Fundamentals of Inhomogeneous Fluids

edited by

Douglas Henderson

Utah Supercomputing Institute/IBM Partnership
and University of Utah
Salt Lake City, Utah

Marcel Dekker, Inc. New York • Basel • Hong Kong

Library of Congress Cataloging-in-Publication Data

Fundamentals of inhomogeneous fluids / edited by Douglas Henderson.
p. cm.
Includes bibliographical references and index.
ISBN 0-8247-8711-0 (alk. paper)
1. Solution (Chemistry) 2. Surface chemistry. 3. Liquids.
I. Henderson, Douglas. II. Title: Inhomogeneous fluids.
QD541.F86 1992
541'.0422--dc20 92-21900
CIP

This book is printed on acid-free paper.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Preface

Around 1960 Joe Hirschfelder wrote an essay outlining what he felt were the three unsolved problems in theoretical chemistry. I think that he called them bottlenecks. One of the bottlenecks was the lack of a satisfactory theory of liquids. If a theoretical physicist were to have made in 1960 a similar list of bottlenecks in theoretical physics, the list would undoubtedly have been different. However, it is quite likely that the need for a satisfactory theory of liquids would have appeared on this list also.

In any case, the past quarter century has seen a dramatic increase in our understanding of homogeneous or bulk liquids. However, our understanding of inhomogeneous liquids (liquids near interfaces or confined liquids) is less satisfactory. Since the feature that distinguishes a liquid from a dense gas is the presence of an interface, one might argue that even our understanding of homogeneous liquids is incomplete without an understanding of inhomogeneous fluids. It is for this reason that many of us who worked earlier on the theory of homogeneous liquids have turned our attention to inhomogeneous liquids. Indeed, van der Waals, after developing his theory of liquids, almost immediately turned his attention to interfaces.

A further reason for an interest in inhomogeneous liquids is that many of the technological processes that involve liquids occur at interfaces. A

few examples that come to mind are corrosion, oil recovery, and colloidal stability.

I have used the term inhomogeneous fluids rather than inhomogeneous liquids in the title so as to include such interesting and important topics as the wetting of a surface by a gas. Furthermore, the distinction between a liquid and a dense gas is somewhat artificial. Also, in lattice models, there is a reciprocity between the liquid and vapor phases. This reciprocity is mirrored, at least approximately, in the real world.

There has been considerable progress in the past decade in the development of the theory of inhomogeneous fluids. Thus, it seems timely to bring out a book outlining this progress. The emphasis here is on theory, although experimental work is referred to in the book. In fact, one chapter is devoted to wetting experiments.

The book is divided into thirteen chapters, each written by a recognized expert. Rowlinson begins with a historical introduction written with the insight that only he can provide. Jim Henderson discusses exact sum rules for inhomogeneous fluids, many of which he derived. Evans and I outline the two main theoretical tools in the theory of inhomogeneous fluids, density functional, and integral equation techniques, respectively. Jancovici discusses his exact solutions for two-dimensional homogeneous plasmas, and Blum and I discuss the recent progress, mostly based on integral equations, in the theory of interfacial electrochemistry.

Franck discusses wetting experiments—the viewpoint of an experimentalist is complementary and of interest. The theory of wetting (the modern theory of adsorption) is also described in some of the other chapters (see the chapters of Jim Henderson and Evans, in particular). Lozada-Cassou describes confined liquids, mostly electrolytes, largely using the integral equations he has obtained.

Next, phase transitions (mostly first-order) are considered. Haymet discusses freezing with an emphasis on quantum systems. His view is that a solid is a highly inhomogeneous fluid. Although this likely would not be a useful basis for solid-state theory, it seems useful as a description of solids in equilibrium with a liquid and provides the first really useful theory of freezing. Oxtoby discusses homogeneous nucleation in liquid–vapor and solid–liquid transitions. Marko completes this trilogy with a discussion of liquid crystal transitions. The treatment of these three chapters uses the density functional approach discussed by Evans.

The final two chapters are those of Dawson and Mundy on self-organizing liquids and of Davis on kinetic phenomena in inhomogeneous fluids using a modified Enskog theory. Unfortunately an author who was invited to write a chapter on computer simulations did not submit his manuscript. Its lack is compensated for by the fact that many of the authors in this

volume make references to and comparisons with computer simulations in their own chapters.

I want to thank the chapter authors, who made many suggestions that improved this volume. I hesitate to single out any of them, but Tony Haymet and Jim Henderson were especially helpful. Much of the organization of the book was done while I was a visiting professor at the Scuola Normale Superiore in Pisa, Italy, in the spring of 1989. I am grateful to this institution and to Dr. Alessandro Tani for many kindnesses.

Lastly I want to thank my wife, Rose-Marie. The production of a book goes through three phases: initial excitement at the enterprise, frustration with the delays and necessary drudgeries, and finally relief that the business is finished. I am grateful for her bemused tolerance as I passed through these phases.

Douglas Henderson

Contents

Contributors

L. Blum, Ph.D. Professor of Physics, Department of Physics, University of Puerto Rico, Rio Pedras, Puerto Rico

H. Ted Davis, Ph.D. Professor and Head, Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, Minnesota

K. A. Dawson, Ph.D. Assistant Professor, Department of Chemistry, University of California, Berkeley, California

R. Evans, Ph.D. H. H. Wills Physics Laboratory, University of Bristol, Bristol, England

Carl Franck Associate Professor of Physics, Laboratory of Atomic and Solid State Physics, and Materials Science Center, Cornell University, Ithaca, New York

A. D. J. Haymet, Ph.D. Department of Chemistry, University of Sydney, Sydney, N.S.W., Australia

Douglas Henderson, Ph.D. Research Scientist, Utah Supercomputing Institute/IBM Partnership, and Adjunct Professor, Department of Chemistry, University of Utah, Salt Lake City, Utah

J. R. Henderson, Ph.D. School of Chemistry, University of Leeds, Leeds, England

B. Jancovici, D.Sc. Professor, Laboratoire de Physique Théorique, Université de Paris-Sud, Orsay, France

Marcelo Lozada-Cassou, Ph.D. Professor, Department of Physics, Universidad Autónoma Metropolitana-Iztapalapa, Mexico City, Mexico

J. F. Marko, Ph.D. James Franck Institute, University of Chicago, Chicago, Illinois

Christopher J. Mundy, B.S. Research Assistant, Department of Chemistry, University of California, Berkeley, California

