\neq j}} d\mu_j \quad \text{constant} \quad T, \tag{2}$$

for any component i at constant T. Here, the summation is over all n components of the solution. We note that all the chemical potentials are independent thermodynamic variables in this open ensemble. The above derivatives can be expressed in terms of KB integrals using the fact that[1,2]

$$\left(\frac{\partial \ln \rho_i}{\partial \mu_j} \right)_{T,\mu_{k\neq j}} = \beta(\delta_{ij} + N_{ij}), \tag{3}$$

which is essentially the starting equation for the KB theory. Here, δ_{ij} is the Kroenecker delta function, $N_{ij}=\rho_j G_{ij} \neq N_{ji}$, $\beta = 1/RT$, and R is the Gas constant. From the these two equations one finds

$$d \ln \rho_i = \beta \sum_{j=1}^{n} (\delta_{ij} + N_{ij}) d\mu_j \quad \text{constant} \quad T. \tag{4}$$

The above expression is valid for changes in the concentration of any component in any multicomponent system and any (thermodynamically reasonable) ensemble with T constant. This is the equation derived by Hall, however, using a much longer route.[28] If one is interested in expressing solution compositions in terms of molalities ($m_i=\rho_i/\rho_1$ to within a conversion factor), then using the fact that $d \ln m_i = d \ln \rho_i - d \ln \rho_1$ one can write

$$d \ln m_i = \beta \sum_{j=1}^{n} (\delta_{ij} + N_{ij} - \delta_{1j} - N_{1j}) d\mu_j \quad \text{constant} \quad T, \tag{5}$$

which is also valid for any constant T ensemble. Clearly, in doing so, we have labeled component 1 as the primary solvent and, therefore, it is unique—as it is also experimentally. The consequences of doing this will be discussed later.

In the traditional derivation of KB theory, the set of equations presented in Eq. (4) are converted into matrix form after taking derivatives with respect to $\ln \rho_j$ with all $\rho_{k\neq j}$ held constant.[4] They can then be solved to obtain a series of expressions involving the quantities

$$\beta \left(\frac{\partial \mu_i}{\partial \ln N_j} \right)_{T,V,N_{k\neq j}} = \beta \left(\frac{\partial \mu_i}{\partial \ln \rho_j} \right)_{T,N_{k\neq j}}. \tag{6}$$

These constant volume derivatives then have to be transformed using a series of thermodynamic relationships into the required and experimentally relevant derivatives at constant P as defined by

$$\mu_{ij} = \beta \left(\frac{\partial \mu_i}{\partial \ln N_j} \right)_{T,P,N_{k\neq j}} = \beta \left(\frac{\partial \mu_i}{\partial \ln m_j} \right)_{T,P,m_{k\neq j}}. \tag{7}$$

This is clearly the most general approach. However, it has long been recognized that the expressions obtained for higher multicomponent systems ($n \geq 3$) involve considerable algebraic manipulation.[4,27] In addition, a significant degree of cancellation of terms in the expressions is often found but not easily recognized in the matrix formulation.

Here, we will adopt a different route which we believe is much simpler for mixtures with a large number of components. Eliminating $d\mu_1$ from Eq. (4) using the corresponding Gibbs–Duhem (GD) relationship at constant T and P, $\sum_{j=1}^{n} \rho_j d\mu_j = 0$, provides

$$d \ln \rho_i = \beta \sum_{j=2}^{n} [\delta_{ij} + N_{ij} - m_j(\delta_{i1} + N_{i1})] d\mu_j \quad \text{constant} \quad T,P. \tag{8}$$

This can be used to obtain an expression for changes in the molalities,

$$d \ln m_i = \beta \sum_{j=2}^{n} (\delta_{ij} + N_{ij}^{+}) d\mu_j \quad \text{constant} \quad T,P, \tag{9}$$

where $N_{ij}^{+} = N_{ij} + m_j(1 + N_{11} - N_{i1} - N_{j1})$ and $i=2,n$. The additional constraint of constant P arises from our use of the corresponding GD expression. This is the equation provided by Hall[28] for changes in molal concentrations at constant T and P. It also appears in the original KB paper without derivation.[1] The above set of equations can be written in a general $(n-1)\times(n-1)$ matrix form for a mixture of n components so that

$$\begin{bmatrix} 1+N_{22}^{+} & N_{23}^{+} & N_{24}^{+} & \cdots & N_{2n}^{+} \\ N_{32}^{+} & 1+N_{33}^{+} & & & \\ N_{42}^{+} & & 1+N_{44}^{+} & & \\ \vdots & & & \ddots & \\ N_{n2}^{+} & & & & 1+N_{nn}^{+} \end{bmatrix} \begin{bmatrix} \beta d\mu_2 \\ \beta d\mu_3 \\ \beta d\mu_4 \\ \vdots \\ \beta d\mu_n \end{bmatrix} = \begin{bmatrix} d \ln m_2 \\ d \ln m_3 \\ d \ln m_4 \\ \vdots \\ d \ln m_n \end{bmatrix}. \tag{10}$$

