*Journal of Statistical Physics, Vol. 75, Nos. 1/2, 1994*

# The Story of Coulombic Criticality*

**Michael E. Fisher**[1]

*Received August 24, 1993*

Recent experiments on phase separation and criticality in ionic fluids are reviewed briefly. The data suggest a sharp distinction between *solvophobic criticality*, displayed by nonionic fluids and some electrolytes, that is associated with Ising-like exponents, $\beta \simeq 0.325$, $\gamma \simeq 1.239$, and $\nu \simeq 0.631$, and *Coulombic* (or *ionic*) *criticality* characterized by classical, van der Waals exponents, $\beta = 0.5$, $\gamma = 1$, and $\nu = 0.5$. Only experiments on the sodium–ammonia system seem to straddle this dichotomy: they show crossover from classical to Ising behavior close to $T_c$ at a characteristic crossover scale $t_\times = |T_\times - T_c|/T_c$. A range of theoretical issues thus raised is discussed, including other conceivable options (spherical model, tricriticality, etc.). Attention is drawn to Nabutovskii's work and various scenarios are illustrated with the aid of schematic phase diagrams containing multicritical points that could, in principle, separate two distinct universality classes of electrolyte criticality. The advantages of examining a basic four-state lattice model that allows for ionic association–dissociation, etc., are reviewed. The issue of the existence, location, and nature of the long-heralded but still elusive gas–liquid transition and critical point in the continuum restricted primitive model (hard spheres carrying charges $+q$ and $-q$) is taken up in further detail. Earlier theoretical work and recent Monte Carlo simulations are summarized. In an effort to obtain a physically transparent, semiquantitative description, the work of Debye and Hückel and its subsequent elaboration via Bjerrum's concept of bound ion pairs is revisited and seen to predict phase separation and criticality. Recent work by Levin and the author is described which repairs serious defects of the earlier theories by including the interaction of the ion-pair dipoles with the screening ionic fluid, following Debye–Hückel methods. The resulting mean field theory agrees quite well with the simulations and appears to embody the most crucial physical effects. However, the role of critical fluctuations, the related interplay of the charge and density correlation functions, the likelihood of Ising-like behavior, and the associated crossover scale $t_\times$ remain important unsettled questions. An Appendix presents a critique

---

* This article is based on a lecture delivered at the First Lars Onsager Symposium held at the Norwegian Institute of Technology, Trondheim, Norway, 2–4 June 1993.

[1] Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742, U.S.A.

of arguments by Stell to the effect that the restricted primitive model should display Ising behavior and that $1/r^4$ effective interactions might be significant.

---

**KEY WORDS**: Criticality; electrolytes; restricted primitive model; Coulombic criticality; ionic criticality; solvophobic criticality; Ising-like; Debye–Hückel theory; Bjerrum ion pairs; lattice model; crossover; multicritical points.

## I. IONIC CRITICALITY. THE CHALLENGE

Many ionic fluids or electrolytes, both single-component systems, like molten NaCl, and solutions of salts dissolved in aqueous or organic solvents undergo phase separation and, subsequently, as the temperature is raised, exhibit gas–liquid or liquid–liquid critical points, which can be described by a coexistence curve

$$\Delta x \sim \Delta\rho \equiv \rho_{\text{liq}} - \rho_{\text{gas}} \sim |t|^{\beta} \tag{1}$$

when $t \equiv (T - T_c)/T_c \rightarrow 0-$, where $x$ is, say, mole fraction or volume fraction and $\rho$ is the number density of the solute/salt or pure electrolyte; by a susceptibility/compressibility/zero-angle-scattering intensity diverging as

$$\chi_T \sim K_T \sim S(0) \sim 1/|t|^{\gamma}, \qquad t \rightarrow 0 \tag{2}$$

and by a correlation length for the density–density fluctuations varying as

$$\xi(T) \approx \xi_0^{\pm}/|t|^{\nu} \qquad \text{as} \quad t \rightarrow 0\pm \tag{3}$$

In the case of nonionic fluids the analogous critical points are well established, both theoretically and experimentally, to be of *Ising-like character* (**Isg**) with

$$\beta \simeq 0.32_5, \qquad \gamma \simeq 1.24, \qquad \nu \simeq 0.63 \tag{4}$$

By contrast, however, ionic fluids, which, of course, entail long-range Coulombic interactions, typically appear, at least on first study, to exhibit *classical* or van der Waals behavior (**Cl**). In particular, their coexistence curves are often well described by $\beta = 1/2$. Notable examples are the metal–ammonia systems (see ref. 1). This raises intriguing fundamental theoretical questions: Q.1: Can ionic fluids exhibit true asymptotic non-Ising-like criticality? Q.2a: If so, what are the characteristics of such *ionic criticality* (**Ion**)? Q.2b: If *not*, why do so many ionic fluids *appear* to display classical criticality?

These issues are of renewed interest since Singh and Pitzer$^{(2)}$ discovered in 1988 that the organic salt triethyl-*n*-hexylammonium triethyl-*n*-

hexylboride $[(C_2H_5)_3(C_6H_{13})N^+(C_2H_5)_3(C_6H_{13})B^-]$ in diphenyl ether $[(C_6H_5)_2O]$ exhibits criticality at $T_c = 317$ K and is susceptible to precise observations down to temperature intervals of $|t| \simeq 10^{-4.5}$: the coexistence curve data strongly indicate $\beta \simeq 0.47_6$, markedly close to $\beta = 1/2$.(3) On the other hand, Japas and Levelt Sengers(4) studied solutions in water of tetra-*n*-pentylammonium bromide $[(C_5H_{11})_4N^+Br^-]$, which has a similar, rather large organic cation but a smaller anion. Down to $|t| \simeq 10^{-4}$ they found excellent fits to the coexistence curve with $\beta = 0.319$–$0.337$ close to $\beta_{\text{Isg}}$. A strong dichotomy is evident!

Further aspects of the experimental background have been well reviewed by Pitzer in 1990,(5) Japas and Levelt Sengers in 1991,(4) and, most recently, by Levelt Sengers and Given.(6) What can be said theoretically?

## 2. THE ISING-LIKE OPTION AND CROSSOVER

A very natural answer to Q.1 is "No" so that **Ion ≡ Isg**. That has, indeed, been argued by Stell and co-workers(7–9); but, to this author at least, the discussions, while plausible, are not very convincing and actually appear somewhat circular. (An informal critique of ref. 9 is appended to this article.) Thus Pitzer's view(5) that the current situation is "inconclusive" seems fair. Certainly, Q.2b has not been adequately addressed: If Ising-like behavior is controlling, why is it not always seen?

That question can be sharpened by examining what at present seem to be the only experiments straddling the classical–nonclassical coexistence-curve dichotomy for ionic systems: those are the 1969–1970 observations of Chieux and Sienko(1) on the sodium–ammonia system ($Na + NH_3$). They saw an effective exponent (i.e., a local slope on a log–log plot of $\Delta x$ vs. $|t|$) of value $\beta_{\text{eff}} \simeq 1/2$ in the range $|t| = 0.1$ down to 0.01; nevertheless, on approaching closer to $T_c$, the log–log plot turned downward and indicated $\beta_{\text{eff}} \simeq 0.33$,(1) so suggesting true asymptotic behavior of Ising character.

To describe such a *crossover* one can appeal to renormalization-group (RG) theory, which, first, recognizes the so-called *corrections to scaling* and, second, allows for distinct (RG) fixed points with, in general, different exponents and differing degrees of relative stability. To allow for the singular correction terms one often writes (see, e.g., refs. 3–5)

$$\Delta x \approx B\,|t|^{\beta}\,(1 + b_{\theta}\,|t|^{\theta} + b_{2\theta}\,|t|^{2\theta} + \cdots) \tag{5}$$

where $\theta$ is the *leading* correction-to-scaling exponent, which, for Ising-like systems, has a value $\theta \simeq 0.54$ (see, e.g., ref. 10) that is frequently *approximated* by $\theta = 1/2$. Physically, however, it is more informative to normalize the correction factor by writing it in the form, say,

$$\left[1 + \frac{1}{4}\left(\frac{t}{t_\times}\right)^{\theta} + B_2\left(\frac{t}{t_\times}\right)^{2\theta} + \cdots\right] \tag{6}$$

where the factor 1/4 is chosen for practical convenience: then $t_\times$ represents the *crossover scale* (or relative temperature deviation at which significant changes in the effective exponent are seen).

If one now *assumes* that the true asymptotic exponent in (5) is $\beta = \beta_{\text{Isg}}$, then the Sienko–Chieux data(1) may be described by $t_\times \simeq 0.6 \times 10^{-2}$. More recently, Chieux *et al.*(11) have performed scattering experiments on $\text{Na} + \text{ND}_3$ above and below $T_c$. They observe quite parallel crossovers from the near-classical values $\gamma \simeq 1.0$ and $\nu \simeq 1/2$ to $\gamma \simeq 1.24$ and $\nu \simeq 0.63$ with essentially the *same* value of $t_\times$.

Now the Singh–Pitzer data(2) give essentially *no* hint of an Ising value for $\beta$. If, nonetheless, one insists on using (5) with $\beta = \beta_{\text{Isg}}$ (and so also adjusts the necessarily fitted value of $T_c$), one finds $t_\times \simeq 1.0 \times 10^{-4}$ and $B_2 \simeq -0.013$ (rewriting the fits presented by Pitzer(3,5)). If this is, indeed, to be the proper theoretical interpretation of these experiments it is clearly mandatory that some reasonable explanation be provided for such a small value of $t_\times$. However, the crossover behavior seen in the sodium–ammonia system certainly points to the correctness of the Ising-like option.

## 3. NON-ISING-LIKE OPTIONS

What about opposing answers such as **Ion ≡ Cl** or **Ion ≡ Other**? Of course, many approximate closed-form theories, from van der Waals onward, yield the classical exponents, $\beta = 1/2$, $\gamma = 1/2$, and $\nu = 1/2$. A favorite theory of this type for electrolytes is the mean spherical approximation (MSA): see, e.g., ref. 12. But, clearly, such theories must be discounted in assessing the universality class of the criticality.

Recently Kholodenko and Beyerlein claimed,(13) on the basis of earlier work,(14) that simple ionic models, in particular the fundamental *restricted primitive model* (or "RPM") having hard spheres of diameter $a$ with charges $\pm q$, should be described by the Kac–Berlin spherical model and hence have coexistence curves truly exhibiting $\beta = 1/2$. However, the analysis in ref. 14 has been severely criticized and it was shown that the claim(13) is without any theoretical foundation.(15) Beyond that, one can argue on rather general grounds(15) that spherical-model criticality is a poor candidate for ionic criticality (confining attention here and below to $d = 3$ dimensions): in particular, it is associated with the divergence of the compressibility *below* $T_c$ as the coexistence curve is approached from either side—something that is never observed in fluids!

Irrespective of theory, however, the issue was soon settled *experimentally*. The spherical model (for $d = 3$) predicts $\gamma = 2$, much larger than both classical and Ising values. But that was ruled out both by Weingärtner *et al.*(16) and Zhang, Briggs, Gammon, and Levelt Sengers.(17) Indeed, for the

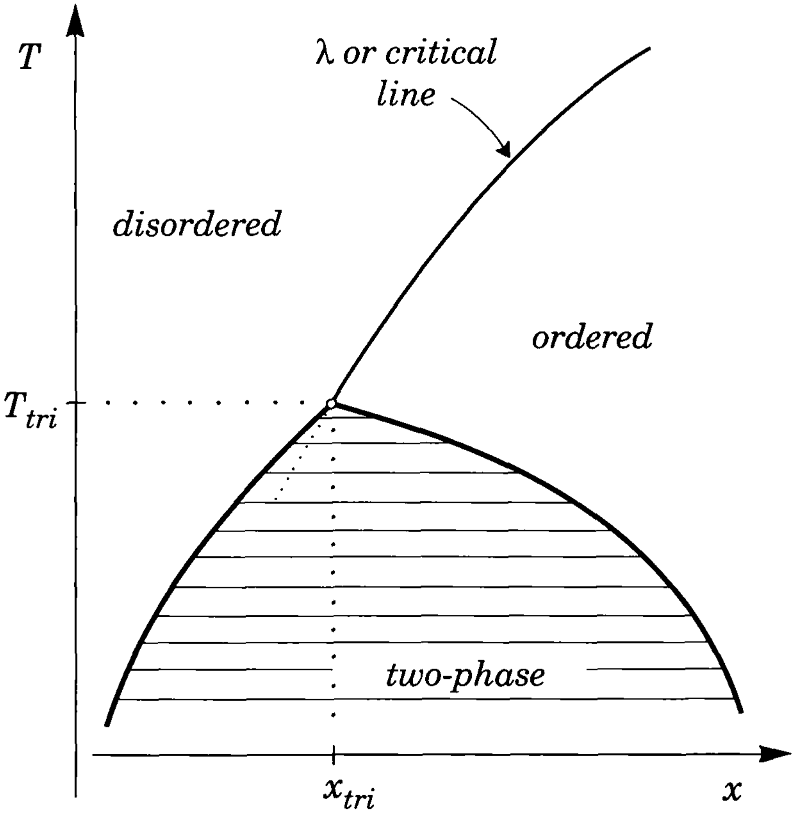

Fig. 1. Typical appearance of a tricritical point, at $(T_{tri}, x_{tri})$, and its vicinity in the plane of symmetry for $d = 3$ dimensions.(19) In the case of helium three-four mixtures, $x$ would denote the mole fraction of $^3$He and the "ordered" phase is a superfluid; for common magnetic materials, the ordered phase is an easy-axis antiferromagnet, the disordered phase is paramagnetic, and $-x$ corresponds to the magnetization induced by a field parallel to the easy axis. The coexistence curve is well described by $\Delta x \sim |T_{tri} - T|^{\bar{\beta}}$ with $\bar{\beta} = 1$ (since the anticipated logarithmic correction factors appear, numerically, only as changes in amplitude).(19) Note the lambda (or critical) line separating ordered and disordered phases: no such feature is seen in the experiments on phase separation in electrolytes.

