PHYSICA A

Criticality in the hard-sphere ionic fluid

Yan Levin*, Michael E. Fisher

Institute for Physical Science and Technology, The University of Maryland, College Park, Maryland 20742, USA

Abstract

A physically based mean-field theory of criticality and phase separation in the restricted primitive model of an electrolyte (hard spheres of diameter a carrying charges $\pm q$) is developed on the basis of the Debye–Hückel (DH) approach. Simple DH theory yields a critical point at $T^* \equiv k_B T a / q^2 = 1/16$, which is only about 15% above the best recent simulation estimates $(T^*_{c,\text{sim}} = 0.052–0.056)$ but a critical density $\rho^*_c \equiv \rho_c a^3 = 1/64\pi \simeq 0.005$ that is much too small $(\rho^*_{c,\text{sim}} = 0.023–0.035)$. Allowing for hard-core exclusion effects reduces these values slightly. However, correction of the DH linearization of the Poisson–Boltzmann equation by including *pairing* of $+$ and $-$ charges improves ρ^*_c significantly. Bjerrum's theory of the (required) association constant is revisited critically; Ebeling's reformulation is strongly endorsed but makes negligible numerical difference at criticality and below. The nature and size of the associated, dipolar ion pairs is examined quantitatively and their solvation free-energy in the residual fluid of free ions is calculated on the basis of DH theory. This contribution to the total free energy proves crucial and leads to a rather satisfactory description of the critical region. The temperature variation of the vapor pressure and of the density of neutral dipolar pairs correlates fairly well with Gillan's numerical cluster analysis. Possible improvements to allow for larger ion clusters and to better represent the denser ionic liquid below criticality are discussed. Finally, the replacement of the DH approximation for the ionic free energy by the mean spherical approximation is studied. Reasonable critical densities are generated but the MSA critical temperatures are all 40–50% too high; in addition, the predicted density of neutral clusters seems much too low near criticality and, along with the vapor pressure, appears to decrease too rapidly by an exponential factor below T_c.

1. Introduction and overview

The theory of ionic systems has a long history going back to the nineteenth century. A major advance in understanding, particularly in regard to solutions of strong electrolytes, was achieved by Debye and Hückel (DH) whose fundamental papers

* Now at Instituto de Fisica, Universidade Federal do Rio Grande do Sul, Caixa Postal 15051, CEP 91501-970 Porto Alegre, RS Brazil.

were published over 70 years ago. Our main concern here will be with the equilibrium thermodynamic properties which, at low densities, rest primarily on the DH discovery of screening by counterions [1]. (For good modern accounts of the theory see Falkenhagen and Ebeling [2] and McQuarrie [3].) The so-called DH *limiting laws*, which are universally valid at low densities, being independent of the size and structure of the ions, etc., are widely appreciated: less well recognized, however, is that the *full* DH theory [1–3] makes allowance for the finite sizes of the ions which are modelled, as we will do here, by hard spheres of uniform diameter a.

Indeed, one of our present purposes is to demonstrate that the full DH theory has merits unsuspected until very recently [4–6], in that it predicts gas–liquid phase separation and an associated critical point. Indeed, when supplemented by Bjerrum's [7] important concept of ion-pairing [2, 8–12] introduced only three years after the DH theory, the critical parameters are found to agree surprisingly well [4–6, 10, 11] with the best recent Monte Carlo estimates for the *restricted primitive model* (RPM) in which, in a medium of dielectric constant D, half the hard spheres carry charges $+q$ and a half charges $-q$ [5, 6, 13–17]. Nevertheless, as we show below, further specific physical ingredients must be incorporated in the theory in order to obtain a fully reasonable account of the phase separation and critical point at a first, mean-field type, level of approximation [5, 6].

A fresh impetus for the study of criticality in electrolytes has resulted from recent work by Pitzer and collaborators [18, 19]. They found that the critical behavior of a number of organic salts in appropriate solvents, most notably, tri-ethyl-n-hexylammonium triethyl-n-hexylboride $(C_2H_5)_3C_6H_{13}N^+(C_2H_5)_3C_6H_{13}B^-$ in diphenyl ether, $(C_6H_5)_2O$, appeared to be closely classical (or van der Waals-like) with a critical exponent [19] $\beta \simeq 1/2$, and, in subsequent experiments [20, 21], $\gamma \simeq 1$ and $\nu \simeq 1/2$. This is in contradiction to the expected Ising-type values [6] $\beta \simeq 0.32$, $\gamma \simeq 1.24$, $\nu \simeq 0.63$ which, however, do satisfactorily characterize other electrolytes such as tetrapentyl ammonium bromide, $(C_5H_{11})_4N^+Br^-$, in water [22] and nonaqueous solutions of other salts [23]. Overviews of the experimental situation have been presented [18, 24, 25]: two seemingly distinct classes of ionic solution exist which have been termed *solvophobic* and *Coulombic*, respectively. In the former, typically characterized by large solvent dielectric constants, the phase separation is mainly driven, as in ordinary nonelectrolyte solutions, by the "dislike" of the ionizable solute for the solvent. The principal interactions are of short range and standard Ising critical exponents are observed. By contrast, in Coulombic systems appropriate solvents of low dielectric constant allow strong electrostatic interactions to drive the phase separation. (In a chemical setting one always has in mind energy scales set by room temperature but must note that the product, Dk_BT is the relevant corresponding-states parameter for electrolytes [9, 12].) The dominating Coulombic forces evidently yield near-classical critical behavior.

The theoretical issues raised by this observational dichotomy and previous arguments pertaining to it [26], have been reviewed recently [6, 27]. It is certainly conceivable that the special long-range attractive/repulsive nature of the Coulombic

interionic forces leads to a universality class of critical behavior which is truly *non*Ising in nature [6]. Establishing that represents a prime theoretical challenge. On the other hand [6], it seems more likely at this juncture that the true asymptotic critical behavior when $t \equiv (T - T_c)/T_c \to 0$ is always of Ising character but that a *crossover* from nonasymptotic, close-to-classical behavior occurs on a scale $t_\times \equiv |T_\times - T_c|/T_c$ that, in the Coulombic systems, is too small to be experimentally accessible [5, 6]. This viewpoint is supported experimentally by the sodium–ammonia systems [28] which seem to be described by such a scenario with $t_\times \simeq 0.6 \times 10^{-2}$. More recent experiments in which the dielectric constant of the solvent is "tuned" [23] also suggest values of $t_\times \simeq 10^{-3}$ to 10^{-2}. If, however, this is the correct conclusion, a good theoretical justification is needed *and* a satisfactory theory must clearly yield an understanding of why $t_\times$ can be as small as $\sim 10^{-4}$ for systems like Pitzer's salt [19, 21]. Examination of the data [23, 28] also suggests that such Coulombic, classical-to-Ising crossovers may be significantly *sharper* than those expected on the basis of previous renormalization group crossover calculations [29, 30]; if sustained, that feature also would demand explanation.

In approaching these theoretical challenges the renormalization group (RG) approach naturally comes to mind because of its successes in resolving many different classes of criticality and multicriticality, particularly in magnetic systems [31]. However, in most cases such RG analyses have been founded on mean-field-based Landau–Ginzburg–Wilson effective Hamiltonians which, it could normally be argued, incorporated all the most essential physical degrees of freedom and the couplings between them. The adequacy of such underlying mean-field theories can normally be gauged by their characterization of the phases arising and the principal transitions they undergo [32], as well as by the resulting estimates of critical temperatures, etc., which are typically correct to within, say, 20 to 30%. The original van der Waals theory for criticality in ordinary molecular fluids with (comparatively) short-range interactions provides a prime example of such a mean-field theory: both reduced critical temperature and density are well represented. Of course, mean-field theories almost invariably predict classical critical exponents owing to their neglect or inadequate treatment of fluctuations: but that is what an RG treatment aims to correct (below the upper critical dimensionality).

Our first aim, therefore [5, 6], has been to uncover a simple mean-field-type treatment of, in particular, the restricted primitive model in the hope of shedding light on the essential physical mechanisms leading to phase separation and criticality in Coulombic systems. While the relative accuracy of our critical-point predictions will be of interest as an overall gauge of reliability, we stress that the transparent elucidation of the various physical effects and their comparative importance is a primary objective. It is, perhaps, appropriate to remark that, for the present purposes, the stress on the hypernetted chain (HNC) integral equation for electrolytes found in much of the literature, is inappropriate since it is now clear (although perhaps not as widely appreciated as it should be) that these and other integral equations (PY, YBG, etc.) give accounts of the correlation functions and derived thermodynamics in

the critical region that are quite pathological (unlike van der Waals or similar mean-field theories) [33]. The mathematical reasons for the unphysical behavior in the critical region are, at present, almost totally obscure [33]. Thus we personally cannot regard the current standard integral equation formulations as in any way physically transparant or truly informative.

The balance of this article is organized as follows: Section 2 first sets out in systematic fashion a variational formulation of *multi*component thermodynamics for chemically reacting fluids that is suitable for deriving quite general [34] mean-field-type approximations. This is needed because, in addition to the two species of positive and negative ions, the theories to be developed invoke distinct bound species, in particular, neutral dipolar pairs in chemical equilibrium with the dissociated or free ions. Beyond the ideal gas terms for each species, various simple approximate representations of the multicomponent hard-core ion–ion, and ion–cluster interactions are briefly discussed [35, 36].

The restricted primitive model for ionic fluids is taken up in Section 3 and current estimates of its critical parameters are summarized [14–17, 37]. The full Debye–Hückel theory, which is "linearized" *and* allows for the hard-core ion–ion exclusion [1–3, 38], is recapitulated concisely for subsequent reference and development. Section 4 demonstrates the existence of a critical point and the associated coexisting fluid phases [4, 5]: see Fig. 1, below. The properties of the phase boundaries and the critical behavior, which is fully classical, are elucidated in detail [5]. The effects of supplementing the pure DH theory with hard-core terms in the total free energy are looked at in Section 4.3.

Bjerrum's proposal [7] of allowing for *ideal* neutral dipole ion pairs is studied in Section 5 accepting, in the first instance, his original ansatz for the association constant, $K(T)$, which plays a vital role. The critical temperature and inverse Debye length κ_c are unchanged from pure DH theory but the critical density becomes more realistic [4, 5]; however, the coexistence curve assumes an unphysical "banana" shape [5, 45]: see Fig. 2, below. Inclusion of hard-core terms (Section 5.3) makes matters worse! The "physical origin" of ionic phase separation is discussed in Section 5.4 [49–51].

The appropriate assignment of the association constant $K(T)$ is critically reviewed in Section 6 and Ebeling's theory [2, 8, 11, 46, 47, 53–60] is endorsed [5]; in the critical region and below, however, Ebeling's expression for $K(T)$ differs in magnitude negligibly from Bjerrum's ansatz [5]. The specific nature of dipolar ionic clusters at low temperatures is considered in some detail in Section 6: see Fig. 4 below [61]. On this basis the solvation free energy of a dipolar pair in the sea of free ions is analyzed in Section 7 using the DH theory [5, 48, 62]. This dipole–ionic (DI) contribution to the total ionic free energy, ignored by Bjerrum and subsequent workers, proves crucial to the proper incorporation of dipolar clusters in the theory [5]. Indeed, as shown in Section 8, the resulting DHBjDI and DHBjDIHC theories, (the latter with hard-core terms included) give an account of the critical region in the RPM that, judging by Monte Carlo studies [5, 6, 13–17], is encouragingly good: see Figs. 6 and 7, below. It

does appear that all the dominant physical effects are sensibly represented in the theory. Beyond the thermodynamic properties, the densities of bound ionic clusters in the near-critical vapor, which is discussed in Section 8.3, also seem to correlate well with Gillan's explicit cluster calculations (for somewhat lower temperatures) [9, 17(c), 44].

An analysis of the low-temperature coexistence curves within DHBjDI theory is given in Appendix A. Physically, however, association into larger neutral *and* charged clusters becomes important at lower temperatures in the *vapor* [12, 44]. How this may be incorporated into the theory is discussed briefly in Section 8.4, while Appendix B addresses tentatively the estimation of the association constants for higher-order clusters. When T drops below T_c the *liquid* density at coexistence increases very rapidly: the task of providing a physically accurate account of this feature of the phase behavior is beyond the main scope of the present work; nevertheless a few remarks are made in Section 8.5.

For $d = 3$ dimensions it is well known that the mean spherical approximation (MSA) can be solved exactly for the RPM [63, 64]. This has attracted the attention of many workers who, the literature suggests, have preferred it to DH theory. For completeness, therefore, we present in Section 9 (following, in fact, earlier work by Ebeling and Grigo [11, 42]) an account of results for criticality, coexistence, etc., that follow by using the MSA expression for the ionic free energy in place of DH theory [72, 73]. For reasons that are not obvious but which we attribute to its more formal and less physical theoretical basis, the MSA seems to provide a significantly poorer description of the critical region and of the vapor and cluster formation. (As regards the HNC and other integral equations, recall our remarks above [33].)

Finally, we reemphasize that although the DH-based (and MSA-based) theories automatically provide classical critical exponents and, indeed, full Landau or van der Waals critical behavior, this result, of itself, has *no bearing* at all on the observations of Pitzer and subsequent experimentalists [6, 18–25, 28] revealing classical or close-to-classical exponents in real electrolyte solutions. To attack that issue theoretically it is essential, as mentioned, to address the statistical mechanics of the density fluctuations and the related correlation functions [6, 27]. In particular, if one knew the density–density correlation length, $\xi(T, \rho)$, one could, in a first approach, employ the Ginzburg criterion [5, 65] to gauge closeness to classical behavior. A *tentative* estimate made by the present authors [5], seems to indicate, however, that little if any insight can be gained by that route. But we leave further investigation of the density fluctuations and that battle plan for subsequent work [66, 67].

## 2. Multicomponent thermodynamics: variational formulation

The thermodynamics of multicomponent fluid systems is a long-established discipline and allowance for chemical equilibria between various species presents no difficulties in principle. However, we have not found at hand a general exposition

convenient for developing mean-field type approximations (and we have come across seeming confusion in the literature concerning van der Waals loops, "wiggles", etc. in various plots). Accordingly, it seems useful to state the fundamentals systematically in the form in which we will use them; this also provides a basis for explaining our notation and general approach.

To this end, consider a system of distinct chemical species, A, B, ... , J, ... (which might be atoms, molecules, ions, ion pairs, clusters, etc.) with chemical potentials μ_a, $\mu_b, ... , \mu_j, ...$ and number densities $\rho_a \equiv N_A/V$, $\rho_b \equiv N_B/V, ..., \rho_j \equiv N_J/V$, with V the total volume and N_J the number of entities of species J. The pressure $p(T; \mu_a, \mu_b, ...)$ is an appropriate thermodynamic potential from which all properties follow, as usual, via

$$\rho_j = \left(\frac{\partial p}{\partial \mu_j}\right)_{T,\mu_k}, \qquad \mathscr{S} \equiv \frac{S}{V} = \left(\frac{\partial p}{\partial T}\right)_{\mu_j}, \tag{2.1}$$

where S is the total entropy. The *phase space* $(T; \mu_a, \mu_b, ...)$ is divided up into single-phase regions by phase-boundary "surfaces" (submanifolds of codimension unity) on which distinct phases can coexist. Equilibrium under a chemical reaction, say,

$$\nu_A A + \nu_B B + \cdots = \nu_X X + \nu_Y Y + \cdots , \tag{2.2}$$
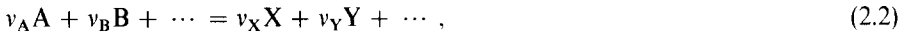
with stoichiometric coefficients ν_A, etc., is ensured by imposing

$$\nu_A \mu_a + \nu_B \nu_b + \cdots = \nu_X \mu_x + \nu_Y \mu_y + \cdots . \tag{2.3}$$

This amounts to selecting a "plane" (linear subspace of codimension unity) within the full phase space.

Now approximate, microscopic, or semimicroscopic descriptions of physical systems are most naturally formulated in terms of local densities (or suitable local "order parameters"). The corresponding thermodynamic potential is then the total Helmholtz free energy, $F(T; N_A, N_B, ... ; V)$, with differential relation

$$dF = -S\,dT - p\,dV + \sum_j \mu_j dN_J . \tag{2.4}$$

In the thermodynamic limit it is convenient to define the reduced Helmholtz free-energy density

$$\bar{f}(T; \rho_a, \rho_b, ...) \equiv f/k_B T \equiv -F/V k_B T . \tag{2.5}$$

Then, via the second law, we find that the pressure is given by the *variational expression*

$$\bar{p}(T; \mu_a, \mu_b, ...) = \max_{\{\rho_j\}} \left(\bar{f}(T; \rho_a, \rho_b, ...) + \sum_j \bar{\mu}_j \rho_j \right), \tag{2.6}$$

where, for convenience and brevity, we use

$$\bar{p} \equiv \beta p, \quad \bar{\mu}_j \equiv \beta \mu_j, \qquad \beta = 1/k_B T . \tag{2.7}$$

Now note that multiple (extremal) solutions $\{\rho_j\}_{j=1,2,\ldots}$, of the maximization problem for fixed T and $\{\mu_j\}$ correspond to the existence of two (or more) potentially coexisting phases. Given $f(T; \rho_a, \rho_b, \ldots)$, such solutions may be found (in principle and, in practice, numerically) by plotting p versus the chemical potentials μ_j *parametrically* using the relation

$$\bar{\mu}_j(T; \rho_a, \rho_b, \ldots) = -(\partial \bar{f}/\partial \rho_j)_{\rho_k}, \tag{2.8}$$

which follow directly from (2.6) as well as from (2.4). The equality of pressure, temperature, and chemical potentials for different densities corresponds to intersections of different branches of the $p(T; \mu_a, \ldots)$ surface and yields the phase boundaries. (There is, of course, no necessity for any "equal areas" or Maxwell construction and, indeed, the associated van der Waals loops need not and, in some interesting cases [34], will not exist.)

Our aim now is to construct $\bar{f}(T; \rho_a, \ldots)$ approximately as a *sum* of terms each embodying contributions from specific degrees of freedom and appropriate physical mechanisms. First, for each species, J, there will be an ideal-gas term

$$\bar{f}_j^{\text{Id}}(T; \rho_j) = \rho_j - \rho_j \ln[\rho_j \Lambda_j^{k_j d}/\zeta_j(T)], \tag{2.9}$$

where $\zeta_j(T)$ represents the *internal configurational partition function*, for the species J [5]. In order to specify precisely the kinetic energy contributions embodied in the overall or mean thermal de Broglie wavelength

$$\Lambda_j(T) = h/\sqrt{(2\pi\bar{m}_j k_B T)}, \tag{2.10}$$

we assume *classical statistical mechanics* in d spatial dimensions: for essentially all chemical applications near and above room temperature this will be fully adequate. Furthermore, in the absence of chemical reactions the assumption has essentially no operational significance for us. It does, however, have some effect in the discussion of association constants for chemical reactions as we comment below. Accepting classical mechanics and performing the momentum integrals separately for each of the k_j particles of masses m_{ji} associated with a single J entity, shows that the average mass $\bar{m}_j$ in (2.10) is just the geometric mean, $(m_{j1} m_{j2} \cdots m_{jk_j})^{1/k_j}$ [5].

Next we wish to represent the short-range repulsive or "hard core" interactions between the various species. Near close-packing constructing an accurate expression for $\bar{f}^{\text{HC}}$ obviously poses a profound problem. However, we will be concerned mainly with low densities, say $\rho/\rho_{\max} \lesssim 0.4$, so that a variety of reasonable approximations are available. The simplest is the familiar *free-volume approximation*

$$p = k_B T \sum_i \rho_i \bigg/ \left(1 - \sum_j B_j \rho_j\right), \tag{2.11}$$

which derives (see below) from

$$\bar{f}^{\text{HC}} = \left(\sum_i \rho_i\right) \ln\left(1 - \sum_j B_j \rho_j\right). \tag{2.12}$$

The free-volume form is generally valid near close packing [35] and with $B_j = v_j$ the "volume" occupied by a particle of species J it is exact for $d = 1$. More generally one may (a) set $B_j = 1/\rho_{j,\max}$ or (b) choose the B_j to obtain the correct, high-T diagonal second virial coefficients noting that (2.11) predicts $B_{ii}^{(2)} = B_i$, $B_{ij}^{(2)} = B_i + B_j$ for the general second virial coefficients. For reference below recall that for spheres of diameter a the values

$$B_j/a^3 = 1/\sqrt{2} \simeq 0.7071, \qquad 4/3\sqrt{3} \simeq 0.7698, \qquad 1, \qquad \tfrac{2}{3}\pi \simeq 2.0944, \qquad (2.13)$$

correspond, respectively, to fcc, bcc, and sc packings, and to the exact ($d = 3$) second virial coefficient.

For our present purposes these alternatives will suffice to indicate the role of hard-core repulsions since the densities of most interest will be quite low ($\rho/\rho_{\max} \lesssim 0.1$). We remark, nevertheless, that the form

$$\bar{f}^{\text{HC}} = \sum_{(i,\,j)} B_{ij}^{(2)} \rho_i \rho_j \left[\ln\left(1 - \sum_k b_k \rho_k\right) \middle/ \sum_k b_k \rho_k \right] \qquad (2.14)$$

incorporates the correct second virial coefficients and if one takes, say, $b_j = 1/\rho_{j,\max}$, approximates those of higher order in free-volume fashion. Beyond that, for a few species representable by hard spheres, one has analytic forms [35] tested by simulation, like the well known Percus–Yevick equation of state for hard spheres [11], from which the approximate expressions for $\bar{f}^{\text{HC}}$ follow by integration via

$$\bar{f}^{\text{HC}} = -\int_0^1 \left(\bar{p}(T; \lambda\rho_a, \lambda\rho_b, \ldots) - \lambda \sum_j \rho_j \right) d\lambda/\lambda^2. \qquad (2.15)$$

Finally, one might use exact virial coefficients, as far as known, and extend the series for $\bar{p}$ by Padé or partial differential approximants [36].

In the absence of crystallization or other high density phase transitions built into the chosen f^{HC}, no phase changes or criticality follow from $f \simeq f^{\text{Id}} + f^{\text{HC}}$. Attractive forces between neutral atoms or molecules are most simply introduced, following van der Waals, by including *second virial terms* of the form

$$\bar{f}^{\text{2V}} = \sum_{(j,\,k)} A_{jk} \rho_j \rho_k / T, \qquad (2.16)$$

with fixed, positive A_{jk}. In the Kac–Baker limit of infinitely weak, infinitely long-ranged attractive potentials this represents an exact expression and the combination $f^{\text{Id}} + f^{\text{HC}} + f^{\text{2V}}$ yields the standard van der Waals theory and its many-component extensions. As remarked in the introduction, such approximate theories have been remarkably successful at the qualitative and semiquantitative levels.

To proceed further for electrolytes we evidently need an expression for f^{Ion}, the excess free energy arising from the ionic interactions. For this we turn to DH theory.