To continue we will choose our required ensemble and then take derivatives with respect to the molality of one component, in this case $\ln m_j$, keeping T, P, and all other $m_{k\neq j}$ constant. This makes the resulting expressions less general than previous approaches, but one can easily recover derivatives with respect to other species by a simple index change. One finds that

$$\begin{bmatrix} 1+N^+_{22} & N^+_{23} & N^+_{24} & \cdots & N^+_{2n} \\ N^+_{32} & 1+N^+_{33} & & & \\ N^+_{42} & & 1+N^+_{44} & & \\ \vdots & & & \ddots & \\ N^+_{n2} & & & & 1+N^+_{nn} \end{bmatrix} \begin{bmatrix} \mu_{2j} \\ \mu_{3j} \\ \mu_{4j} \\ \vdots \\ \mu_{nj} \end{bmatrix} = \begin{bmatrix} \delta_{2j} \\ \delta_{3j} \\ \delta_{4j} \\ \vdots \\ \delta_{nj} \end{bmatrix} \quad \text{or} \quad \mathbf{D_n}\boldsymbol{\mu}=\mathbf{d}, \tag{11}$$

where the matrix elements are given by $D_{\alpha\beta}=\delta_{\alpha\beta}+N^+_{(\alpha+1)(\beta+1)}$, the vector elements of μ_α are given by $\mu_{(\alpha+1)j}$, and the vector elements of d_α are given by $\delta_{(\alpha+1)j}$ with α, $\beta=1$, $n-1$. Hence, we have a set of simultaneous equations which can be solved to give the chemical potential derivatives. Therefore,

$$\boldsymbol{\mu}=\mathbf{D_n^{-1}d}. \tag{12}$$

One can express the inverse in terms of cofactors of the original D_n matrix. The chemical potential derivatives are then given by $\mu_{ij}=D_n^{j-1,i-1}/|D_n|$, for $i,j\neq 1$ and where $D_n^{j-1,i-1}$ is a cofactor of D_n. If the chemical potential derivative of species 1 is required, it can be obtained from the GD equations

$$\mu_{1j}=-\sum_{k=2}^{n} m_k\mu_{kj}=-m_j\sum_{k=2}^{n}\mu_{jk}, \tag{13}$$

using the solutions to Eq. (11). In the above expression $j=1$, n, and $m_1=1$.

There are several advantages of this approach. First, it can be applied directly to any number of solution components in any constant T and P ensemble. Therefore, we do not have to transform the subsequent expressions from a constant T and V to a constant T and P ensemble. Second, we have eliminated the chemical potential of species 1 and therefore the resulting set of equations and the corresponding matrix is reduced. Third, the simplicity of the column vector on the right-hand side of Eq. (12) indicates that each chemical potential derivative expression involves only one element of the inverse matrix in the numerator, together with the determinant of D_n in the denominator. This set of combined factors greatly simplifies the resulting expressions.

In the previous sections, we have focused on derivatives of the chemical potentials taken with respect to molality. Derivatives with respect to mole fraction or molarity can be obtained by noting that

$$d\ln x_i = d\ln\rho_i-\sum_{j=1}^{n}x_j d\ln\rho_j \quad \text{constant} \quad T \tag{14}$$

and

$$d\ln m_i = d\ln\rho_i - d\ln\rho_1 \quad \text{constant} \quad T \tag{15}$$

for any number of components at constant T. One could develop Eq. (14) in terms of the KB integrals. However, it is much easier, especially for closed systems, to convert the molality based derivatives to mole fraction derivatives after the former has been obtained. For closed systems these equations become

$$\left(\frac{\partial\ln x_i}{\partial\ln m_j}\right)_{T,P,m_{k\neq j}}=\delta_{ij}-x_j, \tag{16}$$

with

$$\left(\frac{\partial\ln\rho_i}{\partial\ln m_j}\right)_{T,P,m_{k\neq j}}=\delta_{ij}-\rho_j\bar{V}_j, \tag{17}$$

and therefore

$$\left(\frac{\partial\ln\rho_i}{\partial\ln x_j}\right)_{T,P,m_{k\neq j}}=\frac{\delta_{ij}-\rho_j\bar{V}_j}{1-x_j}, \tag{18}$$

for any number of components at constant T and P.