Pitzer system(3) $\gamma = 1.01 \pm 0.01$ has been found,(17) supporting the hypothesis **Ion = Cl**, namely, that ionic criticality should be classical in character.

In light of these further experiments, Kholodenko and Beyerlein returned to the fray in 1993.(18) Again citing ref. 14 (but with no mention of the criticisms in ref. 15), they now suggested that criticality in symmetric electrolytes (such as the RPM) exhibiting exponents $\beta = 1/2$ and $\gamma = 1$ might be "well understood" in terms of the *tricritical* behavior found in systems such as the antiferromagnetic Blume–Emery–Griffiths (BEG) model (a spin-one Ising model) and observed experimentally in superfluid helium three–four mixtures and certain magnetic materials (see, e.g., ref. 19).

Unfortunately, it is difficult to take this proposal seriously. In the first place, a tricritical point is a *multicritical* point characterized by *four* relevant thermodynamic fields(19) as against a normal critical point with only two. (In the antiferro-BEG model the four fields are temperature, uniform magnetic field, staggered magnetic field, and a more subtle cubic/third-order staggered field.) But there is no experimental evidence for such extra relevant fields beyond temperature and chemical potential. One might, nonetheless, suppose (although for no very good reason in a generic chemical system) that attention could be confined to the tricritical *plane of symmetry*. (This is, e.g., appropriate for helium three–four mixtures.) Then for $d = 3$ dimensions there is, indeed, a coexistence curve and also a diverging susceptibility which can be described by $\gamma = 1$ (although with logarithmic correction factors(19)). Perhaps this seems encouraging: however, as sketched in Fig. 1, this coexistence curve must be described by an exponent $\tilde{\beta} = 1$ (up to small logarithmic corrections)(19) rather than $\beta = 1/2$ (which pertains only to some analog of the staggered magnetization). More importantly, there must necessarily be, *in addition*, a lambda or *critical line* that emanates from the peak of the coexistence curve and separates a disordered "low-density" region from an ordered "high-density" region (or vice versa): see Fig. 1. This is a profound difficulty since the observations of phase separation and criticality under consideration give no hint of any such critical line; nor is there convincing reason for thinking there should be one.

## 4. ISING-LIKE VERSUS CLASSICAL

At present, therefore, there are no serious contenders for **Other** in the equation **Ion ≡ Other**; it is, however, premature to conclude that only the classical option remains open even though, at first hearing, **Ion ≡ Cl** is rather plausible in view of the long-range nature of the electrostatic potential: Recall that in a single-component fluid of, say, hard spheres governed by an attractive Kac potential, $\varphi(r) = u(r/R_0)/R_0^d$ of range $R_0$, classical criticality is attained when $R_0 \to \infty$.(20) However, the long-range Coulombic forces are exponentially screened at finite ionic concentrations (when $d > 2$); furthermore, in the experiments, the Debye screening length $\xi_D \equiv 1/\kappa$ [calculated by the standard formula: see, e.g., Eq. (14) below] seems to be quite small in the critical region, typically say 6–8 Å or less,(16,21) whereas the correlation length amplitude for scattering by the density fluctuations is $\xi_0 \simeq 10$ Å in the Pitzer system.(17)

On the other hand, a newly formulated phenomenological classification(5,21,6) distinguishes *solvophobic phase separation*, driven primarily by short-range repulsive solvent–salt or vacuum–salt net interactions, from *Coulombic phase separation*, driven primarily by the strong electrostatic forces. Systems in the former class are expected to display Ising-like criticality with little contribution from the ions (as in the Japas–Levelt Sengers system(4)); those in the Coulombic class, by contrast, should resemble the RPM via the laws of corresponding states(22): in particular, the relative critical density/concentration/volume fraction $\rho_c/\rho_{\max}$ should be small (see further below, especially Fig. 5). In addition, ionic association plays a significant role, as gauged through the conductivity, etc.,(5,21,23) and, so it appears, they exhibit classical critical behavior.(3,5,16)

At the observational level these distinctions are certainly meaningful and, indeed, seem quite sharp. Clearly, an aim of a more fundamental theory must be to cast light on this important classification by translating it into one or more quantitative (model?) parameters that might be shown in some analytical way to control the type of criticality or to determine the rate of approach to asymptotic behavior, i.e., to control crossover.

To investigate these matters theoretically it is useful to entertain various possible scenarios. For example, strong crossovers are often associated with the close vicinity of a multicritical point of some type. An interesting example is provided by the behavior of solutions of polymers of length $N$ monomers in a poor solvent (see, e.g., ref. 24). Below a critical temperature $T_c(N)$ phase separation occurs into an extremely dilute solution of more or less isolated polymer molecules and a more concentrated, but still dilute solution in which polymer–polymer interactions play a role. A typical coexistence curve for such a system is sketched in Fig. 2: Note the strong asymmetry and the linearity of the high-concentration side of the curve for $t \lesssim 0.8T_c$.

In fact, if one examines the coexistence curve for, say, NaCl or for the RPM, as based on estimates presented by Pitzer(5) (see also recent Monte Carlo studies discussed below(25,26)), one is struck by a strong resemblance! Could this resemblance be more than "skin deep"? (See also ref. 27.) One knows, in fact, that such polymer solutions are controlled by the small

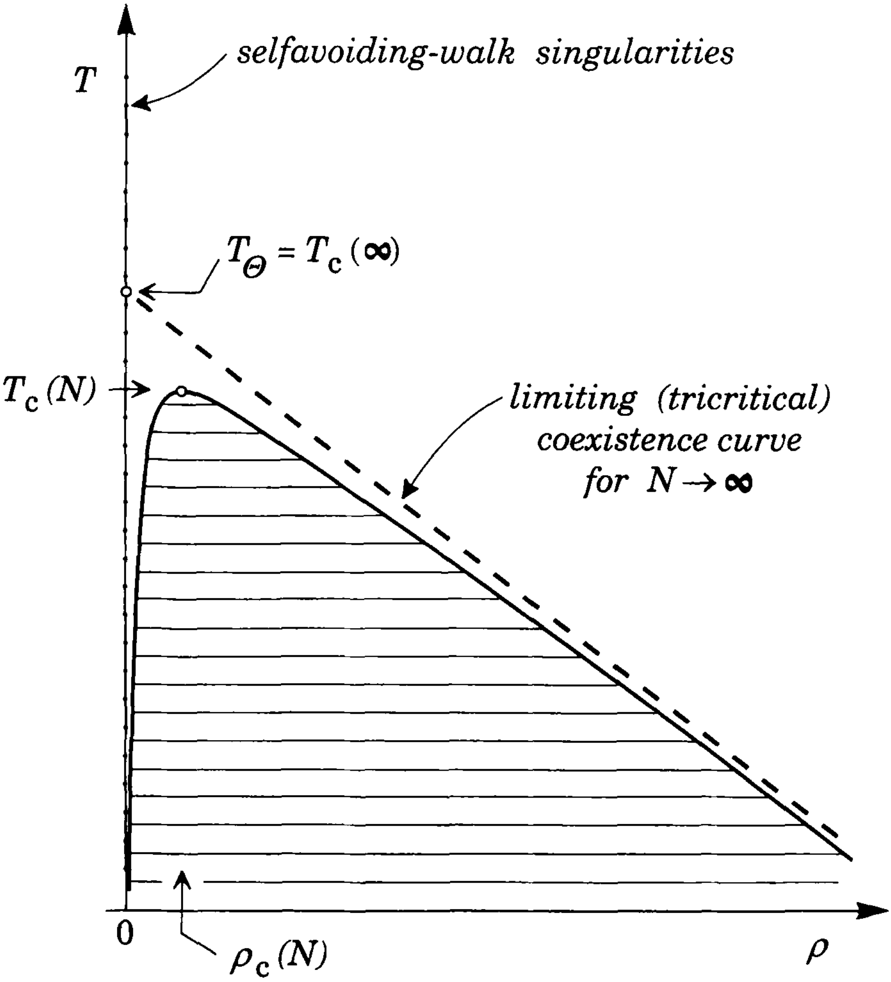

Fig. 2. Sketch of a coexistence curve of a solution of polymer molecules of degree of polymerization $N$ in a poor solvent. In the limit $N \to \infty$, the theta point appears at $T_\theta$ and $\rho_\theta = 0$. The dotted lines on the $\rho = 0$ axis indicate the critical-like singularities associated with the behavior of long ($N \to \infty$) self-avoiding polymer molecules at infinite dilution. Above $T_\theta$ this line is the analog of the lambda line in Fig. 1 and might, speculatively, correspond to the Debye–Hückel limiting-law singularities arising when $\rho \to 0$.

parameter $1/\sqrt{N}$.(24) That sets the scale of $\rho_c(N)$ and of $[T_c(\infty)-T_c(N)]$. In the limit $N \to \infty$ one attains the so-called Flory *theta point* at $T = T_\Theta$ and $\rho_\Theta = 0$: see Fig. 2. This, in turn, may be described in RG theory by a suitable (zero-density) tricritical point with (for $d \geqslant 3$) appropriate classical exponents and a linear ($\beta = 1$) coexistence curve. Furthermore, as $N \to \infty$ the nonclassical/Ising-like regions of the coexistence curves shrink so they appear more classical.

Of course, there is no obvious analog of the crucial parameter $N$ in electrolyte solutions. Is it possible, nonetheless, that a formulation can be found for ionic systems in which $\rho_c/\rho_{\max}$ plays a similar role to $1/\sqrt{N}$? Then $\rho_c/\rho_{\max} \to 0$ would yield a type of tricritical point at $\rho = 0$ with $T_{\mathrm{tri}} = \lim T_c(\rho_c/\rho_{\max} \to 0)$. The role of the nontrivial single-chain, self-avoiding walk statistics(24) in the dilute-limit polymer solutions above $T_{\mathrm{tri}}$ would naturally be played by the Debye–Hückel limiting laws, which entail, as is well known, the $\rho^{3/2}$ singularity in the virial expansion as $\rho \to 0$.(28–30) Crossover to such a tricritical point might conceivably explain the apparent classical character of ionic criticality. But, in the absence of some reasonable theoretical proposal for identifying the analog of polymer length, this speculation is to be regarded as no more than a guide to what *might* be uncovered by a good theory.[2]

## 5. LATTICE MODELS AND MULTICRITICALITY

In light of the solvophobic/Coulombic phenomenological classification it seems essential to study carefully models that encompass features of *both* classes. (Equally, experiments might be designed with a multicomponent and, hence, "tunable" solvent in order to explore the behavior of "interpolating" systems.) In selecting such models one should recall the crucial role played by the lattice gas *alias* the Ising model in understanding criticality in simple, nonionic classical fluids. Accordingly, it seems timely to examine *dissociative lattice* models. The simplest such model embodying the necessary features seems to be the four-state lattice electrolyte specified by(31):

Each lattice site may be (i) empty or (ii) occupied by a (neutral) *molecule* at activity $z$, or (iii) a positive *or* (iv) negative *ion* of charge $\pm q$, respectively.

Nearest-neighbor attractive interactions of strength, say $\varepsilon_0$, operate

---

[2] More recently, with Y. Levin and X.-J. Li, author has shown that *within* Debye–Hückel theory for general dimensionality $d$, this scenario *is* realized when $d \to 2+$ and then $1/\sqrt{N} \sim (d-2)$. The limiting coexistence curve is described by $\rho_c = \rho_{\mathrm{gas}} = 0$ and $\beta$ (really $\tilde{\beta}$, see above) equal to unity with logarithmic corrections. However, improvements in the theory (see below and ref. 60) do *not* sustain this picture.

between molecules; between ions only the (lattice) Coulombic potential acts. The density of ions relative to molecules is controlled by an activity

$$y = \exp(-E_{\text{dissoc}}/k_B T)$$

which can also be regarded as defining a chemical equilibrium constant $K(T)$ for the reaction $M \rightleftharpoons I^+ + I^-$. When $y \to 0$ one has a simple, well-understood (see, e.g., ref. 32) solvophobic lattice gas displaying Ising criticality; when $y \to \infty$ with $yz = \tilde{z}$ fixed, however, one obtains a purely Coulombic lattice version of the (continuum) restricted primitive model, say, the LRPM. The model can obviously be extended to include ion–molecule couplings $\varepsilon_{0+}$ and $\varepsilon_{0-}$ and short-range ionic interactions, including those breaking ionic symmetry (so $\varepsilon_{0+} \neq \varepsilon_{0-}$, etc.): it seems overdue for serious study in its own right.