David W. Oxtoby, Ph.D. Professor of Chemistry, Department of Chemistry, James Franck Institute, University of Chicago, Chicago, Illinois

J. S. Rowlinson Dr. Lee's Professor of Chemistry, Physical Chemistry Laboratory, Oxford University, Oxford, Great Britain

1

Development of Theories of Inhomogeneous Fluids

J. S. Rowlinson
Oxford University
Oxford, Great Britain

I. INTRODUCTION: LAPLACE'S THEORY OF THE SHARP INTERFACE

From the earliest days of physical science it must have been clear that matter in thin films and at surfaces differed in its properties from the same matter in bulk. The "natural philosophers" of the seventeenth and eighteenth centuries were curious about both the optical and mechanical properties of surface films, as, for example, the interference colors of Newton's rings and the damping of the waves on the sea by a film of oil, a subject that interested Benjamin Franklin, among others. The first of these effects to be tackled quantitatively was that of capillarity—the rise of water (and the depression of mercury) in narrow tubes or, in its more practical aspect, the rise of sap in trees. This apparent defiance of the laws of gravity was correctly perceived to arise from cohesive forces between the ultimate constituents of matter that were not gravitational in origin. Francis Hauksbee, Newton's assistant at the Royal Society, showed this convincingly by observing that the capillary rise of water depended only on the internal diameter of the tube and was independent of the thickness of its walls. The force responsible must therefore be of short range and act essentially only between the layers of liquid and solid that are in contact. As well as this cohesive force it was realized that

there were other, presumably weaker, forces that acted within the body of the liquid since, as James Jurin (1719) and A. C. Clairaut (1743) observed, drops of liquid that are too small for their shape to be influenced by gravity are spherical. J. A. Segner (1751) synthesized these and similar observations into the concept of surface tension.

The nature of the cohesive forces was, of course, unknown and was to remain so until this century, but speculation was not lacking. John Keith (1709) proposed an attractive force that varied as r^{-n}, where n exceeded the Newtonian value of 2. More complex schemes were devised by John Rowning (1735) and Rudjer Boskovic, or Roger Boscovich (1758), who suggested alternating zones of attraction and repulsion around each particle, arranged in such a way that the separations at which the force is zero, corresponded to the densities of the material in its solid, liquid, and gaseous forms. Such speculation had outrun the evidence, for there was no theory to link these hypothetical forces with their physical or macroscopic consequences. It was probably because of this lack that the second half of the eighteenth century saw little further advance. At the beginning of the nineteenth century two men independently rescued the subject from its stagnation, Thomas Young [1] in London and Pierre Simon Laplace [2] in Paris. Many of their conclusions were the same, but those of Young are less accessible since in his early papers, he expressed his results in words rather than in symbols and equations, and in his later papers he was still using old methods of handling the calculus which makes his work hard to follow. This short account is, therefore, based substantially on that of Laplace and, indeed, it was his work which, for the same reasons, was more influential later in the century.

Laplace calculated the force between two bodies of liquid with planar parallel walls and separated by vapor of negligible density. He knew that the forces between the molecules must be short-ranged on a human scale but he assumed, nevertheless, that their range was long compared with the mean separation of the molecules. This assumption has now become a formal requirement in some molecular theories of matter and has been given the name of the *mean-field approximation*. It has played an important, and often controversial, role in physics for nearly 200 years.

With this approximation he did not have to worry about the actual distribution of the molecules in the liquid but could assume that it was uniform, that is, that they were distributed at random. His result for the attractive force per unit area between two surfaces in contact, K, is most easily expressed in modern form in terms of the potential energy, $\varphi(r)$, of a pair of molecules at a separation r. This potential energy is related to the force between the pair, $f(r)$, by $-f(r) = d\varphi(r)/dr$. If $\varphi_{\text{att}}(r) < 0$ and $\varphi_{\text{att}}(\infty) = 0$, then $f(r) < 0$; that is, the force is attractive. Laplace's

result for the attractive force per unit area of two slabs of liquid in contact can be expressed as

$$K = -2\pi\rho^2 \int_\sigma^d \varphi_{att}(r) r^2 \, dr \tag{1}$$

where ρ is the number of molecules per unit volume, d the range of the intermolecular potential, and σ a cutoff distance for the attractive potential. He showed also that the work done per unit area to separate the two slabs of liquid to a separation d (or greater) is given by

$$H = -\pi\rho^2 \int_\sigma^d \varphi_{att}(r) r^3 \, dr \tag{2}$$

Since two surfaces have been created, this quantity is twice the excess energy per unit area of the liquid surface and hence twice the liquid–gas surface tension; $H = 2\gamma^{lg}$. The cohesive energy density of the liquid, or energy per unit volume, is $-2K$, and the "internal pressure" inside a spherical drop of radius R is $K + (H/R)$, a quantity that is not directly accessible to experiment. The difference of the observed pressure inside and outside the drop is accessible and is given by what we now call *Laplace's equation*,

$$p^l - p^g = \frac{H}{R} = \frac{2\gamma^{lg}}{R} \tag{3}$$

If a liquid touches a solid, similar arguments can be used to obtain expressions for the two surface tensions, γ^{sg} and γ^{sl}. These are related to the angle θ between the liquid–gas and the liquid–solid surface by *Young's equation*,

$$\gamma^{sg} = \gamma^{sl} + \gamma^{lg} \cos\theta \tag{4}$$

These two equations, of Laplace and Young, suffice to solve all common problems of capillarity, such as the rise (or fall) of liquids in narrow tubes, the shapes of liquid surfaces in contact with solid surfaces, the shapes of systems of bubbles, and so on.

This achievement was an important success in the Laplacian (or Newtonian) program of obtaining the physical properties of matter from the microscopic interparticle forces. The very success of the theory provoked a more detailed examination from which several difficulties emerged. The first was the nature of the repulsive forces that are needed to balance the postulated attractive forces in a system at equilibrium. Laplace, a believer in the caloric theory, at first ascribed the necessary repulsion to heat, a

view that received some support from the observation that most liquids expand on heating. In his later work (1819 onward) he came around to the view that the strength of these caloric-induced repulsions must fall off more rapidly with distance than the attractive forces, as Boskovic and others had deduced earlier. The second mechanical difficulty was the structure of the supposed equilibrium array of stationary particles. It was not difficult to envisage how a crystal might be constructed by packing together in a regular way a large number of molecules of simple shape, but there could be no such geometrical picture of a liquid. This, again, was a problem that could not be tackled within the essentially static model of Laplace, and its solution had to await the development of a kinetic theory of matter later in the century.

