Structure and thermodynamics of a two-dimensional Coulomb fluid in the strong association regime

E. Lomba
Instituto de Química Física Rocasolano, CSIC, Serrano 119, E-28006, Madrid, Spain

J. J. Weis
Laboratoire de Physique Théorique, UMR8627, Université de Paris XI, Bâtiment 210, 91405 Orsay Cedex, France

F. Lado
Department of Physics, North Carolina State University, Raleigh, North Carolina 27695-8202, USA

(Received 30 May 2007; accepted 27 June 2007; published online 16 August 2007)

The behavior of a two-dimensional neutral Coulomb fluid in the strong association regime (low density, high ionic charge) is explored by means of computer simulation and the hypernetted chain integral equation. The theory reproduces reasonably well the structure and thermodynamics of the system but presents a no-solution region at temperatures well above the computer simulation estimates of the metal-insulator transition. In contrast with hypernetted chain predictions for the three-dimensional Coulomb fluid, here the breakdown of the solution is not accompanied by divergences in any physical quantity. © 2007 American Institute of Physics.

[DOI: 10.1063/1.2764478]

I. INTRODUCTION

The study of a two-dimensional (2D), neutral system of charged particles interacting via a logarithmic potential (instead of the well-known $1/r$ dependence which one finds in 3D charged systems) might seem at first to be a rather academic problem with few implications for real physical systems. But, it turns out that this 2D Coulomb fluid is a significant physical model for systems which can be considered two-dimensional and whose relevant thermal excitations are vortices, such as thin-film superconductors/superfluids.$^1$ These systems are known to undergo a transition of the Berezinskii-Kosterlitz-Thouless (BKT) type$^{2-5}$ in which the vortex-antivortex pairs in the superconducting/superfluid phase unbind, giving rise to a “normal” phase, characterized by the presence of isolated vortices that coexist with vortex-antivortex pairs. In the language of the Coulomb fluid, this is nothing but a metal-insulator transition. The conducting phase is formed by isolated ions coexisting with neutral clusters, while the insulating phase corresponds to a system of neutral clusters alone. Various computer simulation studies have provided evidence that this BKT continuous transition transforms into a first-order transition, both in the case of the 2D lattice Coulomb fluid$^6$ and in continuous models,$^7-9$ in accordance with Minnhagen and Wallin’s renormalization group predictions.$^{10}$ In addition to these and other numerical and theoretical studies related to the critical behavior and the existence of various exact sum rules in 2D Coulomb fluids$^{11-16}$ (in most cases for pointlike particles), results from classical integral equation theories such as the hypernetted chain (HNC) equation have also been obtained$^{17-19}$ for 2D charged hard disks. In Ref. 18 it was shown that the HNC approximation furnishes a reasonably good description of the 2D Coulomb fluid away from the phase transition region. In this work, we aim at providing a more thorough analysis of the HNC performance (in terms of both structural and thermodynamic quantities) in the strong association regime (low density, high ionic charge) which lies in the neighborhood of the phase transition region, together with detailed quantitative structural information obtained from Monte Carlo computer simulations. A map of the no-solution region of the HNC is obtained and compared with the phase diagram obtained from computer simulations.$^8,9$ We will see that the nature of the breakdown of the HNC solution in the 2D Coulomb fluid differs from that found in association with phase transitions in 3D fluids,$^{20}$ especially in the restricted primitive model (RPM) of electrolytes.$^{21,22}$

The rest of the paper is organized as follows. The model and a brief account of the HNC equation for two-dimensional charged systems in presented in the next section and details of the simulations are summarized in Sec. III. Finally, in Sec. IV we present our most significant results.

II. THE MODEL. THE HNC EQUATION IN THE TWO-DIMENSIONAL COULOMB FLUID

According to the solution of Poisson’s equation for point charges in two dimensions (more accurately, a cross-sectional view of infinite parallel lines carrying uniform charge distributions), the interaction between two charged hard disks of diameters $\sigma$ in vacuum is given simply by

$$
\beta \phi_i(r) = \begin{cases} 
- \frac{1}{2} \sum_i \ln(r/\sigma) & \text{if } r \geq \sigma \\
\frac{1}{2} \sum_i \ln(r/\sigma) & \text{if } r < \sigma
\end{cases}
$$