Note, first, that by virtue of the lattice character one can gain some explicit theoretical control over the dense phase(s) of the model. One such phase should be an ionic crystal which coexists with an ionic, Debye–Hückel vapor at low $T$ and which then melts into a dense ionic fluid. It is certainly possible to generate low-temperature expansions describing this coexistence that should pin down the low-$T$ region of the phase diagram.

Second, Monte Carlo simulations of this model offer the possibility, as in other lattice models (see, e.g., ref. 33), of making reasonable estimates of the critical behavior itself as, say, $y$ is varied. That should at least reveal if there are regions where *crossover* from **Cl** to **Isg** might arise (as in $Na + NH_3$).

Figure 3 shows a highly *speculative* phase diagram for this model embodying insights from the significant work of Nabutovskii *et al.*,(34) who argued that certain electrolytes near criticality were susceptible to an *instability* into a *charge density wave* phase. Such an instability is, in essence, associated with a Lifshitz multicritical point(35): see Figs. 3 and 4. Nabutovskii's work is essentially based on a Landau–Ginzburg charge and density expansion truncated at the quadratic or Gaussian level. It has been discussed recently by Høye and Stell,(8) but seems to warrant further study as regards the nature of ionic criticality.

Figure 3 also illustrates an important conceptual point concerning plausible scenarios. If one discounts criticality with continuously variable exponents (which are generically implausible in $d = 3$ dimensions), two hypothetically distinct sorts of critical behavior, such as **Isg** and **Ion**, *cannot* transform smoothly into one another in a phase diagram. *Either* the corresponding critical lines (or surfaces) must be finitely separated from one another, as suggested by Fig. 3; *or*, if the critical lines (or surfaces) do

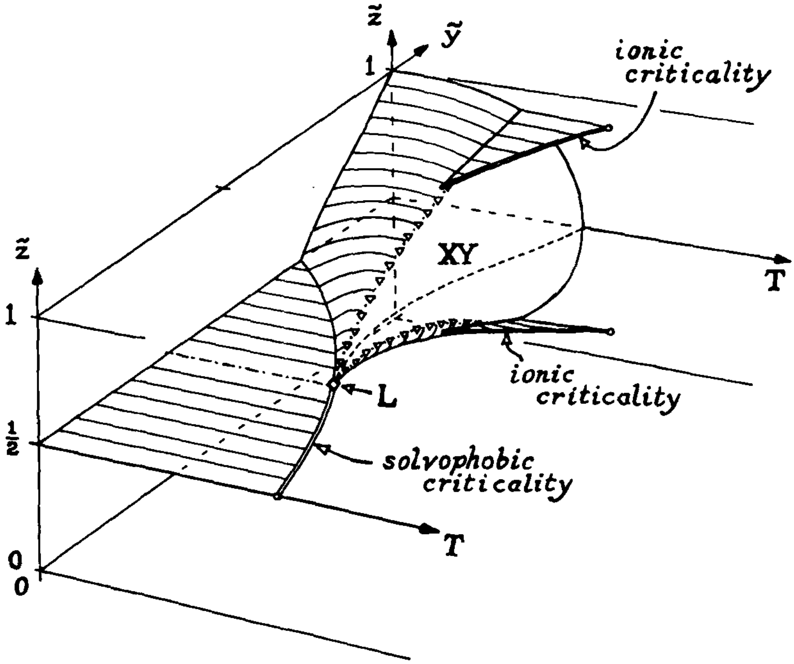

Fig. 3. A speculative and schematic phase diagram for a dissociative lattice gas model electrolyte in terms of the transformed overall activity $\tilde{z} = z/(1+z)$ and transformed dissociative activity $\tilde{y} = y/(1+y)$, where $y = \exp(-E_{\rm dis}/k_B T)$ (see text). Note the Lifshitz multicritical point $L$ terminating the solvophobic critical line at small $y$ and delimiting the Coulombic (large $y$) region. This illustrates a possible scenario uncovered by Nabutovskii and co-workers.(34) Lines of tricritical points (denoted by triangles) and an $XY$-like critical surface emanate from $L$; but in a continuum model or a sufficiently isotropic lattice model this critical surface is expected to be weakly first order in character.(36) Inside the surface appears a charge-density-wave phase somewhat analogous to a liquid crystal.(34) At large enough $y$, however, this should transform into an ionic *crystal*. (The behavior of ANNNI models(37–39) is relevant to the details of the spatially modulated phases in a lattice model.) Growing from the melting surface of the spatially modulated phases are shown two "wings" of first-order transitions (related by particle–hole symmetry in the simplest models). These wings terminate in critical lines that are quite distinct from the low-$y$ solvophobic critical lines and are thus *candidates* for describing ionic criticality with nonstandard exponents. Note that the particle–hole symmetry is artificial as regards the description of the phase diagrams of continuum-space models at high densities: thus, when interpreting this figure for a continuum situation, the upper third, say $\tilde{z} > 2/3$, should be dispensed with. Alternatively, one may introduce lattice hard cores of finite extent—nearest-neighbor repulsions will suffice—which have essentially the same effect.

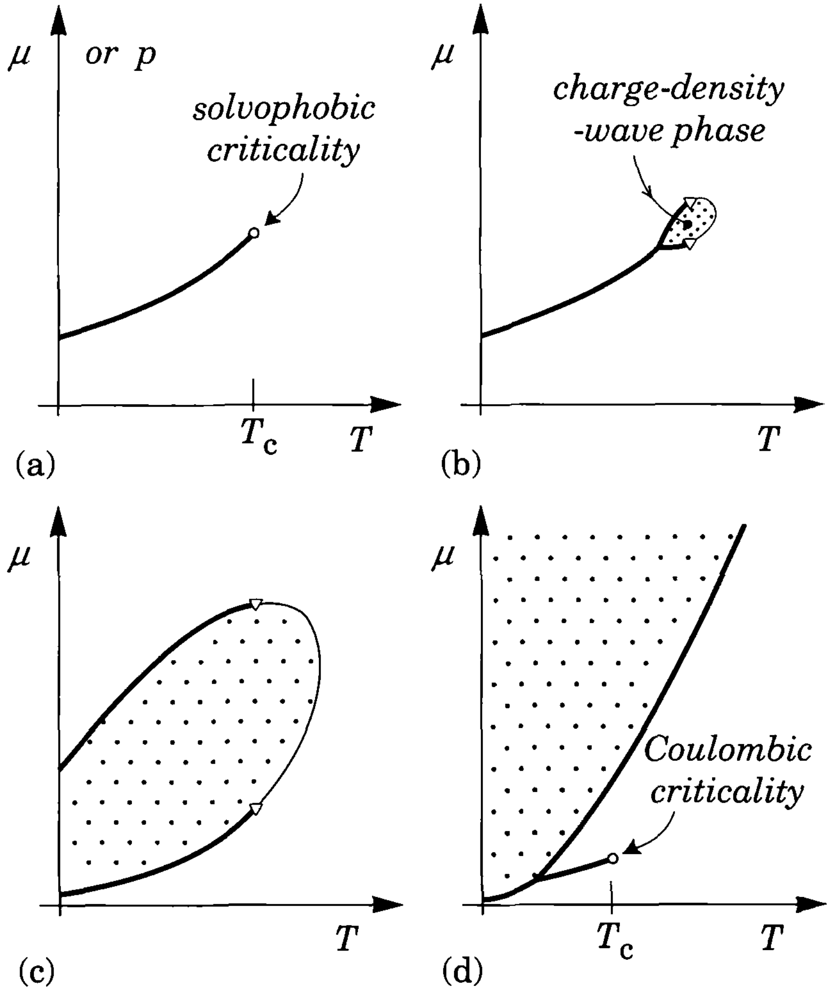

Fig. 4. Successive sections of Fig. 3 at increasing $\tilde{z}$ illustrating schematically (b) the near-critical charge-density instability uncovered by Nabutovskii *et al.*(34) and (c, d) the possible eventual appearance of a disconnected first-order surface terminating in a critical point which, conceivably, could be of non-Ising type.

meet, there must be a separating *multicritical point* (or line) such as the Lifshitz point,(35) shown in Fig. 3, or a bicritical point (see, e.g., ref. 40), etc. In general, the multicritical point must have its own distinct exponents and associated thermodynamic structure. If, in some reasonable model or some real physical system, an Ising-like and an ionic critical line do smoothly join and transform into one another as some parameter, say the pressure or magnetic field, varies, then that fact represents a strong argument in favor of **Ion** ≡ **Isg**. The conclusion that a crossover phenomenon is at the heart of the matter follows and Q.2b ("What fixes $t_\times$?") becomes vital.

These observations demonstrate the value of studying the phase diagrams of a range of dissociative lattice gas models. In particular, one should confirm, refine, or, more probably, *correct* Fig. 3. One approach available again for lattice models (but not for continuum models) is via *real-space renormalization group* approximations, such as the Migdal–Kadanoff method and its extensions.(41 44) These are useful because they generate nonclassical exponents which are distinct for different multicritical points. Frequently, however, especially for $d = 3$ dimensions, the exponent values are not at all accurate. Nevertheless, interesting information about phase diagrams can often be obtained (although one must still be alert to pitfalls: see, e.g., ref. 45).

A notable merit of discrete-state lattice models, which can also be exploited, is the ability, via Kac–Hubbard–Stratonovich transformations, etc., to derive corresponding field-theoretic models with some control over the magnitudes of various pertinent terms in the effective Hamiltonian.(31) Important to this enterprise for ionic models is that the long-range Coulomb interaction can be exactly represented in terms of a local (explicitly, nearest-neighbor) effective Hamiltonian via the so-called sine-Gordan transformation.(46,47) This, indeed, forms the basis of a rigorous proof(47,48) that for small enough $z$ and $q^2/ak_BT$, full Debye screening occurs in the primitive lattice model (in $d = 3$ dimensions); i.e., *all* correlation functions decay exponentially.

Given a field-theoretic Hamiltonian, one is in a position to implement a renormalization-group ($\epsilon = 4 - d$)-expansion. An exploratory investigation(31) suggests that this should clarify more definitively the issue of solvophobic electrolyte criticality remaining Ising-like. That could significantly extend the considerations of Stell and Høye(8,9) and Nabutovskii *et al.*(34) Beyond that, if a distinct type of ionic criticality actually exists, such an RG treatment might reveal its nature, so answering Q.2a. This route was, of course, the way our understanding of the critical effects of long-range power-law potentials decaying as $\varphi(r) \sim 1/r^{d+\sigma}$ ($\sigma > 0$) was originally obtained.(49)

Otherwise, one may reasonably hope that it will point to an answer to Q.2b and suggest why apparent classical behavior is typical. In this respect, the theoretical discussions of crossover to classical behavior in micellar phase separation (where $\gamma < 1$ was originally claimed(50,51)) are both instructive and cautionary(52,53); note, in particular, the calculation of the crossover function for the effective exponent to $O(\varepsilon^2)$,(53) which could be useful in future work.

Finally, lattice models offer some hope of *exact* solution, at least in certain limits. However, for the present problem it is *essential* to examine

Coulombic potentials in dimensions satisfying $2<d<4$. The Bethe lattice, on which many problems have been instructively solved, is therefore probably not so worthwhile. However, *suitably* chosen spherical models *might* prove interesting (compare with ref. 54, where critical endpoints were studied). (One should not, of course, allow confusion with the MSA or mean spherical approximation.$^{(12,55)}$) Smith$^{(56)}$ has, indeed, studied a spherical model with Coulombic interactions: a crossover to asymptotic oscillatory behavior of the charge correlation functions was found and also a phase transition to a charge-ordered or crystalline state. (Compare with Figs. 3 and 4.) Unfortunately, this model lacks, in essence, an overall activity $z$ and it has no charge discreteness. It may, however, be possible to modify or extend the model in useful ways by introducing dissociative aspects, etc., mimicking the discrete four-state model described above.

## 6. THE CRITICAL POINT OF THE RESTRICTED PRIMITIVE MODEL

Although lattice models have distinct theoretical advantages, as outlined above, they have not been studied very seriously. Even if they had been, one would still wish to understand, as far as possible, the behavior of continuum models which are certainly more realistic as regards gas–liquid phase separation (although perhaps hardly more so for solutions with discrete solvent molecules of sizes comparable to those of the ions). The most basic continuum model is undoubtedly that originally considered by Debye and Hückel (DH),$^{(28)}$ namely, the RPM consisting of $i = 1, 2, \ldots, N_+ + N_- = N = \rho V$ hard spheres of diameter $a$, $N_+ = \frac{1}{2}N$ carrying charges $q_i = +q$, and $N_- = \frac{1}{2}N$ charges $q_i = -q$. The solvent is represented merely by a dielectric constant $D$ embodied in the pair potential

$$\varphi_{ij}(r_{ij}) = q_i q_j / D r_{ij} \qquad \text{for} \quad r_{ij} > a \tag{7}$$

The basic energy scale is set by the maximum binding energy for a pair of ions, namely

$$\varepsilon = q^2/Da \tag{8}$$

The natural temperature and density scales are thus

$$T^* = k_B T/\varepsilon \qquad \text{and} \qquad \rho^* = \rho a^3 \tag{9}$$

For the various cubic lattice packings one has $\rho^*_{\max} \equiv \rho^*_{\text{fcc}} = \sqrt{2}$ while $\rho^*_{\text{bcc}} = \frac{3}{4}\sqrt{3}$ and $\rho^*_{\text{sc}} = 1$.