To determine the corresponding expressions for the partial molar volumes (PMVs) in multicomponent systems, it is sufficient to express the PMVs in terms of the chemical potential derivatives. We will continue to treat species 1 as a unique component. Starting with Eq. (4) for the differential of the number density of species 1 and eliminating $d\mu_1$ by using the GD relationship at constant T and P, one finds

$$d\ln\rho_1=-\beta\sum_{j=2}^{n}m_j(1+N_{11}-N_{j1})d\mu_j \quad \text{constant} \quad T,P. \tag{19}$$

Obviously, there is a series of similar expressions depending on the initial choice of i in Eq. (4). Taking the derivative with respect to $\ln m_k$ while keeping T, P, and all other $m_{j\neq k}$ constant provides for $k\neq 1$,

$$\begin{aligned}\rho_k\bar{V}_k&=\sum_{j=2}^{n}m_j(1+N_{11}-N_{j1})\mu_{jk}\\ &=m_k\sum_{j=2}^{n}(1+N_{11}-N_{j1})\mu_{kj},\end{aligned} \tag{20}$$

where the appropriate chemical potential derivatives are provided by Eq. (12). If required, the PMV of species 1 can be obtained from the fact that

$$\sum_{j=1}^{n}\rho_j\bar{V}_j=1, \tag{21}$$

for all mixtures.

Finally, if one starts from Eq. (4) and then takes derivatives with respect to pressure with all m_j and T constant, one can show that for any component i,

$$RT\kappa_T = \sum_{j=1}^{n} (\delta_{ij} + N_{ij})\bar{V}_j = \bar{V}_i + \sum_{j=1}^{n} N_{ij}\bar{V}_j, \tag{22}$$

where κ_T is the isothermal compressibility. This can be used to derive an expression for the compressibility. If we chose $i=1$ and then eliminated the PMV of species 1 by using Eq. (21), then

$$\rho_1 RT\kappa_T = (1 + N_{11}) - \sum_{j=2}^{n} (1 + N_{11} - N_{j1})\rho_j \bar{V}_j, \tag{23}$$

which can be written as

$$\rho_1 RT\kappa_T = (1 + N_{11}) - \sum_{j=2}^{n} (1 + N_{11} - N_{j1}) \times \sum_{k=2}^{n} m_k(1 + N_{11} - N_{k1})\mu_{kj}, \tag{24}$$

using the corresponding chemical potential derivatives.

Before leaving this section we note that the PMV of a species can be considered to involve two contributions.[2] The first relates to the change in volume of the solution due to the volume occupied by the additional molecule located at a fixed position in the system. The second involves the ideal contribution to the PMV which arises due to the fact that the additional molecule will possess a momentum, corresponding to the particular temperature, which contributes to the pressure of the system when the molecule is released. Under constant P conditions this gives rise to a change in volume according to the compressibility of the solution. Therefore, one can isolate the former change by writing

$$V_i = \bar{V}_i - RT\kappa_T. \tag{25}$$

From Eq. (22) we obtain the relationship

$$\rho_i V_i = -N_{1i} - \sum_{j=2}^{n} (N_{ji} - N_{1i}) \sum_{k=2}^{n} m_k(1 + N_{11} - N_{k1})\mu_{kj}, \tag{26}$$

which is now a better measure of the actual volume occupied by each species in solution. Another interesting property of solutions is the pseudochemical potential (μ^*). The pseudochemical potential plays an important role in solution theory and is defined by the equation[2]

$$\mu_i = \mu_i^* + RT \ln \Lambda_i^3 \rho_i, \tag{27}$$

where Λ is the thermal de Broglie wavelength. From this equation it is quite easy to show that

$$\beta\left(\frac{\partial \mu_i^*}{\partial \ln m_j}\right)_{T,P,m_{k\neq j}} = \mu_{ij} - (\delta_{ij} - \rho_j \bar{V}_j), \tag{28}$$

and also

$$\beta\left(\frac{\partial \mu_i^*}{\partial \ln \rho_j}\right)_{T,P,m_{k\neq j}} = \beta\left(\frac{\partial \mu_i}{\partial \ln \rho_j}\right)_{T,P,m_{k\neq j}} - 1, \tag{29}$$

which completes our preliminary analysis.

Hence, we have a general set of Eqs. (12), (20), and (24) which can be used to derive the KB expressions for the chemical potential derivatives, PMVs, and compressibility of any multicomponent mixture.