(1)

where $\beta = 1/k_BT$ (with $T$ the absolute temperature and $k_B$ Boltzmann’s constant), $Z_i$ is the ionic charge of disk $i$ in electron units, and the coupling constant is $\Gamma = 2\beta e^2/\sigma$, with $e$ the electron charge in Gaussian units. Once the interaction
is given, the fluid structure can be obtained from the Ornstein-Zernike equation, which for a mixture reads

\[ h_{ij}(r_{12}) = \sum_l \rho_l \int d\mathbf{r} c_{j\beta}(r_{13}) h_{ik}(r_{32}), \quad (2) \]

where \( h \) and \( c \) are the total and direct correlation functions, respectively, and \( \rho_l \) is the number density of species \( l \); the pair distribution functions are \( g_{ij} = h_{ij} + 1 \). In this paper, we will deal with a fully symmetric system whereby \( \rho_+ = \rho_- \) and \( Z_+ = -Z_- \); we define \( \rho = \rho_+ + \rho_- \), and we will take \( |Z_{i+}| = 1 \). In Fourier space Eq. (2) can be cast into matrix form to yield

\[ \tilde{\mathbf{C}}(k) = [\mathbf{I} - \tilde{\mathbf{C}}(k)]^{-1} \tilde{\mathbf{C}}(k) \tilde{\mathbf{C}}(k), \quad (3) \]

where \( \mathbf{I} \) is the identity matrix, the tilde denotes a 2D Fourier transformation, and

\[ [\tilde{\mathbf{C}}(k)]_{ij} = \sqrt{\rho_+ \rho_-} \tilde{\gamma}_{ij}(k), \quad (4) \]

\[ [\tilde{\mathbf{C}}(k)]_{ij} = \sqrt{\rho_+ \rho_-} \tilde{c}_{ij}(k), \quad (5) \]

with \( \gamma_{ij} = h_{ij} - c_{ij} \). The HNC closure reads

\[ c_{ij}(r) = \exp[-\beta \phi_{ij}(r) + \gamma_{ij}(r)] - 1 - \gamma_{ij}(r). \quad (6) \]

Now Eqs. (3) and (6) can be solved numerically by iteration only if one appropriately handles the long-range behavior of the direct correlation function (and that of the total correlation function under weak screening conditions). In principle one might rewrite the equations in terms of short-range functions following the standard renormalization procedure\(^\text{23,24}\) as applied to this system in Ref. 18. However, since our calculations will be carried out for relatively high ionic couplings, the screening will be large enough for the total correlation functions to be short ranged and we will make use of the somewhat simpler formulation sketched below.

For practical purposes then, we construct a long-range form of the interaction potential and its transform,\(^\text{25}\)

\[ \beta \phi_{ij}^{LR}(r) = -Z_{ij} \Gamma \left[ \ln \left( \frac{r}{\sigma} \right) + 1 + \frac{2}{\rho} E_1 \left( \frac{2}{\rho} \right) \right], \quad (7) \]

\[ \beta \phi_{ij}^{LR}(k) = Z_{ij} \Gamma \frac{2\pi}{(k\sigma)^2} \exp \left( -\frac{1}{2} \frac{(k\sigma)^2}{\sigma} \right), \quad (8) \]

where \( E_1(x) \) is the exponential integral.\(^\text{26}\) These can be used to define a new set of short-ranged correlations and interaction of the form

\[ \beta \phi_{ij}^{SR}(r) = \beta \phi_{ij}(r) - \beta \phi_{ij}^{LR}(r), \quad (9) \]

\[ c_{ij}^{SR}(r) = c_{ij}(r) + \beta \phi_{ij}^{LR}(r), \quad (10) \]

\[ \gamma_{ij}^{SR}(r) = \gamma_{ij}(r) - \beta \phi_{ij}^{LR}(r), \quad (11) \]

and similarly for their transforms.