3. Debye–Hückel theory for the RPM

For convenience let us restate briefly the definition of the restricted primitive model as employed by Debye and Hückel. A total of N hard spherical ions, each of diameter a, specify the *overall density* and *reduced density*

$$\rho = N/V, \qquad \rho^* \equiv \rho a^d, \tag{3.1}$$

where, henceforth, we consider only dimensionality $d = 3$. (For general d see Ref. [34].) Each ion i carries a charge q_i taking only the values of $\pm q$ in the restricted model; overall electroneutrality is always assumed. We will use

$$\rho_1 = \rho_+ + \rho_- \tag{3.2}$$

to denote the density of *free* ions (with, of course, $\rho_+ = \rho_- = \frac{1}{2}\rho_1$). In pure DH theory there is no allowance for ion pairing, etc., so one has $\rho_1 = \rho$, although more generally $\rho_1 < \rho$. The potential of interaction between two ions at separation r is

$$\varphi_{ij}(r) = + \infty, \quad r < a,$$
$$= q_i q_j / Dr, \quad r \geqslant a. \tag{3.3}$$

which sets the energy scale and thence the reduced temperature

$$T^* = k_B T / |\varphi_{ij}(a)| = k_B T D a / q^2. \tag{3.4}$$

Classically, as explained, the masses, m_j, of the ions play no role.

From a rigorous viewpoint it is by no means certain that the RPM actually has a gas–liquid transition. However, the analogy with real molten salts like KCl and NaCl [9, 12, 18], and the improving quality of the simulations [5, 6, 13–16] makes a transition highly plausible (for $d \geqslant 3$). Indeed, the current numerical evidence suggests [17] that the critical point lies in the range

$$T_c^* = 0.052\text{–}0.056, \qquad \rho_c^* = 0.023\text{–}0.035. \tag{3.5}$$

The critical osmotic coefficient is appreciably less studied but probably lies in the range

$$Z_c \equiv p_c / \rho_c k_B T = 0.08\text{–}0.10, \tag{3.6}$$

or *lower* [4(b), 12, 37].

The DH theory [1–3], which we will recapitulate briefly for reference, employs the subsequently widely applied mean-field idea of examining one statistical element, in this case the ith ion $\mathscr{I}_i$, interacting with its neighbors in an averaged manner. It appears to differ, however, from other mean-field theories, particularly those developed for lattice systems, that enjoy a rather direct variational characterization (so that they provide rigorous upper bounds for the true free-energy) or which, as in the Bethe lattice or Kac–Baker infinite-range limits, can be shown to be exact for particular models or limiting circumstances. Such formulations of DH theory, and/or its extensions would be valuable. In default of such results we proceed traditionally. If

$\psi_i(\{q_j\}; T, \{\rho_k\})$ is the mean electrostatic potential at ion $\mathscr{I}_i$ due to the population of other ions (in a canonical ensemble specified by T and the species densities $\{\rho_k\}$) the ionic contribution to the total free energy is given exactly by the DH charging process as [2, 3]

$$F^{\text{ion}}(T, \{\rho_k\}) = \sum_i q_i \int_0^1 d\lambda \, \psi_1(\{\lambda q_j\}; T, \{\rho_k\}) \,. \tag{3.7}$$

Note that the self-energy of the ions is not included.

The theory then aims to obtain an approximation for ψ_i by considering the electrostatic potential $\phi(r)$ in the neighborhood of $\mathscr{I}_i$. By the hard-core conditions other ions are excluded from the sphere $|r| < a$ centered on $\mathscr{I}_i$. Thus $\phi(r)$ satisfies Laplace's equation with spherically symmetric average solution

$$\phi(r) = q_i/Dr + \psi_i, \qquad \text{for } r < a \,. \tag{3.8}$$

Outside the exclusion sphere, which is an essential feature of DH theory, the pair correlation functions $g_{++}(r) = g_{--}(r)$ and $g_{+-}(r)$ specify the mean charge distributions. In DH theory these are approximated by Boltzmann factors leading to the expression

$$\rho_q(r) = \sum_k \rho_k q_k \exp\left[-\beta q_k \phi(r)\right] \,. \tag{3.9}$$

for the mean charge density. Substitution in Poisson's equation, $\nabla^2 \phi = -4\pi \rho_q/D$, yields the Poisson–Boltzmann (PB) equation. As a second approximation, *linearization* yields, with the aid of electroneutrality,

$$\nabla^2 \phi = \kappa^2 \phi \,, \tag{3.10}$$

with inverse squared Debye length given generally by

$$\kappa^2(T, \{\rho_k\}) = 4\pi \sum_k \rho_k q_k^2/D k_{\text{B}} T \,, \tag{3.11}$$

so that for present purposes we have

$$\kappa^2 a^2 = 4\pi \beta q^2 \rho_1/D = 4\pi \rho_1^*/T^* \,. \tag{3.12}$$

Apart from the consequent technical simplifications, linearizing the PB equation is significant since it ensures the satisfaction of important electrostatic self-consistency conditions [3, 38]. At low temperatures, however, the approximation entailed, namely, $\sinh(\beta q\varphi) \approx \beta q\varphi$, is serious for nearby, oppositely charged ions, i.e., when $\exp(\beta q^2/Da) = e^{1/T^*}$ is large.

A different approach to deriving the PB equation is by reformulating the ionic fluid problem in terms of a field theory. A corresponding mean field theory is then generated by the saddle-point approximation to the full field theory. Indeed, it is well known that a Coulomb gas *without* a hard core is equivalent to a sine-Gordon field

theory [39]. The PB equation for the potential of mean force, $\phi(r)$, is precisely equivalent to the sine-Gordon saddle point with the field analytically continued to pure imaginary values. However, this formal approach obscures significant physical insights accessible from DH theory. In particular, it is unclear how to incorporate the hard cores of the RPM in an effective fashion in contrast to the naturalness of DH theory.

In solving the DH equation one must require that both $\phi(r)$ and the electric field $-\nabla\phi$ be continuous across the exclusion sphere $|r| = a$ and that $\phi(r)$ vanishes as $r \to \infty$ [2, 3]. In the exterior region, $r > a$, one obtains the well-known screened potential

$$\phi(r) = q\vartheta(\kappa a)\,\mathrm{e}^{-\kappa r}/Dr, \qquad \vartheta(x) = \mathrm{e}^x/(1 + x); \tag{3.13}$$

for $r < a$ the form (3.8) results with

$$\psi_i(q; T, \rho_1) = -\,q\kappa/D(1 + \kappa a). \tag{3.14}$$

Implementation of the DH charging process embodied in (3.7) finally yields the DH approximation for f^{Ion}, namely [1–3],

$$\bar{f}^{\,\mathrm{DH}} = [\ln(1 + \kappa a) - \kappa a + \tfrac{1}{2}(\kappa a)^2]/4\pi a^3$$

$$= (\kappa^3/12\pi)\,[1 - \tfrac{3}{4}\kappa a + \tfrac{3}{5}(\kappa a)^2 + \cdots]\,. \tag{3.15}$$

The original DH theory [1] amounts to the assertion $f \simeq f^{\mathrm{Id}} + f^{\mathrm{DH}}$. This is now known [2, 3] to be correct to leading order in κa as $\rho \to 0$, the so-called DH *limiting laws* resulting from retaining only $(\kappa^3/12\pi)$ in (3.15). Errors must arise at next order, $(\kappa a)^4 \sim \rho^2$, if only because the loss of phase volume due to the hard-core repulsions is not accounted for; but, as indicated, accuracy must also be lost when $(\kappa a)^3\mathrm{e}^{1/T^*}$ is large. Awareness of these limitations [2, 3, 9, 12], is probably what held earlier workers back from studying the implications of DH theory for phase coexistence, to which we now turn. It must be noted, nevertheless, that experience shows that physically well founded mean-field theories for three-dimensional systems typically exhibit a qualitative and even semiquantitative, validity going significantly beyond their formal level of accuracy.

4. RPM coexistence within Debye–Hückel theory

4.1. Preliminaries

To analyze the DH result (3.15) recall first that we have $\rho_+ = \rho_- = \tfrac{1}{2}\rho_1$ and so, correspondingly,

$$\mu_+ = \mu_- = \mu_1, \qquad \bar{p} = \max_{\rho_1}\,[\bar{f}(T;\rho_1) + \bar{\mu}_1\rho_1]\,. \tag{4.1}$$

Then, combining the ideal-gas terms for cations and anions, the pure DH approximation is

$$\bar{f}(T,\rho_1) = 2\bar{f}^{\,\mathrm{Id}}(T,\tfrac{1}{2}\rho_1) + \bar{f}^{\,\mathrm{DH}}(T,\rho_1)\,, \tag{4.2}$$

where, at this stage, we also have $\rho_1 = \rho$, and $\mu_1 = \mu$. It is convenient here and below to put

$$x = \kappa a \quad \text{with } \rho_1^* = x^2 T^*/4\pi\,. \tag{4.3}$$

Then, with $\Lambda_+ = \Lambda_- = \Lambda_1(T)$ and $\zeta_+ = \zeta_- = 1$, the chemical potential is given by the well known results [2, 3]

$$\ln \gamma_\pm \equiv \bar{\mu}_1 - \bar{\mu}_1^{\mathrm{Id}} = -x/2T^*(1 + x)\,, \tag{4.4}$$

$$\bar{\mu}_1^{\mathrm{Id}} = \ln(x^2 T^*) + \ln(\Lambda_1^3/8\pi a^3)\,, \tag{4.5}$$

while the relation for the pressure can be written

$$p^\dagger \equiv 4\pi a^3 \bar{p} = x^2 T^* + \ln(1 + x) - x + \tfrac{1}{2}x^2/(1 + x)\,. \tag{4.6}$$

4.2. Criticality and coexistence

To discuss the possibility of criticality, consider the isothermal compressibility K_T: one finds

$$\frac{1}{\rho k_\mathrm{B} T K_T} = \rho_1 \frac{\partial \bar{\mu}}{\partial \rho_1} = 1 - \frac{x}{4T^*(1 + x)^2}\,. \tag{4.7}$$

Equating $1/K_T$ to zero specifies the *spinodals*: $\rho_{s\,\pm}(T)$. One easily sees that there are two real roots, $x_{s\,\pm}(T) \gtrless x_c$, whenever $T < T_c$, where

$$x_c = \kappa_c a = 1\,, \qquad T_c^* = 1/16\,. \tag{4.8}$$

This locates the critical point and yields

$$\rho_c^* = 1/64\pi = 0.004973_6\,,$$

$$Z_c \equiv p_c/\rho_c k_\mathrm{B} T_c = 16 \ln 2 - 11 = 0.09035_5\,. \tag{4.9}$$

The two sides of the coexistence curve, $\rho_l(T)$ and $\rho_v(T)$, are bounded above and below by the spinodals following from (4.7). The spinodal curve is evidently parabolic through the critical point and may be expanded in powers of the critical point deviations

$$t \equiv (T - T_c)/T_c\,, \qquad m = (\rho - \rho_c)/\rho_c\,, \tag{4.10}$$

in the form

$$m_{s\,\pm}(T) = \bar{m}_s(T) \pm B_s|t|^{1/2}[1 + O(t)]\,, \tag{4.11}$$

with critical amplitude $B_s = 4$ and diameter $\bar{m}_s(T) = -7t + O(t^2)$. At low temperatures, on the liquid and vapor sides, one finds

$$\rho_{s+}/\rho_c \approx 1/T^*, \qquad \kappa_{s+}a \approx 1/4T^*. \tag{4.12}$$

$$\rho_{s-}/\rho_c \approx \tfrac{1}{16}(T/T_c)^3, \qquad \kappa_{s-}a \approx 4T^*, \tag{4.13}$$

The coexistence curve can be expanded similarly as

$$m_\pm(T) = \bar{m}(T) \pm B|t|^{1/2}[1 + O(t)], \tag{4.14}$$

with $B = 4\sqrt{3}$, and is also parabolic so $\beta = \tfrac{1}{2}$, i.e., the critical exponent is classical. The coexistence diameter, $\bar{m}(T) = -(99/5)t + O(t^2)$, is markedly steeper than $\bar{m}_s(T)$: linear extrapolation to the axes yields intercepts $t_d = \tfrac{5}{99}$ at $\rho = 0$ and $\rho_d^* = 13/40\pi = 0.1034_5$ at $T = 0$. A plot of the coexistence curve, which exhibits its striking asymmetry, is presented in Fig. 1a. At low T the two branches of the coexistence curve are described by

$$\frac{\rho_\ell}{\rho_c} = \frac{4}{T^*}\left\{1 - 8T^*[\ln(1/2T^*) - \tfrac{1}{2}] + \cdots\right\}, \tag{4.15}$$

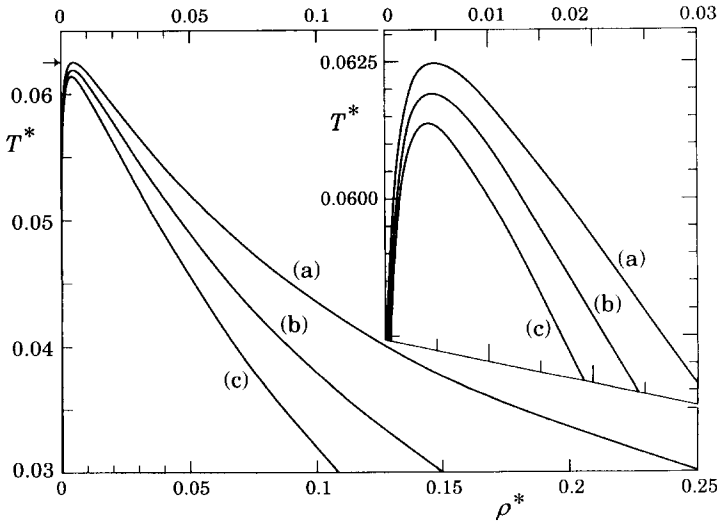$$\frac{\rho_v}{\rho_c} = \frac{4e}{T^*}\,e^{-1/2T^*}[1 - 8T^*\ln(1/2T^*) + \cdots], \tag{4.16}$$

Fig. 1. Plots of the coexistence curve, with $T^* = k_B T D a/q^2$ and $\rho^* = \rho a^3$, for the restricted primitive model (RPM) according to Debye–Hückel (DH) theory for (a) the pure theory, and (b) as supplemented with hard-core repulsive contributions with $B_1 = a^3$, corresponding to the free volume approximation with simple-cubic packing limit, and (c) with $B_1 = \tfrac{2}{3}\pi a^3$, representing the exact second virial coefficient. The inset is on a larger scale. For case (b) we may also quote $x_c = 0.92906_5$, $T_c^* = 0.061908_2$, and $Z_c = 0.093213_5$.

which also differ significantly from the spinodals. The vapor pressure curve, $\tilde{p}_\sigma(T) \equiv \beta p_\sigma a^3$, at low T is similarly given by

$$\ln \tilde{p}_\sigma(T) = -\frac{1}{2T^*} - \ln(16\pi T^*) + 1 + 4T^* \ln(2T^*) + \cdots , \tag{4.17}$$

where here there are neglected terms of order $e^{-1/4T^*}/T^{*2}$ although the leading corrections, as in (4.15) and (4.16), are of order $(T^* \ln T^*)^2$.

Of course, one expects that all the critical behavior, not merely the coexistence curve, is classical. This is most readily seen, and explicit results derived, by constructing a Landau-type expansion. Thus, using (4.6) and (4.8)–(4.10) we find

$$p_0^\dagger(T, \mu) - p^\dagger(T, \mu) = \min_m \mathscr{F}(m; T, \mu) , \tag{4.18}$$

where the "background" $p_0^\dagger(T, \mu)$ is linear in μ and analytic in T. If we define the *ordering field* by

$$h = c_1[\bar{\mu} - \bar{\mu}_\sigma^*(T)] , \tag{4.19}$$

where, with $\tau \equiv \sqrt{T_c/T} = (1 + t)^{-1/2}$, we have

$$\bar{\mu}_\sigma(T) = -4\tau^3/(1 + \tau) + \ln(\Lambda_1^3/2^7 \pi a^3) , \tag{4.20}$$

then we can write the expansion generally as

$$\mathscr{F}(m; T, \mu) = \tfrac{1}{2} c_2 t_2 m^2 + \tfrac{1}{4} u_4 m^4[1 + 18t + \cdots] + \cdots$$
$$- hm - \tfrac{1}{3} c_3 t_3 m^3 - u_5 m^5[1 + \mathrm{O}(t)] + \cdots , \tag{4.21}$$

with explicit coefficients

$$t_2 = 1 - 4\tau^3/(1 + \tau)^2 = t - \tfrac{15}{16} t^2 + \mathrm{O}(t^3) , \tag{4.22}$$

$$t_3 = \tfrac{8}{9}[1 - 2\tau^3(1 + 3\tau)/(1 + \tau)^3] = t + \mathrm{O}(t^2) , \tag{4.23}$$

$$c_1 = c_2 = \tfrac{1}{16}, \qquad c_3 = \tfrac{9}{256}, \qquad u_4 = \tfrac{1}{768}, \qquad u_5 = \tfrac{1}{512} . \tag{4.24}$$

In the standard way one now checks the exponents $\gamma = \gamma' = 1$, $\delta = 3$, and $\alpha = \alpha' = 0$ (discontinuous), cross-checks (4.11) and (4.14), and derives critical amplitudes and leading corrections for the compressibility, the critical isotherm, the constant-volume specific heat, etc.

It is evident that this strategy and the resulting classical behavior will invariably follow from the general approach [40] unless, by some special accident, by construction, or by some unusual mechanism [34], one of the partial free-energy densities, say f^X, has singularities at the critical point already built in or generates them by some coupling mechanism going beyond the basic formulation [41].

4.3. Hard-core contributions

The low-T result (4.16) for the coexistence curve is obviously in error once it predicts $\rho_{\ell}^*(T) \gtrsim 1$, i.e., when $T/T_c \lesssim 1/\pi$. The cause, clearly, is the failure to take account of the mutual excluded volume of the ionic hard cores. However, this is easy to rectify and it proves significant to do so [11, 42]. The simplest option is to add a second virial term, $\bar{f}^{\,HC2V} = B_{11}^{(2)} \rho_1^2$, to (4.2). This can be handled analytically without difficulty. With $\tilde{B} = B_{11}^{(2)}/\frac{2}{3}\pi a^3$, which should be unity for hard spheres, the critical point is determined by

$$T_c^* = \tfrac{1}{8} x_c (1 + 3x_c)/(1 + x_c)^3 \,, \tag{4.25}$$

$$24(1 - x_c)(1 + x_c)^3 = \tilde{B} x_c^3 (1 + 3x_c)^2 \,, \tag{4.26}$$

with solution

$$x_c = \kappa_c a = 0.93249_1, \qquad T_c^* = 0.0613333_4 \,. \tag{4.27}$$

The critical temperature is thus reduced by $1.86_7\%$ while the critical density falls $8.49_1\%$ to $\rho_c^* = 0.004551_3$. The analysis of (4.18) remains valid with, to first order in $\tilde{B}$, the addition of $\tilde{B}/768$ to $\tilde{\mu}_\sigma$ in (4.20) and $\tilde{B}/48$ to t_2 in (4.22).

The plots of the coexistence curves in Figs. 1b and c are constructed using the free-volume forms (2.12), (2.13), with $j = 1$, and (b) $B_1 = a^3$, and (c) $B_1 = \frac{2}{3}\pi a^3$. To three-figure accuracy the critical point for (c) is given by (4.27); in other words, not surprisingly since ρ_c^* is so small, the higher virial coefficients have no significant effect near criticality. For (b), to equal or better accuracy, one may thus use (4.26) with $\tilde{B} = 3/2\pi$.

At low T, of course, the free-volume form respects the maximum density and one finds, in place of (4.17),

$$\ln \tilde{p}_\sigma(T) = -\frac{1}{2T^*} + \frac{1}{2\sqrt{\rho_M^* T^*}} + \left(\frac{1}{2} - \frac{1}{8\pi\rho_M^*}\right) \ln\left(\frac{4\pi\rho_M^*}{T^*}\right) - \ln 8\pi + o(1) \tag{4.28}$$

where $\rho_M = 1/B_1$ corresponds to the maximum density [see (2.12)] while the leading neglected terms are of order $T^{*1/2}$. The vapor side of the coexistence curve is then given by

$$\rho_v^*(T) = \tilde{p}_\sigma(T)\left[1 + O(\tilde{p}_\sigma^{1/2}/T^{*3/2})\right], \tag{4.29}$$

which merely confirms that the low density vapor is, in leading order, ideal. The liquid side approaches ρ_M according to

$$\rho_\ell^*(T) = \rho_M^*\left[1 - 4\sqrt{\pi\rho_M^* T^*} - 2T^* \ln(4\pi\rho_M^*/eT^*) + 8\pi\rho_M^* T^* + o(T^*)\right]. \tag{4.30}$$

Although no longer unphysical, this result for $\rho_\ell(T)$ is, of course, not adequate for the RPM since one must anticipate at least one crystalline phase near close packing which, as stressed, is certainly not included in the free volume expressions for f^{HC}.

Nevertheless, at lower densities (and, hence, intermediate values of T/T_c) the predicted coexistence curve should be more reliable when higher order hard-core virial coefficients are included, even if only approximately.

4.4. Assessment

If we accept the Monte Carlo based critical point estimates (3.5) and (3.6) [17], we see from (4.8) and (4.27) that T_c^* in pure DH theory is probably only 12–20% too high, quite comparable to van der Waals theory. The DHHC theory (DH with hard-core terms) reduces the discrepancy by about 2%. Furthermore, the critical osmotic coefficient Z_c is, within the $\pm 11\%$ uncertainties, in full agreement with the MC estimate. However, the DH prediction for the critical density ρ_c^* is too small by a factor of from 4.6 to 7.0; and inclusion of hard-core terms *increases* this discrepancy to 5.1 to 7.7. This major failure is obviously associated with the anticipated inadequacy of the DH linearization of (3.10) near criticality given that $1/T_c^* \gtrsim 16$ so that $e^{1/T_c^*} \sim 10^7$! One might wish, thus, to discard the DH theory of criticality entirely; but in light of the reasonable estimate of T_c^* and the realistically asymmetric character of the coexistence curve, that is much *too* pessimistic. Rather, following Bjerrum [7], it is reasonable to suppose that the main error incurred by linearization is, physically, due to neglect of closely associated, oppositely charged, and hence neutral, ion pairs and their effects. Accordingly we now consider Bjerrum's theory in some detail [2, 8, 9, 11, 12].

5. Ideal Bjerrum dipoles

5.1. Basic principles

In 1926 Bjerrum proposed [7] to extend DH theory by regarding an electrolyte, specifically the RPM, as composed of "free" $+$ and $-$ ions together with neutral $(+, -)$ associated ion pairs or dipoles that exist in equilibrium with the free ions. The specific recognition of $+$ and $-$ ions at small separations as constituting a distinct chemical species was expected, when appropriately treated, to correct for the correspondingly large Boltzmann factors at low T that are seriously underestimated by the linearization of the PB equation in the DH theory. To develop the concept, let μ_2 and ρ_2 denote the chemical potential and densities of the Bjerrum ion pairs and write, with no approximation at this stage,

$$\bar{p}(T;\mu) = \max_{\rho_1, \rho_2} \left[2\bar{f}_1^{\text{Id}}(T, \tfrac{1}{2}\rho_1) + \bar{f}_2^{\text{Id}}(T, \rho_2) + \bar{f}^{\text{Ex}}(T, \rho_1, \rho_2) + \bar{\mu}_1 \rho_1 + \bar{\mu}_2 \rho_2 \right],$$

(5.1)

where the excess free energy, $\bar{f}^{\text{Ex}}$, includes all hard-core and Coulombic effects. In the ideal-gas terms we may take $\zeta_+ = \zeta_- = 1$ and $\Lambda_+ = \Lambda_- = \Lambda_1$ as before, and then

have $\Lambda_2 = \Lambda_1$. Chemical equilibrium is ensured by

$$\mu_2 = \mu_+ + \mu_- = 2\mu_1 , \tag{5.2}$$

which leads, after some algebra, to the law of mass action in the form

$$\rho_2 = \tfrac{1}{4}\rho_1^2 K(T) \exp(2\bar{\mu}_1^{\mathrm{Ex}} - \bar{\mu}_2^{\mathrm{Ex}}) , \tag{5.3}$$

(recalling that $\rho_+ = \rho_- = \tfrac{1}{2}\rho_1$) where the *association constant* is given by $K(T) = \Lambda_+^3 \Lambda_-^3 \zeta_2 / \Lambda_2^6 \zeta_+ \zeta_- = \zeta_2(T)$, while the excess chemical potentials follow from

$$\bar{\mu}_i^{\mathrm{Ex}}(T, \rho_1, \rho_2) = -\partial \bar{f}^{\mathrm{Ex}}/\partial \rho_i \quad (i = 1, 2) , \tag{5.4}$$

and vanish with ρ_1 and ρ_2.