B. Four component mixtures

As an example of the current approach, we will generate the expressions for a four component system where all components appear at finite concentrations. To our knowledge the explicit KB expressions for a four component system have not been presented in the literature. Using the above approach we find

$$\begin{bmatrix} \mu_{22} \\ \mu_{32} \\ \mu_{42} \end{bmatrix} = D_4^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \tag{30}$$

and where the inverse of D_4 is given by

$$D_4^{-1} = \frac{1}{|D_4|} \begin{bmatrix} (1+N_{33}^{+})(1+N_{44}^{+}) - N_{43}^{+}N_{34}^{+} & N_{24}^{+}N_{43}^{+} - N_{23}^{+}(1+N_{44}^{+}) & N_{23}^{+}N_{34}^{+} - N_{24}^{+}(1+N_{33}^{+}) \\ N_{34}^{+}N_{42}^{+} - N_{32}^{+}(1+N_{44}^{+}) & (1+N_{22}^{+})(1+N_{44}^{+}) - N_{42}^{+}N_{24}^{+} & N_{24}^{+}N_{23}^{+} - N_{34}^{+}(1+N_{22}^{+}) \\ N_{32}^{+}N_{43}^{+} - N_{42}^{+}(1+N_{33}^{+}) & N_{23}^{+}N_{42}^{+} - N_{43}^{+}(1+N_{22}^{+}) & (1+N_{22}^{+})(1+N_{33}^{+}) - N_{23}^{+}N_{32}^{+} \end{bmatrix}. \tag{31}$$

Therefore, the expressions for the chemical potential derivatives in a four component mixture are given by

$$\beta\left(\frac{\partial \mu_2}{\partial \ln m_2}\right)_{T,P,m_3,m_4} = \frac{(1+N_{33}^{+})(1+N_{44}^{+}) - N_{43}^{+}N_{34}^{+}}{|D_4|},$$

$$\beta\left(\frac{\partial \mu_3}{\partial \ln m_2}\right)_{T,P,m_3,m_4} = -\frac{(1+N_{44}^{+})N_{32}^{+} - N_{34}^{+}N_{42}^{+}}{|D_4|},$$

$$\beta\left(\frac{\partial \mu_4}{\partial \ln m_2}\right)_{T,P,m_3,m_4} = -\frac{(1+N_{33}^{+})N_{42}^{+} - N_{43}^{+}N_{32}^{+}}{|D_4|}, \tag{32}$$

with

$$|D_4| = (1+N_{22}^{+})(1+N_{33}^{+})(1+N_{44}^{+}) - (1+N_{22}^{+})N_{34}^{+}N_{43}^{+} - (1+N_{33}^{+})N_{24}^{+}N_{42}^{+} - (1+N_{44}^{+})N_{23}^{+}N_{32}^{+} + 2N_{23}^{+}N_{34}^{+}N_{42}^{+}, \tag{33}$$

and where we have also used the fact that $\rho_i N^{+}_{ij} = \rho_j N^{+}_{ji}$ to simplify the above determinant. The final derivatives (μ_{12} and μ_{11}) can be obtained after application of the GD relationships

$$\beta\left(\frac{\partial \mu_1}{\partial \ln m_2}\right)_{T,P,m_3,m_4} = -m_2\mu_{22} - m_3\mu_{32} - m_4\mu_{42} \tag{34}$$

and

$$\beta\left(\frac{\partial \mu_1}{\partial \ln N_1}\right)_{T,P,N_2,N_3,N_4} = -\mu_{12} - \mu_{13} - \mu_{14}. \tag{35}$$

Derivatives with respect to other species molalities (m_3 and m_4) are obtained quite easily by either inspection, by noting that $m_j\mu_{jk} = m_k\mu_{kj}$, or from the fact that

$$\begin{bmatrix}\mu_{23}\\ \mu_{33}\\ \mu_{43}\end{bmatrix} = D_4^{-1}\begin{bmatrix}0\\ 1\\ 0\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\mu_{24}\\ \mu_{34}\\ \mu_{44}\end{bmatrix} = D_4^{-1}\begin{bmatrix}0\\ 0\\ 1\end{bmatrix}, \tag{36}$$

which are also relatively simple to solve.

Application of Eq. (20) and the expressions found above provides the following expressions for the PMVs in four component mixtures,

$$\begin{aligned}\rho_2\bar{V}_2 &= m_2(1+N_{11}-N_{21})\mu_{22} + m_3(1+N_{11}-N_{31})\mu_{32}\\ &\quad + m_4(1+N_{11}-N_{41})\mu_{42},\\[2ex] \rho_3\bar{V}_3 &= m_2(1+N_{11}-N_{21})\mu_{23} + m_3(1+N_{11}-N_{31})\mu_{33}\\ &\quad + m_4(1+N_{11}-N_{41})\mu_{43},\\[2ex] \rho_4\bar{V}_4 &= m_2(1+N_{11}-N_{21})\mu_{24} + m_3(1+N_{11}-N_{31})\mu_{34}\\ &\quad + m_4(1+N_{11}-N_{41})\mu_{44}.\end{aligned} \tag{37}$$