A third and more subtle problem concerned the size of the postulated molecules and the range of the forces between them. Young observed that since the integral H has a factor of r^3 and the integral K, a factor of r^2, the ratio $2H/K$ is a measure of the effective range of the potential, $\varphi_{att}(r)$. We have seen that H is twice the surface tension, so is easily measured, but that K, the cohesive energy density, is not directly measurable. Young estimated it for water at 25 kbar (in modern units), probably by an argument based on the compressibility of the liquid. He deduced that the effective range of $\varphi_{att}(r)$ is therefore about 1×10^{-10} m, a figure that we now know to be low only by a factor of about 5. He then went on, wrongly, to identify this distance with the mean molecular separation in steam on the point of condensation, since he thought that this was the distance at which the attractive forces must start to act if they were to bring about condensation. Since liquid water is more than 1000 times denser than steam at its normal boiling point, he was able to convince himself that the mean molecular separation in liquid water is many times smaller than the range of the cohesive forces, and so justify the use of the mean-field approximation.

We have seen how the early enthusiasm for Newton's physics in the first half of the eighteenth century led to many attempts to explain the properties of gases and liquids in terms of interparticle forces, and how this enthusiasm faded in the second half of the century when the program failed to produce quantitative results. The work of Laplace and the Arceuil school in France gave the field a renewed vigor that lasted for about 25 years. From about 1830 onward, however, the effort faltered again for the same reason—no significant new results were forthcoming. It was not until the creation of the mechanical theory of heat and of kinetic theory and the development of thermodynamics that the next round in the theory of liquids and their surfaces could be opened.

II. VAN DER WAALS' THEORY OF THE DIFFUSE INTERFACE

In 1869, Dupré [3] recalculated the internal energy density, $-K$, of Laplace's theory. He called this quantity *le travail de désagrégation totale*, a name that reflected the way in which it was calculated—by calculating the force needed to remove the particles in the surface layer of a liquid, one by one. A particle that is within a distance d of the surface experiences a net inward attraction because of the lack of spherical symmetry in its surroundings. Although his result was correct (within the mean-field approximation), his argument is flawed since a system that contains particles on which there are unbalanced forces cannot be at equilibrium, so the liquid cannot have the sharp surface that he and Laplace had supposed. This flaw was, in fact, pointed out by Poisson [4], a follower of Laplace, in 1831, but he had then gone on to deduce, incorrectly, that if a sharp surface were replaced by a diffuse one, of thickness comparable with the range d, the surface tension would vanish. His fault lay in an implicit assumption that it is possible to define unambiguously the local values of a thermodynamic function in terms of the local values of other functions, as in a homogeneous system. A systematic application of this argument does, indeed, lead to the vanishing of the surface tension, but the argument itself is correct only if the attractive forces are of zero range, and it is this circumstance that is responsible for the vanishing of the surface tension, as is seen from Eq. (2) in the limit of $d \to \sigma$. A less restrictive view of the local thermodynamics of inhomogeneous systems was needed before an interface of nonzero thickness and forces of nonzero range could be reconciled with a nonzero value of the surface tension.

Three men, apparently independently, found the natural way of generalizing this restricted view to a more general or nonlocal thermodynamics of an inhomogeneous system. They were Karl Fuchs [5] at Pressburg (now Bratislava, in Czechoslovakia), Lord Rayleigh [6], who worked mainly in his laboratory at his home at Terling in Essex in Britain, and J. D. van der Waals [7] in Amsterdam. The theory put forward by Fuchs and by Rayleigh was still in the tradition of Laplace in that it was purely a mechanical treatment of the problem; it ignored the motions of the molecules or, in thermodynamic terms, it used an energy where a free energy was more appropriate. Van der Waals was the first to realize the importance of this distinction.

Fuchs observed that since the attractive force has a nonzero range, molecules in a surface layer will influence, and be influenced by, other molecules that are in a fluid of quite different local density. Hence the

average energy at a given point, r, so also the average value of the local pressure, will depend not only on the density $\rho(\mathbf{r})$ but on its gradient and higher derivatives. If we have a flat surface in the xy-plane, these derivatives can be written $\rho'(z)$, $\rho''(z)$, and so on. For obvious reasons of symmetry, there cannot be a term linear in $\rho'(z)$, and Fuchs was the first to show that the excess of the tangential or horizontal component of the pressure over the vertical component at height z can therefore be expressed as

$$\Delta p(z) = -\tfrac{1}{60}\{\rho(z)\rho''(z) - [\rho'(z)]^2\}I \tag{5}$$

where

$$I = -\int r^3 f(r)\, d\mathbf{r} \tag{6}$$

and where $f(r) < 0$ is the force between two molecules at separation r. He took this result no further, but a simple integration over z gives for the surface tension,

$$\gamma = \int_{-\infty}^{+\infty} \Delta p(z)\, dz = \tfrac{1}{30} I \int_{-\infty}^{+\infty} [\rho'(z)]^2\, dz \tag{7}$$

This last step was taken by van der Waals in a short communication to the Academy of Sciences that was published in 1889 and in his definitive paper of 1893 [7], and, in 1892, by Rayleigh [6].