Now the definition of a pair of oppositely charged nearby ions as "associated" has an inescapable element of arbitrariness that, in turn, bears on the value of the association constant which will be addressed in detail below. For the moment, however, note that in light of the definition of $\zeta_2(T)$ as a configurational partition function it is natural to take [7]

$$K(T; R) = 4\pi \int_a^R e^{a/rT^*} r^2 \, dr , \tag{5.5}$$

since one has $-\beta\varphi_{+-}(r) = a/rT^*$. The lower limit of integration corresponds just to the hard-core constraint. Ideally one might wish to let $R \to \infty$; but, even for short-range interaction potentials, $K(T)$ would then diverge. Bjerrum observed [7], however, that despite the scale-free nature of the Coulomb potential, a natural scale is set by the minimum of the integrand in (5.5), which, of course, represents the radial Boltzmann distribution. Certainly, once the integrand starts increasing beyond $r_{\min} = \tfrac{1}{2}a/T^*$ it seems unreasonable to regard a pair of ions as associated! Bjerrum chose $r_{\min}$ as his cutoff taking, in traditional notation,

$$R^{\mathrm{Bj}} = \tfrac{1}{2}ba \quad \text{for} \quad b \equiv 1/T^* \geqslant 2 ,$$
$$= a \quad \text{for} \quad b < 2 . \tag{5.6}$$

To evaluate $K(T)$ generally it is useful to put

$$K(T; R = cba) \equiv 4\pi a^3 Q_0(b; c) \, e^b/b , \tag{5.7}$$

since one finds that the numerical factor, given by [2]

$$Q_0(b; c) = \tfrac{1}{6}b^4 e^{-b}[\mathrm{Ei}(b) - \mathrm{Ei}(1/c) + (c + c^2 + 2c^3)e^{1/c}] - \tfrac{1}{6}b(b^2 + b + 2) , \tag{5.8}$$

for $b \geqslant 2$, is slowly varying and of order unity for all T. Here $\mathrm{Ei}(z)$ is the standard exponential integral [43]. Explicitly, for all fixed c one has the low-T asymptotic expansion

$$Q_0(b; c) = 1 + 4T^* + 4{\cdot}5T^{*2} + 4{\cdot}5{\cdot}6T^{*3} + \cdots , \tag{5.9}$$

This is divergent but for the Bjerrum choice $c = \frac{1}{2}$ it gives reasonable results up to $T^* \simeq 0.1$ when truncated at the smallest term. More Concretely, the lower bound

$$Q^{\mathrm{Bj}}(b) \geqslant Q^{\mathrm{Bj}}_{\lessgtr}(b) \equiv b[1 - e^{-(b-4)(b-2)/b}] \, (b-4), \tag{5.10}$$

matches (5.8) when $T \to 0$ and, when positive, mimics the overall behavior of Q^{Bj} moderately well: note, in particular, that $Q^{\mathrm{Bj}}(b)$ *vanishes* for $T^* > \frac{1}{2}$ $(b < 2)$. A convenient approximation, reproducing (5.9) is

$$Q^{\mathrm{Bj}} \simeq Q^{\mathrm{Bj}}_{\lessgtr}(b)\left(1 + \frac{4}{b^2} + \frac{20}{b^3} + \frac{360}{b^4} + \frac{3360}{b^5} + q_6\frac{26-b}{b^6}\right), \tag{5.11}$$

with $q_6 = 14785.06$: the accuracy is better than 1 in 10^4 for $14.3 \leqslant b \leqslant 40$. Finally, numerical values for $b = 10$, 16, and 20 are $Q^{\mathrm{Bj}} = 2.095936$, 1.390497, 1.274680, respectively.

Although, as we will discuss, Bjerrum's choice is not optimal, the behavior of $K(T)$ for $T^* \lesssim 0.1$, which region is our main concern, will not differ appreciably for other reasonable assignments: in particular the factor Q, defined generally as in (5.7), will remain of order unity. Accordingly we tentatively accept Bjerrum's expression and explore the consequences of the theory.

5.2. Bjerrum plus Debye–Hückel theory: criticality

Bjerrum, having in mind primarily low densities, proposed to treat his dipolar pairs as ideal particles, supposing, since they were neutral, that they had no significant Coulombic interactions. Then he adopted DH theory: the pairs being neutral, should not enter into the DH expression (3.12) for $\kappa(T, \rho_1)$, which remains valid as, then, does the DH derivation of the ionic free energy. At this DHBj level of approximation we thus have

$$\bar{f}(T; \mu) = 2\bar{f}^{\mathrm{Id}}_1(\tfrac{1}{2}\rho_1) + \bar{f}^{\mathrm{DH}}(\rho_1) + \bar{f}^{\mathrm{Id}}_2(\rho_2), \tag{5.12}$$

where, of course, $\zeta_2(T) \equiv K(T)$.

However, Bjerrum regarded the ion pairs as more or less literally having a physical diameter up to $R^{\mathrm{Bj}}(= \frac{1}{2}ba)$ and so argued that distinct, free ions should not be allowed to approach closer than R^{Bj} (before changing their nominal identity). Accordingly, in DH theory he replaced the hard-core diameter a by $R^{\mathrm{Bj}}(T) = \frac{1}{2}a/T^*$ for $T^* < \frac{1}{2}$. We will argue in detail below that at the low temperatures of interest for phase separation and criticality, this view of associated pairs is not reasonable. In accord with the conclusions of others [2, 5, 8, 11, 12], therefore, we will, at least for the present, retain the original DH interpretation of a. It should be noted, however, as discovered by Friedman and Larsen [9], that if one does adopt Bjerrum's re-identification of a, then the theory represented by (5.12) predicts *no phase separation* at any T (even though, as Bjerrum himself showed in relation to experiments, it still represents a distinct improvement on pure DH theory at low densities when $T^* \lesssim \frac{1}{3}$.)
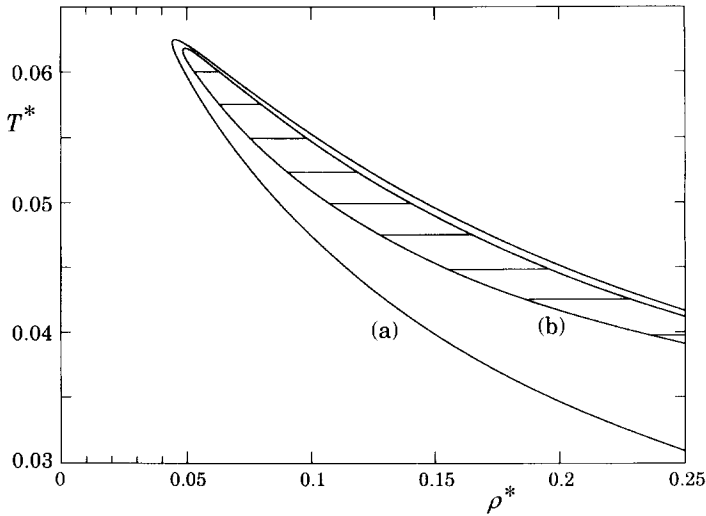
Fig. 2. Coexistence curves for the RPM according to (a) DH theory with ideal Bjerrum pairs and (b) as supplemented with free-volume hard-core terms with $B_1 = a^3$ and $B_2 = 2a^3$. When B_1 and B_2 are sufficiently large, the predicted coexistence curve closes off at a lower critical point.

With the previous interpretation of $\bar{f}^{DH}$, namely, (3.15) in (5.12) one finds that the resulting DHBj theory predicts phase separation with the coexistence curve plotted in Fig. 2a. The critical parameters $x_c = \kappa_c a = 1$, $T_c^* = \frac{1}{16}$ and $\rho_{1c}^* = 1/64\pi$ are *unchanged* from DH theory [See (4.8) and (4.9)] but one finds the (overall) critical density

$$\rho_c^* = \rho_{1c}^* + 2\rho_{2c}^* = (1 + Q_c e^8/2^9)\rho_{1c}^* = 0.0452384 , \tag{5.13}$$

where (4.4), (5.3), (5.6)–(5.8), and the Bjerrum value $Q_c = Q^{Bj}(16)$ have been invoked and, similarly, the osmotic coefficient

$$Z_c = \frac{16 \ln 2 - 11 + Q_c e^8/2^{10}}{1 + Q_c e^8/2^9} = 0.454963 . \tag{5.14}$$

Note that ρ^* in (5.13) denotes, as always, the overall or total density of ions (or hard spheres).

The inclusion of Bjerrum pairs has evidently increased the overall critical density by a factor of about 9.0 relative to pure DH theory, which is in the desired direction. Indeed, the prediction (5.13) is now only 50–70% high compared to the central Monte Carlo estimates (3.5) while DH theory was low by factors of 5 to 7! Thus both T_c and ρ_c are now reasonably well approximated. The mechanism for this improvement is that the formation of the neutral pairs reduces, for a given overall density ρ, the density of free ions, ρ_1, which is all that drives the phase separation (within this DHBj theory). In fact, the ratio of Bjerrum pairs to free ions at criticality is given by $\rho_{2c}/\rho_{1c} = Q_c e^8/2^{10} = 4.047 \ldots$. At first sight this may appear high but Gillan's detailed studies [17(c)] of clusters in the low density vapor shows that a critical ratio

of this magnitude is to be expected [44]. This point was also made by Friedman and Larsen [9].

Although the critical temperature and densities predicted by the DHBj theory may be reasonable, the banana-like shape of the coexistence curve shown in Fig. 2a is certainly not plausible [10, 45]! Note that the crown of the coexistence curve is still parabolic so that $\beta = \frac{1}{2}$ and all the other critical exponents remain classical [41]. Indeed, the Landau expansion description (4.18)–(4.21) remains valid but now with coefficients

$$c_1 \simeq 0.568, \quad c_2 \simeq 5.17077, \quad c_3 \simeq 80.7910, \quad u_4 \simeq 8.91227, \tag{5.15}$$

while one again has [40] $c_2/\sqrt{u_4} = \sqrt{3}$.

Another notable feature, going beyond the identity of T_c^{DH} and T_c^{DHBj} is that, despite the banana shape of the DHBj coexistence curve, the actual gap, $\rho_{liq}^* - \rho_{vap}^*$, is the *same* at given T^* as in the DH theory! On reflection these features are not so surprising since in (5.12) no terms couple ρ_1 and ρ_2 so that from (5.1) we conclude, with an obvious notation,

$$p^{DHBj}(T;\mu) = p^{DH}(T,\mu) + p_2^{Id}(T;\mu). \tag{5.16}$$

In words, then, the DHBj theory simply superimposes the pressure of an ideal gas of Bjerrum pairs on the DH pressure function which, alone, carries all the phase transition information! Only the free ions drive the transition. These observations also explain why the critical osmotic coefficient in (5.14) is so high: as discussed, the critical vapor is about 90% ion pairs which are treated as ideal so that $p_2 = k_B T \rho_2$; since $\rho = \rho_1 + 2\rho_2$ and ρ_{1c} and p_c^{DH} are small one has $Z_c \simeq p_2/2\rho_2 k_B T = \frac{1}{2}$.

5.3. DHBj theory at low temperatures

It is evident from (5.16) that the peculiar banana shape of the coexistence curve is due to a drastic increase in $\rho_{2\sigma}(T)$, the density of Bjerrum pairs on the phase boundary, as T falls. Note that, within this DHBj theory $\rho_2(T, \rho)$ is *continuous* across the phase boundary, clearly an artificial consequence of treating the Bjerrum pairs as ideal. But what causes $\rho_{2\sigma}(T)$ to increase? Might it be due to neglect of the hard-core repulsions?

The latter question is soon disposed of by adding to (5.12) an expression for $\bar{f}^{HC}$, the hard-core free energy, which, incidentally, will couple ρ_1 and ρ_2 so that (5.16) is no longer valid. The simplest choice is the free-volume form (2.11) with, for sc packing, $B_1 = a^3$ and, since an ion pair takes twice the volume of a single ion, $B_2 = 2a^3$. The resulting DHBjHC coexistence curve is shown in Fig. 2b. Evidently the situation is now worse! Not only does ρ_c increase slightly but the coexistence curve narrows strongly and the banana effect grows. Indeed, for large enough values of B_1 and B_2 the coexistence region vanishes entirely.

To answer the crucial first question, then, let us neglect the hard-core interactions and use (5.16): this implies the ideal-gas relation

$$\ln \rho_2^* = \bar{\mu}_2 + \ln \left[K(T)/4\pi a^3 \right] - \ln (\Lambda_1^6/4\pi a^6) , \tag{5.17}$$

where we have $\bar{\mu}_2 = 2\bar{\mu}_1$ by (5.2). The chemical potential $\bar{\mu}_1(T, \rho_1)$, follows from DH theory via (4.4) and (4.5), as

$$\bar{\mu}_1 = -\frac{x}{2T^*(1 + x)} + \ln(x^2 T^*) + \ln(\Lambda_1^3/8\pi a^3) . \tag{5.18}$$

To evaluate this *on* the *phase boundary* at low T we may use (4.15) for the coexisting liquid density $\rho_\ell \equiv \rho_{1,\text{liq}}$ to conclude $x_\ell \sim 1/T^* \to \infty$ and hence

$$\bar{\mu}_{1\sigma}(T) = -1/2T^* + \ln(1/T^*) + O(1) . \tag{5.19}$$

Finally, on using (5.7) for $K(T)$ in (5.17), one discovers that the leading terms in $1/T^*$ exactly cancel to leave

$$\ln \rho_2^* = \ln(1/T^*) + \ln \left[Q(b) e^2/128\pi \right] + 8T^* \ln(2\sqrt{e} T^*) + \cdots , \tag{5.20}$$

where the explicit correction terms follow from a closer analysis. If the association constant factor $Q(b)$ remains bounded when $b \to \infty$, as it does for the Bjerrum $K(T)$ assignment, this shows that $\rho_2^* \sim 1/T^* \to \infty$ when $T \to 0$, in accord with the banana observation. If, however, a better analysis of $K(T)$ led, for example, to $Q(b) \sim b^{-\sigma}$ with $\sigma > 1$ or to $Q(b) \sim e^{-\tau b}$ with $\tau > 0$, the banana effect would *not* appear. (In fact Fuoss [46] has argued for taking $Q(b) = \frac{1}{3}b$, so $\sigma = -1$; but this is not justifiable [47]: see also below.)

Thus, to understand the DHBj theory more fully and to repair it, we must examine the basis of ion association more carefully. In addition, it is clear that the Bjerrum assumption of *electrical* ideality for the ion pairs is dubious at low T since even a tightly bound pair of ions has a long-range dipolar electric field and so should attract a screening cloud of counterions [5, 48].

5.4. The origin of electrolyte phase separation

Before taking up these two issues let us briefly address, in the light of the DH and Bjerrum theories, the question: "What drives phase separation in electrolytes?" The issue is obviously not sharply defined but it can be usefully broached at the level that ascribes ordinary gas–liquid separation to the short-range attractive forces between molecules which energetically favor a condensed, closely packed state over a dispersed, low-density vapor. At sufficiently low temperatures a positive interfacial tension between regions of low and high density states then stabilizes phase separation for spatial dimensionalities $d \geqslant 2$. As T rises, thermal fluctuations reduce the disparity between dense and dilute phases. Since liquid and gas differ in no essential aspect of symmetry or order this eventually leads to a critical *point*, rather than a critical or melting *line*.

Solvophobic phase separation, as seen in typical binary molecular fluids [18, 24, 25, 49, 50] may be understood in full analogy: solute molecules, A, energetically prefer solute near-neighbors over solvent neighbors, B, because, in total, the mean AA and BB short-range attractive forces exceed the unlike, AB attractions, i.e., loosely, $\varepsilon = \varphi_{AB} - \frac{1}{2}(\varphi_{AA} + \varphi_{BB})$ is positive, and sets the scale of T_c. If the binary fluid is an electrolyte the short-range forces may still provide the driving force and largely determine T_c; and if the solvent is of high dielectric constant this is to be expected. In the *Coulombic* electrolytes, however [18, 24, 25, 49, 50], the ionic interactions dominate and ε is too small to specify T_c. Rather, through the law of corresponding states, the RPM should provide a satisfactory description [9, 12, 22].

Now we have seen that significant ion pairing occurs in the vicinity of and below the RPM critical point. Long ago Friedman [51] proposed that the presence of many dipolar ion pairs might provide the driving mechanism for Coulombic phase separation. Specifically, he argued that the relatively large dipole moments could lead to a drastically increasing dielectric constant, $D(\rho_2)$, at higher concentrations. This, in turn, would stabilize the solute dipoles (relative to their dilute state in the pure solvent). A rapid enough increase of D would lead to phase instability and separation into a dilute and a concentrated dipole phase.

How plausible is this scenario in which the free ions play no (significant) role? Surely it must be doubted since we now know that the free ions *alone* will, within DH theory, drive phase separation with a quite reasonable estimate of T_c. Furthermore, following Bjerrum, the introduction of ideal, electrically *inactive* dipoles *preserves* the DH transition and improves the estimate of the critical density. The operative mechanism in DH theory is clearly the effective net ion–ion *attraction* provided by the Debye screening and seen already in the limiting-law pressure: although for randomly distributed ions the overall Coulombic $(+, -)$ attractions are precisely cancelled by the $(+, +)$ and $(-, -)$ repulsions, the statistical screening correlations reduce the free energy of each ion in the field of its neighbors and so favor a denser phase at low T [52]. The mean separation of the free ions at DH (and DHBj) criticality is 12 or more ionic diameters, which does not support a dipole dominated picture. For these reasons it is hard to believe that the dipole–dielectric-constant interaction plays a major role as Friedman suggests [51].

Nevertheless, Friedman's arguments correctly identify a potentially serious gap in the Bjerrum theory, namely, the neglect of the Coulombically active dipolar nature of the ion pairs. In fact we find [5, 6], as explained below, that the interactions of a dipolar ion pair with the screening ionic fluid of free ions is a truly significant effect.

6. Ionic association

6.1. Physical and configurational clusters

The statistical mechanical theory of molecular and ionic association has been of long standing concern to physical chemists [2, 7–9, 11, 12, 17(c), 18]. Nevertheless,

there is no overall consensus on a "best" treatment; but that is not so surprising because there are various reasonable, valid definitions of association into dimers, trimers, etc. However, for our present purposes, namely the elucidation of thermodynamic equilibrium, there is an approach which we judge should be clearly preferred and which leads to an expression for the ionic association constant in the RPM first advocated by Ebeling [2, 8, 11] and later endorsed by Yokoyama and Yamatera [53]. To explain this, let us briefly recall some central issues.

First, in a chemical reaction the amount of some product species, say PQ, is often gauged spectroscopically: it is then necessary for theory to determine those configurational and dynamical states of P and Q entities that exhibit the identifying spectral lines to the requisite precision. In other words, one must delineate those regions of phase space that will, by the relevant observations, be judged as having P and Q bound or associated in a cluster. Note, in particular, that a dynamic P–Q collision at high relative kinetic energy would not, normally, be counted as an associated state even though P and Q particles may approach very closely. Thus Hill [54] has presented a classical statistical mechanical discussion of "physical clusters" of monatomic particles in an imperfect gas in which the criterion of "bound" is that the relative kinetic energy does not exceed the (negative) minimum of the pair interaction potential $\varphi(r)$.

It is clear, however, that if the aim is to understand equilibrium thermodynamic properties (as against, say, the ionic conductivity which is an intrinsically dynamic phenomenon, particularly outside the low-density, kinetic-theory regime) then any condition linking kinetic and configurational degrees of freedom is unhelpful since, in computing the partition function, it effectively precludes one performing all the (classical) momentum-space integrals once and for all at the start. Such considerations lead naturally to *clustering–distance conventions* in which only the position coordinates play a role: specifically, if two particles P and Q have centers closer than a distance R^{Cl} then they belong to the same cluster and are "associated" [44]. Such a viewpoint, particularly in one-dimensional models [55], can be very fruitful.

A clustering–distance criterion seems to have been understood by Bjerrum and must, surely, also [47] have been in Onsager's mind in making his famous dictum [56] that: "The distinction between free ions and associated pairs depends on an arbitrary convention. Bjerrum's choice is good but we could vary it within reason. In a complete theory this would not matter; what we remove from one page of the ledger would be entered elsewhere with the same effect" [47, 57]. But lacking a complete theory, balancing the ledger may be a tricky task! The main practical difficulty arises because the clustering–distance convention implies that the original monomers must now be treated as having enhanced hard cores, potentially of very large size. This enhancement represents a strong perturbation (except at very low densities) whereas the actual interactions just "inside" these enhanced hard cores may actually be rather small. In light of this we cannot accept Onsager's positive verdict on Bjerrum's choice. In particular, when T falls $R^{Bj}(T)$ *increases* without bound while physical intuition, which we will shortly reinforce, tells one that a typical ionic pair will shrink in size,

becoming ever more tightly bound as $T \to 0$! It is also instructive to compare $R^{Bj} = a/2T^*$ with the screening length $\xi_D = 1/\kappa = (Dk_BT/4\pi q^2 \rho_1)^{1/2}$: if $\rho_1^* \geqslant T^{*3}/\pi(1 - 2T^*)^2$ one has ξ_D less than $R^{Bj} - a$ which seems unphysical; furthermore, at the pure DH critical temperature this occurs when ρ_1 exceeds $\frac{1}{49}\rho_{1c}^{DH}$, a remarkably low density. Finally, the Bjerrum choice leads to a *sharp* disappearance of associated pairs at $T_{Bj}^* = \frac{1}{2}$ (where $R^{Bj} - a$ vanishes). That almost inevitably leads to an artificial nonanalyticity in predicted thermodynamic properties at T_{Bj}.

6.2. The association constant

A more enlightening perspective emerges if, instead of focussing on a microscopic definition of association, one asks how to compute the association constant, $K(T)$. To answer this, consider a general imperfect gas mixture of two species, p and q, with pair potentials $\varphi_{pp}(r)$, $\varphi_{pq}(r)$ and $\varphi_{qq}(r)$, which we choose to regard, in a "chemical" picture [2, 42, 55(a)], as a mixture of *three* species, monomers p and q, and dimers, or associated PQ pairs, s, in equilibrium with species p and q. Then let us introduce the decomposition

$$\varphi_{pq}(r) = \varphi_{pq}^\dagger(r) + \varphi_s(r) , \tag{6.1}$$

where $\varphi_{pq}^\dagger$ represents the *residual* pq pair potential that will still operate in the presence of the associated pairs s: it might be envisaged as the repulsive part of φ_{pq}. Then $\varphi_s(r)$, which is naturally regarded as "attractive", constitutes an *internal* potential for an s "molecule" but will not otherwise play a direct role. Of course this split is arbitrary at this point.