Then, Eq. (6) can be rewritten as

\[ c_{ij}^{SR}(r) = \exp[-\beta \phi_{ij}^{SR}(r) + \gamma_{ij}^{SR}(r)] - 1 - \gamma_{ij}^{SR}(r). \quad (12) \]

Now, once an initial estimate of \( \gamma_{ij}^{SR} = h_{ij} - c_{ij}^{SR} \) is made, a new iterate of \( c_{ij}^{SR} \) can be calculated from Eq. (12) and this quantity numerically Fourier-transformed without long-range issues using

\[ \gamma_{ij}^{SR}(k) = 2\pi \int_0^\infty dr r \gamma_{ij}^{SR}(r) J_0(kr), \quad (13) \]

where \( J_0(r) \) is the zero-order Bessel function. One can recover the complete function \( \tilde{\gamma}_{ij}(k) \) by subtracting the analytical transform \( \beta \phi_{ij}^{LR}(k) \) and from Eq. (3) get a new estimate of \( \tilde{\gamma}_{ij}(k) \). The short-range part \( \gamma_{ij}^{SR} = \gamma_{ij} - \beta \phi_{ij}^{LR} \) is then numerically Fourier-inverted to give a new

\[ \gamma_{ij}^{SR}(r) = \frac{1}{2\pi} \int_0^\infty dk k \gamma_{ij}^{SR}(k) J_0(kr). \quad (14) \]

The orthogonality of the numerical Fourier transforms Eqs. (13) and (14) is guaranteed using Lado’s choice for the real and reciprocal grid spacing.\(^\text{27}\) The iteration process can be accelerated using Ng’s scheme.\(^\text{28}\)

Once the correlation functions are known, thermodynamic properties are readily obtained. The pressure is calculated through

\[ \frac{\beta P}{\rho} = 1 + \frac{1}{4} \pi \rho \sigma^2 [g_{++}(\sigma^+) + g_{+-}(\sigma^-)] - \frac{1}{4} \Gamma. \quad (15) \]

(Note that for point charges one obtains the exact sum rule\(^\text{29,30}\) \( \beta P/\rho = 1 - \Gamma/4 \).) The excess internal energy reads

\[ \frac{\beta U^{ex}}{N} = -\frac{1}{2} \pi \rho \Gamma \int_0^\infty dr g_{++}(r) - g_{+-}(r) \ln \left( \frac{r}{\sigma} \right). \quad (16) \]

In Eqs. (15) and (16) the charge symmetry of the model and values \( |Z_i| = 1 \) are implicit. The isothermal compressibility \( \kappa_T \) by

\[ \rho k_B T \kappa_T = 1 + \frac{1}{\rho} \sum_{\rho_{ij}} \rho_i \rho_j \tilde{h}_{ij}(0). \quad (17) \]

Note that \( U^{ex} \) can be shifted up or down according to the normalization length one chooses for the logarithmic potential. If the energy in Eq. (16) with normalization length \( \sigma \) is denoted \( U^{ex}(\sigma) \), then the energy \( U^{ex}(L) \) for an arbitrary normalization length \( L \) is given by

\[ \frac{\beta U^{ex}(L)}{N} = \frac{\beta U^{ex}(\sigma)}{N} - \frac{1}{2} \ln \left( \frac{L}{\sigma} \right). \quad (18) \]

Additionally, it is possible to calculate the excess specific heat at constant volume as

\[ C_v^{ex} = \frac{\pi}{\rho} \sum_{\rho_{ij}} \rho_i \rho_j \int_0^\infty dr \beta \phi_{ij}(r) g_{ij}(r) \times \left[ \beta \phi_{ij}(r) - \beta (\partial \gamma_{ij}(r)/\partial \beta)_{\rho} \right]. \quad (19) \]

where

\[ C_v^{ex} = \left( \frac{\partial U^{ex}/Nk_B}{\partial T} \right)_v. \quad (20) \]

If we define \( f' = \beta \partial f/\partial \beta \), then \( \beta (\partial \gamma_{ij}/\partial \beta)_{\rho} \) can be calculated from the integral equation that results from differentiation of Eqs. (2) and (6), i.e.,
In Eq. (23), $r_{ij} = r_i - r_j$, $E_i(x)$ is again the exponential integral, $y = 0.5772156...$ is Euler’s constant, and $N$ is the number of particles. The prime in the sum over $n = (n_x, n_y)$, with $n_x, n_y$ integers, restricts it to $i \neq j$ for $n = 0$. The dimensionless parameter $\alpha$ controls the relative contributions to the Ewald sum of the direct and reciprocal space terms. With the choice $\alpha = 6$, adopted in our calculations, only terms with $n = 0$ need to be retained in the first sum of Eq. (23). The sum in reciprocal space extends over all lattice vectors $k = 2\pi n/L$ with $|n|^2 \leq 64$. Note that the last term in Eq. (23) arises from the fact that distances in Eq. (1) are scaled by $\sigma$ and not $L$, as seen in Eq. (18). Thermodynamic and structural properties were obtained by averaging over $5 \times 10^6$ to $10 \times 10^6$ trial moves per particle after equilibration of the system.