Van der Waals opens his treatment by introducing at once the condition that the free energy is a minimum in a system of fixed mass, volume, and temperature. He makes free use of the hypothetical continuous isotherm that passes smoothly through the two-phase region. The usefulness of this curve had been guessed first by James Thomson in 1871 [8] after the publication of Andrews's classic experiments on the critical point of carbon dioxide, and it had become accepted as a useful construct after its use in van der Waals' thesis in 1873 [9]. The results of van der Waals can be expressed simply by introducing first the Gibbs dividing surface of zero adsorption. This divides the liquid from the gas at an arbitrary height z_0 defined by

$$\int_{-\infty}^{\infty} [\rho(z) - \rho^{l,g}]\, dz = 0 \tag{8}$$

where $\rho^{l,g}$ is ρ^l for $z < z_0$ and ρ^g for $z > z_0$. A local excess free-energy density $\Psi(z)$ can then be defined as the sum of two terms,

$$\Psi(z) = \Delta\psi[\rho(z)] + \tfrac{1}{2}m[\rho'(z)]^2 \tag{9}$$

where $\Delta\psi$ is the amount by which the free-energy density on the contin-

uous (Thomson–van der Waals) isotherm exceeds that of the two bulk phases at the same overall density, which is

$$\psi^l \frac{\rho(z) - \rho^g}{\rho^l - \rho^g} + \psi^g \frac{\rho^l - \rho(z)}{\rho^l - \rho^g}$$

He showed that the coefficient m is the fourth moment of the attractive part of the intermolecular potential,

$$m = -\frac{2\pi}{3} \int r^4 \varphi_{\text{att}}(r)\, dr \tag{10}$$

and so differs only by a numerical factor of 30 from Fuchs's integral I of Eq. (6). Van der Waals obtained the surface tension from the integral of the excess free-energy density defined with respect to the Gibbs dividing surface of zero adsorption:

$$\gamma = \int_{-\infty}^{\infty} \Psi(z)\, dz = m \int_{-\infty}^{\infty} [\rho'(z)]^2\, dz \tag{11}$$

The two terms in Eq. (9) contribute equally to the surface tension.

This "square-gradient" theory of van der Waals was wholly within the mean-field approximation, all fluctuations are ignored, and it gives rise to a classical critical index:

$$\gamma \sim (T^c - T)^{\mu} \qquad \mu = \tfrac{3}{2} \tag{12}$$

It was, however, the most satisfactory theory of the surface layer between gas and liquid to be put forward until modern times. Its impact was limited by the general loss of vitality of the field of liquid state physics after 1914 [10], so that its results were rediscovered independently many years later, as will be discussed below.

III. INTERFACES BETWEEN SOLUTIONS OF ELECTROLYTES

At the end of the nineteenth century the word *solution* usually meant an aqueous solution of an electrolyte, that is, an acid, a base, or a salt. At the same time as Fuchs, Rayleigh, and van der Waals were developing a theory of the liquid–gas interface, Nernst [11] and Planck [12] were taking up the problem of the junction between two aqueous solutions of electrolytes, which could be either solutions of different substances or of the same substance at different compositions. Such an interface differs fundamentally from the liquid–gas interface, for it cannot be at equilibrium; diffusion of the electrolyte must slowly bring about complete mixing and

the junction will disappear. Nevertheless, diffusion in liquids is slow compared with the time needed to establish an almost stationary distribution of ions, so Nernst and Planck were justified in using quasi-thermodynamic arguments. Their work was contemporary with that of Fuchs and of van der Waals' first note, but there is no evidence that either set of workers knew of or was influenced by the other. Nernst and Planck phrased their argument in terms of the force generated by a gradient of osmotic pressure which is balanced by that of the electric field. A modern route to the same result (e.g., that of MacInnes [13]) considers the flow of ions into and out of an infinitesimal layer of the interface. The potential of the liquid junction, E_L, between two phases, α and β, is given by

$$E_L = -\frac{RT}{F}\int_{\beta}^{\alpha}\sum_i \frac{t_i}{z_i}\, d\ln a_i \tag{13}$$

where t_i, z_i, and a_i are the transport number, number of electronic charges, and activity of the ion of species i.

This expression is not without its problems. There is the question of the meaning to be attached to single-ion activities since these cannot be measured—a restriction that was not fully appreciated until this century. A characteristically forthright expression of our inability to determine single-ion chemical potentials and electrode potentials was given by Guggenheim in 1929 [14]. The restriction arises from the need to preserve electrical neutrality. If, however, the junction is between two solutions of the same electrolyte, for example a uni-univalent electrolyte in a cell in which the electrodes are reversible to the anion, then, since $t_+ + t_- = 1$, we have

$$E_L = \frac{RT}{F}\int_{\alpha}^{\beta} d\ln a_- - 2\frac{RT}{F}\int_{\alpha}^{\beta} t_+\, d\ln a_\pm \tag{14}$$

where $a_\pm^2 = a_+ a_-$, the square of the mean ionic activity, which is measurable. The first term is the contribution to the emf of the cell from the two electrodes, so the last term is the electromotive force (emf) of the whole cell.

Matters are less simple if we have a junction between two different electrolytes. Even if we are willing to make an assumption about single-ion activities (e.g., that $a_+ = a_- = a_\pm$), we still cannot integrate Eq. (13) without knowing more about the concentration gradients of the ions in the interface and how t_i depends on the local concentration. The last problem can be solved by auxiliary experiments, but the problem of the gradients of ionic concentration requires a further ad hoc assumption. Planck took this assumption to be what we now call a boundary zone of

constrained diffusion, which can be envisaged as the zone established by diffusion in the pores of a plug, the two faces of which are washed by the two homogeneous phases sufficiently rapidly for the concentration at each face to be constrained to be that of the appropriate bulk phase. The form of equations to which this gives rise are set out in an appendix in MacInnes's book [13].

A simpler assumption, made later by Henderson [15], seems to agree as well with experiment. He assumed that the composition of each point in the boundary zone was that of a fraction of y of phase α and fraction $(1 - y)$ of phase β. He assumed further that y is a linear function of distance from either of the faces of the zone, although it was later found that this assumption was redundant. (It is a curiosity that many years later, and in ignorance of Henderson's work, I made exactly the same assumptions about the composition profile of the boundary layer between a binary liquid mixture and its vapor in order to calculate the surface tension [16].)

IV. ELECTRICAL DOUBLE LAYER

The surface between an electrode and a solution of an electrolyte is an inhomogeneous system that can be at true equilibrium and so can be treated with more precision than the liquid–liquid junction between two solutions in a common solvent. The treatment of an interface at true equilibrium might, at first sight, be expected to owe something to the earlier work of Laplace or of van der Waals on the liquid–gas surface, but the long range of ionic forces means that the dominant structures are quite different and, at first, the theoretical treatments had nothing in common. It is only in recent times that the theoretical methods have converged, as both have been grounded more thoroughly in modern statistical mechanics.