Now the obvious but crucial point is that one would like both descriptions to agree thermodynamically giving, in particular, the *same* equation of state. This desideratum can be investigated by constructing the Mayer activity expansions within the two pictures (following, say, Hill [54]). To ensure a well defined pressure, $p(T, \mu_p, \mu_q)$, the expansions should converge, at least at low densities. To that end, we assume first that the potentials φ_{pp}, φ_{pq}, and φ_{qq} are sufficiently short-ranged (so excluding ionic couplings). Then one easily sees that the expansions can be matched provided (i) the association constant is given by

$$K(T) = \int (e^{-\beta\varphi_{pq}(r)} - e^{-\beta\varphi_{pq}^\dagger(r)}) \, d^d r , \tag{6.2}$$

and (ii) that appropriate further potentials $\varphi_{ps}, \ldots$ and/or higher order association constants are introduced. The former condition ensures consistency at the (original) second virial coefficient level, since $\rho_s \propto \rho_p\rho_q/K$, while the latter enters only at third and higher levels and so may reasonably be neglected at a first approximation.

To illustrate some consequences suppose, for simplicity, that $\varphi_{pq}(r)$ contains a hard core of diameter a but vanishes for $r > R_0$. If one then adopts a clustering–distance convention one must, as indicated, set $\varphi_{pq}^\dagger(r) = \infty$ for $r < R^{Cl}$, and $\varphi_{pq}^\dagger = 0$ for

$r > R^{Cl}$. Of course, it is natural to take $R^{Cl} \geqslant R_0$. Then for $d = 3$ (6.2) yields the Bjerrum-type expression

$$K(T) = 4\pi \int_a^{R^{Cl}} e^{-\beta\varphi_{pq}(r)} r^2 \, dr \, . \tag{6.3}$$

But, as stressed, that entails strong "unnatural" residual pq interactions that will lead to inaccuracies at high T and moderate densities. It is more natural and should provide a better leading approximation, especially at higher densities, to retain in $\varphi_{pq}^{\dagger}(r)$ just the original hard-core part. Then (for $d = 3$) one finds instead

$$K(T) = 4\pi \int_a^{\infty} (e^{-\beta\varphi_{pq}(r)} - 1) r^2 \, dr \, , \tag{6.4}$$

where, for the short-range forces assumed, no upper limit is now needed. Unsurprisingly these two results agree asymptotically at low T whenever φ_{pq} actually has an attractive well. Rather generally, indeed, one may expect a relative insensitivity to details at low T when the association is strong and $K(T)$ is large.

Now, to improve DH theory for ionic systems one would like to follow a similar course by choosing the residual potential $\varphi_{pq}^{\dagger}(r) \equiv \varphi_{+-}^{\dagger}(r)$ to be the original hard-core repulsions *plus* those parts of the attractive Coulombic potential that are satisfactorily treated via the linearization of the PB equation [11, 12, 53]. Clearly, however, the analysis just given fails because the Mayer-type integrals diverge at large distances when the Coulombic potential is used: And, of course, in the density expansions a $\rho^{3/2}$ term now enters *before* the quadratic or "second virial" term which might be matched by choice of $K(T)$.

Nevertheless, as shown explicitly be Ebeling [2, 8, 11] (and reinforced by Yokoyama and Yamatera [53]) one can choose and, we believe, optimally *should* choose a form of $K(T)$ for the RPM such that (a) the exact density expansion of the equation of state is reproduced up to corrections of order $\rho^{5/2}$ while (b) the hard-core repulsions are treated separately (following section 2 as we have done) and (c) the "direct" ionic interactions are treated by DH theory [58] with $x = \kappa a$ where a is *still* the original ionic hard-core diameter. Fortunately, following Mayer's pioneering work and the studies of Friedman [2, 59], the necessary exact analytical results for $d = 3$ have been known for over 30 years. Thence Ebeling found [2]

$$K^{Eb}(T) = 8\pi a^3 \sum_{n=0} b^{2n+4}/(2n+4)!(2n+1) \, , \tag{6.5}$$

$$= 4\pi a^3 \int_1^{\infty} (e^{b/y} + e^{-b/y} - 2 - b^2/y^2) y^2 \, dy \, , \tag{6.6}$$

which may be compared with (5.5), and which, following (5.7) and (5.8), implies

$$Q^{\text{Eb}}(b) = \tfrac{1}{6}b^4 e^{-b} \left[\text{Ei}(b) - \text{Ei}(-b) + 6b^{-1} + 4b^{-3} \right] - \tfrac{1}{6}b(b^2 + b + 2)$$

$$- \tfrac{1}{6}b e^{-2b}(b^2 - b + 2). \tag{6.7}$$

(For a direct but more heuristic derivation see Yokoyama and Yamatera [53], who also tabulate the sum in (6.5).)

Note, first, that K^{Eb} is an analytic function for all $T > 0$ unlike K^{Bj}. Furthermore [53], $K^{\text{Eb}}(T; a)$ decreases monotonically with increasing ionic diameter a at fixed T. in contrast to Fuoss' expression [46, 60], $K^{\text{Fu}} = \tfrac{4}{3}\pi a^3 e^b$, which has an unphysical minimum in a when $q^2/k_B T D a \equiv b = 3$. It follows from (6.7) that $Q^{\text{Eb}}(b) \to 1$ as $T \to 0$, which again invalidates Fuoss' formula [46, 47, 60]. Furthermore, $Q^{\text{Eb}}(T)$ has precisely the same low T asymptotic expansion, (5.9), as Q^{Bj}. Thus, despite its shortcomings, Bjerrum's result [(5.5) with $R = R^{\text{Bj}}$] becomes increasingly reliable when $T \to 0$. Indeed, the discrepancy $\delta \equiv (K^{\text{Bj}}/K^{\text{Eb}}) - 1$ is less than 0.4% for $T^* \leqslant 0.067$ and

$$10^4 \delta \lesssim 20, \qquad 10, \qquad 4, \qquad 1, \qquad 0.2,$$

$$\text{for } T^* = \tfrac{1}{16}, \qquad 0.059, \qquad 0.055, \qquad \tfrac{1}{20}, \qquad 0.046. \tag{6.8}$$

This justifies the use of $K^{\text{Bj}}(T)$ for numerical purposes in the domain of phase separation $T^* \lesssim \tfrac{1}{16}$: see also Fig. 3 which is discussed further below. It also shows that the unphysical banana effects in Fig. 2 *cannot* be attributed to a poor choice of association constant.
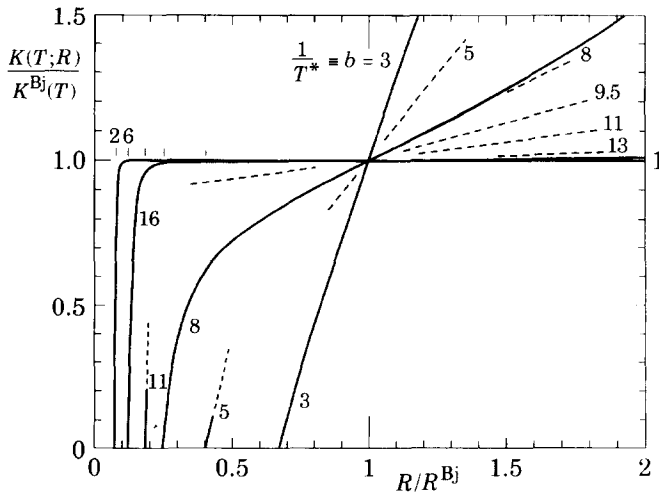

Fig. 3. Plots of the general association constant, $K(T; R) = 4\pi \int_0^R \exp(ab/r) r^2 \, dr$ normalized by the $K^{\text{Bj}}(T)$ versus R/R^{Bj} ($R^{\text{Bj}} = \tfrac{1}{2}ba$) for various values of $b = 1/T^*$. The choice $R = R^{\text{Bj}}$ locates a point of inflection of $K(T; R)$ and, hence, of maximal insensitivity to R. The intersections with the R/R^{Bj} axis correspond to the limits $R \to a+$. Note that for $b \gtrsim 13$ an almost level plateau is reached rapidly for small $R - a$.

It is worth commenting that $K^{Eb}(T)$ can be represented exactly by the general form (5.5) with a suitable choice, say $R^{Eb}(T)$, of the cutoff: to reproduce (6.5) one needs $R^{Eb} - a \approx a/12T^{*4}$ at high T. Alternatively [2], one can take $R \to \infty$ and, by using (6.6), introduce a smooth cutoff function, $W^{Eb}(b/y)$, in place of the sharp Bjerrum cutoff. As we will see, however, the Bjerrum scale $R \sim ba$, which the smooth cutoff retains, bears no useful relation to the microscopic size of a dipolar pair to which issue we now turn.

6.3. The physical size of a dipolar pair

As we have seen, Ebeling's analysis of the association constant, which we accept as optimal, entails *no* commitment as to the size or structure of an "associated dipolar pair". The first point in our approach at which the nature and size of a dipolar ion pair plays a role is in constructing the hard-core free-energy contribution at a level beyond the free-volume approximation (where, as above, we may merely take $B_2 = 2B_1$). Specifically, although the hard-core second virial coefficient $B_{11}^{(2)}$ needs no change, the coefficients $B_{12}^{(2)}$ and $B_{22}^{(2)}$ clearly require some knowledge of the dimensions or effective dimensions of a pair. Nevertheless, at the low densities of main interest, the results for the equation of state will be relatively insensitive to the precise approximations made (as in Fig. 1; see also below).

At this point, however, it is clear that to understand and, hopefully, correct the unphysical banana effect, it is essential to study the electrical interactions of a dipolar pair. Since the dipole moment depends directly on the separation between charges we must address the question of how large a dipolar pair "really" is.

Now, as discussed, Bjerrum took as a measure of the size of a dipolar pair the scale $R^{Bj} = \frac{1}{2}ba$. Granted the basic Boltzmann weight $4\pi r^2 \exp(ba/r)$ in (5.5) this seems, at first sight, the most natural length scale set by the problem. It has the merit, evident in Fig. 3, that it locates that cutoff point at which the value of K is least sensitive to the choice of R. Nevertheless, it is clear from the figure that for $b \lesssim 9$, i.e. $T^* \gtrsim 0.11$, the value of K does depend strongly on R. Conversely, for $b > 13$ or $T^* \lesssim 0.070$, there arises a remarkably broad plateau over which $K(T; R)$ is almost independent of R. In essence this constancy explains why $K^{Bj}(T)$ is an excellent approximation to $K^{Eb}(T)$ for $T^* \lesssim 0.07$: see (6.8).

An immediate corollary of the plateau, obvious from the figure, is that at low T there is *another, shorter* length scale measuring the rate at which $K(T; R)$ rises to its plateau value. This length, say $a_0 \equiv (1 + s_0)a$, may be specified precisely by taking the initial slope of the plots, set by the Boltzmann weight *at contact*, namely $4\pi a^2 e^b$, and using $K^{Bj}(T)$ as a measure of the plateau value. On recalling (5.6)–(5.9) this yields

$$s_0(T) \equiv K^{Bj}(T)/aK'(T; a) = Q_0(b; \tfrac{1}{2})/b. \tag{6.9}$$

Of course one might more satisfactorily replace $Q^{Bj} = Q_0(b; \tfrac{1}{2})$ by $Q^{Eb}(b)$ but, as seen, the differences will be inconsequential in the region of criticality and below that is of

interest here. A moment's consideration shows that a typical Bjerrum pair must have an ion-to-ion surface separation of magnitude $s_0(T)a$, at least at low T (where, again, the definition is insensitive to the cutoff). Thus $s_0(T)$ should provide a reasonable measure of the "true" size of a dipolar pair.

Perhaps more natural is to recognize the fluctuations in pair size and so define an average separation $\langle r \rangle \equiv a_1 \equiv [1 + s_1(T)]a$ using the basic Boltzmann factor; a cutoff is again needed and we adopt R^{Bj}. (A cutoff could be avoided here, but at a cost in analytical complexity, by using the effective, *subtracted* Boltzmann factor appearing in Ebeling's integral in (6.6).] More generally, we may examine the moments of $s = (r - a)/a$. Thus, on introducing

$$Q_l(b; c) = b \int_0^{bc - 1} e^{-bs/(1 + s)} s^l (1 + s)^2 \, ds,$$ (6.10)

which extends the definition of Q_0 in (5.7), we obtain, for $l > 0$,

$$[s_l(T)]^l \equiv \left\langle \left(\frac{r - a}{a} \right)^l \right\rangle = \frac{Q_l(b; 1/2)}{Q_0(b; 1/2)},$$ (6.11)

$$= \frac{l!}{b^l} \left(1 + l \frac{l + 5}{b} + l \frac{l^3 + 12l^2 + 41l + 38}{2b^2} + \cdots \right).$$ (6.12)

where, as previously, the expansion is only asymptotic but is valid for low T when truncated.

From these results and (5.9) one finds $s_0(T) \approx s_1(T) \approx T^*$ and $T \to 0$ with a ratio $s_1/s_0 = 1 + 2T^* + 14T^{*2} + \cdots$. Thus, as anticipated on intuitive grounds, the dipolar pairs rapidly approach hard-core contact at low T. From the behavior of the higher moments, or more directly from (6.10), one sees that the low-T distribution of s approaches Poissonian form, i.e., $\mathcal{P}(s)\, ds \approx e^{-s/s_0}\, ds/s_0$, which is fairly sharp. The behavior at higher values of T^* can be read off from Fig. 4: $s_0(T)$ remains below 0.11 up to $T^* \simeq 0.07$, well above T_c^{DH}; although $s_1(T) = \langle r \rangle/a - 1$, rises more rapidly, as the distribution $\mathcal{P}(s)$ broadens, it remains below 0.2 in the same range [61]. At the estimated critical point (3.4) one has $\langle r \rangle$ less than $1.10a$. This certainly justifies the cutoff convention $R^G = 2a$ which Gillan used in his study of cluster free energies below $T^* = 0.050$ [44]. One may also enquire as to the values of $K(T; R)$ if one uses a cutoff scaled with $s_0(T)$ in place of $R^{Bj}(T)$. On choosing, for example, $R = 3s_0$ one finds a reduction below K^{Bj} of 5.0% at the lowest temperatures rising to 7% at $T^* \simeq 0.03$, and 15% at $T^* \simeq 0.06$.

At temperatures above $b = 10$ or so the values of s_0 and s_1 become increasingly sensitive to the choice of the Bjerrum cutoff. (As mentioned, one could pursue better definitions if this regime were of special interest.) But it remains surprising that even within Bjerrum's own conception, the mean pair separation, ionic-surface-to-ionic-surface never exceeds $0.77a$ and, clearly, bears no relation to the cutoff $R^{Bj} = \frac{1}{2}ba$.
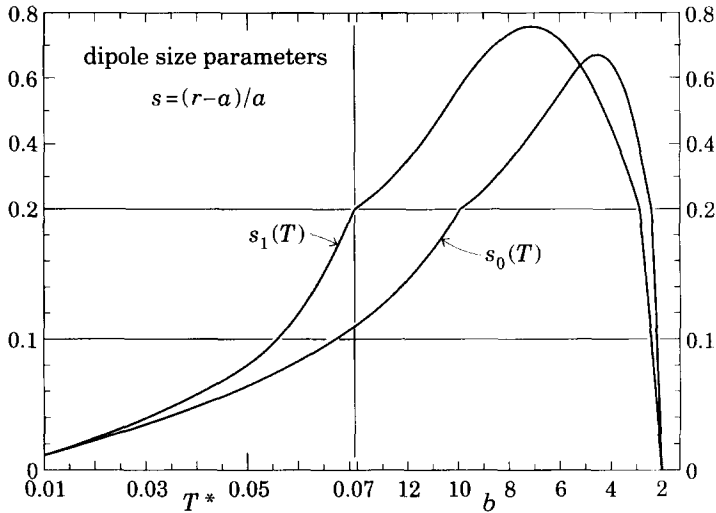
Fig. 4. Temperature variation of the ion-pair size parameters, $s_0(T) = (r_0 - a)/a$, based on the Coulombic distribution at contact normalized by $K^{Bj}(T)$, and $s_1(T) = (\langle r \rangle - a)/a$, using the Coulombic distribution with Bjerrum cutoff. Note the break in the vertical scale at $s = 0.200$ and in the horizontal scale between T^* and $b = 1/T^*$ (at $T^* \simeq 0.071 \simeq 1/14.1$).

Having elucidated the physical sizes of associated ion pairs in the low temperature regime ($T^* \lesssim \frac{1}{10}$, say) we now enquire into the thermodynamics of a dipole in a sea of free ions.

7. Solvation free energy of an ionic pair

An associated ion pair is electrically neutral and hence does not contribute directly to the inverse Debye length defined generally in (3.11). Furthermore, at the temperatures of principal interest to us, say up to $T^* = 1.5T_c^* \simeq 0.08$, a typical associated pair is quite small with center-to-center diameter $a_1 \cong \langle r \rangle \lesssim 1.3a$ (see Fig. 5). Nevertheless, this nonvanishing charge separation results in a significant electric dipole moment, qa_1, which generates a (bare) electric potential decaying only as $1/r^2$. That is nonintegrable so that even at the lowest overall densities one must expect significant interaction with the free ions. These will screen the bare dipolar field and the resulting ionic correlations will reduce the free energy of a pair in the electrolyte. Evidently, this dipole–ionic interaction leads to an additional, solvation free energy contribution $\bar{f}^{DI} \propto \rho_2$ that should be added to the previous DH, Bj, and HC terms. Going beyond the DHBj theory in this way corrects for its major physical defect: the total neglect of the strong Coulombic character of an ion pair. We may reasonably hope that it will eliminate the banana effect in Fig. 2 since the system can lower its free energy by dissolving associated pairs in the ionic liquid so draining away the excessive reservoir of ideal pairs predicted by pure Bjerrum theory.
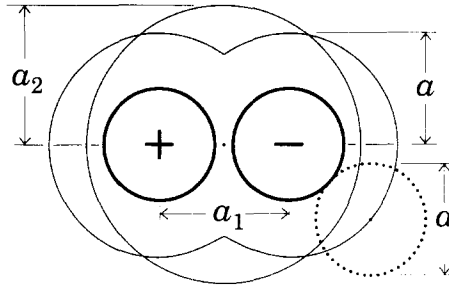
Fig. 5. Depiction of a closely associated pair of oppositely charged ions of diameter a at separation a_1. The true exclusion zone is the bispherical region traced by the center of a contacting free ion (indicated by the dotted circle). The circle of radius a_2 represents the approximating spherical exclusion zone centered symmetrically on the coordinate origin (central dot).

7.1. The approximating sphere

To compute $\bar{f}^{\text{DI}}(T; \rho_1, \rho_2)$ we return to the DH approach; instead of considering the energy of a single ion in the sea of free ions and deriving the linearized-PB or DH equation, (3.10), we now study a closely associated pair of oppositely charged ions in the ionic fluid: see Fig. 5. The mean electric potential, $\phi(r)$, close to each ionic center at $r_+ = 0$ and $r_- = 0$, must still satisfy Laplace's equation and approach the pure-Coulombic-plus-background form (3.8) with $r = |r_\pm|$, $q_i = \pm q$, and, by symmetry, $\psi_i = \pm \psi$. Likewise the potential in the ionic fluid away from the ion pair still satisfies the DH equation. Finally, $\phi(r)$ and $\nabla\phi(r)$ must again be everywhere continuous. The major new feature is that these conditions must now be imposed on the boundaries of the bispherical hard-core exclusion region shown in Fig. 5 and defined by $|r_+|, |r_-| \leqslant a$.

This problem has been analyzed recently [27(b,c)] in a somewhat broader setting but for large separations requiring, in particular, $a_1 > 2a$: then the exclusion zone becomes two separated spheres. But when $a_1 < 2a$, as here, the technical difficulties of obtaining a precise, numerically tractable solution have not been overcome [27(c)]. Accordingly, we adopt a strategy first used, to our knowledge, by Kirkwood [48] in considering zwitterions: specifically, the ideal bispherical exclusion zone will be approximated by a symmetrically centered sphere of radius a_2: see Fig. 5.

The optimal choice of $a_2(T)$ would, clearly, be that value which yields the exact but, alas, unknown answer! Simple geometrical considerations (see Fig. 5) suggest bounds [5], namely, (i) $a_2 \leqslant a + \frac{1}{2}a_1$, for which the approximating sphere can just envelop the bispherical region, and vice-versa (ii) $a_2 \geqslant a(\frac{3}{4} - \frac{1}{2}s_1 - \frac{1}{4}s_1^2)^{1/2}$; for $T^* \lesssim \frac{1}{16}$ these yield $0.825a < a_2 < 1.565a$ and have a mean value 1.19_5a which falls to $1.183a$ as $T^* \to 0$. Matching the *volumes* of bisphere and approximating sphere yields $a_2 \simeq 1.19_1a$ when $a_1 = a$ (as at $T = 0$). However, we opt to set a_2 equal to the angular average radius of the bisphere, which seems somewhat more appropriate: this yields

$$a_2(T = 0) \simeq \bar{a}_2 = 1.1619_8a \quad (a_1 = a). \tag{7.1}$$

It would, clearly, be appropriate to incorporate a temperature dependence by taking, say, $a_2(T) \simeq a_2(0) [1 + c_1 s_1(T)]$. The behavior of the upper bound suggests $c_1 \lesssim \frac{1}{3}$ so that up to T_c^{DH} an increase in a_2 of only 4% is expected. Since, however, the uncertainties in the approximation for $a_2(0)$ are probably at least 3% this refinement is hardly worthwhile. Rather we propose to vary a_2 in subsequent numerical studies (see below) to check that there is no undue sensitivity to the choice made. Note, furthermore, that a full treatment would also require averaging the solvation free energy over a thermal distribution of ionic separations: see Section 6.3 above.

7.2. Dipolar potential

To solve the boundary value problem posed in terms of the approximating sphere, we adopt polar coordinates with origin at the center of symmetry and z axis passing through the centers of the two ions at $z = \pm \frac{1}{2} a_1$. Appealing to the azimuthal symmetry the solution in the region $\frac{1}{2} a_1 < r \leqslant a_2$ can be expanded as

$$\phi_< (r, \theta) = \sum_{l=0}^{\infty} (A_l r^l + c_l r^{-l-1}) P_l (\cos \theta) . \tag{7.2}$$

As indicated above, this solution must consist of two parts: one part is the potential due to the ionic pair alone which must survive (out to $r = \infty$) in the absence of any free ions. This can be written

$$\phi_0(r, \theta) = \frac{q}{D r_+} - \frac{q}{D r_-},$$

$$= \frac{2q}{Dr} \sum_{j \geqslant 0} P_{2j+1} (\cos \theta) (a_1/2r)^{2j+1} , \tag{7.3}$$

the expansion being valid for $r > \frac{1}{2} a_1$. Comparison with (7.2), noting that the A_l must vanish when $\kappa^2 \sim \rho_1 \to 0$, determines the coefficients

$$c_{2j} = 0, \qquad c_{2j+1} = 2(q/D) (\tfrac{1}{2} a_1)^{2j+1} . \tag{7.4}$$

The second part of $\phi_< (r, \theta)$ represents the potential within the exclusion zone due to the external, induced screening charge. This must be regular for $r \leqslant a_2$ and determines the mean electrostatic energy of the associated ion pair as

$$q \psi_2(\{q_j\}; T, \rho_1) = 2q \sum_{j=0}^{\infty} A_{2j+1} (\tfrac{1}{2} a_1)^{2j+1} , \tag{7.5}$$

using $(\pm)^l P_l (\pm 1) = 1$. Note that successive contributions can be associated with the pure multipole moments of the discrete pair of ions at spacing a_1.