Does the RPM undergo gas–liquid phase separation? And, if so, what are the critical parameters $T_c^*$ and $\rho_c^*$? A naive mean-field theory based on uniformly distributed ions must predict *no* gas–liquid transition in the

RPM since a hard-sphere fluid has none and the positive and negative parts of the pure Coulomb potential cancel exactly! Nevertheless, in 1968 Stillinger and Lovett(57) asserted, without discussion, but apparently in analogy to observed criticality in real fused salts (see also McQuarrie(58)) that there was a critical point; but as to *where*, they remained silent. Today no one doubts their conclusion, but "How may one estimate the location of the critical point?"

To answer this, recall first that van der Waals (vdW) theory for simple fluids can be obtained readily by constructing the reduced Helmholtz free energy $f = -F(T; N_a, N_b, \dots; V)/N$ for species $k = a, b, \dots$ with densities $\rho_a = N_a/V$, etc., as a sum of terms. For a single-component system, one needs just

$$f^{\mathrm{vdW}}(T, \rho) = f^{\mathrm{ideal}} + f^{\mathrm{HC}} + f^{2\mathrm{V}} \tag{10}$$

where (i) the usual expression for an ideal gas with molecular partition function $\zeta(T)$ is

$$f^{\mathrm{ideal}} = k_{\mathrm{B}} T\{\rho - \rho \ln[\rho \Lambda^3(T)/\zeta(T)] \tag{11}$$

(ii) the repulsive interactions are represented by the hard-core term $f^{\mathrm{HC}}(T, \rho)$, traditionally approximated by a free-volume expression diverging at $\rho_{\mathrm{max}}$; and (iii) the attractions are introduced via a second virial coefficient, namely,

$$f^{2\mathrm{V}}(\rho) = A\rho^2, \qquad A > 0 \tag{12}$$

At a semiquantitative level this theory is rather successful. Thus in Fig. 5 the coexistence curves (a) and (c) represent vdW theory for a lattice system [using the exact single-site form for $f^{\mathrm{HC}}(T, \rho)$] and for a continuum, respectively. Plots (b) and (d) depict the corresponding "true" results (known through series expansions, experiments, and simulations to very adequate precision). Lattice vdW theory reproduces the exact value of $\rho_c/\rho_{\mathrm{max}} (= 1/2)$ but overestimates $T_c^*$ [using $\varepsilon = |\varphi_{\mathrm{min}}|$, where $\varphi(r)$ is the pair potential] by 20–25%; the original vdW continuum theory predicts $T_c^*$ only some 15% too high and gives $\rho_c/\rho_{\mathrm{max}} = 1/3$ compared with $\sim 0.29$ for typical gas–liquid criticality. Of course, the vdW theories yield $\beta = 1/2$ in place of $\beta_{\mathrm{Isg}} \simeq 0.325$, the latter value reflecting the much flatter maxima of the "true" coexistence curves.

Can one do as well in such a physically clear and direct way for the RPM electrolyte? A natural first thought (see also ref. 59) is to replace $f^{2\mathrm{V}}(\rho)$ in (10) by the well-known Debye–Hückel limiting law(28–30)

$$f^{\mathrm{LL}}(T, \rho) = k_{\mathrm{B}} T\kappa^3/12\pi = C\rho^{3/2}/T^{1/2} \tag{13}$$

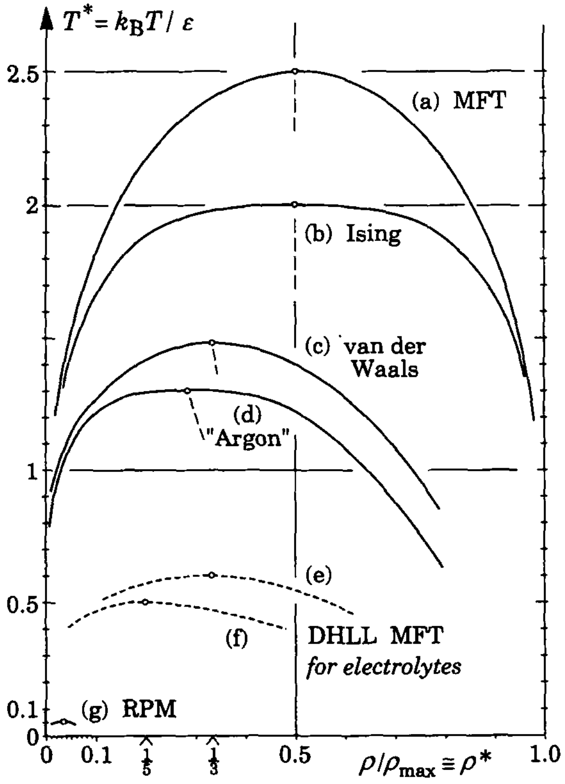

Fig. 5. Semiquantitative coexistence curves illustrating the differences between typical solvophobic and Coulombic critical parameters (see text): $\rho/\rho_{max}$ is plotted versus $T^* = k_B T/\varepsilon$, where, for continuum systems, $\rho_{max}$ may be taken as the density of the crystal or (equivalently, within the precision sought) of the liquid at the triple point, while $\varepsilon = |\varphi_{min}|$ is measured by the minimum of the attractive part of the pair potentials; for the (singly charged) ionic systems one has $\varepsilon = q^2/Da$. Van der Waals (or mean field) critical points (a, e) for *lattice* models and (c, f) for *continuum* systems are compared with the corresponding "true" results for (b) nearest-neighbor lattice gases, (d) Lennard-Jones (12, 6) or argon-like continuum fluids, and (g) the restricted primitive model (as assessed via recent simulations(25,26)). The critical parameter $p_c/\rho_c k_B T_c$ shows similar trends, taking the values (a) 0.386..., (b)·$0.25_4$, (c) 0.3750, (d) $0.29_0$, (e) 0.216..., (f) 0.2083, while the value for the RPM is still subject to considerable uncertainty. The ionic mean field results (e) and (f) use the Debye–Hückel limiting law-form $-C\rho^{3/2}/T^{3/2}$ for the increment to $p/k_B T$ in place of the standard, van der Waals second virial coefficient term $-A\rho^2/T$. Evidently the changes in the critical parameters $T_c^*$ and $\rho_c/\rho_{max}$ away from the solvophobic (or short-range single-component) values are in the correct direction (downward), but by factors that are not nearly small enough to approximate the RPM (g).

where the inverse Debye screening length is given by

$$\kappa^2(T, \rho_+ + \rho_-) = 4\pi q^2(\rho_+ + \rho_-)/Dk_B T \tag{14}$$

For lattice and continuum systems this yields(60) the coexistence curves (e) and (f) in Fig. 5, respectively. Both $\rho_c/\rho_{\max}(=1/3$ and $1/5)$ and $T_c^*(\simeq 0.5\text{–}0.6)$ have dropped significantly. But are these estimates anywhere near correct?

The answer is a resounding "No!": the true values of $T_c^*$ and $\rho_c/\rho_{\max}$ both appear to be an order of magnitude smaller. Indeed, evidence suggests(60)

$$T_c^* \simeq 0.057 \pm 1_5 \qquad \text{and} \qquad \rho_c^* \simeq 0.030 \mp 8 \tag{15}$$

where the signs of the uncertainties are correlated (see also Fig. 6 below). These dramatically small numbers highlight the theoretical problem and indicate the subtlety of phase separation and criticality in pure Coulombic systems. The justification of (14) brings us to the next chapter of the story.

## 7. THE ENTRY OF MONTE CARLO

The power of modern computers suggests that Monte Carlo and related simulation methods should provide a good handle on the RPM. The actual history, however, is not so encouraging. The first serious studies seem to be those of Vorontsov-Veliaminov and co-workers in the Soviet Union in 1970–1976(61–63) (see also ref. 22). They saw evidence of a transition with $T_c^* \simeq 0.094$ and with $\rho_c^*$ in the range 0.3–0.4(63): these estimates differ drastically from (14). Figure 6 shows, in plot (a), a representation of their proposed coexistence curve(22,63): note the logarithmic density scale. Unhappily, the result seems to bear comparatively little relation to the truth—a not so uncommon feature of early Monte Carlo studies where the difficulties of attaining proper equilibrium and allowing for finite-size effects were not well appreciated. Indeed, by 1979 Voronstsov-Veliaminov seems to have concluded that his estimate of $\rho_c^*$ needed to be reduced by a factor of about 3. (See text in ref. 22 associated with ref. 31.)

The first serious *analytical* attack on criticality was reported in a pioneering article by Stell *et al.*(59) in 1976. They brought to bear a wide range of modern approaches to liquid theory involving truncations of different systematic series expansions, various approximate integral equations, etc. All methods confirmed the presence of phase separation and of criticality at small $T^*$, but the quantitative agreement among them was relatively poor. Stell *et al.*, in summary, stated preferred overall estimates of $T_c^* \simeq 0.085$ (with uncertainties of order ±10%, not *so* far from Vorontsov-

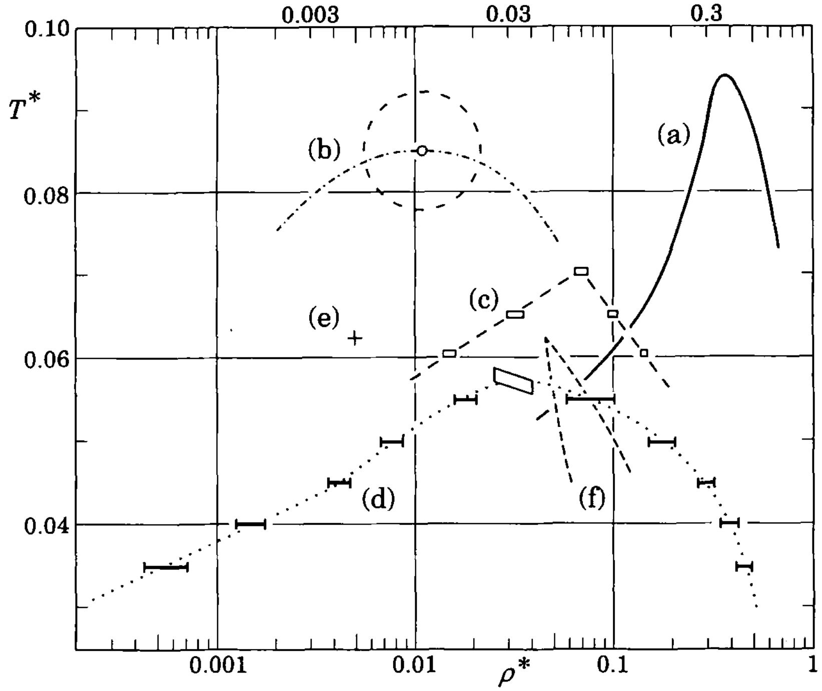

Fig. 6. Coexistence curves (and critical parameters) for the restricted primitive model advanced on the basis of Monte Carlo calculations by (a) Vorontsov-Veliaminov *et al.* (1970–1976),[61, 63] (c) Valleau for 32 ions (1991),[25] and (d) Panagiotopoulos for 512 ions (1991),[26] and (b) schematically, using various systematic approximation schemes by Stell, Wu, and Larsen (1976).[59] The point (e) is marked for comparison with Fig. 7. The plot (f) is reproduced from ref. 22: see text. The curves connecting the data points in cases (c) and (d) are no more than guides to the eye. The sloping box probably contains the true critical point. [Note that a similar figure in ref. 26 misplots the results (a), (b), and (f).]

Veliaminov *et al.*) and $\rho_c^* \simeq 0.011$[59]: see plot (b) in Fig. 6 and note the strikingly low value estimated for $\rho_c^*$ (although it is subject to uncertainties of $\pm 50\%$[59]).

Subsequent work, notably, by Gillan[64] and Pitzer and co-workers[5,65] carefully examined the low-density, low-temperature vapor phase. One quickly discovers, as emphasized by Bjerrum[66] already in 1926—only 3 years after the Debye–Hückel theory was published[28]—that for $T^* \lesssim 1/4$ a significant amount of *ion-pairing* or *neutral dimer formation* appears. Indeed, as $T$ falls further, increasing association of ions into rather tightly bound trimers, tetramers, etc., also occurs. By calculating the thermodynamic properties of the dominant clusters as far as feasible (invoking

Monte Carlo and other aids), the free energy of the electrolyte vapor can be estimated. (However, the inevitable degree of arbitrariness in defining a bound cluster should be noticed.) Combining the vapor data with improved Monte Carlo estimates(67) of the liquid free energy *below* $T_c$ yields estimates for the phase boundary and thence the coexistence curve. Plausible extrapolation(5) suggests, in particular, that $\rho_c^*$ exceeds 0.02 in accord with (14).