Helmholtz [17] was the first to realize that a charged metal plate immersed in a solution of an electrolyte would have as its neighbor a zone in which the ions were predominantly of the opposite charge, thus forming a double layer. He defined the moment of this layer as the product of its thickness D and the density of positive (or negative) electricity within it, e. He envisaged, however, that D was of molecular dimensions, so that the electrolytic part of his double layer comprised more an adsorbed layer of ions than a diffuse zone that is thick by comparison with the size of any one ion. Billiter [18] realized that the double layer need not be confined to one layer, or even few layers, but would spread out into the bulk of the liquid phase. He spoke of "a dissociation of the double layer" [19] but did not develop a theory for calculating its thickness and structure.

That task was undertaken independently by Gouy in Lyon in 1910 [20] and by Chapman in Oxford in 1913 [21]. Each treated the problem in the same way and their two papers are remarkable, even by the standards of the day, in that neither of them has a single reference.

The diffuseness of the layer arises from the opposing effects of the electric force from the electrode, which tends to make the layer thin, and the combination of the repulsive forces between the ions and of their own diffusive motion from places of high density to those of low density (then generally called the osmotic force), which tends to make the layer thick. If the layer is sufficiently thick, we can, as a first approximation, assume that the ions are of negligible size.

Their method of tackling the problem is to assume that the concentration profile of the ions satisfies two equations. The first of these is Poisson's equation of electrostatics,

$$\text{div grad } V(\mathbf{r}) = \frac{\rho(\mathbf{r})}{\epsilon} \tag{15}$$

where ρ is the density of charge at position $\mathbf{r}$, ϵ the permittivity of the solution (taken to be independent of $\mathbf{r}$), and V the electric potential. The second equation is Boltzmann's distribution for ions in an external potential V:

$$n_i(\mathbf{r}) = n_i^{\infty} \exp\left[\frac{-z_i e V(\mathbf{r})}{kT}\right] \tag{16}$$

where $n_i(\mathbf{r})$ is the concentration at $\mathbf{r}$ of ions of species i which carry a charge of $z_i e$, and n_i^{∞} is the concentration at infinite distance from the electrode where V is zero. Since

$$\rho(\mathbf{r}) = \sum_i n_i(\mathbf{r}) z_i e \tag{17}$$

we have

$$\nabla^2 V(\mathbf{r}) = -\frac{1}{\epsilon} \sum_i n_i^{\infty} z_i e \exp\left[\frac{-z_i e V(\mathbf{r})}{kT}\right] \tag{18}$$

which is the Poisson–Boltzmann equation, the most used equation in the theory of inhomogeneous electrolytic solutions. It is not exact because the Boltzmann distribution is correct only for ions of zero size and with no forces between them other than the Coulomb forces. Moreover, the solvent is treated as a continuous dielectric medium, not as a collection of discrete molecules that exert forces on the ions and on each other. Nevertheless, the Poisson–Boltzmann equation embodies much of the essential physics of the problem, at least for dilute solutions.

It is convenient to depart from strict historical order and consider first the linearized equation obtained by assuming that V is small and hence that the exponential can be expanded to give

$$\nabla^2 V(\mathbf{r}) = -\kappa^2 V(\mathbf{r}) \tag{19}$$

where

$$\kappa = e\left(\sum_i \frac{z_i^2 n_i^\infty}{\epsilon kT}\right)^{1/2} \tag{20}$$

The parameter κ^{-1} has the dimensions of length and is a measure of the thickness of the double layer. This equation was used in 1923 by Debye and Hückel [22] to determine the potential and the concentration of ions around a chosen ion in a dilute solution. The problem is one of spherical symmetry for which the operator has the form

$$\nabla^2 = \frac{d^2}{dr^2} + \frac{(2/r)d}{dr} \tag{21}$$

so $V(r)$ has the form that is now usually called a Yukawa potential [23]:

$$V(r) = \frac{ze\kappa}{\epsilon}\frac{1}{\kappa r}e^{-\kappa r} \tag{22}$$

where z is the charge on the central ion and where the constants of integration have been chosen so that $V(r)$ approaches its Coulomb value as r becomes infinite.

For a flat surface, as considered by Gouy and Chapman, it is not necessary to linearize the Poisson–Boltzmann equation. The operator ∇^2 now takes the simple form d^2/dx^2, where x is the distance perpendicular to the planar electrode. For a single symmetrical electrolyte, $z_i = z_+ = z_- = z$, we have the equation

$$\frac{d^2V(x)}{dx^2} = \frac{2zen^\infty}{\epsilon}\sinh\frac{zeV(x)}{kT} \tag{23}$$

of which the integral with the appropriate boundary conditions is

$$\tanh\frac{zeV(x)}{4kT} = \tanh\frac{zeV^\circ}{4kT}e^{-\kappa x} \tag{24}$$

where V° is the potential at the surface of the electrode ($x = 0$). Gouy wrote this as

$$\kappa x = \ln\frac{u_0^{1/2} - 1}{u^{1/2} - 1} - \ln\frac{u_0^{1/2} + 1}{u^{1/2} + 1} \tag{25}$$

where $u = n_+/n_+^\infty$, that is, the relative excess of cations, and u_0 is the value at $x = 0$. Clearly, $u_0 < 1$ if the plate carries a positive charge ($V > 0$), and $u_0 > 1$ if it is negative, and vice versa for the anions since $n_+ n_- = n_+^\infty n_-^\infty = n^2$.

The Gouy–Chapman theory and Poisson–Boltzmann equation on which it is based, like many simplest approximations, have virtues that can easily be lost when attempts are made to improve them. Equations (23) to (25) satisfy, for example, the conditions of mechanical equilibrium at all points in the diffuse layer. Nevertheless, the imperfections of the Gouy–Chapman theory and, more particularly, of its linearized version, the Debye–Hückel theory, have been recognized from the earliest days. The best known formal analyses of the Debye–Hückel theory were those of Fowler [24], Onsager [25], and Kirkwood [26].

The first significant attempt to improve the Gouy–Chapman theory was based on the recognition by Stern [27] that the nonzero size of the ions prevented their centers from reaching the actual surface of the electrode, and so prevented the surface concentration of counterions reaching the impossibly high figures that could result from the Gouy–Chapman equation. He assumed that there would be an adsorbed layer of counterions whose density he calculated from arguments similar to those used earlier by Langmuir [28] for calculating the extent of adsorption of molecules at uncharged surfaces. This adsorbed layer of counterions is now usually known as the Stern layer. Much later, Bikerman [29] proposed a correction for the nonzero size of the ions in the diffuse layer by introducing a co-volume, b, but did not attempt to calculate the consequences of his proposal. Since the two assumptions of a linearized Boltzmann distribution and a co-volume are reminiscent of the two assumptions behind the van der Waals equation of state of homogeneous fluids, it is of interest to see what is the corresponding result for an inhomogeneous electrolyte solution. This point is explored briefly in an appendix to this chapter. Dutta and Baggchi [30] and, independently, Eigen and Wicke [31] proposed similar modifications of the Debye–Hückel equation.