The above expressions obey Eq. (21) as required. If necessary, the PMV of 1 can then be obtained using Eq. (21). Finally, for the isothermal compressibility we obtain

$$\begin{aligned}\rho_1 RT\kappa_T = 1 + N_{11} &- (1+N_{11}-N_{21})[(1+N_{11}-N_{21})m_2\mu_{22} + (1+N_{11}-N_{31})m_3\mu_{32} + (1+N_{11}-N_{41})m_4\mu_{42}]\\ &- (1+N_{11}-N_{31})[(1+N_{11}-N_{21})m_2\mu_{23} + (1+N_{11}-N_{31})m_3\mu_{33} + (1+N_{11}-N_{41})m_4\mu_{43}]\\ &- (1+N_{11}-N_{41})[(1+N_{11}-N_{21})m_2\mu_{24} + (1+N_{11}-N_{31})m_3\mu_{34} + (1+N_{11}-N_{41})m_4\mu_{44}],\end{aligned} \tag{38}$$

which can be simplified to provide

$$\begin{aligned}\rho_1 RT\kappa_T = 1 + N_{11} &- (1+N_{11}-N_{21})^2 m_2\mu_{22} - (1+N_{11}-N_{31})^2 m_3\mu_{33} - (1+N_{11}-N_{41})^2 m_4\mu_{44}\\ &- 2(1+N_{11}-N_{21})(1+N_{11}-N_{31})m_3\mu_{32} - 2(1+N_{11}-N_{21})(1+N_{11}-N_{41})m_4\mu_{42}\\ &- 2(1+N_{11}-N_{41})(1+N_{11}-N_{31})m_3\mu_{34}.\end{aligned} \tag{39}$$

C. Three component mixtures

While three component systems had been studied before,[4,27] it is interesting and informative to compare the limiting expressions provided here with those currently in the literature, especially due to the different notations involved. In addition, this will aid in the development of a general recursive relationship for the derivatives. The limiting forms are quite easy to obtain as we have $N^{+}_{ij} \to 0$ as $\rho_j \to 0$, and $m_i\mu_{ij} = m_j\mu_{ji} \to 0$ as $\rho_i \to 0$ or $\rho_j \to 0$. Therefore, as m_4 tends to zero, one obtains the derivatives for a ternary system of 1, 2, and 3. The chemical potential derivatives are then given by

$$\begin{aligned}\beta\left(\frac{\partial \mu_2}{\partial \ln m_2}\right)_{T,P,m_3} &= \frac{(1+N^{+}_{33})}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}},\\ \beta\left(\frac{\partial \mu_3}{\partial \ln m_2}\right)_{T,P,m_3} &= -\frac{N^{+}_{32}}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}}.\end{aligned} \tag{40}$$

The corresponding PMV expressions reduce to

$$\begin{aligned}\rho_2\bar{V}_2 &= m_2\frac{(1+N_{11}-N_{21})(1+N^{+}_{33}) - (1+N_{11}-N_{31})N^{+}_{23}}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}},\\[2ex] \rho_3\bar{V}_3 &= m_3\frac{(1+N_{11}-N_{31})(1+N^{+}_{22}) - (1+N_{11}-N_{21})N^{+}_{32}}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}},\end{aligned} \tag{41}$$

with a compressibility given by

$$\begin{aligned}\rho_1 RT\kappa_T = 1 + N_{11} &- m_2\frac{(1+N_{11}-N_{21})^2(1+N^{+}_{33})}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}}\\ &+ 2m_2\frac{(1+N_{11}-N_{21})(1+N_{11}-N_{31})N^{+}_{23}}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}}\\ &- m_3\frac{(1+N_{11}-N_{31})^2(1+N^{+}_{22})}{(1+N^{+}_{22})(1+N^{+}_{33}) - N^{+}_{23}N^{+}_{32}}.\end{aligned} \tag{42}$$

D. Two component mixtures

Two component mixtures have clearly been studied before, but not using the present notation. After taking an additional $\rho_3 \to 0$ limit, one finds that for the two component case,

$$\beta\left(\frac{\partial \mu_2}{\partial \ln m_2}\right)_{T,P,N_1} = \frac{1}{1+N_{22}^{+}} \tag{43}$$

and

$$\rho_2 \bar{V}_2 = m_2 \frac{1+N_{11}-N_{21}}{1+N_{22}^{+}}, \tag{44}$$

with

$$\rho_1 RT\kappa_T = 1 + N_{11} - m_2 \frac{(1+N_{11}-N_{21})^2}{1+N_{22}^{+}}, \tag{45}$$

which further reduces to the required compressibility equation when $\rho_2 \to 0$.

