Journal of Statistical Physics, Vol. 87, Nos. 5/6, 1997

A General Approach to Association Using Cluster Partition Functions

E. M. Hendriks,[1] J. Walsh,[2] and A. R. D. van Bergen[3]

Received December 5, 1995; final November 8, 1996

A systematic and fundamental approach to associating mixtures is presented. It is shown how the thermodynamic functions may be computed starting from a partition function based on the cluster concept such as occurs in chemical theory. The theory provides a basis for and an extension of the existing chemical theory of (continuous) association. It is applicable to arbitrary association schemes. Analysis of separate cases is not necessary. The assumptions that were made to allow the development were chosen such as to make the principle of reactivity valid. It is this same principle that links various theories: the chemical theory of continuous association, the lattice fluid hydrogen bonding model, and first-order perturbation theory. The equivalence between these theories in appropriate limits is shown in a general and rigorous way. The theory is believed to provide a practical framework for engineering modeling work. Binary interaction parameters can be incorporated. The association scheme is accounted for by a set of generic equations, which should facilitate robust implementation in computer programs.

KEY WORDS: Association; phase equilibrium; statistical mechanics; partition function; chemical engineering; equation of state; chemical equilibrium; cluster distribution.

1. INTRODUCTION

In fluids such as water, methanol, or acetic acid, clusters of molecules form due to strong intermolecular attractive forces. The lifetime of these clusters is long enough, depending on conditions, to be characterized by a variety

---

[1] Shell Research and Technology Center, Amsterdam, P.O. Box 38000, 1030 BN Amsterdam, The Netherlands; e-mail: hendrik2@siop.shell.nl.

[2] Shell Westhollow Technology Center, Houston, Texas; e-mail: walsh@shell.com.

[3] Shell International Exploration and Production B. V., P.O. Box 60, 2280 AB Rijswijk-ZH, The Netherlands; e-mail: vanbergena@ksepl.nl.

0022-4715/97/0600-1287$12.50/0

of spectroscopic techniques. The thermodynamic properties (e.g., boiling temperature, heat of vaporization) of associating fluids are profoundly different from nonassociating fluids, and mixtures of associating and non-associating fluids often form separate phases.

Three main approaches have been developed for understanding the thermodynamic properties of associating fluids.

In *perturbation theory* the intermolecular potential is split into a reference part and a short-range, directional part which gives rise to association. Using the grand canonical ensemble, an equation of state can be derived which resembles Mayer's virial equation of state for non-associating fluids, the difference being that, for an associating fluid, two sets of virial coefficients occur, one set containing the nonassociating reference potential and the other set containing the full potential. The equations do not converge at low temperatures where most of the molecules are involved in clusters. As shown by Locket(15) and Wertheim,(23, 24) when the density of monomers is split into free and bonded monomer densities, the low-temperature convergence of the series is repaired. Though perturbation theory is used, the bonding state of the system is incorporated at an early stage. Using this approach, Wertheim obtained accurate equations for the thermodynamic functions and for the cluster size distribution. His theory was later simplified by Gubbins and co-workers(1, 13) using only first-order thermodynamic perturbation theory (TPT-1), which allows tree-like clusters but no cycles. In this form, the theory has been incorporated into equations of state like SAFT (statistical associating fluid theory(1, 11)) which have been applied successfully to phase equilibria of associating mixtures.

In the second approach, usually referred to as *chemical theory*, the associating fluid is considered to be a mixture of various clusters (species) and free monomers. Chemical equilibrium constants are used to specify the relative densities of the various clusters,

$$K_j = \frac{\rho_{j+1}}{\rho_j \rho_1} \tag{1}$$

Heidemann and Prausnitz(8) used the van der Waals equation of state to account for the various physical (non-association) interactions. Donohue and coworkers(4, 12) used perturbation chain theory for these same interactions and obtained an equation of state in the form

$$\beta P = \sum \rho_i + \beta P^{(0)} \tag{2}$$

in which ρ_i is the number density of i-clusters and $P^{(0)}$ is the equation of state for a fluid of free monomers interacting with each other through the reference potential.

In the third approach,(17) referred to as the *lattice fluid hydrogen bonding model* (LFHB), the physical interactions between monomers are represented by a lattice fluid theory, which is modified to account for association by considering all ways of adding association (hydrogen) bonds between available donor and acceptor sites. The resulting model has been applied successfully to phase equilibria of associating mixtures as well.(16)

In this paper, we will use the cluster concept, as it occurs in the chemical equilibrium theory, to derive a partition function in which contributions from physical and association interactions are combined, but in a way different from LFHB. Based on assumptions of a nature similar to those made in other theories of association, we will derive all thermodynamic functions required to perform phase equilibrium calculations. This will turn out to be possible for arbitrary association schemes, and perhaps surprisingly, without the need to explicitly enumerate the possible clusters. The theory provides a fundamental basis for the chemical equilibrium theory of continuous association, which is thereby extended to arbitrary association schemes, i.e., no longer restricted to linear chains. We will also show that this theory is closely related to the other theories of association (TPT-1, LFHB), through the principle of equal reactivity(5) which also provides a link with the mathematical theory of gelation.(2, 5, 9) The expressions for the free energy of association in these theories are in fact equivalent, except for the determination of the association constants and their density dependence. This equivalence is shown here more generally and more rigorously than before.(3)

2. A GENERAL THEORY OF ASSOCIATION

2.1. Description of the System

We think of our system as a mixture of different substances (monomers) which may form clusters through association. Monomers—free and bonded—are labeled through a Roman index, i. Monomers that are not bonded are called free monomers. The number of monomers in the mixture is N_i, which variables constitute, besides volume V and temperature T, all the degrees of freedom of the system. The corresponding mole fractions $x_i = N_i/N$, where N is the total number of monomers, are referred to as *apparent* mole fractions.