These attempts at improvement are, however, all without formal theoretical foundation. Modern work has developed in two directions. The first is based on the critical work of Kirkwood and others and led in 1951 to the so-called modified Poisson–Boltzmann equation of Loeb [32], which has been developed further by Levine and Outhwaite [33]. The second has been the introduction into these fields of the integral equations used first for homogeneous and uncharged liquids, such as the Yvon–Born–Green and hypernetted chain integral equations. With these developments we reach the modern era, and so subjects of later chapters of this book.

One field of application of the theory of the diffuse layer should, however, not go unnoticed. This is the theory of the forces between colloidal particles, which is based, in part, on the behavior of two interacting double layers. Earlier work on this subject, such as that of Langmuir [34], Derjaguin and Kussakov [35], and Levine [36], was, as has been said, "conflicting on some essential points, even on the important question of whether the double layer interaction produces attraction or repulsion" [37]. The matter was due to have been thrashed out at the Faraday Society Discussion in Cambridge in September 1939, but that meeting was never held, although many of the papers and a little of the discussion by correspondence were published [38]. The classic work on this problem, which resolved many of these difficulties and paved the way for more recent work that is firmly grounded in statistical mechanics, was that carried out in the Netherlands during World War II by Verwey and Overbeek and published by them as a monograph in 1948 entitled *Theory of the Stability of Lyophobic Colloids* [37].

V. VIRIAL EXPRESSIONS AND FUNCTIONAL EXPANSIONS: THE MODERN ERA

From its earliest development in the 1880s up to World War II the properties of solutions of electrolytes were at the center of what had become known as physical chemistry. The field has developed continuously since the war, although it has never regained the leading position it had in that subject during the 1920s and 1930s. In contrast, the theory of uncharged liquids and solutions, which was a flourishing branch of physics until World War I, became something of a backwater until after World War II [10]. In consequence, the theory of inhomogeneous liquids did not make any significant advance beyond the square-gradient theory of the liquid–gas interface of van der Waals for nearly 60 years. This theory itself was reinvented by Landau and Lifshitz in 1935 [39] to treat the formally similar problem of the interface between two magnetic domains, and by Mitsui and Furuichi in 1953 [40] for that between two ferroelectric domains. In 1958, Cahn and Hilliard [41] independently derived the results again for the liquid–gas surface in a treatment of nucleation, but although they quote, in a different context, results from the relevant papers of Rayleigh and van der Waals, they did not recognize at first that their results were essentially those of van der Waals.

The foundations of the modern treatment of interfaces can be seen in two developments that it is not too farfetched to regard as the completion of the programs of Laplace and van der Waals. The first was the development of an exact virial expression for the surface tension. It had long

been recognized that the tension depended on a difference in the interface between the normal and the tangential components of the pressure tensor. For a plane interface,

$$\gamma = \int_{-\infty}^{\infty} [p_N(z) - p_T(z)]\, dz \tag{26}$$

The use of this expression requires an exact statistical mechanical expression for the components of the pressure tensor. This was given first in 1949 by Kirkwood and Buff [42] and, in a different form, by Irving and Kirkwood [43]. There is more than one way of tackling this problem, so $p_T(z)$ is not a unique function, although the integral (26) is invariant to this lack of uniqueness. One of the versions of $p_T(z)$ is now usually associated with the name of Harasima [44]. The normal component, $p_N(z)$, is, by the condition of mechanical equilibrium, a constant at a planar interface and equal to both p^l and p^g, the pressures in the liquid and gaseous phases. The result obtained by substituting these statistical mechanical expressions for the pressure tensor into the integral in (26) is the so-called virial expression for the surface tension. It can be regarded as the exact result to which the treatment initiated by Laplace led, even if the path to this result took over 140 years to accomplish.

A natural corollary of this result is the Yvon–Born–Green (YBG) equation, which is obtained by differentiating with respect to the position of particle 1 the statistical mechanical expression for $\rho(\mathbf{r}_1)$ in an inhomogeneous system:

$$-kT\nabla_1\rho(\mathbf{r}_1) = \rho(\mathbf{r}_1)\nabla_1 V(\mathbf{r}_1) + \int \rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2)\nabla_1\varphi(\mathbf{r}_{12})\, d\mathbf{r}_2 \tag{27}$$

where $V(\mathbf{r}_1)$ is the external potential at $\mathbf{r}_1$, and $\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ is the two-body distribution function. In a homogeneous fluid each term vanishes (the integral by symmetry) and the first nontrivial YBG equation is the second member of the hierarchy in which the left-hand side is $-kT\nabla_1\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ and the integral is over the three-body distribution function $\rho^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$. For obvious reasons, early users of these equations paid more attention to this equation than to Eq. (27), the one-body equation.

Yvon [45] was the first, in 1935, to use such equations, although Kirkwood [46] obtained an equivalent two-body equation the same year. There were later independent derivations in 1946 by Born and Green [47] and by Bogoliubov [48]. Because of these many independent works the hierarchy of equations is sometimes known as the YBG and sometimes as the BBGKY hierarchy.

The second foundation of the modern treatment of interfaces has been the method of functional expansions of the distribution functions. It is this approach that was described above as the natural end of the line of development opened by Fuchs, Rayleigh, and van der Waals with the square-gradient approximation. It is a necessary feature of this early work that it postulates the existence of thermodynamic functions of states of the fluid that have no independent equilibrium existence, such as a homogeneous fluid state of a density between that of the orthobaric gas and liquid states. A similar freedom in statistical mechanical theory to suppose the existence of functionals of distribution functions of states not at equilibrium came much later, but has proved to be equally fruitful. Such functionals were introduced in 1960 by Lee and Yang [49] and by Green [50], and were soon applied to the statistical mechanics of inhomogeneous systems by Morita and Hiroike [51], De Dominicis [52], Stillinger and Buff [53], Lebowitz and Percus [54], Mermin [55], and Ebner and Saam [56].