E. Open and semiopen systems

Traditionally, KB expressions for open and semiopen systems have been derived starting from the fully closed system results.[29,30] This can be quite tedious. Recently, we suggested starting from expressions for the fully open system and transforming to the required ensemble in a stepwise manner.[24,31] This made the manipulations easier although several steps were still required. However, it is clear from Eqs. (4) and (5) that results for open and semiopen systems become almost trivial. As an example we will derive an expression for the preferential binding parameter $(\partial m_3/\partial m_2)$ for ternary mixtures in the T, μ_1, and μ_3 ensemble, where 1 is the primary solvent, 2 is the biomolecule of interest, and 3 is an additive. Starting from Eq. (5) one immediately finds

$$\left(\frac{\partial \ln m_2}{\partial \mu_2}\right)_{T,\mu_1,\mu_3} = \beta(1+N_{22}-N_{12}) \tag{46}$$

and

$$\left(\frac{\partial \ln m_3}{\partial \mu_2}\right)_{T,\mu_1,\mu_3} = \beta(N_{32}-N_{12}), \tag{47}$$

which can be solved to yield

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,\mu_1,\mu_3} = \frac{N_{23}-m_3 N_{21}}{1+N_{22}-N_{12}}, \tag{48}$$

and is in agreement with previous results.[24] It is clear from Eq. (5) that the same expression is obtained if we have any number of additional components at a constant chemical potential.

Alternatively, one can start from Eq. (9) to obtain an expression for the equivalent property in the T, P, and μ_3 ensemble. Hence,

$$(1+N_{22}^{+})\beta\left(\frac{\partial \mu_2}{\partial \ln m_2}\right)_{T,P,\mu_3} = 1 \tag{49}$$

and

$$N_{32}^{+}\beta\left(\frac{\partial \mu_2}{\partial \ln m_2}\right)_{T,P,\mu_3} = \left(\frac{\partial \ln m_3}{\partial \ln m_2}\right)_{T,P,\mu_3}, \tag{50}$$

which can be solved quite easily to give

$$\left(\frac{\partial m_3}{\partial m_2}\right)_{T,P,\mu_3} = \frac{N_{23}^{+}}{1+N_{22}^{+}}, \tag{51}$$

and is also in agreement with previous results.[24] Again, the same expression is valid in the presence of any number of additional species as long as their chemical potentials remain constant.

F. A general recursion relationship for the chemical potential derivatives

Analysis of the chemical potential derivatives for two, three, and four component systems enables a general recursion relationship to be established. It is clear that the denominator will always contain the determinant $|D_n|$ for a general n component system. If we focus on the expressions in the numerator, one immediately observes that the numerator of μ_{ii} is just the determinant of the D matrix for the corresponding $n-1$ system in which component i has been eliminated. Hence, we have

$$\mu_{nn}^{n} = \frac{|D_{n-1}|}{|D_n|}, \tag{52}$$

where the superscript indicates a derivative defined in an n component system. A simple change in indices provides equivalent expressions for any μ_{ii}^{n} where $i \neq 1$. The numerators of the other derivatives ($\mu_{nj}, j \neq 1$, or n) also follow a simple pattern. The chemical potential derivatives for the nth component obey the recursive relationship

$$\mu_{nj}^{n} = -\frac{|D_{n-1}|}{|D_n|}\sum_{k=2}^{n-1} \mu_{kj}^{n-1} N_{nk}^{+} = -\mu_{nn}^{n}\sum_{k=2}^{n-1} \mu_{kj}^{n-1} N_{nk}^{+}, \tag{53}$$

which is just a factorization of the D_n matrix that one observes due to the relative simplicity of Eq. (12). Expressions for the μ_{ij} derivatives where $i \neq j \neq 1$ can then be found by inspection.

G. Five component systems

Using the recursion relationship developed above, one can generate expressions for the chemical potential derivatives in five component solutions. For simplicity, we will only consider the chemical potential derivatives. The following expression is obtained from Eqs. (52) and (32) followed by a simple index change ($5 \leftrightarrow 2$) in the numerator,

$$\begin{aligned}\mu_{22}|D_5| &= (1+N_{33}^{+})(1+N_{44}^{+})(1+N_{55}^{+}) \\ &\quad - (1+N_{33}^{+})N_{45}^{+}N_{54}^{+} - (1+N_{44}^{+})N_{35}^{+}N_{53}^{+} \\ &\quad - (1+N_{55}^{+})N_{34}^{+}N_{43}^{+} + 2N_{34}^{+}N_{45}^{+}N_{53}^{+}.\end{aligned} \tag{54}$$

Using Eq. (53) and the set of derivatives for a four component solution provided in Eq. (32), one finds

$$\begin{aligned}\mu_{52}|D_5| = &-(1+N_{33}^{+})(1+N_{44}^{+})N_{52}^{+}+(1+N_{33}^{+})N_{42}^{+}N_{54}^{+}\\ &+(1+N_{44}^{+})N_{32}^{+}N_{53}^{+}-N_{32}^{+}N_{43}^{+}N_{54}^{+}\\ &-N_{34}^{+}N_{42}^{+}N_{53}^{+}+N_{34}^{+}N_{43}^{+}N_{52}^{+}.\end{aligned} \tag{55}$$