On the monomers are functional groups or association sites. These sites, labeled A_i, may form weak bonds with similar sites on other molecules. The number of different sites of a given type on a monomer is denoted as $d(A_i)$. The association scheme specifies the possible bonds $\overline{A_i B_j}$ that may form between different sites on different monomers. In a quantitative theory, a set of association parameters is required for each possible bond, basically corresponding to the enthalpy and entropy of association.

Through formation of bonds, the monomers will combine to form clusters or "species." Clusters are labeled using Greek letters, which are supposed to represent the corresponding cluster in a unique way, also distinguishing isomers. A free monomer is considered to be a cluster as well, and is referred to as M_i. The number of monomers of type M_i in any cluster of type α is $k(\alpha, i)$ and the number of bonds of type B is $b(\alpha, B)$. Note that, due to the existence of isomers, these two numbers by themselves do not uniquely specify the cluster. The number of clusters of type α is denoted as n_α. This is also referred to as the cluster distribution. The equilibrium cluster distribution, i.e., the set $\{n_\alpha\}$, should be a well-defined (vector) function of the composition, volume, and temperature.

For simplicity, we refer to bond formation as association throughout this paper, also if the active sites are on different types of monomer, where the term cross-association or solvation is commonly used.

2.2. Partition Function

We start from the canonical partition function, which should represent a sum over all attainable physical states. We assume now that the clusters are well defined, i.e., that a criterion exists with which it is possible to decide, in principle, when a monomer is part of a cluster. The internal motions of any cluster will be captured by its internal partition function, which is assumed not to be affected by the interactions between different clusters. These latter interactions are accounted for in a partition function for a mixture of "frozen" clusters. In the complete partition function, all states that lead to the same frozen cluster distribution are collected together, giving rise to a combinatorial factor. The process is analogous to that applied in lattice fluid theory,(6) where the first step is to put the molecules on the vertices of a lattice in all the possible different ways, and in the second step the motion of the molecules about the lattice points is considered.

We first write the partition function for a fixed cluster distribution, i.e., we fix the set of numbers $\{n_\alpha\}$. This must then be summed over all possible cluster distributions, under the appropriate constraints of material balance. In the thermodynamic limit, the series may be replaced by the maximum term.

Following this procedure, we recognize a number of different factors in the partition function:

- A factor $\{1/\prod_i N_i!\}$ to account for the nondistinguishability of the monomer particles. The remaining cluster counting may be treated classically—as opposed to quantum mechanics—i.e., both the clusters and the association sites on them are to be treated as distinguishable.

- A factor for the number of ways to select the monomeric units to supply the material to build all the clusters. This number may be computed by first ordering all sets of monomers which may be done in $\prod_i N_i!$ ways. In each set, assign n_α groups of $k(\alpha, i)$ monomers to the clusters of type α. Any permutation of monomers within each group will not lead to a new arrangement. Also, any permutation of complete sets of monomers making up a cluster will not lead to a new arrangement. We have to divide by the numbers of these various types of permutations. We arrive at

$$\prod_i N_i! \prod_{\alpha, i} \frac{1}{\{k(\alpha, i)!\}^{n_\alpha}} \prod_\alpha \frac{1}{n_\alpha!} \tag{3}$$

- For each cluster α, a factor W_α, for the number of ways the cluster may be built from its building blocks, where all the monomers and all the association sites are to be considered distinguishable.

- For each cluster α, a factor representing its internal partition function ω_α; ω_α is a product of contributions from rotational and vibrational modes.

- A configurational integral for a fixed cluster distribution.

Putting all the factors together, we obtain for the complete partition function

$$\Omega = \max_{\{n_\alpha | \mathrm{m.b}\}} \prod_\alpha \frac{(V s_\alpha)^{n_\alpha}}{n_\alpha!} \left[V^{-N_c} \int e^{-\mathcal{W}/k_B T} \, d\mathbf{r}\{n_\alpha\} \right] \tag{4}$$

where in s_α we combined contributions from the internal partition functions and from the combinatorial factors:

$$s_\alpha = \frac{\omega_\alpha W_\alpha}{\prod_i k(\alpha, i)!} \tag{5}$$

As indicated before, to compute the actual value of the partition function, the maximum must be taken over all possible cluster distributions $\{n_\alpha\}$, satisfying the conditions of material balance (m.b)

$$\sum_\alpha n_\alpha k(i, \alpha) = N_i \tag{6}$$

N_c is an abbreviation for the total number of clusters,

$$N_c = \sum_{\alpha} n_{\alpha} \tag{7}$$

The factor V^{N_c} was separated out in Eq. (4), so that the configurational integral that appears between brackets in the partition function corresponds to the residual Helmholtz free energy for the mixture of frozen clusters. If the clusters would have no physical interactions, that factor would be unity. So, we may rewrite the partition function as follows:

$$\Omega = \max_{\{n_{\alpha} \mid \mathrm{m.b}\}} \prod_{\alpha} \frac{(V s_{\alpha})^{n_{\alpha}}}{n_{\alpha}!} \exp\left(-\frac{\tilde{A}^{\mathrm{res}}(V, T, \{n_{\alpha}\})}{RT}\right) \tag{8}$$

where the residual Helmholtz free energy is defined as the difference of the actual value and the value in the ideal gas state at the same volume V and temperature T,

$$\tilde{A}^{\mathrm{res}}(T, V, \{n_{\alpha}\}) = \tilde{A}(T, V, \{n_{\alpha}\}) - \tilde{A}^{\mathrm{i.g.}}(T, V, \{n_{\alpha}\}) \tag{9}$$

We may view this way of writing the partition function as one way of separating the association interactions from the cluster–cluster interactions. The justification of this should come from the different time and length scales on which the various phenomena take place. Apart from this, the theory is still very general. In the next sections, we will make a connection with both Wertheim's theory (TPT-1) and with the existing chemical theory of association, but to be able to do this we need to make further approximations.