The low value of $T_c^*$ and the consequent strong clustering constitute, in fact, a major challenge to direct Monte Carlo simulation in the broad critical neighborhood: thus the ion pairs, in particular, can be very tightly bound, since $\exp(-\varepsilon/k_B T) = \exp(-1/T^*) < 4 \times 10^{-6}$ for $T^* \lesssim 0.08$. Consequently, it is hard to be sure that equilibrium is achieved in the time available for simulation. The long-range character of the Coulomb force and the associated finite-size boundary conditions pose further problems. Nevertheless, two recent studies inspire somewhat increased confidence.

In 1991 Valleau(25) simulated a system of $N = N_+ + N_- = 32$ ions at three low temperatures: see Fig. 6, plot (c). He concluded that $T_c^*$ was close to 0.070, significantly *lower* than the results of Stell *et al.*,(59) whereas $\rho_c^*$ appeared much higher at about 0.07. The implied coexistence curve would then appear rather sharp: see Fig. 6. If one supposed that it should instead be somewhat rounded off, one might rather conclude $T_c^* \simeq 0.068 \pm 1_5$ and $\rho_c^* = 0.06 \pm 1$.

Very soon after Valleau's work Panagiotopoulos(26) introduced a novel Monte Carlo method—a Gibbs-ensemble technique—especially designed to facilitate the investigation of phase transitions by running, simultaneously, *two* distinct simulations at the same temperature and chemical potential $\mu$, but with differing densities. His results for $N = 512$ ions at $T^* = 0.035, \ldots, 0.055$ are displayed in Fig. 6, plot (d).(26)

The disagreement with the data of Valleau is rather large and perhaps surprising. While differences in the electrostatic boundary conditions[3] certainly play a role, it seems likely that the larger value of $N$ is also important. Indeed, finite-size scaling theory for critical points(68,69) shows that the shift or rounding of a critical point should scale as

$$[T_c(N) - T_c(\infty)]/T_c(\infty) \sim \Delta T/T_c(\infty) \sim (\xi_0^d \rho_c/N)^{1/d\nu} \qquad (16)$$

If we take $\nu \fallingdotseq 1/2$, the change in $N^{-1/d\nu}$ is from 0.10 to 0.016, a factor 6.3 improvement; this falls to about 4.3 if $\nu \simeq 0.63$; but, either way, it seems that the Panagiotopoulos data(26) should be significantly closer to the true behavior of the RPM.

[3] See especially J. M. Caillol, *J. Chem. Phys.* **100**, 2161 (1994).

Panagiotopoulos himself stated $T_c^* = 0.056$ and $\rho_c^* = 0.040$ without any discussion(26): but those values would seem to imply an unusually flat coexistence curve with a rather small value for the exponent $\beta$. If, however, one neglects further finite-size effects and accepts $1/2 \leqslant \beta \leqslant 1$, Panagiotopoulos' data can be roughly extrapolated—a linear $(\rho, T)$ plot seems preferable to the logarithmic plot of Fig. 6: this suggests that the sloping box shown in Fig. 6 probably contains the critical point of the RPM. The estimate (15) roughly represents this conclusion: however, insofar as the Valleau and Panagiotopoulos data are comparable, extrapolation on $N$ using (16) suggests that $T_c^*$ could well be from 3 to 6% lower than indicated by (15). Likewise, there will be shifts $\Delta\rho_c(N)$ in the critical density: their scaling properties are somewhat more subtle (entailing the breaking of the expected asymptotic gas–liquid symmetry). Evidently, (15) cannot be regarded as definitive until further, more precise simulations are performed that in particular also allow systematic extrapolation on system size. It may be some time before such calculations prove feasible. By the same token the likelihood that one could convincingly distinguish classical from Ising-like critical behavior by Monte Carlo studies seems remote at present.

## 8. MEAN-FIELD THEORIES FOR ELECTROLYTES

We may accept the estimates (15) for the location of the critical point of the RPM and note that, via the law of corresponding states,(21,22,70) Coulombic electrolytes seems to match the model reasonably well. Unfortunately, however, the Monte Carlo calculations leading to (15) provide little if any physical insight into the small magnitudes of $T_c^*$ and $\rho_c^*$. Hardly more really seems to be gained from the analytic truncation schemes,(59) nor from the solution of various approximate integral equations, such as the HNC and PY equations (see, e.g., refs. 22, 71, and 72). Indeed, as observed some time ago,(73) most of the standard approximate integral equations for fluids when applied in the critical region seem to give highly unphysical, not to say disastrous, accounts of the thermodynamics and correlation functions (on which they are based).(71–73) By comparison, van der Waals equations, despite their inability to reproduce the Ising-like exponents, appear as paragons of virtue! [Recall Fig. 5, plots (a)–(d).]

The most hopeful route to resolving theoretically the Coulombic–solvophobic dichotomy would seem to lie in a renormalization group treatment. However, many, if not most, successful RG theories of particular types of criticality and multicriticality have rested on crucial physical insights into the causes and character of the transition in question.

Frequently these in turn have been embodied in some sort of mean field theory. Accordingly we return to the question of Section 6: "Is there a simple, physically transparent, semiquantitative theory of criticality in the RPM?"

In fact, the answer seems to be "Yes!"(60) To explain this, recall, first, the famous theory of Debye and Hückel (DH),(28) perhaps the first theory to embody fluctuating particle–particle correlations in a fundamental way. DH focus on the free energy of an ion $i$ embedded in a sea of other ions $j$, positive and negative, distributed about it at distance $r$ according to the Boltzmann factor $\exp[-q_j\phi_i(r)/k_B T]$, where $\phi_i(r)$ is the mean local electrostatic potential around $i$.(28-30) This leads directly to the Poisson–Boltzmann (PB) equation for $\phi_i(r)$. Linearizing the PB equation facilitates its solution, but *also* ensures the satisfaction of important statistical consistency conditions.(29,30) Thence follows the crucial DH screening of the bare Coulomb interaction by the factor $e^{-\kappa r}$, where, for the RPM, $\kappa(T, \rho)$ is given by (14) above. The finite diameter $a$ of the ions enters via a natural boundary condition which simply asserts that any screening charge vanishes for $r \leqslant a$.(28-30) Finally, a simple charging process—switching on the electrostatic interactions by putting $q_i \Rightarrow \lambda q_i$ and integrating from $\lambda = 0$ to 1—yields the Coulombic contribution to the free energy as

$$f^{\mathrm{DH}} = k_B T[\ln(1+\kappa a) - \kappa a - \tfrac{1}{2}(\kappa a)^2]/4\pi a^3 \tag{17}$$

Note that when $\rho$, and so $\kappa a$, becomes small this reproduces the limiting law (13): that, however, is *independent* of the ionic diameter $a$ and hence lacks an important part of the physics. (Note, indeed, that a point-charge classical electrolyte with no repulsive cores would actually collapse, having a free energy unbounded below.)

The combination $f = f^{\mathrm{ideal}} + f^{\mathrm{DH}}$ [see Eq. (11)] then constitutes the DH theory—advanced over 70 years ago. It was, of course, highly successful in accounting for the basic properties of real electrolytes. And, indeed, the limiting laws, Eq. (13), etc., were subsequently shown to be exact consequences of statistical mechanics.(29,30) However, only in the last few years has it been realized(60,70) that the DH theory itself predicts a critical point for the RPM. The critical parameters turn out to be(60,70)

$$T_c^* = \frac{1}{16} = 0.0625, \qquad \rho_c^* = \frac{1}{64\pi} = 0.00497... \tag{18}$$

with, furthermore,

$$\kappa_c a = 1 \qquad \text{and} \qquad p_c/\rho_c k_B T_c = 16 \ln 2 - 11 = 0.09035... \tag{19}$$

Evidently, then, there is indeed a very physical theory that predicts a $T_c^*$ for the RPM just 7–13% above the Monte Carlo-based estimate (15)! It would seem to parallel vdW theory for simple fluids. The corresponding coexistence curve is shown in Fig. 7. Naturally, all the critical exponents are classical.

Following vdW theory, one may also add a free-volume term $f^{\mathrm{HC}}$ to account for the hard-core exclusions. If the volume parameter is chosen to get the correct second virial coefficient, $T_c^*$ is reduced, but only by about 1.9% (while $\rho_c$ falls by 8.5%). On the other hand, the DH prediction for the critical density, which can be written, invoking (14), as $\rho_c^* = (\kappa_c a)^2\, T_c^*/4\pi$, is too low by a factor of 5–7. Can this not so trivial defect be understood and corrected?

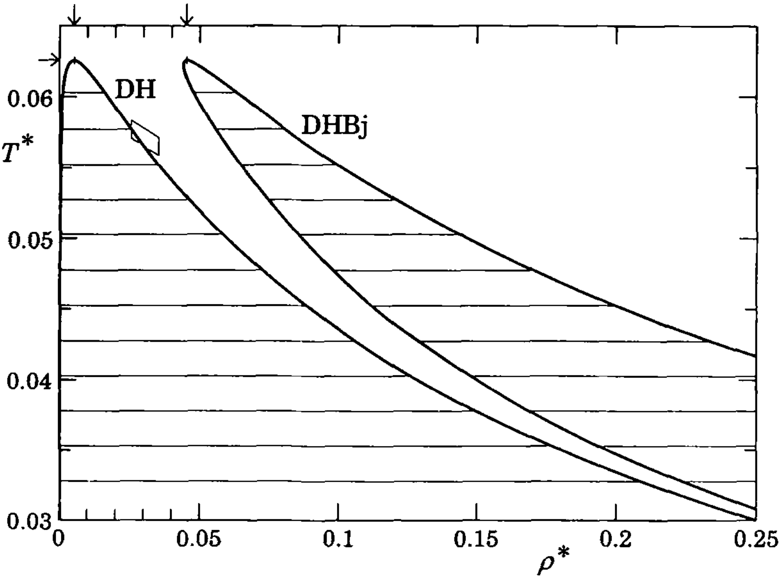

Fig. 7. Coexistence curves predicted for the restricted primitive model: according to the original theory of Debye and Hückel[28] (DH) and by the theory as modified by inclusion of neutral, ideal ion pairs according to the proposal of Bjerrum[66] (DHBj). The peculiar "banana" shape implied by the DHBj theory is a direct consequence of the assumption that the strongly dipolar ion pairs have no interactions with the ionic fluid of free ions or with each other: see text. The sloping box represents the estimates for the true critical point of the RPM: see also Fig. 6, in which (e) represents the pure DH critical point, while (f) somewhat *mis*represents the DHBj theory.[22]

## 9. BJERRUM ASSOCIATION

Now a major approximation of DH theory is the linearization of $\exp[-q_j\phi_i(r)/k_BT]$. This is clearly most serious (i) when $q_i = -q_j$ and $r \simeq a$, so that $q_j\phi_i \approx -q^2/Da$, and (ii) when $T^* = k_BTDa/q^2$ is small (which it is near criticality). But, as discussed in Section 8, these are just the circumstances leading to the significant association of + and − ions into strongly bound neutral pairs. By allowing for such dimers one can, in fact, go a long way to improve DH theory as shown by Bjerrum,(66) who treated the neutral bound pairs as a distinct, ideal (i.e., noninteracting) chemical species in equilibrium with the DH ions.

Theoretically it is advantageous to consider the pressure(60)

$$p(T; \mu_1, \mu_2) = \max_{\rho_1, \rho_2} \{f(T; \rho_1, \rho_2) + \mu_1\rho_1 + \mu_2\rho_2\} \tag{20}$$

where now $\rho_1 = \rho_+ + \rho_-$ is the density of free ions, while $\rho_2$ is the density of the uncharged dimers; of course, $\mu_1$ and $\mu_2$ are the corresponding chemical potentials. Chemical equilibrium is ensured by imposing $\mu_2 = 2\mu_1$ $(\equiv \mu_+ + \mu_-)$. In addition, one needs to know the *association constant*

$$K(T) \equiv \zeta_2(T)/\zeta_+(T)\,\zeta_-(T) \tag{21}$$

where the $\zeta_k$ are the dimer and monomer internal partition functions: see Eq. (11) above.

By considering, say, a positive ion attracted by the direct field of a negative ion, Bjerrum was led to the form(30,66)

$$K(T) = 4\pi \int_a^d [\exp(+q^2/k_BTDr)]\, r^2\, dr \tag{22}$$

The lower cutoff clearly expresses the hard-core condition: without the upper cutoff, however, the integral would diverge. Bjerrum proposed to take $d = \frac{1}{2}a/T^*$ (for $T^* \leqslant 2$), which corresponds to the minimum of the integrand. At first sight this introduces a most unwelcome degree of arbitrariness. In consequence, many later authors have revisited the question. It transpires, however, that Bjerrum's choice was most judicious! Careful considerations by, in particular, Ebeling(30,74,75) have led to a more elaborate expression for $K(T)$ which is not restricted to $T^* < 2$ and which, moreover, ensures that DH theory is exactly corrected to order $\rho^2$ (i.e., *beyond* the limiting-law $\rho^{3/2}$). (See also refs. 76 and 77.) Nevertheless, Bjerrum's form is reproduced *asymptotically* to *all* orders at low temperatures, giving

$$K(T) = 4\pi a^3 T^* e^{1/T^*}[1 + 4T^* + 20(T^*)^2 + \cdots] \tag{23}$$

while for $T^* \lesssim 1/10$ the accuracy of (22) is already better than 2 or 3% and drops to less than 0.2% for $T^* \lesssim 1/16$.(60) [It should, incidentally, be remarked that the form for $K(T)$ advocated by Fuoss(78) and used in ref. 70 cannot be considered satisfactory, especially for the RPM: see refs. 30 and 77.]