Consequently, using a simple index change of $5 \leftrightarrow 3$ and $5 \leftrightarrow 4$, one finds

$$\begin{aligned}\mu_{32}|D_5| = &-(1+N_{44}^{+})(1+N_{55}^{+})N_{32}^{+}+(1+N_{44}^{+})N_{35}^{+}N_{52}^{+}\\ &+(1+N_{55}^{+})N_{42}^{+}N_{34}^{+}-N_{42}^{+}N_{54}^{+}N_{35}^{+}\\ &-N_{45}^{+}N_{52}^{+}N_{34}^{+}+N_{45}^{+}N_{54}^{+}N_{32}^{+}\end{aligned} \tag{56}$$

and

$$\begin{aligned}\mu_{42}|D_5| = &-(1+N_{33}^{+})(1+N_{55}^{+})N_{42}^{+}+(1+N_{33}^{+})N_{45}^{+}N_{52}^{+}\\ &+(1+N_{55}^{+})N_{32}^{+}N_{43}^{+}-N_{32}^{+}N_{53}^{+}N_{45}^{+}\\ &-N_{35}^{+}N_{52}^{+}N_{43}^{+}+N_{35}^{+}N_{53}^{+}N_{42}^{+},\end{aligned} \tag{57}$$

respectively. Derivatives with respect to other species can be obtained by inspection. The above expressions are in agreement with those obtained directly via Eq. (12). If required, the corresponding PMVs and compressibility can be obtained from Eqs. (20) and (24).

III. DISCUSSION

We have provided general relationships which can be used to develop explicit expressions for chemical potential derivatives, PMVs, and the isothermal compressibility of any mixture of n components in terms of KB integrals. Our choice of the molality concentration scale makes species 1 unique. Consequently, some of the "symmetry" in the expressions that might be observed for molarity or mole fraction based derivatives is lost using the current notation. We consider this to be an acceptable sacrifice in many cases. It is therefore informative to compare and relate the expressions generated here with those developed by other approaches, especially when considering symmetric ideal solutions. To do this we will refer to the expressions of Smith for ternary solutions which represent the most condensed form for the chemical potential derivatives.[24] They are easily expanded to provide the expressions of Ruckenstein and Shulgin[27] and Ben-Naim.[4] Smith provided the following expressions:

$$\begin{aligned}\beta\left(\frac{\partial \mu_1}{\partial \ln N_2}\right)_{T,P,N_1,N_3} &= -\frac{\rho_2 A_3}{\rho_1 A_2 A_3+\rho_2 A_1 A_3+\rho_3 A_1 A_2},\\ \beta\left(\frac{\partial \mu_2}{\partial \ln N_2}\right)_{T,P,N_1,N_3} &= \frac{\rho_1 A_3+\rho_3 A_1}{\rho_1 A_2 A_3+\rho_2 A_1 A_3+\rho_3 A_1 A_2},\\ \beta\left(\frac{\partial \mu_3}{\partial \ln N_2}\right)_{T,P,N_1,N_3} &= -\frac{\rho_2 A_1}{\rho_1 A_2 A_3+\rho_2 A_1 A_3+\rho_3 A_1 A_2},\end{aligned} \tag{58}$$

where the A's are given by

$$\begin{aligned}A_1 &= 1+\rho_1(G_{11}+G_{23}-G_{12}-G_{13}),\\ A_2 &= 1+\rho_2(G_{22}-G_{23}-G_{12}+G_{13}),\\ A_3 &= 1+\rho_3(G_{33}-G_{23}+G_{12}-G_{13}).\end{aligned} \tag{59}$$

Comparison with the expressions provided in Eq. (40) indicates that

$$\begin{aligned}A_1 &= \rho_1 N_{32}^{+}/\rho_2 = \rho_1 N_{23}^{+}/\rho_3,\\ A_2 &= 1+N_{22}^{+}-N_{32}^{+} = 1+N_{22}^{+}-\rho_2 A_1/\rho_1,\\ A_3 &= 1+N_{33}^{+}-N_{23}^{+} = 1+N_{33}^{+}-\rho_3 A_1/\rho_1.\end{aligned} \tag{60}$$

Specific combinations of KB integrals often appear repeatedly in other formulations. For instance, one can define for $i \neq j$,

$$\eta_{ij} = \rho_i+\rho_j+\rho_i\rho_j(G_{ii}+G_{jj}-2G_{ij}) = \rho_i A_j+\rho_j A_i. \tag{61}$$

In the current notation it is found that

$$\eta_{i1} = \rho_1(1+N_{ii}^{+}). \tag{62}$$

We attempted to find a similar factorization and relationships as found in Eqs. (58) and (59) for four component systems, but were unsuccessful.