2.3. Basic Assumptions and Approximations

We now list the set of assumptions that will put the theory into a form that can be applied practically in the computation of phase properties and equilibria for associating systems. These assumptions are:

1. Monomers of a given type contribute equally to the ideal-gas free energy of any cluster containing them.

2. Association bonds of a given type contribute equally to the ideal-gas free energy for any cluster in which they are present.

3. The physical interactions, excluding association, between the N_c clusters $\{n_{\alpha}\}$ may be represented by a residual free energy function of the form

$$\tilde{A}^{\mathrm{res}}(T, V, \{n_{\alpha}\}) = RTF(N_c, \mathbf{Q}, V, T) \tag{10}$$

where the quantities $\mathbf{Q} = (Q_1, Q_2, \ldots)$ depend only on composition, i.e., $\{n_\alpha\}$ and temperature, but not on volume. One may think of these as equation-of-state parameters. The precise form of the function F is immaterial, as long as it has the mathematical properties of an extensive, residual function.

4. The composition dependence of the equation-of-state parameters is through *linear* mixing rules:

$$Q_l = \sum_\alpha q_{l\alpha} n_\alpha \tag{11}$$

5. The pure cluster parameters $q_{l\alpha}$ are equal to the sum of the contributions from the monomeric units.

$$q_{l\alpha} = \sum_i k(\alpha, i)\, q_{l,\,\mathrm{M}_i} \tag{12}$$

6. The clusters contain no cycles. This is not an independent assumption; it can be derived from the other ones, at least for finite (microscopic) clusters, as explained in the next section.

Assumptions 1, 2, and 5 are consequences of the group contribution concept.(19) Assumptions 1 and 2 imply that any cluster internal partition function can be written as a product of factors

$$\omega_\alpha = \prod_i \omega_{\mathrm{M}_i}^{k(\alpha, i)} \prod_B P_B^{b(\alpha, B)} \tag{13}$$

The B-sum is over all possible bonds. The parameter P_B corresponds to an intrinsic probability of bond formation for bonds of the type B.

The assumption of linear mixing rules, which is also made in the development of Heidemann and Prausnitz,(8) is motivated by the chemical similarity of the clusters. Although the form of the equation of state is not important, the mixing rules are essential in the development. Later, we shall show that it will still be possible to introduce binary interaction parameters to account for the chemical differences between different monomers, albeit in a nonrigorous way.

A consequence of assumptions 4 and 5 is that the equation-of-state parameters $\mathbf{Q}$ depend on the monomer amounts only, i.e.,

$$Q_l = \sum_i N_i q_{l,\,\mathrm{M}_i} \tag{14}$$

This implies that these parameters do not depend on the detailed cluster distribution. They can be computed on the basis of the apparent mole numbers alone.

In order to be able to perform computations, we need to derive an expression for the free energy of the full, associating system. This we will do in three steps: first, we apply the condition of equilibrium to obtain an analytical expression for the cluster distribution; second, we use the principle of equal reactivity to determine the parameters in that expression; third, we substitute the result back into the partition function and compute the free energy by taking the logarithm.

2.4. The cluster Distribution in Equilibrium

In order to obtain the detailed cluster distribution, we must impose the condition for the maximum term, under the constraints of material balance, Eq. (6). There is one such constraint for each type of monomer. A convenient way to do this is to apply Lagrange's method, introducing a function

$$\mathscr{F} = \ln \Omega + \sum_i \lambda_i \left(\sum_\alpha n_\alpha k(i, \alpha) - N_i \right) \tag{15}$$

The quantities λ_i are Lagrange multipliers. If we differentiate $\mathscr{F}$ with respect to these quantities, we recover the constraints. Therefore, we may determine the maximum term by setting the derivative of $\mathscr{F}$ with respect to each n_α to zero. Using the full partition function and applying Stirling's formula to $\ln(n_\alpha!)$, we obtain the condition

$$\mathscr{G}_\alpha \equiv \ln \left(\frac{Vs_\alpha}{n_\alpha} \right) - \frac{\partial F}{\partial N_c} \frac{\partial N_c}{\partial n_\alpha} - \sum_l \frac{\partial F}{\partial Q_l} \frac{\partial Q_l}{\partial n_\alpha} + \sum_i \lambda_i k(\alpha, i) = 0 \tag{16}$$

The term involving the derivatives of F with respect to the equation-of-state parameters can be worked out further, making use of assumptions 4 and 5 of the previous section. The result is a sum over all monomer types i, in which each term is proportional to $\sum_i k(\alpha, i)$. These terms may simply be absorbed into the terms containing the Lagrange multipliers, leading to a mere redefinition (indicated by a prime below) of these multipliers. We obtain

$$n_\alpha = (V \mathscr{D} s_\alpha) \prod_i e^{\lambda_i' k(\alpha, i)} \tag{17}$$

where $\mathscr{D}$ is defined through the relation

$$-\ln \mathscr{D} = \left(\frac{\partial F(N_c, \mathbf{Q}, V, T)}{\partial N_c} \right)_{\mathbf{Q}, V, T} \tag{18}$$

The next step is to substitute the expression for s_α, Eq. (5) into Eq. (17), using Eq. (13) for the pure cluster internal partition functions. We combine each Lagrange multiplier with a monomer internal partition function into a new quantity z_i,

$$z_i \equiv \omega_{\mathrm{M}_i} e^{\lambda'_i} \tag{19}$$

to yield

$$n_\alpha = (VW_\alpha \mathscr{D}) \prod_B P_B^{b(\alpha, B)} \prod_i \frac{z^{k(\alpha, i)}}{k(\alpha, i)!} \tag{20}$$

Note that n_α occurs also on the right-hand side, implicitly via the dependence of $\mathscr{D}$ and z_i on N_c. The natural next step would be to determine the values of the quantities z_i by imposing all the material balance constraints. This, however, leads to infinite sums that can be evaluated in closed form only for special cases, such as, for example, the case of continuous linear association.$^{(8)}$ We will proceed in a different way, and derive a closed set of equations which completely fixes the cluster distribution, in the next section.