IV. RESULTS

The HNC equation is solved using 1000 grid points and a choice of grid spacing in $r$ such that the 50th grid point exactly corresponds to $r = \sigma$ (note that Lado’s method$^{27}$ in 2D leads to unequally spaced grid points). In these conditions, the maximum range in the correlation functions is approximately $20\sigma$, which has proven to be adequate for the state points here considered.

A first question to be addressed refers to the existence of a region of no solution for the HNC equation in the case of the 2D Coulomb fluid, and if it exists, how far from the phase transition region predicted by simulation does it lie? In Fig. 1 we have plotted the boundaries of the solution region of the HNC equation together with the BKT transition line simulation estimates from Orkoulas and Panagiotopoulos,$^8$ the vapor-liquid (VL) equilibrium curve as obtained from the Monte Carlo simulations of the same authors,$^8$ and the critical end point estimate of Lidmar and Wallin.$^9$ At first sight, one can appreciate that the no-solution curve of the HNC is more or less parallel to the envelope of the BKT transition line and the VL curve, whose transition temperatures are overestimated in the HNC approximation by a factor of 2 (when compared to simulation estimates) for densities below $\rho \sigma^3 = 0.2$. The agreement improves for higher densities. In three dimensions the HNC overestimates the critical temperature (inverse coupling) to a larger degree$^{33}$ ($\Gamma^{-1} = 0.12$, $\rho_{HNC} \sigma^3 = 0.007$, versus the Monte Carlo estimates$^{31} \Gamma^{-1} = 0.0487$, $\rho \sigma^3 = 0.065$). In the inset of Fig. 1 one can appreciate the low density behavior of the HNC no-solution boundary and that of simulation estimates of the BKT transition in logarithmic scale (note that the latter and the zero density limit of the BKT transition$^{4,5,35}$ ($\Gamma^{-1} = 1/4$) have been connected by a dotted line for clarity). Interestingly, one observes a maximum in the HNC solution boundary curve, absent in the BKT transition line, but at low density (and within numerical accuracy), the boundary curve seems to converge toward $\rho \rightarrow 0$ and $\Gamma^{-1} \rightarrow 1/4$. From the analysis of Fig. 1 one might then infer that the HNC equation does indeed provide evidence on the existence of a phase transition, which might be of the BKT type. We can further investigate the nature of the transition if we analyze the behavior of the isothermal compressibility and heat capacity when approaching the breakdown of the numerical solution. This is particularly illustrative if the analysis is carried out in the light of the results for the 3D RPM model. To that purpose we have solved the HNC equation for this model of electrolyte along a supercritical ($\Gamma = 7.3$) and a subcritical ($\Gamma = 9.0$) isotherm. Results for the isothermal compressibility and excess heat capacity for this three-dimensional system are presented in Fig. 2. At first sight one observes that the isothermal compressibility before the breakdown of the solution exhibits a rather vertical slope, in accordance with the presence of
square-root branch point signaling the onset of complex solutions.\textsuperscript{21,22} In contrast, the heat capacity presents a true divergence (within numerical accuracy) on the high density side of the no-solution boundary and a square-root branch point on the low density side. This is in perfect agreement with the behavior of the HNC for simple fluids, as found by Sarkisov and Lomba.\textsuperscript{20,36}