Such functionals are at extrema when their arguments, the distribution functions, take their equilibrium values, and the values of the functionals then correspond to the equilibrium value of a thermodynamic potential such as F or Ω. If $\hat{\rho}^{(N)}(\mathbf{r}^N)$ denotes an arbitrary N-body distribution function in the canonical ensemble which is normalized so that

$$\int \hat{\rho}^{(N)}(\mathbf{r}^N)\, d\mathbf{r}^N = N! \tag{28}$$

thus we can define a functional of $\hat{\rho}^{(N)}$ that becomes the free energy when $\hat{\rho}^{(N)}$ becomes $\rho^{(N)}$, the equilibrium distribution function; this functional is

$$\mathcal{F}[\hat{\rho}^{(N)}] = \frac{1}{N!} \int \hat{\rho}^{(N)}[\Phi^{(N)} + kT \ln(\Lambda^3 \hat{\rho}^{(N)})]\, d\mathbf{r}^N \tag{29}$$

The first two functional derivatives of $\mathcal{F}$ can be obtained by using two equations obtained by Yvon in 1958 [57]. The equations are

$$h^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{-kT}{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)} \frac{\delta\rho(\mathbf{r}_1)}{\delta V(\mathbf{r}_2)} \frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \tag{30}$$

and its inverse,

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{kT} \frac{\delta V(\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} + \frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \tag{31}$$

where $h^{(2)}$ and $c^{(2)}$ are the total and direct correlation functions between

points $\mathbf{r}_1$ and $\mathbf{r}_2$. The functional derivatives of $\mathcal{F}$ are

$$\frac{\delta\mathcal{F}[\hat{\rho}^{(N)}]}{\delta\hat{\rho}(\mathbf{r}_1)} = \mu - V(\mathbf{r}_1) \tag{32}$$

$$\frac{\delta^2\mathcal{F}[\hat{\rho}^{(N)}]}{\delta\hat{\rho}(\mathbf{r}_1)\,\delta\hat{\rho}(\mathbf{r}_2)} = \frac{kT\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - kTc^{(2)}(\mathbf{r}_1,\mathbf{r}_2) \tag{33}$$

where μ is the chemical potential. The derivatives are taken in the equilibrium state, so these results can be used to expand $\mathcal{F}$ about its value for a uniform system in powers of $\Delta\hat{\rho}(\mathbf{r}) \equiv \hat{\rho}(\mathbf{r}) - \rho_u$. A first approximation to the excess free energy per unit area of a liquid–gas interface follows from this expansion [58,59] and leads again to the square-gradient approximation to the surface tension of Eq. (11) but with the coefficient m now given not by Eq. (10) but by

$$m(\rho) = \frac{2\pi}{3} kT \int r^4 c^{(2)}(r;\rho_u)\, dr \tag{34}$$

where $c^{(2)}(r;\rho_u)$ is the direct correlation function in a uniform fluid of density ρ_u. This result is the modern justification for the square-gradient approximation. It differs from the original version in that m becomes a function of density, and, moreover, one that is not well defined, for two reasons. The first is that $c^{(2)}(r;\rho)$ is an unknown function if ρ is a hypothetical uniform density between ρ^l and ρ^g, and second, because (34) diverges at the critical point. The two results, old and recent, are, in fact, close in their predictions for the surface tension since a well-known estimate for $c^{(2)}(r)$,

$$-kTc^{(2)}(r) = \varphi_{\text{att}}(r) \qquad \text{at large } r \tag{35}$$

is familiar as the mean-spherical approximation [60].

In view of these uncertainties it is fortunate that there are other lines of argument that lead to an exact expression for the surface tension in terms of the direct correlation function of the real inhomogeneous fluid. This expression can be obtained in several ways, such as the change in the grand potential Ω that arises from an increase in surface area caused by a fluctuation in density, or from the change of pressure in the liquid needed to transform a planar interface into a spherically curved one. This last method uses Eq. (31), the second of Yvon's two equations. The result is

$$\gamma = \tfrac{1}{4}kT \int_{-\infty}^{\infty} \rho'(z_1)\, dz_1 \int (x_2^2 + y_2^2)\rho'(z_2)c^{(2)}(\mathbf{r}_1,\mathbf{r}_2)\, d\mathbf{r}_2 \tag{36}$$

This result was obtained first by Yvon in 1948 and reported by him at a

meeting in Brussels that year but was never formally published. It was obtained independently by Triezenburg and Zwanzig in 1972 [61] and again by Lovett et al. in 1973 [62]. Like most exact equations it suffers from the disadvantage that the essential function $c^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ is generally not known, and attempts to guess it are apt to lead to a result close to the approximate equation, (11) with (34).

The two expressions for the surface tension, Eq. (26) with the appropriate expression for the pressure tensor, and Eq. (36), are very different. The first is not restricted to a sharp interface, although it is easy to show that for such an interface and with the mean-field approximation for the liquid the virial expression of Kirkwood and Buff reduces to that of Laplace. The second obviously takes into account explicitly the diffuseness of the interface, but its relation to the virial expression is not obvious. Schofield [63] was the first to show the equivalence of the two expressions; Waldor and Wolf [64] have recently rederived the same result by a different method.

It is, perhaps, appropriate to close this chapter by noting that the essential feature of a theory of inhomogeneous fluids is the need to account explicitly for the "nonlocal" character of all properties. By this is meant that the property of the fluid at any point $\mathbf{r}$ is determined not only by the local density at $\mathbf{r}$, $\rho(\mathbf{r})$, and by the local temperature T and the local chemical potential μ (the last two being constant throughout a system at equilibrium) but also by the properties of the fluid near but not at $\mathbf{r}$. We can distinguish three levels of accuracy. First the nonlocal character is ignored, so that the free-energy density Ψ (for example) is held to be a function only of the two independent variables T and $\rho(\mathbf{r})$:

$$\Psi(\mathbf{r}) = \Psi[\rho(\mathbf{r});T] \tag{37}$$

This was the assumption made by Poisson which led him to conclude that a diffuse interface has a vanishing surface tension. It was made again in modern times by Tolman [65] and by Ono and Kondo [66].

The next level of accuracy is that of Fuchs, Rayleigh, and van der Waals, in which

$$\Psi(\mathbf{r}) = \Psi[\rho(\mathbf{r}),\nabla\rho(\mathbf{r});T] \tag{38}$$

This leads to the square-gradient approximation and is still a useful approximation for some purposes.