2.5. Application of the Principle of Equal Reactivity

Under the assumptions made, the formation of one particular bond of type B is accompanied by a weight factor $P_B/(V\mathscr{D})$ in the canonical distribution. Apart from the monomers and the association sites involved in the bonding, this factor is independent of any specific cluster property. This implies that the *principle of equal reactivity*$^{(5)}$ should hold, which states that the probability of bond formation between any two sites is independent of the actual shape and size of the particular clusters on which they lie. This enables us to express the cluster size distribution in an alternate way, in terms of the fractions X_{A_i}, of association sites of type A_i that are unbonded. Applying the principle of equal reactivity to monomers and dimers, we will be able to determine the parameters in the cluster distribution.

A consequence of the principle of equal reactivity is that the only (finite) clusters, built from a given set of monomers, that will have a finite probability of formation are those that have the least number of bonds. These are the tree-shaped clusters, with no cycles. The number of other types of clusters will vanish in the thermodynamic limit, since the formation probability for each additional bond must be inversely proportional to the size of the system.

The number of free monomers of type $\mathrm{M_i}$ must be equal to the number of monomeric units of that type, times the probability that all association sites on a given one are unbonded, i.e.,

$$n_{\mathrm{M_i}} = N_i \prod_{\mathrm{A}_i \in \mathrm{M_i}} X_{\mathrm{A}_i}^{d(\mathrm{A}_i)} \tag{21}$$

The product is over all types of association sites on the monomers. This number of free monomers must be equal to the value obtained from Eq. (20) for the same type of monomer. Equating the two yields a relation between the parameters z_i and the fractions of association sites that are unbonded

$$z_i = \frac{N_i}{V\mathscr{D}} \prod_{\mathrm{A}_i \in \mathrm{M_i}} X_{\mathrm{A}_i}^{d(\mathrm{A}_i)} \tag{22}$$

Next we consider a dimer, consisting of two monomeric units $\mathrm{M_i}$ and $\mathrm{M_j}$, connected through a bond of type $\overline{\mathrm{A}_i\mathrm{B}_j}$. We denote this particular dimer as "2." The idea is to compute the number of such dimers in two different ways.

As for the monomers, we apply a probabilistic argument. The number of dimers n_2 must equal the number of monomeric units N_i times the probability $\mathscr{P}$ that any such monomeric unit is part of a dimer of the indicated type. This latter probability can be determined as the product of conditional probabilities,

$\mathscr{P} =$ (Number of ways $d(\mathrm{A}_i)$ to select an A_i site on monomer $\mathrm{M_i}$)

× (Probability that all sites except the selected A_i are unbonded)

× (Probability that the selected A_i site is bonded)

× (Probability $p(\mathrm{A}_i, \mathrm{B}_j)$ that the selected,

bonded A_i site is bonded to a B_j site)

× (Probability that the remaining B_j sites are unbonded) (23)

Combination of these factors gives

$$n_2 = N_i\, d(\mathrm{A}_i)\, p(\mathrm{A}_i, \mathrm{B}_j) \left(\frac{1 - X_{\mathrm{A}_i}}{X_{\mathrm{A}_i} X_{\mathrm{B}_j}}\right) \prod_{\mathrm{A}'_i \in \mathrm{M_i}} X_{\mathrm{A}'_i}^{d(\mathrm{A}'_i)} \prod_{\mathrm{B}'_j \in \mathrm{M_j}} X_{\mathrm{B}'_j}^{d(\mathrm{B}'_j)} \tag{24}$$

The products occurring in this expression are over all types of sites.

As before, we equate this to the value obtained from Eq. (20) for the same type of dimer. We note that the combinatorial factor W_α is equal to the product of $d(A_i)$ and $d(B_j)$. If i and j correspond to different types of monomers, Eq. (20) gives

$$n_2 = V\mathscr{D}\, d(A_i)\, d(B_j)\, P(\overline{A_iB_j})\, z_i z_j \tag{25}$$

Using the expression obtained for z_i in Eq. (22), we find by equating the two expressions for n_2, after some rearrangement and cancellation of common factors,

$$N_j X_{B_j} d(B_j) \left(\frac{P(\overline{A_iB_j})}{V\mathscr{D}}\right) = \left(\frac{1-X_{A_i}}{X_{A_i}}\right) p(A_i, B_j) \tag{26}$$

This expression is also valid if i and j correspond to the same monomer, as the reader may verify. We sum this over all possible monomer types M_j and association sites B_j on them; the probabilities $p(A_i, B_j)$ sum up to unity. After some rearrangement, we obtain a system of equations for the fractions of association sites that are unbonded,

$$X_{A_i} = \frac{1}{1 + (V\mathscr{D})^{-1} \sum_j N_j \sum_{B_j} d(B_j)\, P(\overline{A_iB_j})\, X_{B_j}} \tag{27}$$

This is a closed set of equations in the unknowns X_{A_i}. It is identical to the equation appearing in Wertheim's theory (TPT-1) for the same quantities, except for the density-dependent association constants, which are related here to the cluster internal partition functions and to the residual free energy function F representing the cluster–cluster interactions. For certain simple association schemes, Eq. (27) may be solved in closed form, but generally a numerical procedure is required. Through the parameters z_i, the solution then completely fixes the cluster distribution.