What are the predictions of the combined Debye–Hückel and Bjerrum (DHBj) theory? First, one finds(22,60,70,74) that the critical temperature $T_c^* = 1/16$ is unchanged; so is the product $\kappa_c a = \kappa(\rho_{1c}, T_c) = 1$. However, the overall critical density is increased to

$$\rho_c^* \equiv \rho_{1c}^* + 2\rho_{2c}^* = \left(1 + Q_{16}\frac{e^8}{2^9}\right)\frac{1}{64\pi} = 0.0452... \tag{24}$$

where $Q_{16} = 1.39049...$ is the critical value of the last factor in (23). It is clear physically that the neutral Bjerrum pairs, of density $\rho_2$, reduce the available number of free ions, of density $\rho_1 = \rho_+ + \rho_-$, so that a larger overall ionic density is needed before the free ions achieve criticality. Comparison with the estimate (15) shows that agreement with expectations is much improved: relative to the pure DH result (18) the target has been overshot, but only by 50%, more or less. This is really rather satisfactory. In retrospect, furthermore, since the formation of ion pairs is such a strong nonlinear effect, it is not so surprising that it apparently escapes the systematic high-temperature and other series-based approximation schemes brought to bear by Stell, Wu, and Larsen.(59)

But all is not well. The DHBj coexistence curve is shown in Fig. 7: it is most peculiar and, surely, unphysical. The rounded "banana" shape still exhibits the anticipated classical critical exponents (although in ref. 22 it was reported as a sharp "marlin spike" corresponding to $\beta = 1$: see Fig. 6(f)]. What is the cause of this unexpected behavior?

The reason is not far to seek: as $T$ is lowered further below $T_c$ the association constant $K(T)$ increase exponentially rapidly and it becomes more and more favorable for the ions to pair off into the noninteracting Bjerrum pairs: hence the overall density needed to reach the DH phase boundary of the remaining, essentially independent, free ions increases sharply. This is what distorts the coexistence curve. Actually it is interesting to note that the width

$$\Delta\rho(T) \equiv \rho_{\text{liq}}(T) - \rho_{\text{gas}}(T) \tag{25}$$

of the coexistence curve is the *same* in the DHBj theory as in the pure DH theory (as can be checked graphically in Fig. 7): it is only the diameter, $\bar{\rho}(T) = \frac{1}{2}[\rho_{\text{liq}}(T) + \rho_{\text{gas}}(T)]$ that gets increasingly displaced. At a deeper level, the main defect of the theory is clear: even though the Bjerrum pairs are neutral, they do have relatively large *dipole moments*, which certainly interact with the ionic fluid as well as with one another. Bjerrum's approximation that the ion-pairs are ideal is thus not acceptable at the low temperatures in question.

## 10. BEYOND DEBYE–HÜCKEL AND BJERRUM

To go further than the DHBj theory, in the hope of obtaining a satisfactory semiquantitative description, Levin and the author have recently proposed(60) that the primary interaction of concern when ion pairs begin to form is that of the individual $(+, -)$ dipolar pairs with the screening electrolyte fluid of free ions. Now this, in turn, can be handled by DH theory. More concretely,(60) consider a closely associated pair of oppositely charged ions forming a dipole of magnitude $qa_1$, where the mean ionic pair separation is $a_1(T) \equiv [1 + s_1(T)]\, a$. This is clearly bounded below by the ionic diameter $a$, but must in general increase with $T$. Although there is inevitably some arbitrariness in the definition of $a_1$, the ambiguity is small for $T^* \lesssim 1/16$: explicitly one can estimate $s_1(T) \lesssim 0.13$ in this region.(60) For a first approach, therefore, one may simply take $a_1 = a$.

Then, using the linearized PB equation one can compute the screening of the dipoler field, which is, of course, controlled by the Debye factor $e^{-\kappa r}$. To simplify the calculations, the boundary conditions are imposed on the surface of an *effective spherical cavity* surrounding the dipolar pair and of radius $a_2 \equiv (1 + s_2)\, a_1 \simeq (1 + s_2)\, a$.(60) Then, following the DH procedure,(28–30) one finds that the dipole-ionic-fluid (DI) contribution to the free energy can be written

$$f_{2,1}^{\text{DH}} = \frac{\kappa^2 a_1^2 q^2}{D a_2}\, \rho_2 \omega_2(\kappa a_2) \sim \frac{\rho_1 \rho_2}{a_2 T} \sim \frac{K(T)\, \rho_1^3}{a_2 T} \tag{26}$$

where, with $x = \kappa a_2$, one has

$$\begin{aligned} \omega_2(x) &= 3\left[\ln\left(1 + x + \tfrac{1}{3}x^2\right) - x + \tfrac{1}{6}x^2\right]\Big/ x^4 \\ &= \tfrac{1}{12}\left(1 - \tfrac{4}{5}x + \tfrac{4}{9}x^2 - \cdots\right) \end{aligned} \tag{27}$$

All higher-order multipoles arising from the finite size of the ion-pair dipole also contribute to the free energy and can be likewise evaluated(60); however, their total contribution proves quite negligible.

Geometrical considerations show that the effective cavity radius $a_2$ is bounded when $a_1 = a$ via $\frac{1}{2} \geqslant s_2 \geqslant \sqrt{\frac{3}{4}} - 1 \simeq -0.13$. Elementary arguments further suggest that $0.25 \gtrsim s_2 \gtrsim 0.10$; more precisely, an angular average gives $\bar{s}_2 \simeq 0.16198$,(60) which it is reasonable to use pending a more accurate solution that embodies the proper boundary conditions (determined by the surfaces of two overlapping spheres of radius $a$ with centers separated by $a_1$).

Augmenting the DH and Bj contributions to the free energy with the dipole-ionic term (26) yields what we may call a DHBjDI theory. The coexistence curve then predicted is shown in Fig. 8: it has, once again, a perfectly reasonable and physically sensible shape and, naturally, is described by classical exponents. The new critical parameters are(60)

$$T_c^* \simeq 0.057_4, \qquad \rho_c^* = 0.028_0, \qquad \kappa_c a \simeq 1.13 \tag{28}$$

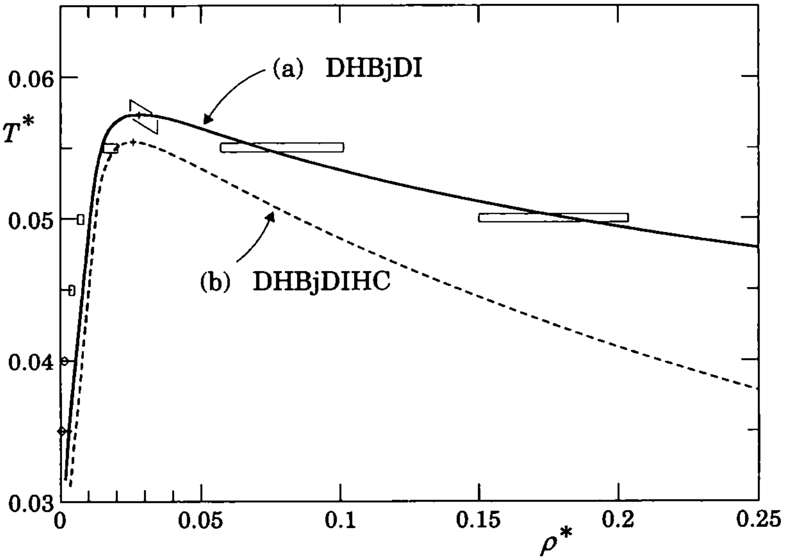

Fig. 8. Coexistence curves for the RPM predicted by the Debye–Hückel–Bjerrum theory augmented by the dipole-ionic (DI) fluid coupling as computed, using DH methods, by Levin and Fisher(60): (a) the DHBjDI theory alone; (b) with the addition of free-volume terms (saturating at close packing on a CsCl structure) representing the hard-core, excluded-volume effects. As in Figs. 6 and 7, the sloping box indicates the likely location of the true RPM critical point.

As can be seen, these fall well within the Monte Carlo-based range (14): that is most gratifying. However, the close agreement must be considered somewhat fortuitous. Certainly, it is reasonable to allow for the direct hard-core repulsions by adding a free-volume term. If one uses parameters corresponding to packing into a bcc (or CsCl) crystal, one obtains the second plot shown in Fig. 8.(60) The critical point now lies somewhat lower at $T_c^* \simeq 0.055_4$, $\rho_c^* \simeq 0.026_0$, with $\kappa_c a \simeq 1.03$ [but is still encompassed by (15)]. Note also that $T_c^*$ is almost proportional to $1/a_2$ [see (26)], so that improved estimates for $s_2$ could also change the critical point estimates. Nevertheless, considering the conceptual simplicity of this theory, the fact that it matches the current Monte Carlo simulations so reasonably suggests that it correctly embodies the most significant physical effects.

Refinements of the theory are readily contemplated. First, perhaps, is the issue of the dipole–dipole interactions (which enter at order $\rho_2^2 \sim \rho_1^4$). These may be included in a purely ad hoc way, as in vdW theory, via a term like (12) with an amplitude, say, $A_{2,2}$. Numerically, in reduced units, a value $A_{2,2}^* \simeq 0.1$ seems reasonable—but this small value is hard to justify theoretically.(60),[3] Rather, the relatively strong dipole–dipole attractions should probably be best accounted for by allowing for the formation of bound $(2+, 2-)$ tetramers and further higher-order clusters. Clearly this can be accomplished along similar lines by introducing additional chemical potentials and appropriate new association constants, etc.

The odd and hence charged clusters, $(2+, 1-)$ and $(1+, 2-)$ trimers, etc., are less stable.(5,64,65) They will also contribute to the screening, but can be included in the DH ionic fluid theory in the standard way.(28–30) Some allowance for their larger effective diameters $a_3$, etc., would, however, be needed. An ingenious (but evidently approximate) suggestion of Debye and Hückel(28) might be invoked in that regard. According to DH theory the neutral clusters do not contribute to the screening directly; the additional free energies of the tetramers, etc., in the ionic fluid can be estimated along the lines used for the dimers which led to (26) and (27) (this will clearly be essential to avoid a repetition of the "banana effect").

---

[3] At higher densities and hence lower temperatures on the coexistence curve the resemblance of the concentrated ionic fluid to the expanded ionic crystal should be noted.(28,58) This can be represented by a free-energy density contribution $\sim (e^2/D)\,\rho^{4/3}$. With a suitably chosen amplitude(60) such an extra term also improves the agreement of the coexistence curve with the simulations, etc., for $T \lesssim 0.8T_c$.(15,26) Many years ago, indeed, McQuarrie(58) combined this term, using its full Madelung amplitude, with a free-volume term inspired by the cell picture of liquids. Thereby, indeed, he obtained a passable description of RPM phase separation and criticality. However, the quite incorrect singular behavior of his free-energy functional at low densities, taking no account of Debye screening, etc., makes this approach basically unacceptable even though at intermediate and high densities it seems to embody a correct physical picture.

It has been suggested by Friedman[23,78] that an important role is played, once a significant number of dipoler ion pairs form, by changes in the dielectric constant and the associated lowering of the free energy, which should favor the liquid phase. For more or less permanent dipoles in a nonelectrolyte, such an effect may well be important. In the presence of strong screening by free ions, however, it seems open to serious question as a distinct mechanism. Indeed, the success of the DHBjDI calculations just described suggests that cooperative, longer-range dipole–dipole couplings between associative ion pairs are not quantitatively important in the RPM: most of the effect may be contained in the polarization of the local ionic fluid embodied in the expression (26). However, this could well be an overoptimistic assessment in light of the large predicted density of ion pairs in the critical region relative to the always rather low density of free ions (controlled by the value $\kappa_c a \simeq 1.1$).

## 11. SOME LAST WORDS

The mean-field approaches to the restricted primitive model reviewed in the previous few sections seem to have uncovered, and to represent semi-quantitatively, the principal physical effects leading to Coulombic criticality in simple, symmetric model electrolytes, namely (i) screening by the free ions as described by Debye and Hückel, (ii) the association of + and − ions to form neutral, tightly bound pairs, following Bjerrum, and (iii) the interaction of the dipole moments of the bound ion pairs with the screening ionic fluid.

On the other hand, at this stage of the theory no account has been taken of the critical fluctuations. But if, as in truth seems likely (see Section 2) and as argued specifically by Stell[9] (but see also the Appendix), the RPM displays asymptotic Ising-like criticality, then the fluctuations must be crucial. Nor do the calculations presented yet supply any estimate of the size of the putative nonclassical regime, i.e., of the vital crossover temperature $t_\times$ which is needed to understand the experiments (see Section 2).