To expand the exterior solution it is expedient to introduce the spherical Bessel functions [62]

$$k_l(x) = g_l(x) e^{-x} / x^{l+1} . \tag{7.6}$$

The first few polynomials introduced here are $g_0 = 1$, $g_1 = x + 1$, $g_2 = x^2 + 3x + 3$, ..., and the general coefficient of x^{l-m} ($m = 0, 1, ..., l$) is

$$g_{lm} = (2m - 1)!! \binom{l + m}{2m}, \tag{7.7}$$

adopting the convention $(-1)!! = 1$ [62]. The coefficients increase uniformly from $g_{l0} = 1$ to $g_{ll} = (2l - 1)!!$ so that $g_l(x)$ is bounded by $(l + 1)(2l - 1)!!$ when $x \lesssim 1$. Conversely, for $x \gtrsim 1$ one has $g_l(x) < (l + 1)(2l - 1)!! x^l$, while $g_l(x) \approx x^l$ as $x \to \infty$.

Now the solution for $r \geq a_2$ vanishing when $r \to \infty$ can be written

$$\phi_>(r, \theta) = \sum_{l=0}^{\infty} B_l k_l(kr) P_l(\cos \theta). \tag{7.8}$$

Matching $\phi_<$ and $\nabla \phi_<$ to $\phi_>$ and $\nabla \phi_>$ on $r = a_2$ yields the relations

$$A_l = - C_l k_{l-1}(x_2) / a_2^{2l+1} k_{l+1}(x_2), \tag{7.9}$$

$$B_l = (2l + 1) C_l / a_2^{l+1} x_2 k_{l+1}(x_2), \tag{7.10}$$

with $x_2 \equiv \kappa a_2$. Finally, the energy of a dipole in the ionic atmosphere can be written

$$q\psi_2 = - \frac{q^2 \kappa^2 a_1^2}{D a_2} \sum_{j=0}^{\infty} \frac{g_{2j}(\kappa a_2)}{g_{2j+2}(\kappa a_2)} \left(\frac{a_1}{2a_2} \right)^{4j}. \tag{7.11}$$

7.3. Dipole–ion interaction free energy

Let us examine the result (7.11) for the energy of interaction of a dipolar pair with the induced screening charge. First note that for temperatures of interest the multipolar sum converges rapidly, the terms decreasing as $1/\chi^j$ with $\chi = (2a_2/a_1)^4 \simeq 19$ to 30. Furthermore, from (7.7) one sees that the coefficients $h_j \equiv g_{2j}/g_{2j+2}$ are reliably estimated by $[1 + O(x_2^2)]/(4j + 3)(4j + 1)$ even up to $x_2 \simeq 1$, which is appropriate for the critical region, while for $x \to \infty$ one has $h_j(x) \approx 1/x^2$. It follows that the sum in (7.11) is very well approximated for all relevant T and x by its first term which represents just the ideal dipole contribution ($l = 1$). In an explicit test the first term fell short of the full sum by about 0.7% for $x_2 = 1$ and by less than 2% for $x_2 = 5$. Accordingly, we will henceforth retain only the first term (although there are no obstacles to working with the full expression (7.11) should it be desired).

Finally, to obtain the free energy one must carry out the Debye charging process (3.6). The result can be conveniently written

$$\bar{f}^{\text{DI}}(T; \rho_1, \rho_2) = \frac{\kappa^2 a a_1^2}{T^* a_2} \rho_2 \omega_2(\kappa a_2) \approx \frac{\pi a_1^2 \rho_1^*}{3 T^{*2} a a_2} \rho_2, \tag{7.12}$$

where one has

$$\omega_2(x) = 3[\ln(1 + x + \tfrac{1}{3} x^2) - x + \tfrac{1}{6} x^2] / x^4$$

$$= \tfrac{1}{12}(1 - \tfrac{4}{5} x + \tfrac{4}{9} x^2 - \tfrac{4}{21} x^3 + \cdots). \tag{7.13}$$

The prefactor a_1^2 in (7.12) is recognized as deriving from the square of the dipole moment, qa_1. The appearance of the factor $1/a_2$ is hard to understand at first sight but it can be justified with the aid of the following heuristic picture of the Debye screening of the charges "seen" by the ionic fluid on the boundary surface of an excluded zone.

To be specific, suppose an effective surface charge is q_σ; then we postulate that an effective localized screening charge $q_s = -q_\sigma$ is induced in the electrolyte at a distance $\xi_D = 1/\kappa$ from the surface. To check this, consider a single, spherical test ion of charge q_0 and diameter a_0. The exclusion *radius* is then $a_2 = \frac{1}{2}(a_0 + a)$. The total surface charge q_σ must, clearly, be spread uniformly over the surface of the exclusion sphere and, in order to create the same, bare long-distance field, should equal q_0. The localized screening charge $q_\sigma = -q_0$ then resides uniformly spread on a sphere of radius $a_2 + \xi_D$. At the origin, where the original charge q_0 is located, the screening charge generates a potential $\psi_1 = q_s/D(\xi_D + a_2)$. Thence the screening energy is $U_1 = q_0\psi_1 = -q_0^2\kappa/D(1 + \kappa a_2)$. This is, indeed, the *exact* DH result [27(c)] as can be checked from (3.14) for the standard case: $q_0 = q$, $a_0 = a$.

Now consider the dipole pair with charges $\pm q$ on the z axis at $z = \pm\frac{1}{2}a_1$ and an exclusion sphere of radius a_2. The surface charges must reproduce the same dipole moment qa_1 (to preserve the long distance field). To simplify the calculation we suppose they also are localized on the axis at $z = \pm a_2$. Then they have values $q_\sigma = \pm qa_1/2a_2$. The complementary screening charges of magnitude $q_s = qa_1/2a_2$ may likewise be localized on the axis at $z = \pm(a_2 + \xi_D)$. Finally, their interactions with the original charges yield the electrostatic energy

$$U_2 = \frac{q}{D}\left(\frac{-2q_s}{a_2 + \xi_D - \frac{1}{2}a_1} + \frac{2q_s}{a_2 + \xi_D + \frac{1}{2}a_1}\right)$$

$$= -\frac{q^2\kappa^2 a_1^2}{Da_2(1 + x_2)^2}\left[1 + \left(\frac{a_1}{2a_2}\right)^2\frac{x_2^2}{(1 + x_2)^2} + \cdots\right]. \tag{7.14}$$

We see that the leading factor *precisely* reproduces the $j = 0$ term in (7.11) when x is small apart from a missing "geometrical" factor of $\frac{1}{3}$. That presumably results from the diffuse and off-axis nature of the true screening charge distribution. [Furthermore, the quadratic prefactor in x_2 in (7.14) and the correction terms, of order $(a_1/2a_2)^2$ and entailing a rational function of x of degree zero, closely resemble the structure of (7.11)!]

In summary, then, the leading dependence on a_2 in the solvation free energy (7.12) arises because the screening charge is of magnitude qa_1/a_2. These heuristic considerations also suggest other types of average, specifically invoking electrostatic weighting, for estimating a_2 optimally. However, it does not seem worthwhile at this stage to pursue such possibilities.

8. Coexistence and criticality in DHBjDI theory

We define DHBjDI theory by adding the dipole–ionic coupling term, $\bar{f}^{DI}(T; \rho_1, \rho_2)$, just evaluated in (7.12) and (7.13), to the original free energy (5.12) for

the DHBj theory and adopting Ebeling's association constant $K^{Eb}(T)$ as given in (6.5)–(6.7), although we may recall that for low temperatures $K^{Bj}(T)$ [see (5.5)–(5.11)] represents a more than adequate approximation which has been used for convenience in the calculations quoted below. Note also that the full specification of $\bar{f}^{DI}$ requires the values of $a_1(T)$ and of $a_2(T)$ as discussed in Sections 6.3 and 7.1. The coexistence curve may now be calculated by matching chemical potentials and pressure, expressed as functions of the densities in the vapor and liquid phases, using the mass action law (5.3) to relate ρ_1 and ρ_2. Iterative processes are involved in the computer programs employed and precision may suffer near criticality and at low temperatures ($T^* \lesssim 0.03$). Since, on the one hand, one knows, as explained previously, that the predicted critical behavior must be classical and, on the other hand, accuracy of greater than $\pm 1\%$ is hardly needed, no special effort has been made to optimize the numerical analysis.

8.1. Criticality with dipole–ion interactions

For a first calculation [5] consider pure DHBjDI theory, with no hard-core terms, and let us set

$$a_1^* \equiv a_1/a \Rightarrow 1, \qquad a_2^* \equiv a_2/a \Rightarrow \bar{a}_2^* = 1.1619_8, \tag{8.1}$$

both corresponding to $T = 0$. The results [5] are shown in Fig. 6a: for a plot with a larger horizontal scale, see Fig. 3 in Ref. [5]. Note that the open horizontal "boxes" and square "diamonds" represent the Monte Carlo simulation estimates of Panagiotopoulos [14]; as discussed [14–17], however, they are subject to some revision in light of later studies.

Evidently the overall shape of the coexistence curve is now, once again, perfectly reasonable! The critical parameters are found to be

$$x_c = \kappa_c a = 1.123, \qquad T_c^* = 0.0574_0,$$

$$\rho_c^* = 0.0277_5, \qquad Z_c = 0.2236, \tag{8.2}$$

These values (some of which differ slightly from those quoted in Ref. [5]) may be contrasted with the DH and DHBj results in Figs. 1 and 2 and Eqs. (4.8), (4.9), (5.13) and (5.14), and compared with the simulation-based estimates (3.5) and (3.6), the former being represented by the vertical box in Fig. 6. Inclusion of the dipole–ionic interactions clearly yields a significant improvement in the estimates of both T_c^* and ρ_c^*. The value of $\kappa_c a \propto \rho_{1c}$ implies $\rho_{2c}/\rho_{1c} = 1.91$ for the critical-point ratio of dipole to free ion densities: that seems quite reasonable [44]. The critical osmotic coefficient Z_c is seen to be smaller than in DHBj theory by a factor of about 2 and so is much closer to the estimate (3.6); that, however, must be regarded as rather uncertain [37].

The natural next step [5] is to include the hard-core repulsions. As before, it is simplest to invoke the free-volume form (2.12) with $B_2 = 2B_1$. With $B_1 = (4/3\sqrt{3})a^3$,
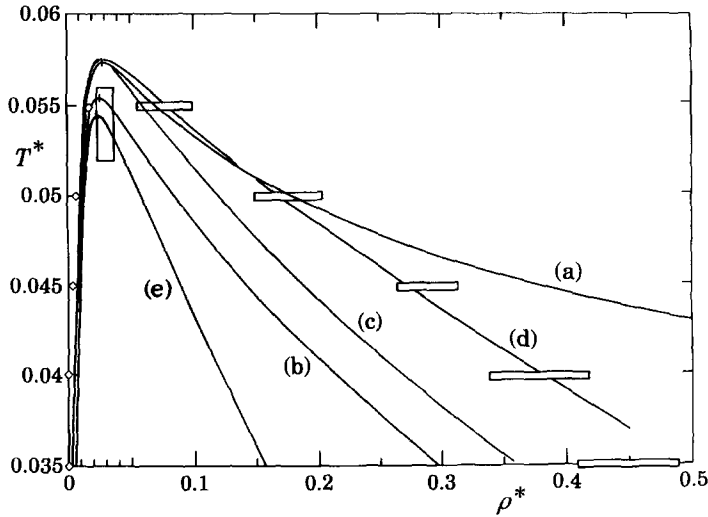
Fig. 6. Coexistence curves predictions for the RPM within DHBj theory augmented with dipole–ion (DI) interactions: (a) with no hard-core terms and $a_1 = a$, $a_2 = \bar{a}_2$; (b) with added free-volume hard-core terms with $B_1 = \frac{1}{2}B_2 = (4/3\sqrt{3})a^3$; (c) with the same hard-core terms but a_1 increased to $1.070a$ (corresponding to $s_0 \simeq 0.070$ in Fig. 4); (d) with $a_1 = a$ but including a further phenomenological term varying as $\rho_2^{4/3}$ with a coefficient chosen to fit the Monte Carlo results of Ref. [14]; and (e) with $a_1 = 1.070a$, $a_2 = \bar{a}_2$, and $B_1 = \frac{1}{2}B_2 = \frac{2}{3}\pi a^3$, only. The Monte Carlo estimates are represented by the open diamonds and horizontal boxes but are subject to revision: see Refs. [16, 17]. The vertical box locates the current simulation-based critical point estimates (3.5). Note, in comparison with Figs. 1 and 2, the differences in vertical and, especially, horizontal scales.

corresponding to bcc close packing, the coexistence curve narrows and T_c falls: see plot (b) in Fig. 6. The critical parameters are now

$$\kappa_c a = 1.028_5, \quad T_c^* = 0.0554_3, \quad \rho_c^* = 0.0259_0, \quad Z_c = 0.248_4, \tag{8.3}$$

corresponding to percentage shifts of -8.4, -3.4, -6.7 and $+11$, respectively. The critical dipole-to-ion ratio increases by 19% to $\rho_{2c}/\rho_{1c} = 2.28$. The coexistence curve evidently now falls below the original Panagiotopoulos estimates but the critical point, in particular, lies closer to the more recent estimate of Orkoulos and Panagiotopoulos [16, 17].

The bcc choice of B_1 underestimates the true ion–ion hard-sphere second virial coefficient by a factor around 2.7 so that significantly larger shifts in critical parameters must be anticipated. Indeed, $B_1 = \frac{1}{2}B_2 = \frac{2}{3}\pi a^3$ yields

$$\kappa_c a = 0.922_7, \quad T_c^* = 0.0522_9, \quad \rho_c^* = 0.0243_8, \quad Z_c = 0.282_2, \tag{8.4}$$

so the critical point is still consistent with the region of uncertainty currently estimated from the simulations [17]. The ratio ρ_{2c}/ρ_{1c} rises to 2.94. The free-volume form with the correct second-virial coefficient, as used here, significantly overestimates the further repulsive terms so $\kappa_c a$, T_c^*, and ρ_c^* here are effectively lower bounds (within the theory).
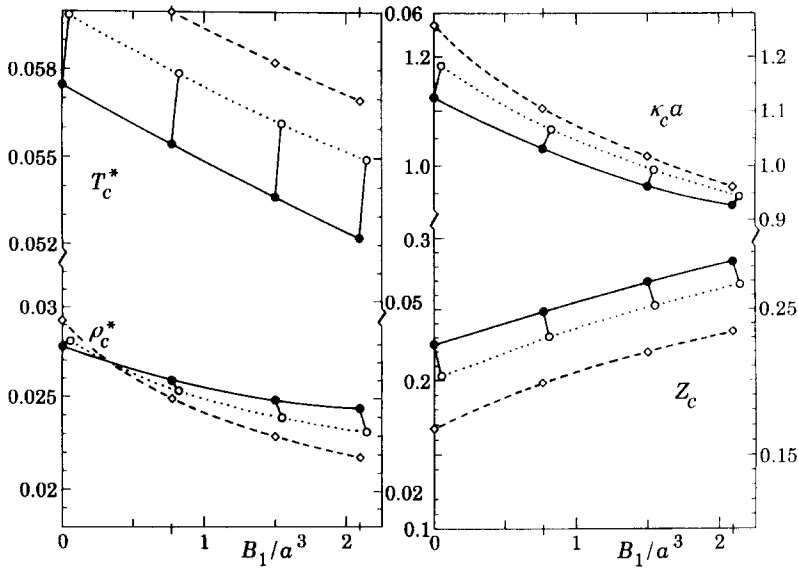
Fig. 7. Some data illustrating the sensitivity of the critical parameters T_c^*, ρ_c^*, $\kappa_c a$, and $Z_c = p_c/\rho_c k_B T_c$ within DHBjDI theory to the reduced dipole dimensions $a_1^* \equiv a_1/a$ and $a_2^* \equiv a_2/a$ for different values of the hard-core, free-volume, second-virial coefficients $B_1 = \frac{1}{2}B_2$. Note $B_1/a^3 = 0, 0.769_8$, and 2.094_4 correspond to *no* hard-core terms, bcc close packing, and the true ion–ion hard-core second virial coefficient, respectively. In the plots the curves are mainly guides to the eye: (a) solid circles and solid curves denote $a_1^* = 1$, $a_2^* = \bar{a}_2^* = 1.162_0$; (b) open diamonds and dashed curves, $a_1^* = 1.150$, $a_2^* = 1.162_0$; and (c) open circles, displaced slightly to the right, and dotted curves, $a_1^* = 1$, $a_2^* = 1.100$. The most appropriate values for a_1^* are probably 1.05–1.10: see Fig. 4; but a_2^* is less certain.

However, we must still recognize the uncertainties in a_2 and allow for the temperature variation of a_1 and a_2. Fig. 7 displays the estimates (8.2)–(8.4) graphically and is designed to illustrate the sensitivity to the expected range of variations in a_1 and a_2. Note that $a_1^* = 1.150$ (open diamonds and dashed lines) corresponds, via Fig. 4 with $s_1(T) \cong a_1^* - 1$, to $T^* \simeq 0.065$ which is somewhat above criticality. The coexistence curve for the lower assignment $a_1 = 1.070a$ with the bcc free-volume hard-core terms is represented by plot (c) in Fig. 6. Evidently, T_c increases (to $T_c^* = 0.0575$) and the coexistence curve is in somewhat better accord with the Monte Carlo estimates [14–17] than plots (a) and (b). If, further, the true second-virial coefficient is used, as in plot (e), both T_c^* and ρ_c^* lie in the Monte Carlo range but the coexistence curve below $0.95T_c$ narrows significantly.

When the free-volume approximation is adopted with the true second-virial coefficient, i.e., $B_1 = \frac{1}{2}B_2 = \frac{2}{3}\pi a^3$, one can extend the set of values (8.4) roughly via

$$T_c^*(a_1, a_2) \simeq T_0^*[1 + c_1(a_1^* - 1) + c_2(a_2^* - 1)],\qquad (8.5)$$

with $T_0^* = 0.05897$, $c_1 = 0.528$, and $c_2 = -0.700$. Clearly, the estimates are fairly sensitive to the adopted value of a_2. Indeed, at the anticipated extreme $a_2^* = 1.50$, the

critical parameters become $\kappa_c a \simeq 1.048$, $T_c^* \simeq 0.0413$, $\rho_c^* \simeq 0.0400$ and $Z_c \simeq 0.288$ for $a_1^* = 1$ and $B_1 = \frac{1}{2}B_2 = (4/3\sqrt{3})a^3$. Thus a more sophisticated analysis of the mean exclusion radius of a dipolar bisphere would be worthwhile.

8.2. DHBjDI coexistence at low temperatures

As seen in Sections 7.2 and 5.3, the predictions of the various theories at low temperatures are quite revealing. Indeed, to be certain that inclusion of the DI coupling always banishes the banana effect, a low-T analysis is really called for. It also serves to exhibit further features of the DHBjDI theory.

Now $\bar{f}^{DI}$ varies inversely with a_2 and, since qa_1 is the dipole moment, it is also proportional to a_1^2. Consequently it is convenient to use the parameters

$$x_2 \equiv \kappa a_2, \qquad \Delta_k \equiv a_1^2 a^{k-2}/a_2^k . \tag{8.6}$$

The dipole–ion contributions to the chemical potentials and pressure are then

$$\bar{\mu}_2^{DI} = -4\pi\Delta_1 \omega_2(x_2)\rho_1^*/T^{*2} , \tag{8.7}$$

$$\bar{\mu}_1^{DI} = -2\pi\Delta_1 \omega'(x_2)\rho_2^*/T^{*2} , \tag{8.8}$$

$$\tilde{p}^{DI} \equiv (\beta pa^3)^{DI} = \bar{\mu}_1^{DI}\rho_1^* , \tag{8.9}$$

where $\omega_2(x)$ is given by (7.13) while

$$\omega'(x) \equiv \frac{\partial}{x \, \partial x}[x^2\omega_2(x)] = \tfrac{1}{6} - \tfrac{1}{5}x + O(x^2),$$

$$\approx 3x^{-3} - 12x^{-4}\ln x + \cdots, \quad (x \to \infty). \tag{8.10}$$

Finally, from (5.7), (5.9), and (6.7) et seq., it is helpful to recall the low-T behavior of the association constant, namely,

$$\ln[K(T)/a^3] = T^{*-1} + \ln 4\pi T^* + 4T^* + \cdots . \tag{8.11}$$

To obtain the phase boundary the pressures and chemical potentials of the vapor (v) and liquid (ℓ) must be matched. This yields four rather complex equations for ρ_{1v}, ρ_{2v}, $\rho_{1\ell}$, and $\rho_{2\ell}$. To make progress when $T \to 0$ we suppose the solutions then satisfy:

A1: $\rho_{1v}/T \to 0$,   **A2**: $\rho_{2v}/T^2 \to 0$, \tag{8.12}

A3: $x_\ell^2 \propto \rho_{1\ell}/T \to \infty$. \tag{8.13}

Having obtained the solutions, these assumptions must be checked. Note that **A1** and **A2** imply $\rho_{vap}(T) \to 0$ and, hence, the absence of any banana effect. Using **A1**, 2, and 3 the matching chemical potential equations can be solved successively in terms of $\rho_{2\ell}$, which is then determined by the pressure equation. The derivation is quite informative but rather intricate: it is outlined in Appendix A while the results are described here.

Consider, first, pure DHBjDI theory with *no hard-core* contributions. On the vapor side of the coexistence curve one has

$$\rho_{1v}^* = \frac{3\Delta_3^{1/2}\Delta_4}{8\pi T^{*3/2}}\, e^{-(2+\Delta_3)/4T^*}\left(1 - \frac{6}{\Delta_3}T^* + \cdots\right), \tag{8.14}$$

$$\rho_{2v}^* = \frac{9\Delta_3\Delta_4^2}{64\pi T^{*2}}\, e^{-\Delta_3/2T^*}\left(1 - \frac{12}{\Delta_3}T^* + \cdots\right), \tag{8.15}$$

where the ellipses denote terms of order $T^{*3/2}$ and $\exp[-(2+\Delta_3)/4T^*]/T^{*7/2}$. Clearly A1 and A2 are valid and the vapor density vanishes exponentially with T in standard fashion.

The vapor pressure curve follows in leading order simply from

$$p_\sigma(T) \approx (\rho_{1v} + \rho_{2v})k_B T \sim e^{-\Delta_3/2T^*}/T^*. \tag{8.16}$$

On the liquid side one finds

$$\rho_{1\ell}^* = \frac{9\Delta_4^2}{16\pi T^*}\left(1 - \frac{4\sqrt{\pi}}{9\Delta_4}T^*\ln T^* + O(T^*)\right), \tag{8.17}$$

$$\rho_{2\ell}^* = \frac{9\Delta_3\Delta_4^2}{64\pi T^{*2}}\left(1 - \frac{12}{\Delta_3}T^* + \frac{4}{\Delta_3\Delta_4}T^{*3/2} + O(T^{*2}\ln T^*)\right), \tag{8.18}$$

so that $\rho_{1\ell}/\rho_{2\ell} \approx 4T^*/\Delta_4$ and A3 is verified. These results may be compared with the pure DH predictions (4.15)–(4.17).