2.6. Thermodynamic Functions

2.6.1. Helmholtz Free Energy. We are now in the position to compute the Helmholtz free energy for the complete system, and from that, all thermodynamic functions. To do this, we substitute Eq. (10) into the expression for the partition function (8) and take the logarithm to obtain

$$\frac{A}{RT} = -\ln \Omega = F - N_c + \sum_\alpha n_\alpha \ln\left(\frac{n_\alpha}{Vs_\alpha}\right) \tag{28}$$

The next step is to substitute the expression for the cluster distribution (20) into the argument of the logarithm and use the material balance constraints, Eqs. (6) and (7). This leads to

$$\frac{A}{RT} = -\sum_i N_i \ln(\omega_{\mathrm{M}_i}) + \sum_i N_i \ln z_i + F - N_c + N_c \ln \mathscr{D} \tag{29}$$

Next, we use the expressions for z_i, Eq. (22), and $\mathscr{D}$, Eq. (18) to get

$$\frac{A}{RT} = \sum_i N_i \left\{ \ln\left(\frac{N_i}{V\omega_{\mathrm{M}_i}}\right) + \sum_{\mathrm{A}_i \in \mathrm{M}_i} d(\mathrm{A}_i) \ln(X_{\mathrm{A}_i}) \right\} + F - N_c + (N - N_c)\frac{\partial F}{\partial N_c} \tag{30}$$

All quantities in this formula, except the number of clusters N_c, can be computed using equations given earlier. We still need to find a relation between N_c and the state variables. This is achieved by realizing that, for a system of clusters without cycles, the total number of clusters is equal to the total number of monomers minus the total number of bonds, $N_c = N - N_b$. The total number of bonds N_b in turn must be equal to half the total number of bonded association sites, since each bond involves two such sites. This leads to

$$\frac{N_c}{N} = 1 - \frac{1}{2} \sum_j x_j \sum_{A_j \in \mathrm{M}_i} d(\mathrm{A}_j)(1 - X_{\mathrm{A}_j}) \tag{31}$$

Thus the value of total number of clusters N_c may be computed from the solution of Eq. (27).

Note that the combinatorial factors W_α do not appear in the final result for the free energy, which is fortunate, since in general they cannot be computed analytically. It is not needed even to enumerate the clusters. However, some care is required when applying Eq. (31). If the monomers contain more than two functional groups, then it is possible that at a finite value of the density one cluster of macroscopic size will form. This "sol–gel" phase transition is predicted by the *mathematical theory of gelation*(2, 5, 9) and it occurs if the numbers W_α grow fast enough with increasing cluster size. Beyond the gelation threshold, the assumption that the clusters have no cycles cannot hold, and therefore Eq. (31) may not be applied.

Known results of chemical theory correspond to special cases of Eq. (30). For example, for APACT,(3, 12) where linear chains are considered, Eq. (27) can be solved in closed form. The mixing rules in this model are such that F does not explicitly depend on N_c, which leads to an equation-of-state of the form of Eq. (2).

2.6.2. Equation of State. We derive the equation of state by differentiating, as usual, the free energy with respect to volume, keeping T and $\{N_i\}$ constant. We do not use the final expression for the free energy, but go back one stage, to Eq. 29. We will use the following auxiliary relation for the derivative of N_c:

$$\left(\frac{\partial N_c}{\partial V}\right)_{\{N_i\},T} = N_c\left(\frac{\partial \ln(V\mathscr{D})}{\partial V}\right)_{\{N_j\},T} + \sum_i N_i\left(\frac{\partial \ln z_i}{\partial V}\right)_{\{N_j\},T} \tag{32}$$

which relation is obtained by substituting Eq. (20) into the definition of N_c performing the differentiation under the summation sign, and applying the material balance constraints, Eq. (6). Using this relation in the derivative of the free energy (29), we find that the terms involving the parameters z_i cancel, and we obtain

$$\frac{p}{RT} = \frac{N_c}{V} - \left(\frac{\partial F}{\partial V}\right)_{\{N_j\},T} + \left(\frac{\partial F}{\partial N_c}\right)_{Q,V,T}\left(\frac{\partial N_c}{\partial V}\right)_{\{N_j\},T} \tag{33}$$

The last two terms may be combined to yield the final result for the equation of state

$$\frac{P}{RT} = \frac{N_c}{V} - \left(\frac{\partial F}{\partial V}\right)_{N_c,Q,T} \tag{34}$$

This result is both simple and appealing. Observe that the *form* of the equation of state is equal to that of a frozen system of clusters without association. For the full, associating system, N_c will, however, vary with temperature and density, and we have given closed sets of equations governing this dependence. The result is consistent with other theories when applied to corresponding examples, but it is more general. Note that to compute the pressure, it is not needed to compute density derivatives of X_A.

2.6.3. Chemical Potentials. The computation of the chemical potential of component i is completely analogous to the derivation of the equation of state. We merely state the result here:

$$\frac{\mu_i}{RT} = \ln\left(\frac{N_i}{V\mathscr{D}\omega_{M_i}}\right) + \sum_I \left(\frac{\partial F}{\partial Q_I}\right)_{N_c}\frac{\partial Q_I}{\partial N_i} + \sum_{A_i \in M_i} d(A_i)\ln X_{A_i} \tag{35}$$

As should be the case, this value is equal to λ_i, the Lagrange multiplier associated with conservation of monomers M_i. This may be explicitly verified on the basis of Eqs. (16)–(19). Note that to compute the chemical potentials—or the fugacity coefficients—it is not needed to compute density derivatives and primitives of X_A.

2.6.4. Binary Interaction Parameters. The theory can be used as a basis for modelling, which involves, besides a specific choice for the residual function *F*, the identification of all the parameters in the theory and their determination from experimental data. It is well known(18) that in a practical engineering model one needs to introduce binary interaction parameters in the equation-of-state parameters to account for nonconformal interactions between substances, caused by differences in size, shape, and polarity, even for systems where association does not play a role. Our assumption of linear mixing rules seems to preclude this possibility.