In Fig. 3 we observe the same quantities for the 2D Coulomb fluid for various isotherms above and below the pseudocritical temperature (the maximum of the no-solution boundary). Two features are readily seen in this figure. Neither the compressibility nor the excess heat capacity shows any sign of divergence or square-root branch points (end points with quasivertical slope). Additionally, the excess heat capacity exhibits a clear minimum on the low density side of the no-solution curve with negative values (note however that the net specific heat capacity $C_v^e = C_v^+ + 2/2$ remains positive, and thus thermodynamic stability is preserved). This latter feature is completely absent from the 3D RPM behavior. As to the lack of divergence and its presence in the three-dimensional model, we must take into account that the phase transition in the RPM is a standard liquid-vapor equilibrium (i.e., first order), by which one should expect both heat capacities and compressibilities to diverge at the spinodal boundary. In contrast, in the region $1/4 < \Gamma^{-1} < 0.056$ to 0.032 (these last values taken from Refs. 8 and 9), the 2D model presents a BKT transition, in which the heat capacity does not diverge\textsuperscript{2-5} (simply exhibits a cusp) and no divergence should be expected in the isothermal compressibilities. At lower temperatures (higher couplings) one should encounter a first-order transition (a condensation). The slope of the HNC compressibility at lower temperatures is appreciably steeper near the breakdown, but still far from a vertical slope or a divergence. It might well be that at much higher couplings one finds the signature of a first-order transition (either a square-root branch point or divergence in the compressibility, or zero pressure states), but unfortunately we have been unable to reach convergent solutions for couplings above $\Gamma = 15$.

Focusing on the fluid structure, in Figs. 4 and 5 we present the like and unlike distribution functions ($g_{++} = g_{--}$ and $g_{+-} = g_{-+}$) and the corresponding structure factors for moderate and low densities and two values of the coupling ($2.5$ and $5.0$), respectively above and below the zero density limit of the BKT transition coupling ($\Gamma_{BKT}(\rho = 0) = 4$). As found by Urbič and Vlachy,\textsuperscript{18} the HNC performs fairly well away from the transition temperature. The unlike pair distribution is relatively well reproduced even at low density, high coupling ($\rho \sigma^2 = 0.05$ and $\Gamma = 5.0$). In the region $2.5 \sigma < r$...
<4\sigma$, the simulated $g_{+−}$ exhibits a shoulder that reflects the presence of clustering. This shoulder is captured by the HNC solution, although somewhat underestimated. The largest discrepancy is found in $g_{++}$ for the same state point. In this case one sees that the height of the first maximum which reflects the formation of clusters of more than two particles is quite underestimated by the HNC equation. Moreover, one observes that the HNC $g_{++}$ for $\Gamma=2.5$ and $\Gamma=5.0$ are rather similar, somehow making evident the inability of the approximation to properly account for the increase of clustering induced by the decrease of density (see Fig. 6 and the discussion below). The structure factors depicted in Fig. 5 are defined by

$$S_{ij}(k) = x_i \delta_{ij} + x_i x_j \bar{r}_{ij}(k),$$

where $x_i$ is the mole fraction of the $i$th component. In Fig. 5, once again one observes the excellent performance of the HNC equation, except for the highest coupling at low $k$ values, especially in the case of $S_{++}$ (which was also the case of $g_{++}$). The much higher value of $S_{++}(0)$ exhibited by the MC results indicates that not only is the first maximum of $g_{++}$ not well reproduced by the HNC, but also this function is much longer ranged in the case of the simulation. The rise of the like structure factors might well point to a tendency of the sample to segregate into clusters of alternating charges. Additionally, the presence of long-ranged oscillations in $S_{−−}$ (well captured by the HNC equation) result from the dominant pairing of unlike charges.

A complementary view of the fluid structure is offered by the cluster size distribution. In order to define the clusters one might use a simple energetic criterion, as proposed by Hill, or a geometrical connectivity criterion based on proximity. The latter approach is simpler and is the one we will follow. Note that, whatever criterion is adopted, there will be a certain degree of arbitrariness as to the distance or energy threshold required to consider two particles as linked. Thus, here we have calculated the cluster distributions from the MC configurations using a proximity criterion with a connectivity cutoff $r_c=1.5\sigma$. This implies that two particles will be deemed linked whenever their mutual separation is less than $1.5\sigma$. A series of analyses was performed using different values of $r_c$; we found that $r_c=1.5\sigma$ is the smallest connectivity distance for which the cluster distribution is more or less converged except for the ratio of free ions. This latter quantity is obviously strongly dependent on the arbi-

---

FIG. 4. (Color online) Pair distribution functions for the 2D Coulomb fluid as calculated in the HNC approximation (lines) and by MC simulation (symbols). Solid curves and circles correspond to $\rho\sigma^2=0.05$ and dotted curves and squares to $\rho\sigma^2=0.15$. Results for $\Gamma=2.5$ are presented in the upper graph and those for $\Gamma=5.0$ in the lower.