Of course, the divergence of $\rho_{1\ell}$ and $\rho_{2\ell}$ as $1/T$ and $1/T^2$ is not realistic and is suppressed by the inclusion of hard-core terms. However, at intermediate temperatures when ρ_{liq}^* is not too large, (8.17) and (8.18) are probably numerically reasonable. In Appendix A the effects of including the *hard cores* within the *free-volume* approximation (2.12) with $B_2 = 2B_1 = 2B^*$ are studied: but note that only the form of divergence of the pressure near $\rho_{\max}$ plays a role [35]. On the liquid side of the coexistence curve one now finds

$$B^*\rho_{1\ell}^* = k_0 T^{*1/3} - \tfrac{2}{3}k_0^2 T^{*2/3} + O(T^*), \tag{8.19}$$

$$2B^*\rho_{2\ell}^* = 1 - k_0 T^{*1/3} - (k_1 - j_0)T^{*2/3} + O(T^*), \tag{8.20}$$

with coefficients depending on $\rho_{\max}^* \equiv 1/B^*$ and a, a_1, and a_2 according to

$$k_0 = \frac{9a^2}{\pi a_2^2}\left(\frac{B^*}{4}\right)^{1/3}, \qquad k_1 = \frac{27a^4}{2\pi^2 a_2^4}(2B^*)^{4/3}, \tag{8.21}$$

$$j_0 = 4\sqrt{\pi}/3\Delta_4(2B^*)^{1/3}. \tag{8.22}$$

Not surprisingly, the free-ion density, $\rho_{1\ell}(T)$, vanishes when $T \to 0$ and so $\rho_{2\ell}$ approaches its close-packing value $1/2B^*$.

On the vapor side of coexistence the inclusion of hard-core terms does not change the leading exponential behavior; however, the correction factors become more

complex and reflect the approach to close packing. Specifically, following the Appendix, one finds for the free ions

$$
\rho_{1v}^* = e^{-1/2T^*} \sqrt{\rho_{2v}^*/\pi T^*} \left[1 + O(\rho_{2v}^*/T^{*2})\right]
$$
$$
\sim \exp\left[-(2 + \Delta_3)/4T^*\right], \tag{8.23}
$$

while the dipolar-pair density varies as

$$
\rho_{2v}^* = \Gamma_0 \exp\left[-\Delta_3/2T^* + \Gamma_1/T^{*2/3} - \Gamma_2/T^{*1/3}\right]
$$
$$
\times \left[T^{*(2/3)+(B^*/6\pi)}\right]^{-1} \left[1 + O(T^{*1/3}\ln T^*)\right], \tag{8.24}
$$

where the various amplitudes are given by

$$
\Gamma_0 = \exp\left[\tfrac{1}{3}(j_0 - k_0 + 8\tfrac{2}{3}k_0^2) - 1\right]/2j_0 B^*, \tag{8.25}
$$

$$
\Gamma_1 = 1/j_0, \qquad \Gamma_2 = 5k_0/3j_0. \tag{8.26}
$$

The vapor pressure curve follows as in (8.16), yielding $p_\sigma \sim T^\sigma \exp(-\Delta_3/2T^*)$ with $\sigma = \tfrac{1}{3} - (B^*/6\pi)$. If the exact second virial coefficient is used for B^* one has $\sigma = \tfrac{2}{9}$.

8.3. Cluster densities in the vapor

As mentioned previously, Gillan [17(c), 44] has studied quantitatively the densities of free ions and clusters in the vapor phase of the RPM for $0.05 > T^* \gtrsim 0.25$. Specifically, he presents plots of $\log_{10} p^*$ (with $p^* = \tilde{p}T^* = pa^4D/q^2$) and $\log_{10}\rho_{(r,s)}^*$ versus $b \equiv 1/T^*$ for the neutral clusters $(r,s) = (+, -), (2+, 2-)$, and $(3+, 3-)$ and for the charged clusters $(+, 0), (2+, 1-), (3+, 2-)$ and, by symmetry, for their charge conjugates [44]. Higher order clusters have negligible relative densities over the range considered. Within graphical accuracy Gillan's results may be represented by

$$
\ln p_\sigma^* \approx -0.47_8/T^*, \qquad \ln \rho_{vap}^* \approx -0.44_5/T^*, \tag{8.27}
$$

where additive constants have been dropped and ρ_{vap} denotes the total vapor density [17(c)]. Since, to excellent approximation, one has $p^* \propto \rho^* T^*$ in the vapor, the two amplitudes here may appear inconsistent until one notices that $\ln T^* \simeq \text{const.} + 0.033\Delta b$ for the range of T^* (or b) in question.

It is of interest to compare the results with our DH based theories using the low-T forms derived. Allowance may be made for the correction factors involving T^* but errors will still be entailed since we actually desire results around $T^* \simeq 0.03$ to 0.04. With those provisos pure DH theory (see Section 4.2) indicates amplitudes of about 0.55 and 0.51_5 in (8.27). These are distinctly higher than Gillan's values and, we believe, this is significant. Furthermore, within DH theory we cannot separately estimate the neutral and charged cluster densities.

By using DHBjDI theory we can, of course, go further. To that end note, first, that Gillan's result for the *total* neutral and charged cluster densities can be represented

(again to graphical accuracy and dropping additive constants) as

$$\ln \rho_0 \approx -\varepsilon_0/T^*, \qquad \ln \rho_\pm^* \approx -\varepsilon_\pm/T^*, \tag{8.28}$$

with $\varepsilon_0 \simeq 0.44_2$ and $\varepsilon_\pm \simeq 0.71_0$. Now within DHBj theories it is natural to identify ρ_0 with the neutral, dipole, ion-pair density ρ_2 and, similarly, to identify $\rho_\pm$ with the total free-ion density ρ_1. If that is valid ρ_0 and $\rho_\pm$ and, hence, ε_0 and $\varepsilon_\pm$, should be related in the vapor phase through the law of mass action in terms of the Bjerrum/Ebeling association constant. This statement is embodied explicitly in the first line of (8.23): allowing for the factor T^*, that leads to $2\varepsilon_\pm \simeq 1 + \varepsilon_0 - 0.03_3$. Indeed, the quoted values satisfy the relation to within better than 1% which confirms both the ρ_1 and ρ_2 identifications and the temperature variation of the association constant.

Now, via (8.14) and (8.15), DHBjDI theory yields the following asymptotic, low-T amplitudes corresponding to (8.27): $\varepsilon_2 = \frac{1}{2}\Delta_3 \equiv a_1^2 a/2a_2^3$ and $\varepsilon_1 = \frac{1}{2} + \frac{1}{4}\Delta_3$. If we accept $a_2 = 1.162_0 a$, as previously, and, following Fig. 4 for $T^* \simeq 0.033$, use $a_1 \simeq 1.05a$, we find $\varepsilon_2 \simeq 0.35$ and $\varepsilon_1 \simeq 0.68$. These in turn imply amplitudes $\simeq 0.39$ and 0.36 in (8.27). (In principle, since $a_1 = 1.05a$ exceeds a, one should recompute the angular average used to estimate a_2: this actually yields $a_2 \simeq 1.164_9 a$.) The changes from pure DH theory are in the correct direction but the values now lie below the Gillan estimates. They would increase slightly (by 0.02 and 0.01) with T^* since $a_1(T)$ increases. However, as *effective* slopes the values should be decreased (by about 0.08 and 0.06) in view of the correction factors displayed in (8.14) and (8.15). On the other hand, if a_2 is decreased to, say, 1.04a (which is low but, perhaps, not unreasonably so) one is led to $\varepsilon_2 \simeq 0.41$ and $\varepsilon_1 \simeq 0.69$ which fall below Gillan's values by only 0.03 and 0.02.

Inclusion of the hard-core term changes the correction factors significantly as seen in (8.24). In this case the exhibited corrections serve to decrease ε_2 by only about 0.04; but since the magnitude of the $T^{*1/3} \ln T^*$ and subsequent corrections is not known, this cannot be considered reliable. Nevertheless, accepting that estimate and the "low" value $a_2 \simeq 1.04a$ yields essentially perfect agreement with the values of $\varepsilon_0 \cong \varepsilon_2$ and $\varepsilon_\pm \cong \varepsilon_1$ derived from Gillan's data.

8.4. Assessment and theory for larger clusters

Even though the value $a_2/a \simeq 1.04$ needed to precisely match the explicit cluster calculations in DHBjDIHC theory is rather low, the possibility of agreement is nevertheless quite encouraging. To go further, two issues must be addressed (beyond the question of the optimal value of a_2). The first is to improve the description of the liquid phase at low and intermediate temperatures. Since, as we have seen, $\rho_1/\rho_2 \to 0$ in the low-T liquid, it is clearly essential to address the residual dipole–dipole pair or collective interactions. This question is taken up briefly below. The other task is to account more directly in the theory for the *higher-order clusters*. Within the framework we have set out, this is a relatively straightforward matter in principle,

although the complications obviously multiply and judicious approximations will be needed in assessing the appropriate equilibrium constants and multiple-ionic interaction terms.

As regards the *association constants*, we will sketch a plausible approach as follows. First we may assume (as can be checked in various ways: see e.g., Ref. [44]) that the most important clusters will be of the form $(k+, l-)$ with $l = k$ or $k \pm 1$, i.e., neutral or singly charged. Then for an $n = l + k \geqslant 2$ cluster we can extend the mass action law (5.3) to the form

$$\rho_n = (\tfrac{1}{2}\rho_1)^n K_n(T) \exp(n\bar{\mu}_1^{\text{Ex}} - \bar{\mu}_n^{\text{Ex}}) . \tag{8.29}$$

Of course, $K_2(T)$ will just be the previous Bjerrum/Ebeling association constant. Following (5.6)–(5.11) and generalizing to $d > 2$ dimensions [34] we can write that as

$$K_2(T) = C_d a^d \mathcal{Q}_2(b) \, e^b / b(d - 2) , \tag{8.30}$$

where C_d is the area of a unit d-sphere,

$$b \equiv q^2 / D a^{d-2} k_B T \equiv 1/T^* , \tag{8.31}$$

while the important point is that $\mathcal{Q}_2(b \to \infty) = 1$ is finite.

To make progress we need the form of $K_n(T)$. In Appendix B we advance a tentative argument based on picturing the most typical clusters as $+ - + - \cdots$ wandering "polymer" chains. We suspect that for n not too large and T^* not too small this will be a rather reasonable representation. The conclusion of Appendix B is that one then has

$$K_n = [C_d a^d / (d - 2)]^{n-1} \mathcal{Q}_n(b) \, e^{v_n b} / b^{(1/2)(d+1)n - d} , \tag{8.32}$$

where the amplitude $\mathcal{Q}_n(b)$ approaches a finite positive value when $T^* \to 0$ or $b \to \infty$. The parameters v_n determine the ground states of an n-cluster via

$$E_n = -v_n E_0, \qquad E_0 = q^2 / D a^{d-2} . \tag{8.33}$$

One has $v_2 = 1$, $v_3 = 2 - 2^{d-2}$ while the polymer picture suggests the bound $v_n \leqslant n - 1$ (based on neglecting all interactions beyond nearest-neighbor attractions).

To calculate the constants $\mathcal{Q}_n(\infty)$ needed at low T^*, a harmonic mode expansion is proposed in Appendix B: explicit results should be computable for small $n > 2$. We may note, however, that even in the critical region, the results of Gillan [44] suggest rather small relative populations for larger clusters. At present, therefore, it does not seem especially worthwhile to pursue this line of calculation.

8.5. The ionic liquid at higher densities

Our primary aim has been to understand the critical region of the RPM. It is evident, however, both from our various calculations, as illustrated in Fig. 6, and from the simulations, that when T falls significantly below T_c the liquid density increases extremely rapidly. Indeed, judging by the simulations, a 15% drop below T_c leads to

a ten-fold increase in density! The critical region is, thus, limited on the liquid side in particular. Two issues then arise in the liquid regime: first, the free-volume approximations used for the hard-core contributions to the free energy become suspect; second, since the density of free ions falls — see (8.19) — the adequacy of representing the ionic energy through the DH plus DI terms alone must also be questioned.

As regards the free-volume approximation, notice that setting B_1 equal to the *exact* second virial coefficient badly overestimates the pressure which, indeed, must then diverge at a false maximum density of $\rho^* = 3/2\pi \simeq 0.477$ compared with the true maximum density of $\rho^*_{\max} = \sqrt{2} \simeq 1.414$. This must cause the computed liquid-side of the coexistence curve to fall much *too* rapidly: see (8.20). This observation provides the argument in favor of using the bcc (or fcc) close-packing value for B_1 in the free-volume expressions as we have done in Fig. 6. In that case, however, the pure hard-core fluid pressure is *under*estimated by a factor of about 2.0 once ρ^* reaches 0.5. (For knowledge of the pressure of a hard-sphere fluid reliable estimates exist up to the solidification density: see Chapter 13 of Ref. [3]). On the other hand, choosing $B_1^* \simeq 1.50$ in the free-volume approximation reproduces the correct pure hard-core pressure to within $\pm 9\%$ up to $\rho^* \lesssim 0.45$. As seen from Fig. 7, the critical parameters do not change dramatically with the choice of B_1 although the effects on the liquid side of the coexistence curve are appreciable.

As discussed originally in Section 2, one could well investigate various improved approximations, such as (2.14). But one must recall that results for *mixtures* of hard spheres or, to better represent the dipolar pairs, of hard spheres and bispherical dumbells, are really called for. At this stage some further calculations might be worth pursuing, using, for example, the Percus–Yevick approximation for mixtures of hard spheres [68], or work by Mansoori et al. [69].

When the density of Bjerrum pairs becomes large with respect to the density of free ions one should expect *residual* dipole–dipole interactions to play a role. It seems very hard to identify such contributions explicitly since the major part of the ionic interactions is accounted for by the DH treatment of the free ions and our DI analysis of the dipole–ionic contributions. Nevertheless, one might attempt to incorporate dipole–dipole interactions at a purely phenomenological level. One way of doing so would be in terms of van der Waals (vdW) contributions to the total free energy as specified in (2.16). However, in the absence of a systematic theory for calculating the dipole–dipole amplitude, A_{22}, it can be treated only as a fitting parameter. (We neglect A_{12} and, of course, A_{11}.) We find that values of A_{22}^* (using reduced units for ρ_2 and T) around 0.1 are sufficient to improve the fit of the liquid side of the coexistence curve to the Monte Carlo data shown in Fig. 6, although the nearly linear shape for $\rho^* \gtrsim 0.1$ is not very well reproduced.

The failure of a vdW approach to provide a good match to the liquid branch suggests that a quadratic density contribution to the free energy does not capture all the relevant physics of the low-temperature ionic fluid. Indeed, it seems possible that at low temperatures and densities well above ρ_c, the ionic liquid starts to resemble, in

local structure, an ionic crystal (frozen from the "fused salt") with, in effect, a periodically varying charge density. Such a picture seems to have been suggested many years ago by J. Ch. Ghosh (cited in Ref. [1]) in discussions of electrolyte solutions antedating the analysis of Debye and Hückel. It was, also, in effect, adopted in an interesting treatment by McQuarrie [70] on the basis of a cell model of the liquid.

If this proposal has validity, the "organized dipoles" will no longer contribute to the free energy quadratically in vdW fashion. Rather, the appropriate density scaling can be deduced by considering some ionic crystalline form, say, for simplicity, an NaCl structure. If the interionic spacing of the "expanded crystal" approximating the liquid is R, the electrostatic energy for N ions is $U_N = -N\alpha q^2/2DR$ where α is the Madelung constant which, for NaCl, is $\alpha \simeq 1.747565$ [71]. The overall ionic density is $\rho = 1/R^3$ and $N = \rho V$ so that on neglecting the entropic contributions because $T^* \ll 1$, the reduced free energy density becomes [70] $f = \alpha q^2 \rho^{4/3}/2D$. In the limit that the density ρ_1 of free ions can be neglected relative to the dipolar density ρ_2, the leading dipole–dipole contribution would thus be

$$\bar{f}_{DD} = \frac{\alpha q^2 2^{1/3} \rho_2^{4/3}}{D k_B T} \equiv \frac{A^* \rho^{*4/3}}{a^3 T^*} . \qquad (8.34)$$

We suppose that the dominant dipole–ionic effects are described as previously.

At this level of approximation one should not take the coefficient in (8.34) at face value. The best that might be hoped is that the $\rho_2^{4/3}/T$ form reasonably represents some of the essential physics at higher densities: then the amplitude A^* could be adjusted to fit the Monte Carlo data. One such trial is represented by the plot (d) in Fig. 6: this corresponds to $A^* = 0.030$ ($\simeq 1.26\,\alpha_{\mathrm{eff}}$) which evidently provides good agreement with the data up to $\rho^* \simeq 0.5$. (The critical parameters are $\kappa_c a \simeq 1.1$, $T_c^* \simeq 0.057_4$, $\rho_c^* \simeq 0.03_2$.)

One particular defect of the form (8.34) should be mentioned. At low overall densities ρ_2 varies as ρ_1^2, via the law of mass action, and hence as ρ^2. The approximation for $\bar{f}_{DD}$ thus generates a term varying as $\rho^{8/3}$: but no such power occurs in the true low-density expansion for the RPM [2, 3, 59]. One could, however, argue that this term lies beyond the "second virial" term by a factor $\rho^{2/3}$ and that it is, therefore, of small consequence granted that, in any case, other terms that *are* to be expected [2, 59] will be absent within our approximations.

Finally, we remark that the plot (c) in Fig. 6, which embodies a value of the dipolar $+/-$ separation a_1 exceeding the closest contact value a (see Fig. 5), mirrors the Monte Carlo liquid-side data moderately well, especially in the light of the more recent simulations [15, 16]. Furthermore, although $\rho_1(T)$ falls to zero when $T \to 0$, it does so, according to (8.19) only as $T^{*1/3}$, that is, relatively slowly at intermediate values of T^*. One can well imagine, therefore, that an improved calculation of a_2, which would depend on T, as should a_1, combined with a better approximation for the hard-core terms might effectively account for all the important electrostatic contributions so obviating the need for a separate dipole–dipole term except at

temperatures below the *triple point* which, it is reasonable to suppose, is present in the true behavior of the RPM.

9. The mean spherical approximation and variants

The mean spherical approximation (MSA) [63] relies on a simple and appealing microscopic approximation for the direct correlation functions combined with the Ornstein–Zernike relation between the direct and standard correlation functions. It is beloved since the defining integral equations can be solved exactly for the RPM in $d = 3$ dimensions [63] (although we remark that DH theory and its extensions can be handled analytically [34] for *all d*). When the charges vanish the theory reduces to the Percus–Yevick (PY) equation for hard spheres: that is reasonably accurate (see Ref. [2], Chapter 13). The ionic contribution to the reduced Helmholtz free energy density, derived via the energy, is given by [2, 63]

$$\bar{f}^{\text{MSA}} = [2 + 6x + 3x^2 - 2(1 + 2x)^{3/2}]/12\pi a^3$$
$$= (\kappa^3/12\pi) \left[1 - \tfrac{3}{4}x + \tfrac{3}{4}x^2 - \cdots \right], \tag{9.1}$$

with $x = \kappa a$. This may be compared with the $\bar{f}^{\text{DH}}$ result (3.15); they differ only in order x^2.

The full MSA theory then corresponds to $f \simeq f^{\text{HCPY}} + f^{\text{MSA}}$ where the first term denotes (one of) the PY expressions for the hard-core free energy (differing according to the thermodynamic route used to obtain the pressure [3]). However, following Ebeling [11, 42], we may regard (9.1) as an approximation for f^{Ion} in its own right: then we may analyze it along lines parallel to those we have used to study f^{DH}. This will provide an instructive comparison with the DH-based results obtained above. Because the analysis follows closely that explained for f^{DH}, only few details of the derivation will be presented.

9.1. Pure MSA (without hard-core terms)

The simplest MSA-based theory, as for DH, is obtained by combining (9.1) with the ideal-gas term (2.9). This can be handled analytically by studying the spinodal curves as in Section 4.2. The critical parameters are found to be

$$x_c = \kappa_c a = 1 + \sqrt{2} = 1.414\,213\cdots, \tag{9.2}$$

which is much higher than any of the DH values, as seen in Fig. 7, and

$$T_c^* = 1/2(1 + \sqrt{2})^2 = 0.0857\,864\cdots, \tag{9.3}$$

$$\rho_c^* = 1/8\pi = 0.0397\,887\cdots, \tag{9.4}$$

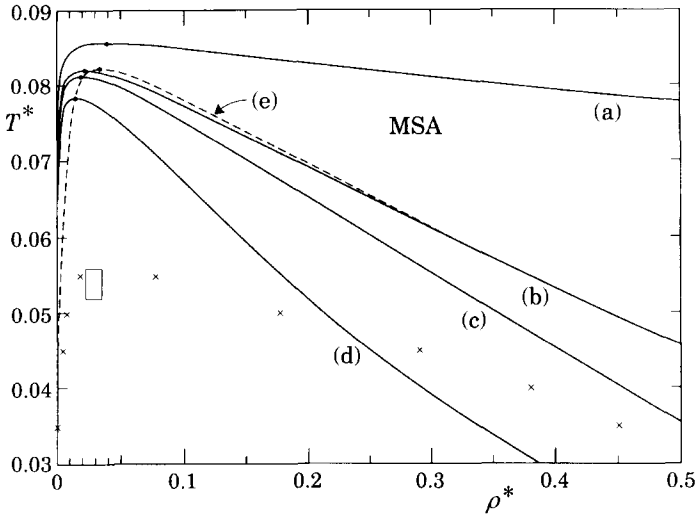$$Z_c \equiv p_c/\rho_c k_B T_c = 1 - \tfrac{2}{3}\sqrt{2} = 0.0577\cdots. \tag{9.5}$$

Fig. 8. Coexistence curve predictions for the RPM from the MSA and its variants. The crosses represent the central values of the Monte Carlo estimates of Ref. [14] (subject to some revision [16, 17]) while the box locates current critical-point estimates: see also Fig. 6 but note the change in vertical scale. Plot (a) derives from the MSA ionic free energy term supplemented only by ideal gas terms. For (b) and (c) hard-core interactions have been included within the free-volume approximation with $B_1^* = 4/3\sqrt{3}$ and $B_1^* = 1$, respectively. The standard MSA with a Percus-Yevick hard-core contribution (embodying the exact second virial coefficient) is reproduced in (d) (after Ref. [64(b)]). The dashed curve, (e), is as for case (b) but with Bjerrum association into (ideal) dipoles also included.

The corresponding coexistence curve is shown in Fig. 8a. By contrast with the parallel DH theory [see Fig. 1 and (4.8), (4.9)] the critical temperature is significantly high – by about 60%. This proves to be an undesirable feature of all the MSA-based theories.

Furthermore, the liquid side of the coexistence curve, with only the ideal gas terms included, is quite unrealistic, dropping extraordinarily slowly. Explicitly, the low-T spinodal on the vapor side is described by the relations

$$\rho_{s-}/\rho_c \approx 4T^{*3}/\pi, \qquad \kappa_{s-} a \approx 4T^*, \tag{9.6}$$

which asymptotically *coincide* with the DH results (4.13), while on the liquid side one has

$$\rho_{s+}/\rho_c \approx 1/32T^{*2}, \qquad \kappa_{s+} a \approx 1/8T^{*2}, \tag{9.7}$$

which diverge as $1/T^2$ when $T \to 0$ in place of $1/T$ in DH theory. The coexistence curves at low T are similarly given by

$$\rho_\ell^* = (81\pi T^{*3})^{-1} [1 - 6\sqrt{2}T^* + O(T^{*2})], \tag{9.8}$$

$$\rho_v^* = \frac{1}{18T^*} e^{-1/T^*} [1 - 6(2\sqrt{2} - 1)T^* + \cdots]. \tag{9.9}$$

Since the vapor rapidly becomes ideal the vapor pressure curve obeys $\bar{p}_\sigma(T)a^3 \approx \rho_v^*$ with exponentially vanishing corrections of relative order $(\rho_v^*/T^{*3})^{1/2}$. By comparison with (4.16) we see that the MSA vapor pressure vanishes with T faster than according to DH theory by a factor $\exp(-1/2T^*)$. By the same token the reduced chemical potential on the phase boundary, $\bar{\mu}_\sigma(T)$, diverges as $-1/T^* - \ln T^*$ according to the MSA while DH theory predicts the weaker divergence $\approx -1/2T^* - \ln T^*$. Various significant consequences of this behavior will be mentioned below.