There is, however, a way to introduce binary interaction coefficients with some foundation. They are typically required in an energy parameter, as follows:

$$a = \sum_{\alpha} \sum_{\beta} n_{\alpha} n_{\beta} \sqrt{a_{\alpha} a_{\beta}} \, (1 - k_{\alpha\beta}) \tag{36}$$

In mixtures of many chemically similar components, the majority of the interaction parameters required will be zero or very small. If they were exactly zero, assumption 4 of Section 2.3 would hold, and a could be expressed in terms of monomer properties alone. So it is not unreasonable that a may be approximated using fewer interaction parameters:

$$a = \sum_{i} \sum_{j} N_i N_j \sqrt{a_{\mathrm{M}_i} a_{\mathrm{M}_j}} \, (1 - k_{\mathrm{M}_i, \mathrm{M}_j}) \tag{37}$$

Note that for systems of inert (nonassociating) components, this reduces to the van der Waals classical mixing rule which for many of those systems works adequately.

Additional support to this idea is provided by the *truncated spectral method*, described in ref. 10. Using this method, using the property that the interaction parameters in Eq. (36) are generally small, the quadratic mixing rule can be effectively transformed to a small number of linear mixing rules, so that the whole theory described here may be applied. At the end, the transformation is applied in the reverse direction, leading to the approximation Eq. (37).

If a simple model such as the Soave–Redlich Kwong equation of state(22) is used to model binary systems where association plays a role, a strong composition dependence has to be introduced in binary interaction parameters such as appear in Eq. (37) to be able to fit the data, and experience shows that this is possible with limited success only. If association is accounted for explicitly, on the other hand, a small constant interaction parameter should suffice, since the (strong) composition dependence caused by association is already present in the model, here through

Eq. (27). Examples showing this clearly can be found for example, in refs. 14, 21, and 25, where an equation of state is used that consists of a very simple "physical" part combined with, essentially, Wertheim's equation, Eq. (27). The authors show that phase diagrams of binaries involving water, hydrocarbons, and alcohols can be represented using a single, binary interaction parameter, and that it is possible to represent vapor–liquid and liquid–liquid data simultaneously using a single set of parameters.

3. RELATION TO EXISTING THEORIES

In order to establish the validity and the value of the results of this paper we want to make a link with existing theories. We will show that and why the theory is consistent with the three types of theories mentioned in the introduction. Our theory may be viewed as a bridge between these other theories, and also as an extension of the current chemical theory of association to include arbitrary association schemes.

3.1. Chemical Theory

The results of this paper are equivalent to the results of Heidemann and Prausnitz,(8) for the cases to which these latter results can be applied, but the theory is not restricted to specific association schemes such as linear chains. This can be seen by writing the left-hand side of Eq. (16) as

$$\mathscr{G}_\alpha = \ln\left(\frac{Vs_\alpha}{n_\alpha}\right) + \frac{\tilde{\mu}_\alpha^{\text{res}}}{RT} + \sum_i \lambda_i k(\alpha, i) \tag{38}$$

where $\tilde{\mu}_\alpha^{\text{res}}$ is the residual chemical potential of the cluster α, as computed from the residual free energy for a mixture of "frozen" clusters. Note that this chemical potential does not contain contributions from association. Now consider a "virtual reaction" $\alpha + \beta \rightarrow \gamma$ by which two clusters produce a third one. This corresponds to a change

$$\mathscr{G}_\alpha + \mathscr{G}_\beta - \mathscr{G}_\gamma = \ln\left(\frac{Vs_\alpha Vs_\beta n_\gamma}{Vs_\gamma n_\alpha n_\beta}\right) - \frac{\tilde{\mu}_\alpha^{\text{res}}}{RT} - \frac{\tilde{\mu}_\beta^{\text{res}}}{RT} + \frac{\tilde{\mu}_\gamma^{\text{res}}}{RT} = 0 \tag{39}$$

which must be zero due to Eq. (16). This equation may be rewritten in terms of fugacity coefficients and cluster mole fractions $x_\alpha \equiv n_\alpha / N_c$ as

$$\frac{\tilde{\phi}_\gamma x_\gamma}{p \tilde{\phi}_\alpha x_\alpha \tilde{\phi}_\beta x_\beta} = \frac{s_\gamma}{RT s_\alpha s_\beta} \tag{40}$$

This equilibrium condition is the starting point of the chemical theory of association. Note that reactions between arbitrary clusters, satisfying the conditions of material balance, are allowed. The effective equilibrium constant is expressed in terms of the number s_α, which contain the pertinent internal partition functions and the combinatorial factors. The cluster distribution Eq. (20) is the solution to the infinite set of equations (40).

3.2. First Order Perturbation Theory

The correspondence with Wertheim's theory (TPT-1) becomes clear from a comparison of the TPT-1 equations with Eq. (27) for the fractions of sites that are unbonded and Eq. (30) for the free energy. The two sets of equations are equivalent, except for the determination of the density-dependent association parameters, which are related to integrals involving the intermolecular potential in Wertheim's theory.