FIG. 5. (Color online) Partial structure factors for the 2D Coulomb fluid as calculated in the HNC approximation (lines) and by MC simulation (symbols). Solid curves and circles correspond to $\Gamma=2.5$ and dotted curves and squares to $\Gamma=5.0$. MC results for $\rho\sigma^2=0.05$ are represented by filled symbols and for $\rho\sigma^2=0.15$ by hollow symbols (HNC results follow closely the MC data).
trary choice of \( r_c \), but, for the purpose of qualitative comparison between different thermodynamic states, this does not pose any inconvenience so long as the same cutoff is used in all instances.

Results for various representative states are presented in Fig. 6. One sees that even at \( \Gamma = 5 \) there is still a large number of unpaired ions, both for \( \rho \sigma^2 = 0.05 \) and 0.15. Clustering increases when the density is lowered, but the effect of augmenting the coupling strength (or lowering the temperature) is more dramatic. Thus, one sees that for the lowest density and highest coupling the number of free ions diminishes considerably and neutral clusters (especially ion pairs) become dominant. In contrast, at \( \rho \sigma^2 = 0.15 \), increasing the coupling hardly changes the cluster size distribution. As an illustration of the effect of increasing coupling strength in the low density case, we show in Fig. 7 two snapshots from MC simulations in which one can clearly see how free ions practically vanish for \( \Gamma = 10 \). Finally, we note in passing that our system is closely related to the 3D restricted primitive model confined to a plane, studied by one of us in collaboration with Levesque and Caillol.\(^{39}\) In this connection, one finds that the cluster size distributions of both systems are very similar. This indicates that the use of a \( 1/r \) or a \( \log r \) functional dependence in the Coulomb interaction is not the leading factor that determines the clustering topology. One can reasonably infer then that the charge neutrality condition plays the central role in this respect.

We can take now a look at the thermodynamics. The excess internal energy and heat capacity obtained from the HNC equation and our MC simulation are plotted in Fig. 8. One observes that the HNC excess internal energy grows rapidly as the density decreases and then, for \( \Gamma < 4 \), falls to zero at very low densities—too low to be seen on Fig. 8—on the left of the no-solution region (cf. Fig. 1). The overall agreement with the simulation data is acceptable, even if it clearly worsens as the density is lowered, which once again reflects the limited ability of the HNC to account for clustering. In this respect, one observes a certain saturation in the Monte Carlo internal energy values at very low densities.
(particularly apparent for $\Gamma=5$). Aside from possible size effects, this nonmonotonic behavior is most likely due to the leading role of dipole-dipole (dipole-quadrupole, quadrupole-quadrupole, ...) interactions at low densities and high ionic strengths, a situation in which unlike charges form long-lived neutral clusters. The bare logarithmic Coulomb interaction would thus be almost completely screened. This behavior is only reproduced by the HNC approximation at very low densities.

As to the heat capacity, the peculiar HNC low density behavior seen in Fig. 3 (the presence of negative excess heat capacities) could not be confirmed by the simulation since it appears at densities too low to obtain reliable estimates. Note that for high couplings the simulated heat capacity exhibits maxima. Nonetheless, these results are simply rough estimates and one cannot expect to reproduce the cusp in $C_v$ characteristic of BKT transitions.

In summary, we have presented a detailed analysis of the HNC behavior in the low density, high ionic coupling region of the 2D Coulomb gas which suggests that the equation somehow captures the presence of the BKT transition. The quality of the HNC results at moderate couplings is excellent and clearly worsens as density and temperature are lowered (i.e., as clustering effects become dominant). Once more, one sees the need to incorporate the effects of association into the integral equation, following the ideas of Kalyuzhnyi and coworkers for asymmetric electrolytes, but this is certainly a highly nontrivial problem in the symmetric case.

ACKNOWLEDGMENTS

E.L. acknowledges support from the Dirección General de Investigación Científica y Técnica under Grant FIS2004-02954-C03-01 and the Dirección General de Universidades e Investigación de la Comunidad de Madrid under Grant S0505/ESP/0299, Program MOSSNOHO-CM.