9.2. MSA with hard-core terms

One may suspect intuitively that the anomalous behavior of the liquid side of the (ideal gas) + MSA plot (a) in Fig. 8 arises from neglect of hard-core contributions even though their absence in zeroth-order DH theory was not so harmful. At a first level of improvement consider adding the free-volume hard-core terms to the MSA ionic free energy. With $B_1^* = 4/3\sqrt{3} = 0.769\ldots$ (as for bcc close packing) the solid curve (b) in Fig. 8 is obtained: the critical parameters are

$$\kappa_c a \simeq 1.8389_9, \qquad T_c^* \simeq 0.08212_7,$$

$$\rho_c^* \simeq 0.02210_2, \qquad Z_c \simeq 0.07369_8. \qquad (9.10)$$

Evidently the critical temperature has fallen somewhat, the critical density significantly. Increasing B_1^* to unity yields further reduction in T_c and ρ_c as seen from plot (c) in Fig. 8.

Finally, the "full" MSA with PY hard-core contributions is presented in Fig. 8d [64(b)]. The critical temperature has fallen significantly to $T_c^* \simeq 0.078_5$, that is 45% higher than the current estimates, while the critical density becomes $\rho_c^* \simeq 0.015$ which is about 50% smaller than the best estimates. Note also that although the PY form of the hard-core term is more accurate than the free-volume approximation, the liquid side of the coexistence curve is not well represented in *any* range of density.

9.3. MSA with Bjerrum pairing

The MSA sets the direct correlation function for $r > a$ proportional to the bare Coulomb interactions scaled by k_BT. This approximation is strongly reminiscent of the DH linearization of the Poisson–Boltzmann equation. (Both approaches also impose the obvious relation $g_{ij}(r) = 0$ for $r < a$ on the ordinary correlation functions: screening then appears in the DH approach via solution of the linearized PB equation while, within the MSA, it comes about through the imposition of the OZ relation.) The similarity of MSA and DH "linearizations" suggests that the MSA also may be improved [11, 42] by inclusion of Bjerrum pairing [72, 73].

To do this, consider first the ideal-gas MSA with (ideal) Bjerrum pairs. As in DH theory, this approximation is analytically tractable. Again there can be *no change* in

$\kappa_c a$ or T_c so that (9.2) and (9.3) still apply. For the present exercise the Bjerrum form (5.5)–(5.7) of the association constant suffices and we record the value $Q^{Bj}_{6+4\sqrt{2}} \simeq 1.758\,795_6$ corresponding to the temperature (9.3). Then the ratio of dipolar pairs to free ions at criticality is

$$\frac{\rho_{2c}}{\rho_{1c}} = \frac{e^2 Q^{Bj}_{6+4\sqrt{2}}}{16(1+\sqrt{2})^2} \simeq 0.1393_6 \,. \tag{9.11}$$

This is very much smaller than the corresponding DH value of 4.047_9: see (5.13) and following discussion. As explained previously [44], Gillan's work [17(c)] shows that the DH prediction is probably of the correct order of magnitude. By the same token the MSA result (9.11) seems *too small* by a factor of 10 or more. From (9.11) the critical density follows as

$$\rho_c^* = \rho_{1c}^* + 2\rho_{2c}^* \simeq 0.05087_9 \,, \tag{9.12}$$

which is only 28% higher than (9.4), while the osmotic ratio increases by a factor of 2.7 to

$$Z_c = p_c/\rho_c k_B T \simeq 0.1537_1 \,. \tag{9.13}$$

Now comes a surprise! In DH theory inclusion of ideal Bjerrum pairs led to "banana" coexistence curves as illustrated in Fig. 2. But that does *not* happen here. On the contrary, the liquid side of the coexistence curve of plot (a) in Fig. 8 hardly changes visibly; the vapor side does rise to much larger *relative* values of the density but the *overall* changes in density remain quite small!

The quantitative nature of the changes induced by Bjerrum pairs can be judged well from the dashed curve (e) in Fig. 8. This represents the result of adding ideal Bjerrum pairs to the MSA with free-volume hard-core terms using $B_1^* = \frac{1}{2}B_2^* = 4/3\sqrt{3}$. In this case the critical parameters, namely

$$\kappa_c a \simeq 1.8485_8, \qquad T_c^* \simeq 0.08237_4 \,,$$

$$\rho_c^* \simeq 0.03307_4, \qquad Z_c \simeq 0.2227_4 \,, \tag{9.14}$$

are all enhanced with respect to those reported in (9.10) even though the rise in T_c is only 0.3%. The pairing ratio ρ_{2c}/ρ_{1c} also increases somewhat but only to 0.2382_5. Evidently the plots (b) and (c) in Fig. 8 merge closely on both liquid and vapor branches for $T^* \lesssim 0.05$. The only significant differences are on the vapor side of the coexistence curve which, however, still approaches the zero-density axis in perfectly standard fashion when T falls! In other words the banana pathology is absent [42].

What is the origin of this striking difference from DH theory? In the ideal (no hard core) approximation it is quite easy to calculate the density of Bjerrum pairs at coexistence: one finds

$$\rho_{2\sigma}^* = \frac{e^{-1/T^*}}{324\pi T^*} [1 - 8(3\sqrt{2} - 2)T^* + \cdots], \quad (\text{MSA}) \,. \tag{9.15}$$

This decays rapidly to zero confirming analytically the no-banana feature; the exponential factor is, indeed, the same as in (9.9). By contrast, in DH theory $\rho_{2\sigma}^*(T)$ *diverges* like $e^2/128\pi T^*$, as follows from (5.20). The crucial difference can now be traced by returning to the derivation of the DH result (5.20): there the ionic chemical potential is given by

$$\bar\mu_1^{\text{Ion}} = \ln \gamma_\pm = \frac{-x}{2T^*(1+x)}, \quad \text{(DH)} \tag{9.16}$$

with, of course, $x \propto \sqrt{\rho_1^*/T^*}$: see (4.4). On the phase boundary, σ, the chemical potential is the same in both phases while in the liquid, the density increases without bound or attains some hard-core limit. Hence one has $x_\sigma(T) \to \infty$ when $T \to 0$ and so $\bar\mu_{1\sigma}^{\text{Ion}}$ diverges as $-1/2T^*$ [see also (5.19)]. In computing $\rho_{2v}^*(T)$ this leads to a cancellation as explained in Section 5.3.

In place of (9.16) the MSA yields

$$\bar\mu_1^{\text{Ion}} = \frac{-x}{T^*(1+x+\sqrt{1+2x})}, \quad \text{(MSA)} \tag{9.17}$$

which, when $x_\sigma \to \infty$, yields the larger-amplitude divergence $\bar\mu_{1\sigma}^{\text{Ion}} \approx -1/T^*$, as mentioned above. Then no cancellation occurs in calculating $\rho_{2v}^*(T)$. Said more physically, the law of mass action asserts

$$\rho_2 \propto K(T)\rho_1^2 \sim T^* e^{1/T^*} [\rho_1(T)]^2, \tag{9.18}$$

where we appeal to (5.7)–(5.9), etc. On the phase boundary in the MSA we found $\rho_1 \sim e^{-1/T^*}/T^*$ which thus leads to the exponential decay recorded in (9.15); but in DH theory one has $\rho_1 \sim e^{-1/2T^*}/T^*$ so the exponential factors cancel in (9.18) leaving only a power variation for the density $\rho_2(T)$ of ideal Bjerrum pairs.

9.4. Assessment of the MSA

At first sight the absence of the banana effect in the MSA appears to be a positive feature of the approximation. But a few moments' reflection shows that, on the contrary, it is a misleading symptom. The physical point is that we know from Gillan's studies [12, 17(c), 44] that the *relative* density of dipolar pairs in the vapor *must increase* as T falls; but by (9.18) this is at best a weak effect in the MSA. In the DH theory with *ideal* pairs the growth of ρ_2/ρ_1 is vastly overestimated: but this is repaired in a physical way by inclusion of the dipole–ionic solvation effects. Then, *either* from (8.14)–(8.15), with no hard-core terms, *or* from (8.23)–(8.24), which allow for hard-cores, one finds

$$\rho_{2v}/\rho_{1v} \sim e^{(2-\Delta_3)/4T^*} = e^{(1-\varepsilon_2)/2T^*}, \tag{9.16}$$

in the vapor phase. The "activation energy" ε_2 was discussed in Section 8.3: values of 0.35 to 0.41 were in accord with the DI analysis. These lead to $(1-\varepsilon_2)/2 \simeq 0.31 \pm 0.02$. On the other hand, Gillan's estimates of the relative

populations, summarized to leading order in (8.28), yield the activation energy $(\varepsilon_0 - \varepsilon_+) \simeq 0.27$ in rather satisfactory accord in light of the neglect of higher cluster formation (see Section 8.4).

Furthermore, Gillan's estimate for the total vapor pressure [see (8.27)] corresponds to an activation energy $\varepsilon_\sigma \simeq 0.48$. This is clearly consonant with the DH factor $e^{-1/2T^*}$ (up to relatively small corrections of the sort discussed in Section 8.3 and associated with the parameter Δ_3 in Section 8.2).

On the other hand, it is clear that within MSA theory the more rapidly falling factor e^{-1/T^*} will dominate the vapor pressure and the population densities $\rho_{2v}(T)$ and $\rho_{1v}(T)$. Allowance for hard-core contributions cannot change this leading behavior, as can be seen from the analysis of the coexistence curves in Appendix A, even though more complex *corrections* will arise as in (8.24). The inclusion of dipole–ionic terms in the MSA analysis presumably can change the behavior somewhat, as seen in Section 8.2; but the numerics discussed in Section 8.3 show that the differences are unlikely to alter the activation energies by more than about 0.10. We conclude that the MSA factor e^{-1/T^*} and the associated divergence of the chemical potential on the phase boundary as $1/T^*$ (in place of $1/2T^*$), represent defects of the approximation relative to the DH approach which necessarily hinder its improvement. The small ratios of ρ_{2c}/ρ_{1c} predicted by the MSA point to a closely related defect.

It is hard to pinpoint the reasons for the relatively poor performance of the MSA. Like other integral-equation-based treatments of the theory of liquids, it primarily embodies low-density virial-type information; but even then, of course, it fails at order ρ^2 since, like DH theory, it evidently contains no description of Bjerrum pairing. On the other hand, the DH theory, being based on a physical picture in real space seems, like other mean-field-theories, to give a more evenly balanced overall picture; in addition, further physical effects, like the pairing and dipole–ion solvation, can be incorporated rather transparently.

Nevertheless, currently missing from the DH formulation is some overall variational principle for the ionic free energy, or some special limiting process (resembling, say, the Baker–Kac theory of weak long-range forces) that would provide some independent basis for confidence in its reliability beyond agreement of the critical parameters, etc., with (current!) simulation estimates. Be that as it may, we believe that the account we have given of criticality in the restricted primitive model of an ionic fluid on the basis of the Debye–Hückel approach has identified the crucial physical effects and accounted for them in at least semiquantitative fashion. It should provide a groundwork on which to attack the still outstanding challenge of determining the universality class of the critical behavior of ionic fluids [6, 18–25, 27].

Acknowledgements

Appendix A. Low-temperature coexistence in DHBjDI theory

In this appendix we sketch the derivation of the results (8.14)–(8.26) for the phase boundary and the densities $\rho_1(T)$ and $\rho_2(T)$ in the coexisting vapor (v) and liquid (ℓ) phases at low temperatures according to DHBjDI theory, first with no hard-core corrections, then with free-volume terms. We use the parameters (8.6), the basic expressions (4.6), (5.1)–(5.4) and (5.16) for the pressure and the chemical potentials, together with the dipolar–ionic terms (8.7)–(8.10), and the low-T form (8.10) for $K(T)$.

The necessary equilibrium equations may be taken as $\mu_{2v} = 2\mu_{1v}$, $\mu_{2\ell} = 2\mu_{1\ell}$, $\mu_{2v} = \mu_{2\ell}$, and $p_v = p_\ell$: these are four equations determining ρ_{1v}, ρ_{2v}, $\rho_{1\ell}$, and $\rho_{2\ell}$. As mentioned in the text, we employ the assumptions A1–3, in (8.12) and (8.13), to solve the equations successively. Thus using A1 and A2 for the vapor we may rewrite $\mu_{2v} = 2\mu_{1v}$ as

$$\rho_{1v}^* \approx \sqrt{\frac{b}{\pi}}\, e^{-b/2}\, (\rho_{2v}^*)^{1/2}\, [1 + j_1 b^2 \rho_{2v}^* - j_2 (b^5 \rho_{2v}^*)^{1/2}\, e^{-b/2}]\,, \tag{A.1}$$

with $b \equiv 1/T^*$, $j_1 = 2\pi\Delta_1/3$ and $j_2 = 2\pi/3$. For the liquid, using A3 converts $\mu_{2\ell} = 2\mu_{1\ell}$ to

$$\ln\left(\frac{\rho_{2\ell}^*}{\pi T^* \rho_{1\ell}^{*2}}\right) \approx \frac{\Delta_3}{2T^*} - \frac{3\Delta_4}{2\sqrt{\pi}}\frac{\rho_{2\ell}^*}{(T^* \rho_{1\ell}^{*3})^{1/2}}\,. \tag{A.2}$$

Now Δ_3 and Δ_4 cannot vanish so the terms on the right-hand side here must balance asymptotically when $T^* \to 0$: this leads to

$$\rho_{1\ell}^* \approx \left(\frac{9a^2 T^*}{\pi a_2^2}\right)^{1/3} (\rho_{2\ell}^*)^{2/3}\, (1 - j_3 T^* \ln T^*)\,, \tag{A.3}$$

with $j_3 = 4/3\Delta_3$.

Next note that $K(T)$ cancels from the equation $\mu_{2v} = \mu_{2\ell}$ which then yields

$$\rho_{2v}^* \approx e^{-\Delta_3 b/2} \rho_{2\ell}^*\, [1 + j_1 (b^5 \rho_{2\ell}^*)\, e^{-\Delta^* b}]\,, \tag{A.4}$$

with $\Delta^* = \frac{1}{2} + \frac{1}{4}\Delta_3$. This relation together with (A.1) and (A.3) expresses all four densities in terms of $\rho_{2\ell}^*$ alone.

Equating the pressures, using (8.9) for p^{DI}, (5.16) for p^{DHBj}, and (4.6) for p^{DH}, gives

$$\rho_{1v}^* + \rho_{2v}^* - j_4(b\rho_{1v}^*)^{3/2} - j_5 b^2 \rho_{1v}^* \rho_{2v}^* + \cdots$$

$$\approx \rho_{1\ell}^* + \rho_{2\ell}^* - j_6(b\rho_{1\ell}^*)^{1/2} + \frac{1}{8\pi}\ln\frac{4\pi\rho_{1\ell}^*}{T^*} - \frac{3\Delta_4}{4\sqrt{\pi}}\frac{\rho_{2\ell}^*}{(\rho_{1\ell}^* T^*)^{1/2}} + \cdots \qquad (A.5)$$

with $j_4 = \sqrt{\pi}/6$, $j_5 = \pi\Delta_1/3$, and $j_6 = 1/4\sqrt{\pi}$. Now one may substitute for ρ_{1v}^*, ρ_{2v}^*, and $\rho_{1\ell}^*$ in terms of $\rho_{2\ell}^* \equiv b^2 z$ and match the terms which diverge most strongly when $b = 1/T^* \to \infty$: these turn out to be the second and fifth (or last displayed) on the right hand side of (A.5). One may then solve recursively for $z^{1/3}$ and thence obtain (8.18) for $\rho_{2\ell}^*$. Using (A.3) yields (8.17) whence A3 is verified. Likewise (A.4) and (A.1) give (8.15) for ρ_{2v}^* and (8.14) for ρ_{1v}^*. Consequently A1 and A2 are valid. Finally, as in the text, the vapor pressure curve follows asymptotically simply from the ideal gas laws: $p_\sigma \approx k_{\text{B}}T(\rho_{1v} + \rho_{2v})$.

The inclusion of hard-core terms changes these results since $\rho_{\text{liq}} = \rho_{1\ell} + 2\rho_{2\ell}$ is now bounded above. The same procedures, however, prove effective. Note that A3 is now obviously valid; A1 and A2 will be verified *post facto*. On the vapor side only the second virial coefficients need be considered. In (A.1) the coefficients j_1 and j_2 gain factors no worse than $[1 + O(T^*)]$, and further quadratic terms in ρ_{1v}^* and ρ_{2v}^* appear in (A.5); but all these changes have negligible effect for small T^*.

As regards the liquid phase, the simplest case to analyze is the *free-volume approximation* (2.12) with $B_2 = 2B_1 \equiv 2B^* a^3$. An additive term $\ln\Theta$ with

$$\Theta(T) \equiv 1 - B^*\rho_{1\ell}^* - 2B^*\rho_{2\ell}^*, \qquad (A.6)$$

then appears on the left-hand side of (A.2). To handle this we adopt, and later check, the further ansatz A4: $\Theta(T)/T^\psi \to \infty$ for some $\psi > 0$, i.e. Θ does not vanish too rapidly. Then the balance of the two right-hand terms is not spoiled and (A.3) is reproduced except for the replacement of $\ln T^*$ by $\ln(T^*/\Theta)$. This shows that $\rho_{1\ell}(T) \to 0$ when $T \to 0$ so that $\Theta(T) \approx 1 - 2B^*\rho_{2\ell}^*$.

On equating $\mu_{2\ell}$ to μ_{2v} as before we now find

$$\ln\rho_{2v}^* = \ln\rho_{2\ell}^* + \frac{\pi\Delta_1}{3T^{*2}}\rho_{1v}^* - \frac{\Delta_3}{2T^*} + \frac{1}{\Theta} - \ln\Theta - 1. \qquad (A.7)$$

Because of the terms in Θ this can no longer be solved for ρ_{2v}^* in terms of $\rho_{2\ell}^*$ as in (A.4). Accordingly, we consider instead the pressure-matching relation (A.5): the terms $\rho_{1\ell}^*$ and $\rho_{2\ell}^*$ on the right-hand side now gain factors Θ^{-1} which, with the aid of (A.3), can be expressed in terms of $\rho_{2\ell}^*$ alone. The last term in (A.5) varies as $\rho_{2\ell}/(\rho_{1\ell}T)^{1/2} \sim (\rho_{2\ell}/T)^{2/3}$ and so still diverges when $T \to 0$. Now, however, its divergence must be balanced by the $\rho_{2\ell}^*/\Theta$ term. Thence $\Theta \sim T^{2/3}$ follows directly.

Further careful analysis yields $\rho_{1\prime}^{*}$ and $\rho_{2\prime}^{*}$ as given in (8.19) and (8.21) and

$$\frac{1}{\Theta(T)} = \frac{T^{*-2/3}}{j_0} - \frac{5k_0}{3j_0}\, T^{*-1/3} - \frac{B^*}{6\pi}\ln T^* + \tfrac{26}{9}k_0^2 - \tfrac{1}{3}k_1$$
$$+ \tfrac{1}{3}j_0 + \mathrm{O}(T^{*1/3}\ln T^*),\tag{A.8}$$

where the B^*-dependent coefficients are given in (8.21) and (8.22). This validates the ansatz A4 for any exponent $\psi > \tfrac{2}{3}$.

Finally, one can return to (A.7), substitute for $\rho_{2\prime}^{*}$ and $\Theta(T)$, and so find ρ_{2v}^{*} and then ρ_{1v}^{*} as given in (8.23) and (8.24).

Appendix B. Higher-order cluster association

In this appendix we present a tentative, heuristic argument yielding the expression (8.32) for the n-cluster association constant, $K_n(T)$, which enters the general mass-action law (8.29). As explained in the text, we wish to extend the $n = 2$ ($d > 2$)-dimensional form (8.30). To that end, recall from (5.5) or (6.2)–(6.4) that $K_2(T)$ may be regarded as a partition-function-like restricted or cutoff configurational integral. Furthermore, at low-T^*, on which we focus here, the precise form of the cutoff becomes irrelevant.

As argued in the text, we regard a typical n-cluster as a wandering "polymer" chain of alternating $+$ and $-$ ions. Accepting this picture, there will be a series of $n - 1$ nearest-neighbor breathing (or bond) modes. If we neglect all couplings to further neighbors each bond yields a factor like

$$I_0 \equiv C_d \int_a^R e^{ba^{d''}/r^{d''}} r^d\, \mathrm{d}r = c_d a^d e^b \int_0^S e^{-d''bs + \mathrm{O}(bs^2)}\, \mathrm{d}s,\tag{B.1}$$

where $C_d = 2\pi^{d/2}/\Gamma(\tfrac{1}{2}d)$, $d'' \equiv d - 2$, $r = a(1 + s)$, and $S = (R/a) - 1$. Now let us suppose the cutoff $R(T)$ does not vanish faster than T so that $bS(T) \to \infty$ when $b \to \infty$ (or $T \to 0$).

Under change of variable to $y = d''bs$, the correction term $\mathrm{O}(bs^2)$ then becomes $\mathrm{O}(y^2/b) \to 0$. Thus we conclude

$$I_0(T) \approx C_d a^d e^b / d''b \quad \text{as } b \to \infty.\tag{B.2}$$

In this expression the factor e^b is clearly recognized as deriving simply from the ground-state pair configuration with energy $E_2 = -E_0 := -q^2/Da^{d''}$. Following the definition (8.31) for E_n, the full n-cluster configurational integral must entail the exponential factor $e^{v_n b}$ [rather than $e^{(n-1)b}$ as would naively follow from (B.2)].

The total nd cluster degrees of freedom may be separated into d free translational modes, $d - 1$ free rotational modes, plus the $n - 1$ breathing or bond modes, which leaves $(d - 1)(n - 2)$ internal angular coordinates, say, θ_j. These coordinates will be

subject to strong exclusion forces arising from the hard-core repulsions and, of course, will also interact with the breathing modes. It is plausible, however, that one can make a Taylor expansion of the full Coulombic energy about a "thermally expanded ground state", with nearest-neighbor separations r_i along the chain only slightly exceeding the hard-core diameter a. In leading order this should yield a well-defined quadratic form, $\mathscr{E}$, in the deviation angles $\psi_j \equiv (\theta_j - \theta_j^0)$, where the θ_j^0 specify the thermal ground state. In particular, at low enough T^* the hard-core constraints can probably be ignored (beyond the extent to which they are already captured in the integrals I_0). In the standard way, the quadratic form $\mathscr{E}$ can be diagonalized yielding $(d-1)(n-2)$ Gaussian angular integrals each reducing to $(2\pi/\lambda_r b)^{1/2}$, where the λ_r are the appropriate eigenvalues. Should it transpire that an n-cluster has a number of distinct thermal ground states of comparable energies they may, at low enough T^*, be treated as distinct species.