As regards the magnitude of $t_\times$, it may nevertheless be that the mean field values of $\kappa_c a$ play a role.[60] (Recall that $\kappa_c a = 1$ in DH and DHBj theory, while $\kappa_c a \simeq 1.0$–$1.15$ according to the augmented, DHBjDI theories.) For simple fluids with attractive *density–density* forces of range $R_0$, the Ginsburg criterion for the validity of Landau theory indicates $t_\times \propto (a/R_0)^6$ (for $d = 3$ dimensions); RG theory confirms this[80] and shows that $R_0^2$ is appropriately defined via the second spatial moment of the pair interaction potential. Now it is tempting to guess that the *effective* attractive density–density potential in the RPM, say $\varphi_\rho(r)$, mirrors the charge–charge pair correlations and so decays as $e^{-\kappa r}/r$. [In general, of course, the simple DH formula (14) for $\kappa$ must require modification in the critical region: see Appendix.] Then one finds $R_0 \propto 1/\kappa_c$ with an appropriate proportionality coefficient. On comparison with a Lennard-Jones (6, 12) fluid, for which the coexistence curve hardly displays a classical region, one might thereby conclude $t_\times \lesssim 0.15 t_\times^{LJ}$.(60) Such an inequality suggests that the RPM and hence its real Coulombic analogs(1–5,21,70) should display somewhat enhanced regions of classical-like behavior. However, on this basis the effect would not be very strong. Furthermore, the argument is incomplete, on various counts: In particular, the full expression for $t_\times$ involves not only the potential range $R_0$, but also other model (or real) parameters which could well be important.

Second, and clearly of most importance at a fundamental level, is the question of the nature and detailed origin of the density or number correlation functions $g_{\rho\rho}(\mathbf{r})$ and their relation to the charge–charge correlation functions $g_{qq}(\mathbf{r})$. For models lacking charge symmetry, the cross correlation functions $g_{\rho q}(\mathbf{r})$ may well also prove of importance: However, standard RG considerations (see, e.g., ref. 80) suggest that effective number and charge critical densities (or "operators") may then result simply as particular, nonuniversal, linear combinations of the different ionic (and, in general, solvent) densities.

Of course, the various approximate integral equations for fluids are formulated directly in terms of the ionic correlation functions $g_{ij}(\mathbf{r})$. This should make the careful study of these integral equations potentially rewarding; unfortunately, as mentioned above, the predicted behavior in the critical region seems to be strongly unplysical.(71–73) Nevertheless, various calculations,(22,34,59) including "mathematically exact but physically approximate" studies of spherical models,(56) etc., do indicate the onset of oscillatory behavior in the charge–charge correlation functions at overall ionic densities comparable to the critical density. This effect is, indeed, consistent with the DH, mean-field values $\kappa_c a$ (= 1.0–1.15), which suggest a large degree of local charge ordering near criticality. Conceivably, indeed, it could be the comparatively longer range of the decaying envelope of the charge density oscillations that sets the effective range $R_0$. See also the discussion in the Appendix [where $g_{qq}(r)$ is to be identified with $h_D(r)$]. However, the integration of such results on the correlation functions into a coherent picture of the critical region of model electrolytes is a challenging and far from completed task.

Needless to say, real electrolytes have further complications, beyond those represented in the RPM or even in the extended four-state lattice model introduced in Section 5: the polarizability of the individual molecular ions is but one example. Before theorists focus too strongly on such more elaborate features, however, the experience of the last three or four decades of studies of critical phenomena should not be forgotten. In particular, the careful and systematic theoretical study of the simplest models, especially lattice models, combined with complementary simulations (allowing properly for finite-size and boundary effects), and allied with precise and thorough experimental studies on well-chosen real systems, has provided the firmest foundations for progress. Along such lines, perhaps, we may look forward to the writing of the next few chapters in the Story of Coulombic Criticality!

## APPENDIX. COMMENTS ON "CRITICAL BEHAVIOR OF IONIC FLUID MODELS" BY G. STELL(9)

Stell's article builds on and extends earlier work(7,8) including a study(81) applicable to long-range power-law potentials

$$\varphi(r) \sim 1/r^{d+\sigma} \qquad (\sigma > 0) \tag{A.1}$$

in $d$ dimensions. That study anticipated some of the results later established by RG methods,(49) in particular, that classical thermodynamic criticality arises when $d > 2\sigma (\leqslant 4)$. There are two principal claims: (a) it is shown that the restricted primitive model (RPM) electrolyte has Ising-like critical behavior; (b) in more realistic models, $1/r^4$ repulsive ion–ion terms (arising from solvent-averaged ion-dipole-ion interaction) will prevent such a critical point except that competing (attractive) terms that suppress its effect may also be present.

Now, as explained in Section 2, the conclusion advanced in (a) has a good chance of being correct: the issue is thus, "How convincing are the arguments actually presented?" (Note that the same conclusion was stated in ref. 7, but there the "published remarks were brief" and "gave few technical details."(9)) For the RPM, Stell decomposes the net ion–ion correlation functions $h_{ij}(r) \equiv g_{ij}(r) - 1$ (where $i, j = +, -$) into density–density and charge–charge correlation functions

$$h_S = \tfrac{1}{2}(h_{++} + h_{+-}) \qquad \text{and} \qquad h_D = \tfrac{1}{2}(h_{++} - h_{+-}) \tag{A.2}$$

respectively. Symmetry dictates $h_{++} = h_{--}$ and $h_{+-} = h_{-+}$.

The strategy then is to argue that $h_S(r)$ and the associated direct correlation function $c_S(r)$ obey essentially the same Ornstein–Zernike equation and same closure equation, namely,

$$c(r) - \beta\varphi(r) = R[h(\cdot); \rho; r] \tag{A.3}$$

in the critical region as does a standard, single-component fluid with net pair correlation function $h(r)$ and direct correlation function $c(r)$. It is presumed that these equations suffice to determine the universal critical exponents (although, in contrast to RG methods, no scheme for actually computing their values is presented).

The correlation functions $h(r)$, etc., are decomposed into (a) leading scaling parts of the form $H(\kappa r)/r^{1+\eta}$, etc., where $\xi \equiv 1/\kappa$ is the bulk density–density correlation length whose divergence signals the critical point, plus (b) remainders whose singularities are presumed not to play a crucial role. A lemma [ref. 9, Eq. (11)], based on a graphical, cluster-sum representation of the functional $R[h(\cdot)]$ then asserts that, on dropping the nonscaling/nonleading remainders, the scaling pieces solve the same functional equations. Hence the same exponents are expected.

A crucial ingredient in the application of this strategy, however, is the *assumption*, not justified in ref. 9, beyond bare assertion, that the charge correlation function $h_D(r)$ decays as $e^{-\Gamma r}$ where the inverse charge correlation length $\Gamma$ does *not vanish at the critical point*. [See also remarks after Eq. (6.18) in ref. 7. At low densities, of course, $\Gamma$ is just the reciprocal of the expected Debye screening length (which diverges when $\rho \to 0$). However, Stell's expectation, not independently supported, that "$\Gamma$ is only weakly sensitive to the singularity in density fluctuation associated with $\kappa \to 0$"(9) seems tantamount to assuming the desired result, namely, that only the density fluctuations go critical. Then, since the density is a simple scalar field, standard, Ising-like behavior certainly becomes very plausible.

Of course, one might well suspect that the statement: "In particular $\Gamma \neq 0$ at the critical point we are considering"(9) is correct, perhaps(82) because of the strong $+/-$ symmetry; if so, it is surely significant. But, against this, one must recognize that at most well-understood critical points (and their associated RG fixed points) a principle of "infection by the critical singularities" applies to *all* thermodynamic functions, to transport coefficients, and to the underlying correlation functions. Typically, one finds that all correlation functions display slow, power-law decays even though certain combinations (such as, e.g., the charge–charge correlations) may decay more rapidly than others and hence may control only subdominant singular behavior. Thus, in binary fluid-phase separation, the overall density and compressibility display singularities although the order parameter is the concentration difference. Furthermore, as Stell says after Eq. (5) of ref. 9, the internal energy $U(\rho, T)$ is/can be wholly determined from $h_D(r)$, the charge correlation function. A temperature derivative of $U$ yields the specific heat which (at $\rho \equiv \rho_c$) diverges as $(T-T_c)^{-\alpha}$ with $\alpha \simeq 0.11$; but the specific heat can be expressed as an integral over the four-point ionic/charge correlation function, which must therefore certainly have some components decaying *more slowly* than any exponential. At very least, therefore, the assertion $\Gamma_c > 0$ seems to require rather careful and detailed justification.

If, contrary to the expectations, $\Gamma$ does vanish at the critical point itself, Stell's line of argument could survive if the charge correlation function $h_D(r)$ decayed sufficiently rapidly, specifically, faster than $1/r^{1+\eta}$ $(d=3)$, which characterizes the density correlation function $h_S(r)$. However, the behavior of the charge correlation function is restricted by the Stillinger–Lovett sum rules,(57) which should be expected to hold everywhere in the single-phase regions away from criticality and, by continuity, even at the critical point itself.(82) These rules imply that both $\int h_D(r)\, d\mathbf{r}$ and $\int r^2 h_D(r)\, d\mathbf{r}$ remain finite at criticality (with, in fact, well-defined negative values(57)). Now if $h_D(r)$ decays *monotonically* for large $r$, this would mean that $h_D(r)$ must decay more rapidly than $1/r^5$. In other words, if we write

$$|h_D(r)| \sim 1/r^{1+\eta_q} \tag{A.4}$$

then we could conclude $\eta_q \geqslant 4 \gg \eta$, which would satisfy Stell's desiderata sufficiently well. Unfortunately, however, it is possible, if not in fact *likely*,(22,34,59) that $h_D(r)$ displays an *oscillatory* decay as $r \to \infty$. If, for example, one thus supposes

$$h_D(r) = J(\sin k_0 r)/r^{1+\eta_q} \qquad (d=3) \tag{A.5}$$

with only the restriction $0 < \eta_q < 3$, it is not hard to verify that both the Stillinger–Lovett sum rules can be satisfied however small $\eta_q$. [In fact one requires, not surprisingly, $k_0 \simeq \kappa(T_c, \rho_c)$, up to factors of order unity, since the second sum rule entails the inverse Debye length $\kappa$.] This example hence demonstrates that the sum rules do *not* provide sufficient power to ensure that $h_D(r)$ decays harmlessly fast at criticality. If one had $\eta_q \leqslant \eta$, then Stell's arguments would certainly fail. But even if one had instead $\eta_q > \eta$, the presence of a slow, oscillatory critical-point decay like (A.5) would suggest quite strongly that the corresponding renormalization-group fixed point was not of standard Ising character (although it is possible, in principle, that $\eta_q$ could be identified with some known, marginal or irrelevant Ising critical operator).

Other queries about Stell's discussion concern the cluster-sum representations of the functional $R[h(\cdot)]$, which are essentially perturbative expansions entailing powers of the density $\rho$. As such they are likely to fail at the critical point or, more concretely, when $r/\xi$ is small (even if $\rho_c$ is small); but Stell seemingly assumes some validity remains in the scaling region: see ref. 9, Eq. (8). At best this is hard to assess.

To explain a related issue, recall that the presence of the long-range

Coulombic forces causes the usual density (or virial) expansions to fail even at low densities since (for $d = 3$) the Debye–Hückel limiting laws yield a $\rho^{3/2}$ power. Of course, this singular behavior can be accommodated by a resummation of graphical contributions, using screened $h$-bonds, etc.(29, 30) Nevertheless, the differences in behavior from a simple single-component fluid are drastic. Thus, while the comparison of perturbative functional representations might be quite satisfying if the ions interacted only via short-range couplings, it seems much less convincing when Coulombic forces are present. Furthermore, in the RPM, these are the only forces acting, so they alone must generate the effective density–density attractions which lead to phase separation, to criticality, and to large-scale fluctuations!

If criticality in the RPM is truly Ising-like, the size of the critical region is still an open question which is not addressed by Stell here. As regards the appearance of classical coexistence curves in real electrolytes, however, he suggests [claim (b) above] that $1/r^4$ effective ion–ion potentials might play a role. Certainly, attractive forces of this character should, since they correspond to $\sigma = 1$ and $d = 3$ in (A.1), give rise to classical critical behavior for the thermodynamic quantities. It is important to note, however, that the scattering function at criticality should then vary, with wavenumber $k = |\mathbf{k}|$ as $S_c(k) \sim 1/k^{2-\eta}$ with $\eta = 1$.(49) This contrasts with the value $\eta = 0$ which characterizes standard classical or van der Waals behavior (and with $\eta \simeq 0.03_6$ for Ising-like systems). That in turn implies, in consonance with the scaling relation $(2 - \eta)\,\nu = \gamma$, that a correlation length exponent $\nu = 1$ should be seen (rather than the standard classical value $\nu = 1/2$). Significantly, that prediction does not accord with the recent scattering experiments on the Coulombic systems.(16, 17) This scenario does not, therefore, appear realistic (although perhaps the various data sets should be reexamined to check if they could possibly be consistent with $\eta = 1$).

On the other hand, Stell observes that in certain models (with dielectric "cavities" rather than standard hard spheres) an effective repulsive $1/r^4$ ion–ion potential may appear. If that is not overwhelmed by similar attractive terms, it would seem to suppress $\mathbf{k} = 0$ critical fluctuations and, instead, lead to a charge-density-wave instability, reminiscent of Nabutovskii *et al.*'s scenario: see Figs. 3 and 4. No experimental evidence for such behavior is known.