We note that the form of Eq. (27) can be derived in a very simple way from the principle of equal reactivity alone, as follows. This principle implies the existence of a type of chemical equilibrium, not between species but between free and bonded pairs of association sites,

$$\mathrm{A}_i + \mathrm{B}_j \rightleftarrows \overline{\mathrm{A}_i\mathrm{B}_j} \tag{41}$$

In terms of number concentrations, denoted by square brackets, an equilibrium equation may be written for each pair of association sites capable of bonding

$$\frac{[\overline{\mathrm{A}_i\mathrm{B}_j}]}{[\mathrm{A}_{i0}][\mathrm{B}_{j0}]} = K(\mathrm{A}_i\mathrm{B}_j) \tag{42}$$

in which the equilibrium constant $K(\mathrm{A}_i\mathrm{B}_j)$ is specific for the types of association sites involved, but does not depend on their concentrations. The notation A_{i0} denotes an unbonded association site of the type A_i. We now introduce the *fraction* of A_i sites bonded to sites of type B_j, $X(\mathrm{A}_i, \mathrm{B}_j)$. The corresponding *concentrations* may be expressed then as

$$[\overline{\mathrm{A}_i\mathrm{B}_j}] = x_i \rho X(\mathrm{A}_i, \mathrm{B}_j) \qquad \text{and} \qquad [\mathrm{A}_{i0}] = x_i \rho X_{\mathrm{A}_i} \tag{43}$$

Substitution of these expressions into Eq. (42) and multiplication of both sides by a common factor yields

$$X(\mathrm{A}_i, \mathrm{B}_j) = \rho x_j K(\mathrm{A}_i\mathrm{B}_j)\, X_{\mathrm{A}_i} X_{\mathrm{B}_i} \tag{44}$$

If we sum both sides of this equation over all possible monomeric units j and association sites B_j on them and use the condition of overall conservation of A_i sites

$$\sum_j \sum_{B_j} X(A_i, B_j) = 1 - X_{A_i} \tag{45}$$

we obtain, after simple rearrangement, Eq. (27).

3.3. Lattice Fluid Hydrogen Bonding Model

In the lattice fluid hydrogen bonding model (LFHB)$^{(17)}$ one also starts from the canonical partition function, in which the contributions from the association interactions are separated from the remaining interactions. Interestingly, this is applied in the reverse order compared to our approach: the physical interactions are accounted for first, using a lattice fluid concept; then the association interactions are added by randomly placing bonds between sites. In this second step, the principle of equal reactivity is applied implicitly. As a result, the contribution to the free energy due to association is equal to that in the other theories, except for the precise value of the (density-dependent) association constants, as we will show now.

Panayiotou and Sanchez derived the free energy of association of the mixture by counting the number of ways a fixed number of bonds can be distributed over the functional groups. To avoid a notation overloaded with indices, we suppress the subscripts i and j. The equations are, however, valid for pure components and mixtures. The *a priori* partition function can then be written as

$$\Omega^{\text{assoc}} = \prod_A \frac{N_A!}{N_{A0}!} \prod_B \frac{N_B!}{N_{B0}!} \prod_A \prod_B \frac{(p_{AB}/V)^{N_{AB}}}{N_{AB}!} \tag{46}$$

where a formal distinction between acceptor (A) and donor (B) sites is made. N_A is the total number of (acceptor) sites of type A, and N_{A0} the number of these that are not bonded (m.m. for the donor sites). N_{AB} is the number of $\overline{AB}$ bonds. The quantity p_{AB} is related to the probability of bond formâtion. It is divided by the total volume, which expresses the probability of encounter in an ideal gas. In a real fluid, p_{AB} may itself depend more weakly on density.

At given values of the external conditions of total volume, temperature, and composition (number of N_A and N_B sites), the actual value of the partition function is to be obtained from maximization with respect to the free variables, viz. N_{A0}, N_{B0} and N_{AB}, under the appropriate constraints. This maximization procedure may be carried out quite simply by comparing the value of Ω^{assoc} before and after a particular "virtual reaction" of the kind displayed in Eq. (41), which must be equal at the maximum. The following condition is obtained:

$$\frac{N_{AB}}{N_{A0}N_{B0}}=\frac{p_{AB}}{V} \tag{47}$$

which is equivalent to Eq. (42) if the identification $K(AB)=p_{AB}$ is made. From what we showed earlier in Section 3.2, Eq. (47) is then equivalent to Eq. (27).

The free energy of association is obtained from the standard relation

$$\frac{A^{\text{assoc}}}{NkT}=-\frac{1}{N}\ln \Omega^{\text{assoc}} \tag{48}$$

This can be worked out, using Stirling's identity again, and applying the conditions expressed in Eq. (46) together with the conservation laws for association sites

$$\sum_{B} N_{AB}=N_A-N_{A0}, \qquad \sum_{A} N_{AB}=N_B-N_{B0} \tag{49}$$

After some algebra, a rather simple form for the free energy of association is obtained,

$$\frac{A^{\text{assoc}}}{NkT}=\sum_{j}\frac{N_J}{N}\ln\left(\frac{N_{J0}}{N_J}\right)+\frac{1}{2}\sum_{J}\frac{N_J}{N}\left(1-\frac{N_{J0}}{N_J}\right) \tag{50}$$

where the summation is over all association sites J (rather than types of sites), donor and acceptor. Remembering that we suppressed the indices i, indicating monomeric units, it is easy to see that the contribution from association to the free energy is equal to that in the other theories except for the values of the equilibrium constants. To make the connection with Eq. (30), in the latter equation F must be set to zero, Eq. (31) substituted, and the ideal gas reference terms removed. The result is identical to Eq. (50).

4. CONCLUSION

We have presented a systematic approach to associating mixtures, which, while being similar in spirit to the hydrogen bonding lattice theory of Panayiotou and Sanchez,(17) provides a basis to and an extension of the chemical theory of continuous association as formulated by Heidemann and Prausnitz(8) and at the same time constitutes a bridge between these theories and perturbation theory, TPT-1.

A strong and practical point is that the method may be applied for *arbitrary* continuous association schemes, i.e., there may be different monomers with an arbitrary number of association sites on them. The association scheme is accounted for in a general way through Eq. (27) so that it is not necessary to analyze each case separately, such as in ref. 7 and 17. The assumptions that had to be made to allow the development were chosen such as to make the principle of reactivity(5) valid. In the case where the cluster–cluster interactions are ideal, the probability of bond formation is independent of the clusters on which the participating active sites are. The result of the physical interactions is that these probabilities change, but all with the same factor, so that they are still independent of the clusters. This is a result of the special mixing and combining rules used.