In total there are thus $(n-1) + \frac{1}{2}(d-1)(n-1) = \frac{1}{2}(d+1)n - d$ inverse factors of b in the expression for K_n: this finally justifies the form (8.32) quoted in the text. Although in the case $n = 2$ one simply has $\mathscr{Q}_2(\infty) = 1$, the values of $\mathscr{Q}_3(\infty)$, $\mathscr{Q}_4(\infty)$, etc. will be nontrivial since, as evident from our arguments, they vary inversely with the square root of the determinant, $\mathscr{D} \equiv \Pi_r \lambda_r$, of the quadratic form $\mathscr{E}$ describing the angular parts of the Coulombic energy. Because the net repulsive interactions reduce the phase space estimated via I_0^{n-1}, we expect $\mathscr{Q}_n(\infty) < 1$ for $n > 2$. Explicit calculations should be feasible for $n = 3$ and, probably, for $n = 4$.

References

[1] P. Debye and E. Hückel, Phys. Z. 24 (1923) 185.

[2] H. Falkenhagen and W. Ebeling, in: Ionic Interactions, ed. S. Petrucci (Academic Press, New York, 1971), Vol. 1; but note slips in Eqs. (49) and (51).

[3] D.A. McQuarrie, Statistical Mechanics (Harper and Row, New York, 1976) Chap. 15.

[4] V. McGahay and M. Tomozawa: (a) J. Non-Cryst. Solids 109 (1989) 27;
(b) J. Chem. Phys. 97 (1992) 2609. To our knowledge, these are the first authors to clearly state that the full DH theory (without further supplement) predicts phase separation and a critical point.
See also, however, (c) W. Ebeling, Z. Phys. Chem. (Leipzig) 247 (1971) 340;
(d) W. Ebeling, W.D. Kraeft and D. Kremp, Theory of Bound States and Ionization Equilibrium in Plasmas and Solids (Akademie Verlag, Berlin, 1976) Sections 6.2, 6.4, in relation to the "plasma phase transition" of
(e) G.É. Norman and A.N. Starostin, Teplofiz. Vys. Temp. 6 (1968) 410; 8 (1970) 413; and
(f) M. Gitterman and V. Steinberg, Phys. Rev. A 20 (1979) 1236; but note that the peculiar "schematic" phase diagrams presented, with two types of critical point, do not appear to follow from the theory set out.
For comments on the theoretical discussion of McGahay and Tomozawa see Refs. [37, 45, 60] below.

[5] M.E. Fisher and Y. Levin, Phys. Rev. Lett. 71 (1993) 3826; note that in the expression for βf^{Id} given before Eq. (3) in this reference the factor Λ_j should appear to the power $3k_j$ when an entity of species j contains k_j ($\geqslant 2$) particles.

[6] M.E. Fisher, J. Stat. Phys. 75 (1994) 1.

[7] N. Bjerrum, Kgl. Dan. Vidensk. Selsk. Mat.-Fys. Medd. 7 (1926) 1.

[8] W. Ebeling, Z. Phys. Chem. (Leipzig) 238 (1968) 400.

[9] H.L. Friedman and B. Larsen, J. Chem. Phys. 70 (1979) 92.

[10] Friedman and Larsen reported briefly in their important review of the field (Ref. [9]) that the full DH theory *supplemented* by Bjerrum ion-pairs treated as an *ideal species* (as discussed further here) leads to a critical point with, in retrospect, rather plausible values of both critical temperature and density. They presented the coexistence curve predicted by this "modified Bjerrum theory" graphically as what appears to be a sharply pointed "marlin spike" with, furthermore, densities $\rho_{\text{liq}}(T) > \rho_{\text{vap}}(T) > \rho_c$ for $T < T_c$, so that both liquid *and* vapor branches always lie *above* the critical density. Friedman and Larsen commented that such a shape "is not realistic". These remarks and observations were a prime stimulus for our own studies but, as we find (see Refs. [5, 6] and Fig. 2 here) the sharp marlin-spike feature is not correct: it seems to have resulted from numerical inadequacies in computing the coexistence curve. (Private communication from Professor H.L. Friedman.) In retrospect it is perhaps surprising that Friedman and Larsen did not also examine the original DH theory which, as we find (Refs. [5, 6]), has a coexistence curve of perfectly acceptable shape.

[11] W. Ebeling and M. Grigo, Ann. Phys. (Leipzig) 37 (1980) 21.

[12] K.S. Pitzer, J. Phys. Chem. 88 (1984) 2689, concisely reviews the statistical mechanics and thermodynamics of real ionic fluids and critically discusses the conceptual aspects of ion pairing, etc.

[13] J.P. Valleau, J. Chem. Phys. 95 (1991) 584.

[14] A.Z. Panagiotopoulos, Fluid Phase Equil. 76 (1992) 97; 92 (1994) 313.

[15] J.M. Caillol, J. Chem. Phys. 100 (1994) 2161.

[16] G. Orkoulas and A.Z. Panagiotopoulos, J. Chem. Phys. 101 (1994) 1452.

[17] On the basis of simulation data reported in Refs. [13, 14] we previously [5] concluded that $T_c^* \equiv k_B T_c Da/q^2 = 0.057 \pm 1_5$ and $\rho_c^* \equiv \rho_c a^3 = 0.030 \mp 8$ represented optimal critical point estimates for the restricted primitive model but cautioned that the T_c estimate might well be 3–6% high (in light, mainly, of finite-size effects). On balance, the more recent studies (Refs. [15, 16]) continue to support these conclusions although Orkoulas and Panagiotopoulos (Ref. [16]) prefer the lower range $T_c^* \simeq 0.053$ and $\rho_c^* \simeq 0.025$. We believe that Caillol's (Ref. [15]) stated conclusion $T_c^* \gtrsim 0.057$ should not be accepted since it is based on an assumption (about the finite-size, canonical density distributions near criticality) which is not trustworthy. This may be seen from the exact finite-size canonical distribution function for a soluble lattice model reported by

(a) D.A. Huse and M.E. Fisher in: Melting, Localization and Chaos, Proc. 9th Midwest Solid State Theory Symp., eds. R. Kalia and P. Vashishta (North-Holland, New York, 1982) p. 259:

distinct "ordered" (e.g., vapor/liquid) peaks appear together with a central disordered peak *above* the true critical point of the infinite system, i.e., in the limiting single-phase regime. Indeed we feel Caillol's data are not inconsistent with $T_c^* \lesssim 0.055$.

Nevertheless, the relatively large discrepancies between the various simulations testify to the technical and interpretative difficulties of performing definitive calculations for near-critical ionic systems even with current powerful computers. For that reason let us also recall the pioneering series-based estimate of

(b) G.R. Stell, K.C. Wu and B. Larsen, Phys. Rev. Lett. 37 (1976) 1369;

namely, $T_c^* \simeq 0.085$, $\rho_c^* \simeq 0.011$, and the subsequent extrapolations made by Pitzer [12], namely, $T_c^* = 0.058_2$–0.059_2 and $\rho_c^* = 0.04_2$–0.05_0, who used estimates for the coexistence curve for $T \lesssim 0.9 T_c$ derived via ion-cluster calculations by

(c) M.J. Gillan, Molec. Phys. 49 (1983) 421;

See also the review Ref. [6].

[18] K.S. Pitzer, Accts. Chem. Res. 23 (1990) 373.

[19] R.R. Singh and K.S. Pitzer, J. Chem. Phys. 92 (1990) 6775.

[20] H. Weingärtner, S. Wiegand and W. Schröer, J. Chem. Phys. 96 (1992) 848.

[21] K.C. Zhang, M.E. Briggs, R.W. Gammon and J.M.H. Levelt Sengers, J. Chem. Phys. 97 (1992) 8692.

[22] M.L. Japas and J.M.H. Levelt Sengers, J. Phys. Chem. 94 (1990) 5361.

[23] T. Narayanan and K.S. Pitzer, J. Phys. Chem. 98 (1994) 9170; Phys. Rev. Lett. 73 (1994) 3002.

[24] H. Weingärtner, T. Merkel, U. Maurer, J.-P. Conzen, H. Glasbrenner and S. Käshammer, Ber. Buns. Phys. Chem. 95 (1991) 1579.

[25] J.M.H. Levelt Sengers and J.A. Given, Molec. Phys. 80 (1993) 899.

[26] See, particularly: (a) G. Stell, Phys. Rev. A 45 (1992) 7628;
      (b) A. Kholodenko and A.L. Beyerlein, Phys. Lett. A 132 (1988) 347; J. Chem. Phys. 93 (1990) 8405;

(c) M.E. Fisher, J. Chem. Phys. 96 (1992) 3352;

(d) A. Kholodenko and A.L. Beyerlein, Phys. Lett. A 175 (1993) 366.

[27] See also (a) G. Stell, J. Stat. Phys. 78 (1994) 197;
Some more recent theoretical studies which tend to discredit one possible source of near-classical critical behavior are reported in
(b) X.-J. Li, Y. Levin and M.E. Fisher, Europhys. Lett. 26 (1994) 683;
(c) M.E. Fisher, Y. Levin and X.-J. Li, J. Chem. Phys. 101 (1994) 2273.

[28] P. Chieux and M.J. Sienko, J. Chem. Phys. 53 (1970) 566; see also Ref. [6].

[29] M.E. Fisher, Phys. Rev. Lett. 57 (1986) 1911; but note, as regards the experimental data stimulating this work, the later experiments by
G. Dietler and D.S. Cannell, Phys. Rev. Lett. 60 (1988) 1852.

[30] C. Bagnuls and C. Bervillier, Phys. Rev. Lett. 58 (1987) 435.

[31] See, e.g., M.E. Fisher, Rev. Mod. Phys. 46 (1974) 597; and, as a particular example not unrelated to the present problem, consider the question of the effects of long-range dipolar interactions on the critical behavior of ferromagnets and antiferromagnets with short-range exchange couplings:
M.E. Fisher and A. Aharony, Phys. Rev. Lett. 30 (1973) 559;
A. Aharony and M.E. Fisher, Phys. Rev. 8 (1973) 3323;
A. Aharony, Phys. Rev. B 8 (1973) 3342, 3349, 3358, 3363.

[32] An interesting case of possible conceptual relevance here, in which the derivation of an adequate LGW effective Hamiltonian poses a more difficult problem, is the critical wetting transition in bulk, three-dimensional fluid systems with short-range forces: see M.E. Fisher and A.J. Jin, Phys. Rev. Lett. 69 (1992) 792;
A.J. Jin and M.E. Fisher, Phys. Rev. B 48 (1993) 2642.

[33] See, e.g., (a) M.E. Fisher and S. Fishman, Phys. Rev. Lett. 47 (1981) 421; and references therein;
(b) J.S. Høye, E. Lomba and G. Stell, Molec. Phys. 75 (1992) 1217;
(c) L. Belloni, J. Chem. Phys. 98 (1993) 8080;
(d) F. Bresme, E. Lomba, J.J. Weis and J.L.F. Abascal, Phys. Rev. E 51 (1995) 289.

[34] See, e.g., Y. Levin, X.-J. Li and M.E. Fisher, Coulombic criticality in general dimensions, Phys. Rev. Lett. 73 (1994) 2716.

[35] See M.E. Fisher, J. Chem. Phys. 42 (1965) 3852 for the validity of the free-volume form near close packing: few approximate analytic forms embody this.

[36] For differential and partial differential approximant techniques that might be used to extend virial series, etc., see, e.g., M.E. Fisher and H. Au-Yang, J. Phys. A 12 (1979) 1677; 13 (1980) 1517;
A.J. Guttmann, in: Phase Transitions and Critical Phenomena, Vol. 13, eds. C. Domb and J.L. Lebowitz (Academic Press, New York, 1989) p. 1.

[37] The most careful consideration of the critical pressure, p_c, seems to lie in the extrapolations made by Pitzer [12]. In quoting a range for Z_c here we have noted (i) that the ratio falls quite sharply as the estimate for T_c^* is lowered and (ii) that the likely values of T_c^* now lie on the low side of Pitzer's estimates. The only recent direct simulation estimates bearing on p_c appear to be those of Valleau [13]. His data most plausibly suggest $T_c^* \simeq 0.065$ (rather than 0.070 as quoted): accepting this and the corresponding lower p_c^* value leads to $Z_c \simeq 0.09$. McGahay and Tomozawa [4(b)] tabulate most of the earlier critical parameter estimates and, on a corresponding states basis quote "experimental averages", including $T_c^* \simeq 0.066$, that lead to $Z_c \simeq 0.098$. Residual attractions in real systems, relative to the RPM, are probably responsible for the high value of T_c^*, which again suggests a lower value of Z_c for the RPM.

[38] L. Onsager, Chem. Revs. 13 (1933) 73.

[39] S.F. Edwards, Philos. Mag. 4 (1959) 1171; and see:
P. Minnhagen, Rev. Mod. Phys. 59 (1987) 1001.

[40] Thus for van der Waals theory the Landau expansion (4.21) still applies with $t_2 \approx t$ but now for $p^\dagger = \beta p / \rho_{max}$ with coefficients $c_1 = \frac{1}{3}$, $c_2 = \frac{3}{4}$, $c_3 = 0$, $u_4 = \frac{3}{16}$, $u_5 = \frac{3}{64}$, $u_6 = \frac{27}{320}$, $\cdots$. Similarly, mean-field theory for a lattice gas with single-site hard core is described by $c_1 = \frac{1}{2}$, $c_2 = 1$, $c_3 = u_5 = u_7 = \cdots = 0$, $u_4 = \frac{1}{3}$, and $u_{2k} = 1/(2k - 1)$. It is worth remarking that in all cases, including pure DH theory, the scale factor $c_2 / \sqrt{u_4}$, which arises when one rescales m to make $u_4 = 1$, takes the value $\sqrt{3}$ [but note the different prefactors defining $p^\dagger$ here and in (4.6)].

[41] It was for this reason that the apparent finding of a "marlin spike", $\beta = 1$ coexistence curve by Friedman and Larsen [9, 10] was so startling.

[42] M. Grigo and W. Ebeling, J. Soln. Chem. 13 (1984) 321. Unfortunately, various slips and deficiencies of presentation make this interesting paper, like Ref. [11], hard to assess.

[43] See, e.g., Handbook of Mathematical Functions, eds. M. Abramowitz and I.A. Stegun (National Bureau of Standards, U.S. Govt. Printing Office, Washington, 1964) Section 5.

[44] Gillan [17(c)] adopts as his (arbitrary) cluster definition the convention that two ions with centers separated by less than $R^G = 2a$ belong to the same cluster. He finds, naturally, that the densities of various clusters in the dilute, low T (coexisting) vapor phase vary exponentially rapidly with $1/T$. At $T^* \simeq \frac{1}{20}$, which is the highest temperature for which his analysis is effective, Gillan finds the total density of neutral clusters $(+, -)$, $(2+, 2-)$, and $(3+, 3-)$ to be $\rho_0^* \simeq 1.15 \times 10^{-3}$, while the total density of positive ions $(3+, 2-)$, $(2+, 1-)$ and $(+)$, in decreasing order of concentration, plus complementary negative ions is only $\rho_1^* \simeq 1.37 \times 10^{-5}$. Owing to the neglect of higher order clusters these estimates should probably be regarded as lower bounds. Nevertheless, the ratio $\rho_0/\rho_1 \simeq 85$ is probably reasonably accurate for $T^* \simeq 0.05$. Rough extrapolation to $T^* \simeq 1/16$ yields a ratio $\rho_0/\rho_1 \simeq 23$. However, that seems likely to be an overestimate for the critical region since the densities of charged ions rise increasingly rapidly with temperature above $T^* \simeq 0.05$ and the total vapor pressure on the phase boundary is increasingly uncertain.

[45] It should be mentioned that although McGahay and Tomozawa [4] considered simple Bjerrum theory, they did not notice the unphysical banana effect of Fig. 1 since they did not compute the coexistence curves.

[46] (a) R.M. Fuoss, J. Am. Chem. Soc. 80 (1958) 5059.
See also J.E. Prue in: Chemical Physics of Ionic Solutions, eds. B.E. Conway and R.G. Barrados (Wiley, New York, 1966).

[47] M.-C. Justice and J.-C. Justice, J. Soln. Chem. 5 (1976) 543; 6 (1977) 819.

[48] J.G. Kirkwood, J. Chem. Phys. 2 (1934) 351.

[49] H. Xu, H.L. Friedman and F.O. Raineri, J. Soln. Chem. 20 (1991) 739.

[50] H. Weingärtner, T. Merkel, S. Käshammer, W. Schröer, and S. Wiegand, Ber. Buns. Phys. Chem. 97 (1993) 970;
W. Schröer, S. Wiegand and H. Weingärtner; Ber. Buns. Phys. Chem. 97 (1993) 975.

[51] H.L. Friedman, J. Phys. Chem. 66 (1962) 1596.

[52] Thus even truncating the pure DH pressure expression at order ρ^2 leads to criticality with $x_c = \kappa_c a = \frac{1}{4}$, $T_c^* = \frac{1}{32}$ and $\rho_c^* = 1/2048\pi \simeq 1.55 \times 10^{-4}$: see (4.6) and (4.7).

[53] H. Yokoyama and H. Yamatera, Chem. Lett. 1973 (1973) 337; Bull. Chem. Soc. Jpn. 48, (1975) 1770.

[54] T.L. Hill, Statistical Mechanics: Principles and Selected Applications (McGraw-Hill, New York, 1956) Section 27.

[55] See, e.g., (a) M.E. Fisher, Physics 3 (1967) 255;
(b) M.E. Fisher and B.U. Felderhof, Ann. Phys. (NY) 58 (1970) 176, 217;
(c) B.U. Felderhof and M.E. Fisher, Ann. Phys. (NY) 58 (1970) 268.

[56] L. Onsager, J. Chim. Phys. Special Issue, 19th Meeting of the Société de Chimie Physique (Montpellier, September 1968) p. 86.

[57] See also Ref. [2] p. 47.

[58] Ebeling's association constant equally corrects the mean spherical approximation (MSA) up to $O(\rho^{5/2})$ since that reproduces DH theory at order ρ^2. See Ebeling and Grigo [11] and the discussion of the MSA below in Section 9.

[59] H.L. Friedman, Ionic Solution Theory (Interscience, New York, 1962) Sections 13, 15.

[60] In their application of Bjerrum's theory, McGahay and Tomozawa [4] employed Fuoss' [46] expression for the association constant being unaware of its unsatisfactory nature [2, 47, 53].

[61] A rough approximation for $s_1(T)$, mentioned in Ref. [5], is $s_1 \simeq Q^{Bj} T^* (1 + 8T^*)$. Used with (5.11) this is around $\pm 10\%$ accurate up to $T^* \simeq 0.07$.

[62] For the necessary properties of the spherical Bessel functions and other technicalities see Ref. [27(c)]; and e.g., A. Erdélyi, W. Magnus, F. Oberhettinger and F.G. Tricomi, eds., Higher Transcendental Functions, Bateman Manuscript Project (McGraw-Hill, New York, 1953), Vol. II, Chap. 7.

[63] E. Waisman and J.L. Lebowitz, J. Chem. Phys. 56 (1972) 3086, 3094.

[64] See also the GMSA proposed by (a) J.S. Høye, J.L. Lebowitz and G. Stell, J. Chem. Phys. 61 (1974) 3253;
For earlier work showing coexistence curves and spinodals graphically, (b) M.M. Telo da Gama, R. Evans and T.J. Sluckin, Molec. Phys. 41 (1980) 1355; and
(c) M. Rovere, R. Miniero, M. Parrinello and M.P. Tosi, Phys. Chem. Liquids 9 (1979) 11.

[65] (a) A.P. Levanyuk, Zh. Eks. Teor. Phys. 36 (1959) 810 [Sov. Phys. JETP 36 (1959) 571];
(b) V.L. Ginzburg, Fiz. Tverd. Tela 2 (1960) 2031 [Sov. Phys. Solid State 2 (1960) 1824];
(c) D. Thouless, Ann. Phys. (NY) 10 (1960) 553;
(d) D. Thouless, Phys. Rev. 181 (1969) 954;
(e) L.P. Kadanoff, et al., Rev. Mod. Phys. 39 (1967) 395;
(f) P.C. Hohenberg in: Fluctuations in Superconductors, eds. W.S. Goree and F. Chilton (Stanford Research Inst., Menlo Park, 1968) p. 305, Section IV; and also see:
e.g. (g) J. Als-Nielsen and R.J. Birgeneau, Amer. J. Phys. 45 (1977) 554;
(h) M.E. Fisher, in: Critical Phenomena, ed. F.J.W. Hahne, Lect. Notes Phys. 186 (Springer-Verlag, Berlin, 1983) pp. 106–107;
(j) M. Plischke and B. Bergersen, Equilibrium Statistical Physics (Prentice Hall, New Jersey, 1989) p. 88.

[66] R.J.F. Leote de Carvalho and R. Evans, Molec. Phys. 83 (1994) 619; J. Phys. Cond. Matt. 7 (1995) L 575.

[67] B.P. Lee and M.E. Fisher, to be published.

[68] J.L. Lebowitz, Phys. Rev. 133 (1964) A 895;
R.J. Baxter, J. Chem. Phys. 52 (1970) 4559.

[69] G.A. Mansoori, N.F. Carnahan, K.E. Starling and T.W. Leland, J. Chem. Phys. 54 (1971) 1523.

[70] D.A. McQuarrie, J. Phys. Chem. 66 (1962) 1508. This approximation for the RPM yields $T_c^* \simeq 0.074_5$ and $\rho_c^* \simeq 0.06_1$.

[71] C. Kittel, Introduction to Solid State Physics, 4th Ed. (Wiley, New York, 1971) p. 118.

[72] See also Y. Zhou, S. Yeh and G. Stell, J. Chem. Phys. 102 (1995) 5785, who invoke a more elaborate, recently formulated type of association theory;
Y. Zhou and G. Stell, J. Chem. Phys. 96 (1992) 1504, 1507;
G. Stell and Y. Zhou, Fluid Phase Equil. 79 (1992) 1.
B. Guillet and Y. Guissani, Molec. Phys. (1995), in press.

[73] Attention might also be drawn to work by Outhwaite and collaborators, (a) L.B. Bhuiyan, C.W. Outhwaite, M. Molero and E. González-Tovar, J. Chem. Phys. 100 (1994) 8301, based on a modified Poisson–Boltzmann (MPB) theory;
(b) C.W. Outhwaite, in: Statistical Mechanics (The Chemical Society, London, 1975) Vol. 2, pp. 188–255; J. Chem. Phys. 50 (1969) 2277. This introduces a "fluctuation potential energy".
(c) C.W. Outhwaite, M. Molero and L.B. Bhuiyan, J. Chem. Soc. Faraday Trans. 87 (1991) 3227, which further improves a symmetrized Poisson–Boltzmann theory (SPB).
Recently, E. González-Tovar and C.W. Outhwaite, Molec. Phys. 83 (1994) 1273 have used the SPB and MPB theories to find $T_c^* \simeq 0.0715$ and 0.0635, respectively.
Another approach due to (d) G.A. Martynov and A. Schmidt, Molec. Phys. 42 (1981) 1493, aims to improve the Poisson–Boltzmann (PB) form perturbatively: their VEX-0 approximation is equivalent to the PB equation and yields $T_c^* \simeq 0.0735$ with $\rho^* \simeq 0.017$, while the first improvement, VEX-1, gives $T_c^* \simeq 0.704$ and $\rho_c \simeq 0.020$ [A. Schmidt, private communication]. All these approaches seem to entail, even when improved, estimates for T_c^* lying above the simplest DH value $\frac{1}{16} = 0.0625$ and, hence, remaining significantly higher than the current simulation estimates $T_c^* = 0.052$–0.056 [see Eq. (3.5)].