The application of the KB theory to symmetrical ideal solutions is also of interest. Ben-Naim[4] has shown that for a general n component mixture to display symmetric ideal behavior, one must have $\Delta G_{ij}=G_{ii}+G_{jj}-2G_{ij}=0$ for all ij pairs. Therefore, in the current notation, one must have $\rho_1 N_{ii}^{+}=\rho_i$ for symmetric ideal solutions. In addition, one finds $A_1=A_2=A_3=1$ for symmetric ideal ternary solutions.

IV. CONCLUSIONS

Using a new approach we have developed explicit relationships for KB integrals in four and five component solution mixtures. In our opinion, the use of molalities as concentration variables provides the simplest path to expressions in multicomponent solutions. We are currently using this type of approach to study biologically relevant systems containing five or more components.

For a general n component system there are $n(n+1)/2$ unique G_{ij} integrals. Determining the integrals from experimental data using the KB inversion approach requires one isothermal compressibility value, $n-1$ independent PMVs, and $n(n-1)/2$ independent μ_{ij} values as a function of composition. This has been achieved for ternary systems.[32] As one moves beyond ternary systems the experimental data becomes increasingly more difficult to obtain. Consequently, we envision that the major use for the expressions provided here will involve either theoretical estimates of the KB integrals or simulated values of the integrals. In either case, the exact KB expressions provide a solid foundation for investigating these complicated solution mixtures.

ACKNOWLEDGMENTS

The project described was supported by Grant No. R01GM079277 from the National Institute of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health.

[1] J. G. Kirkwood and F. P. Buff, J. Chem. Phys. **19**, 774 (1951).
[2] A. Ben-Naim, *Statistical Thermodynamics for Chemists and Biochemists* (Plenum, New York, 1992).
[3] E. Matteoli and G. A. Mansoori, *Fluctuation Theory of Mixtures* (Taylor & Francis, New York, 1990).
[4] A. Ben-Naim, *Molecular Theory of Solutions* (Oxford, New York, 2006).
[5] J. P. O'Connell, AIChE J. **17**, 658 (1971).
[6] A. A. Chialvo, J. Phys. Chem. **97**, 2740 (1993).
[7] P. E. Smith, J. Phys. Chem. B **103**, 525 (1999).
[8] R. Chitra and P. E. Smith, J. Phys. Chem. B **105**, 11513 (2001).
[9] E. Ruckenstein and I. Shulgin, Ind. Eng. Chem. Res. **41**, 4674 (2002).
[10] I. L. Shulgin and E. Ruckenstein, Biophys. Chem. **118**, 128 (2005).
[11] R. M. Mazo, J. Phys. Chem. B **110**, 24077 (2006).
[12] A. Ben-Naim, J. Chem. Phys. **63**, 2064 (1975).
[13] P. E. Smith, J. Phys. Chem. B **108**, 18716 (2004).
[14] S. Shimizu and C. L. Boon, J. Chem. Phys. **121**, 9147 (2004).
[15] S. Shimizu, W. M. McLaren, and N. Matubayasi, J. Chem. Phys. **124**, 234905 (2006).
[16] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith, Cell Biochem. Biophys. **50**, 1 (2008).
[17] A. Ben-Naim, Cell Biophys. **12**, 255 (1988).
[18] Y. Marcus, Monatsch. Chem. **132**, 1387 (2001).
[19] E. Tronel-Peyroz, J. M. Douillard, R. Bennes, and M. Privat, Langmuir **5**, 54 (1989).
[20] F. Chen and P. E. Smith, "Theory and computer simulation of solute effects on the surface tension of liquids," J. Phys. Chem. B (in press).
[21] S. Weerasinghe and P. E. Smith, J. Phys. Chem. B **118**, 10663 (2003).
[22] M. Aburi and P. E. Smith, J. Phys. Chem. B **108**, 7382 (2004).
[23] P. E. Smith, J. Phys. Chem. B **108**, 16271 (2004).
[24] P. E. Smith, Biophys. J. **91**, 849 (2006).
[25] J. P. O'Connell, Mol. Phys. **20**, 27 (1971).
[26] R. Ramirez, M. Mareschal, and D. Borgis, Chem. Phys. **319**, 261 (2005).
[27] E. Ruckenstein and I. Shulgin, Fluid Phase Equilib. **180**, 345 (2001).
[28] D. G. Hall, Trans. Faraday Soc. **67**, 2516 (1971).
[29] J. M. Schurr, D. P. Rangel, and S. R. Aragon, Biophys. J. **89**, 2258 (2005).
[30] I. L. Shulgin and E. Ruckenstein, J. Chem. Phys. **123**, 054909 (2005).
[31] P. E. Smith, J. Phys. Chem. B **110**, 2862 (2006).
[32] E. Matteoli and L. Lepori, J. Chem. Soc., Faraday Trans. **91**, 431 (1995).