Theoretically, either sign of a $1/r^4$ effective ion–ion potential seems questionable. While undoubtedly such a form may appear in a zero-density limit, it seems likely (depending on the model details, which, physically, should not be too "rigid") that the presence of free ions at nonzero density would always screen this effective potential in the same way that the bare $1/r$ electrostatic potential is screened by the Debye–Hückel factor $e^{-\Gamma r}$

Certainly this issue would need to be addressed before such $1/r^4$ terms or, in Fourier space, $|\mathbf{k}|$ terms were seriously invoked. Indeed, recent calculations(83) of the interaction between two charged dielectric spheres in an ionic medium indicate that the "bare" $1/r^4$ term is strongly screened, leaving only a contribution $e^{-2\Gamma r}/r^2$.

## ACKNOWLEDGMENTS

Dr. J. M. H. Levelt Sengers embroiled the author in questions which led directly to this review and to the work described here. She and Prof. M. M. Telo da Gama helped significantly in the navigation of the vast sea of literature concerning the theory of electrolytes. Particular thanks are due to Dr. Yan Levin, who has collaborated closely in discussing Coulombic criticality and whose joint work is described in Sections 6–10. Correspondence and interactions with Prof. H. L. Friedman, Dr. J. A. Given, Prof. K. S. Pitzer, Prof. G. Stell, Prof. J. P. Valleau, and Dr. H. Weingärtner have been much appreciated. The support for the work described here by the National Science Foundation (under Grants DMR 90-07811 and CHE 93-1729) is gratefully acknowledged. The Aspen Center for Physics provided hospitality in July 1993 when this article was written.

## REFERENCES

1. P. Chieux and M. J. Sienko, *J. Chem. Phys.* **53**:566 (1970).
2. R. R. Singh and K. S. Pitzer, *J. Am. Chem. Soc.* **110**:8723 (1988).
3. R. R. Singh and K. S. Pitzer, *J. Chem. Phys.* **92**:6775 (1990).
4. M. L. Japas and J. M. H. Levelt Sengers, *J. Phys. Chem.* **94**:5361 (1990).
5. K. S. Pitzer, *Acc. Chem. Res.* **23**:333 (1990).
6. J. M. H. Levelt Sengers and J. A. Given, *Mol. Phys.* **80**:899 (1993).
7. B. Hafskjold and G. Stell, in *The Liquid State of Matter*, E. W. Montroll and J. L. Lebowitz, eds. (North-Holland, Amsterdam, 1982), p. 175, esp. Section 6.4.
8. J. S. Høye and G. Stell, *J. Phys. Chem.* **94**:7899 (1990).
9. G. Stell, *Phys. Rev. A* **45**:7628 (1992).
10. J.-H. Chen, M. E. Fisher, and B. G. Nickel, *Phys. Rev. Lett.* **48**:630 (1982).
11. P. Chieux, J.-F. Lal, L. Hily, J. Dupuy, F. Leclercq, and P. Damay, *J. Phys.* (Paris) *Ser. IV* **1991**:C5-3; P. Chieux, *J. Phys.* (Paris) *Ser. IV* **1991**:C5-373.
12. C. Caccamo and A. Giacoppo, *Phys. Rev. A* **42**:6285 (1990).
13. A. Kholodenko and A. L. Beyerlein, *J. Chem. Phys.* **93**:8405 (1990).
14. A. Kholodenko and A. L. Beyerlein, *Phys. Lett. A* **132**:347 (1988).
15. M. E. Fisher, *J. Chem. Phys.* **96**:3352 (1992).
16. H. Weingärtner, S. Wiegand, and W. Schröer, *J. Chem. Phys.* **96**:848 (1992).
17. K. C. Zhang, M. E. Briggs, R. W. Gammon, and J. M. H. Levelt Sengers, *J. Chem. Phys.* **97**:8692 (1992).
18. A. Kholodenko and A. L. Beyerlein, *Phys. Lett. A* **175**:366 (1993).
19. M. E. Fisher, in *Magnetism and Magnetic Materials 1974*, AIP Conf. Proc. 24 (American

Institute for Physics, New York, 1975), p. 273; M. E. Fisher and S. Sarbach, *Phys. Rev. Lett.* **41**:1127 (1978); S. Sarbach and M. E. Fisher, *J. Appl. Phys.* **49**:1350 (1978); **50**:1802 (1979).

20. C. J. Thompson, *Mathematical Statistical Mechanics* (Macmillan, New York, 1972), pp. 94–95, 104–105, Appendix C.
21. H. Weingärtner, T. Merkel, U. Maurer, J.-P. Conzen, H. Glasbrenner, and S. Käshammer, *Ber. Buns. Phys. Chem.* **95**:1579 (1991).
22. H. L. Friedman and B. Larsen, *J. Chem. Phys.* **70**:92 (1979).
23. H. Xu, H. L. Friedman, and F. O. Raineri, *J. Solution Chem.* **20**:739 (1991).
24. P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, New York, 1979), Sect. IV.3 *et seq.*
25. J. P. Valleau, *J. Chem. Phys.* **95**:584 (1991).
26. A. Z. Panagiotopoulos, *Fluid Phase Equilibria* **76**:97 (1992).
27. J. F. Douglas, private communication cited in ref. 6.
28. P. W. Debye and E. Hückel, *Physik. Z.* **24**:185 (1923).
29. D. A. McQuarrie, *Statistical Mechanics* (Harper and Row, New York, 1976), Chapter 15.
30. H. Falkenhagen and W. Ebeling, in *Ionic Interactions*, Vol. 1, S. Petrucci, ed. (Academic Press, New York, 1971), but note slips in Eqs. (49) and (51).
31. M. E. Fisher and M. M. Telo da Gama, unpublished.
32. A. J. Liu and M. E. Fisher, *Physica A* **156**:35 (1989).
33. K. Binder, in *Phase Transitions and Critical Phenomena*, Vol. 5b, C. Domb and M. S. Green, eds. (Academic Press, London, 1976).
34. V. M. Nabutovskii, N. A. Nemov, and Yu. G. Peisakhovich, *Sov. Phys. JETP* **52**:111 (1980) [*Zh. Eksp. Teor. Fiz.* **79**:2196 (1980)]; *Phys. Lett.* **79A**:98 (1980); *Mol. Phys.* **54**:979 (1985).
35. R. M. Hornreich, M. Luban, and S. Shtrikman, *Phys. Rev. Lett.* **35**:1678 (1975).
36. S. A. Brazovskii, *Sov. Phys. JETP* **41**:85 (1975) [*Zh. Eksp. Teor. Fiz.* **68**:175 (1975)].
37. M. E. Fisher, *J. Appl. Phys.* **52**:2014 (1981).
38. M. E. Fisher and W. Selke, *Phys. Rev. Lett.* **44**:1502 (1980).
39. W. Selke and M. E. Fisher, *Physica* **15–18**:403 (1982); *Phys. Rev. B* **20**:257 (1979).
40. M. E. Fisher, *Phys. Rev. Lett.* **34**:1634 (1975).
41. A. A. Migdal, *Sov. Phys. JETP* **42**:743 (1976) [*Zh. Eksp. Teor. Fiz.* **69**:1457 (1975)].
42. L. P. Kadanoff, *Ann. Phys.* (N.Y.) **100**:359 (1976).
43. Th. Niemeijer and J. M. J. van Leeuwen, in *Phase Transitions and Critical Phenomena*, Vol. 6, C. Domb and M. S. Green, eds. (Academic Press, London, 1976), p. 425.
44. H. E. Stanley, P. J. Reynolds, S. Redner, and F. Family, in *Real Space Renormalization Groups*, T. W. Burkhardt and J. M. J. van Leeuwen, eds. (Springer-Verlag, Berlin, 1982).
45. M. Kaufman, R. B. Griffiths, J. M. Yeomans, and M. E. Fisher, *Phys. Rev. B* **23**:3448 (1981); J. M. Yeomans and M. E. Fisher, *Phys. Rev. B* **24**:2825 (1981).
46. S. F. Edwards, *Phil. Mag.* **4**:1171 (1959).
47. J. Glimm and A. Jaffe, *Quantum Physics: A Functional Integral Point of View*, 2nd ed. (Springer-Verlag, Berlin, 1987), Section 23.6.
48. D. Brydges, *Commun. Math. Phys.* **58**:313 (1978); D. Brydges and P. Federbush, *Commun. Math. Phys.* **73**:197 (1980).
49. M. E. Fisher, S.-K. Ma, and B. G. Nickel, *Phys. Rev. Lett.* **29**:917 (1972).
50. V. Degiorgio, R. Piazza, M. Corti, and C. Minero, *J. Chem. Phys.* **82**:1025 (1985); M. Corti and V. Degiorgio, *Phys. Rev. Lett.* **55**:2005 (1985).
51. G. Dietler and D. S. Cannell, *Phys. Rev. Lett.* **60**:1852 (1988).
52. M. E. Fisher, *Phys. Rev. Lett.* **57**:1911 (1986).
53. C. Bagnuls and C. Bervillier, *Phys. Rev. Lett.* **58**:435 (1987).

54. M. E. Fisher and M. C. Barbosa, *Phys. Rev. B* **43**:11117 (1991); M. C. Barbosa and M. E. Fisher, *Phys. Rev. B* **43**:10635 (1991).
55. P. T. Cummings and G. Stell, *J. Chem. Phys.* **78**:1917 (1983).
56. E. R. Smith, *J. Stat. Phys.* **50**:813 (1988).
57. F. H. Stillinger, Jr., and R. Lovett, *J. Chem. Phys.* **48**:3858 (1968).
58. D. A. McQuarrie, *J. Phys. Chem.* **66**:1508 (1962).
59. G. R. Stell, K. C. Wu, and B. Larsen, *Phys. Rev. Lett.* **37**:1369 (1976).
60. M. E. Fisher and Y. Levin, *Phys. Rev. Lett.* **71**:3826 (1993); Y. Levin and M. E. Fisher, to be published.
61. P. N. Vorontsov-Veliaminov, A. M. El'yashevich, L. A. Morgenshtern, and V. P. Chasovshikh, *Teplofiz. Vys. Temp.* **8**:277 (1970) [*High Temp.* (USSR) **8**:261 (1970)].
62. V. P. Chasovshikh, P. N. Vorontsov-Veliaminov, and A. M. El'yashevich, *Dokl. Akad. Nauk. Tadzhiksko SSR* **16**(10):23 (1973).
63. V. P. Chasovskikh and P. N. Vorontsov-Veliaminov, *Teplofiz. Vys. Temp.* **14**:199 (1976) [*High Temp.* (USSR) **14**:174 (1976)].
64. M. J. Gillan, *Mol. Phys.* **49**:421 (1983).
65. K. S. Pitzer and D. R. Schreiber, *Mol. Phys.* **60**:1067 (1987).
66. N. Bjerrum, *Kgl. Danske Vidensk. Selsk. Mat.-Fys. Medd.* 7:1 (1926).
67. B. Larsen, *Chem. Phys. Lett.* **27**:47 (1974); *J. Chem. Phys.* **65**:3431 (1976).
68. M. E. Fisher, in *Critical Phenomena*, Proc. 1970 E. Fermi, *Int. Sch. Phys. Varenna*, M. S. Green, ed. (Academic Press, New York, 1971), p. 1.
69. V. Privman, ed., *Finite Size Scaling and Numerical Simulation of Statistical Systems* (World Scientific, Singapore, 1990); J. L. Cardy, ed. *Finite-Size Scaling* (North-Holland, Amsterdam, 1988).
70. V. McGahay and M. Tomozawa, *J. NonCryst. Solids* **109**:27 (1989); *J. Chem. Phys.* **97**:2609 (1992).
71. J. S. Høye, E. Lomba, and G. Stell, *Mol. Phys.* **75**:1217 (1992).
72. L. Belloni, *J. Chem. Phys.* **98**:8080 (1993).
73. M. E. Fisher and S. Fishman, *Phys. Rev. Lett.* **47**:421 (1981), and references therein.
74. W. Ebeling, *Z. Phys. Chem.* **238**:400 (1968).
75. W. Ebeling and M. Grigo, *Ann. Phys.* (Leipzig) **37**:21 (1980).
76. H. Yokoyama and H. Yamatera, *Bull. Chem. Soc. Japan* **48**:1770 (1975); *Chem. Lett.* **1973**:337.
77. M.-C. Justice and J.-C. Justice, *J. Solution Chem.* **5**:543 (1976); **6**:819 (1977).
78. R. M. Fuoss, *J. Am. Chem. Soc.* **80**:5059 (1958).
79. H. L. Friedman, *J. Phys. Chem.* **66**:1595 (1962).
80. M. E. Fisher, in *Critical Phenomena*, F. J. W. Hahne, ed. (Springer-Verlag, Berlin, 1983), p. 1.
81. G. Stell, *Phys. Rev. B* **1**:2265 (1970), Section 4.
82. G. Stell, private communication.
83. Xiaojun Li, Y. Levin, and M. E. Fisher, submitted for publication; reported at the 2nd Liquid Matter Conference, Florence, Italy, September 1993.