It is this same principle of equal reactivity that links the various theories. The equivalence between these theories in appropriate limits has been discussed before(3) but not to the same generality. In ref. 3 it was observed that equivalent or similar expressions result for a number of corresponding cases. In ref. 7 the question was raised if the equivalence would be valid also for the case of multifunctional association. We have now shown that the answer is affirmative, that it includes mixtures, association, and solvation, and that the equivalence becomes exact in the limit of ideal cluster–cluster interactions. This, of course, also means that the chemical potentials must have the same value, despite the apparently very different expressions.(7)

We have displayed systematically what the assumptions and the various steps are. The canonical partition function, Eq. (4), which captures the phenomenon of association is very general. It will be interesting to investigate which of the subsequent assumptions may be relaxed or removed without losing the capability of computation. Of course, the assumptions are very crude and cannot rigorously hold for real fluids. However, the success of existing equation of state theories of (continuous) association(1, 3, 8, 17, 20) shows that this type of theory can be applied with success to the problem of representation of phase equilibria and corresponding data reduction.

We believe that the theory provides a framework for engineering modeling work. Our suggestion for the inclusion of binary interaction parameters, in fact in the standard way, should increase the applicability to practical problems, as is seen in examples in the literature.(21, 25) The fact that the association scheme is accounted for by a set of equations in generic form, Eq. (27) will facilitate robust and generic implementation in computer programs. Cumbersome density derivatives and integrals applied to the association terms, which seem to be required in the original formulation of the chemical theory, are not needed.

REFERENCES

1. W. G. Chapman, K. E. Gubbins, G. Jackson, and M. Radosz, New reference equation of state for associating liquids, *Ind. Eng. Chem. Res.* **29**:1709–1721 (1990).
2. P. C. J. van Dongen and M. H. Ernst, *J. Stat. Phys.* **37**:301 (1984).
3. I. G. Economou and M. D. Donohue, Chemical, quasi-chemical and perturbation theories for associating fluids, *AIChE J.* **37**:1875 (1991).
4. I. G. Economou and M. D. Donohue, *Ind. Eng. Chem. Res.* **31**:2388 (1992).
5. P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, New York, 1953).
6. A. J. Guggenheim, *Mixtures* (Clarendon Press, Oxford, 1952).
7. R. B. Gupta and K. P. Johnston, Lattice fluid hydrogen bonding model with a local segment density, *Fluid Phase Equilibria* **99**:135–151 (1994).
8. R. A. Heidemann and J. M. Prausnitz, A van-der-Waals-type equation of state for fluids with associating molecules, *Proc. Natl. Acad. Sci. USA* **73**:1773 (1976).
9. E. M. Hendriks, Analytic expression for the distribution of gel species in spatially homogeneous systems, *J. Phys. A: Math. Gen.* **18**:L579–L583 (1985).
10. E. M. Hendriks and A. R. D. van Bergen, Application of a reduction method to phase equilibria equations, *Fluid Phase Equilibria* **74**:17–34 (1992).
11. S. H. Huang and M. Radosz, Equation of state for small, large, polydisperse, and associating molecules, *Ind. Eng. Chem. Res.* **29**:2284–2294, (1990).
12. G. D. Ikonomou and M. D. Donohue, Thermodynamics of hydrogen-bonded molecules: The associated perturbed anisotropic chain theory, *AIchE J.* **32**:1716 (1986).
13. G. Jackson, W. G. Chapman, and K. E. Gubbins, Phase equilibrium of associating fluids: Spherical molecules with multiple bonding sites, *Mol. Phys.* **65**:1 (1988).
14. G. M. Kontogeorgis, E. C. Voutsas, I. V. Yakoumis, and D. P. Tassios, An equation of state for associating fluids, *Ind. Eng. Chem Res.* **35**:4310 (1996).
15. A. M. Lockett III, *J. Chem. Phys.* **72**:4822, (1980).
16. D. Missopolinou and Panayiotou, Thermodynamic systems of (oxy)alcohols and (poly)ethers, *Fluid Phase Equilibria* **110**:73–88 (1995).
17. C. Panayiotou and I. C. Sanchez, Hydrogen bonding in fluids: An quation-of-state approach, *J. Phys. Chem* **95**:10090–10097 (1991).
18. J. M. Prausnitz, R. N. Lichtenthaler, and E. Gomez de Azevedo, *Molecular Thermodynamics of fluid phase equilibria* (Prentice-Hall, Englewood Cliffs, New Jersey, 1986).
19. R. C. Reid, J. M. Prausnitz, and B. E. Poling, *The Properties of Gases and Liquids* (McGraw-Hill, New York, 1987).
20. J. Suresh and E. J. Beckman, Prediction of liquid–liquid equilibria in ternary mixtures from binary data, *Fluid Phase Equilibria* **99**:219–240 (1994).
21. E. C. Voutsas, G. M. Kontogeorgis, I. V. Yakoumis, and D. P. Tassios, Correlation of liquid–liquid equilibria for alcohol/hydrocarbon mixtures using the CPA equation of state, *Fluid Phase equilibria*, submitted.
22. S. M. Walas, *Phase Equilibria in Chemical Engineering* (Butterworth, Stoneham, Massachusetts, 1985).
23. M. S. Wertheim, *J. Stat. Phys.* **35**:19 (1984).
24. M. S. Wertheim, *J. Stat. Phys.* **42**:459 (1986).
25. I. V. Yakoumis, G. M. Kontogeorgis, E. C. Voutsas, and D. P. Tassios, An equation of state for associating fluids: Part ii. Vapour–liquid equilibria for alcohol/hydrocarbon mixtures, *Fluid Phase Equilibria*, submitted.