Finally, we have the formally exact results, exemplified by Eq. (36), in which the free-energy density is a function of the properties of the fluid at two points and the correlation between these points:

$$\Psi(\mathbf{r}_1) = \Psi[\rho(\mathbf{r}_1),\rho(\mathbf{r}_2),c^{(2)}(\mathbf{r}_1,\mathbf{r}_2);\, T] \tag{39}$$

The range over which the nonlocality must be considered is the so-called correlation length in the fluid, which is generally a measure of the range of the intermolecular forces. Near the critical point, however, it grows to macroscopic size, so a classical treatment that does not account accurately for the extent of this growth leads to a significant error in its description of how the surface tension vanishes at that point.

APPENDIX

Bikerman [29] proposed that Eq. (16) be replaced by

$$\frac{n_i(\mathbf{r})}{1 - bn(\mathbf{r})} = n_i^* \exp[-z_i e V(\mathbf{r})/kT] \tag{A1}$$

where

$$n(\mathbf{r}) = \sum_i n_i(\mathbf{r}) \tag{A2}$$

and n_i^* is a constant which, for a single symmetrical electrolyte, can be written

$$n_+^* = n_-^* = \tfrac{1}{2} n^\infty (1 - bn^\infty)^{-1} \qquad n_+^\infty = n_-^\infty = \tfrac{1}{2} n^\infty \tag{A3}$$

The charge density is given by

$$\rho(\mathbf{r}) = ze[n_+(\mathbf{r}) - n_-(\mathbf{r})] \tag{A4}$$

$$= -\frac{zen^\infty \sinh[y(\mathbf{r})]}{1 + 2bn^\infty \sinh^2[y(\mathbf{r})/2]} \tag{A5}$$

where $y(\mathbf{r})$ (or y, for simplicity) is $zeV(\mathbf{r})/kT$. In the Gouy–Chapman treatment, ρ_0, the charge density at the plate rises exponentially with $|V_0|$, the surface potential, but here it is restricted by the sizes of the ions to maximum value of $|ze/b|$. One integration of the Poisson–Boltzmann equation gives

$$\left(\frac{dy}{dx}\right)^2 = \frac{2\kappa^2}{bn^\infty} \ln\left(1 + 2bn^\infty \sinh^2 \frac{y}{2}\right) \tag{A6}$$

The total charge in the diffuse layer is proportional to the potential gradient at the electrode:

$$\begin{aligned} \sigma &= \frac{\epsilon kT}{ze}\left(\frac{dy}{dx}\right)_0 \\ &= \left[\frac{2\epsilon kT}{b} \ln\left(1 + 2bn^\infty \sinh^2 \frac{y}{2}\right)\right]^{1/2} \end{aligned} \tag{A7}$$

Since $x^{-1} \ln(1 + x) \leq 1$, this charge is lower than that predicted by the Gouy–Chapman treatment. There seems, however, to be no way of integrating (A6) analytically to obtain $V(x)$. On linearization, as in the Debye–Hückel regime, the leading term is independent of b.

REFERENCES

1. Young, T. (1805). An essay on the cohesion of fluids, *Philos. Trans. R. Soc. 95*:65. Reprinted with some additions (1807). In *Lectures on Natural Philosophy*, Johnson, London, Vol. 2, p. 649.
2. Laplace, [P. S.] (1806). *Traité de mécanique céleste, Supplément au dixième livre, sur l'action capillaire*, Courcier, Paris. English translation with a commentary, Bowditch, N. (1839). *Mécanique céleste by the Marquis de la Place*, Little, Brown, Boston, Vol. 4; reprinted in 1966 by Chelsea Publishing Co., New York.
3. Dupré, A. (1869). *Théorie mécanique de la Chaleur*, Gauthier-Villars, Paris.
4. Poisson, S. D. (1831). *Nouvelle théorie de l'action capillaire*, Bachelier, Paris.
5. Fuchs, K. (1888). Ueber Verdampfung, *Rep. Phys. 24*:141; Ueber den Zusammenhang von Oberflächenspannung, Oberflächendichte und oberflächlicher Wärmeentwicklung. Ein Beitrag zur Capillaritätstheorie, p. 298. Fuchs, K. (1889). Directe Ableitung einiger Capillaritätsfunctionen, *Sitzungsber. Kais. Akad. Wiss. Wien Sekt. IIa 98*:1362.
6. Rayleigh, Lord (1892). On the theory of surface forces. II. Compressible fluids, *Philos. Mag. 33*:209.
7. van der Waals, J. D. (1893). Thermodynamische Theorie der Capillariteit in de Onderstelling van Continue Dichtheidsverandering, *Verh. K. Akad. Wet. Amsterdam (Sect. 1) 1*(8). English translation, van der Waals, J. D. (1979). The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density, *J. Stat. Phys. 20*:197.
8. Thomson, J. (1871). Considerations on the abrupt change at boiling or condensing in reference to the continuity of the fluid state of matter, *Proc. R. Soc. 20*:1.
9. van der Waals, J. D. (1873). *Over de Continuiteit van den Gas- en Vloeistoftoestand*, Sijthoff, Leiden. English translation, van der Waals, J. D. (1988). On the continuity of the gaseous and liquid states, in *Studies in Statistical Mechanics* (J. S. Rowlinson, ed.), North-Holland, Amsterdam, Vol. 14, p. 121.
10. Rowlinson, J. S. (1988). Van der Waals and the physics of liquids, in *Studies in Statistical Mechanics* (J. S. Rowlinson, ed.), North-Holland, Amsterdam, Vol. 14, p. 1.
11. Nernst, W. (1888). Zur Kinetik der in Lösung befindlichen Körper. I. Theorie der Diffusion, *Z. Phys. Chem. 2*:613.
12. Planck, M. (1890). Ueber die Erregung von Electricität und Wärme in Electrolyten, *Ann. Phys. 39*:161. Planck, M. (1890). Ueber die Potentialdifferenz

zwischen zwei verdünnten Lösungen binärer Electrolyte, *Ann. Phys. 40*:561.
13. MacInnes, D. (1939). *The Principles of Electrochemistry,* Reinhold, New York, Chapters 8 and 13 and Appendix.
14. Guggenheim, E. A. (1929). The conception of electrical potential difference between two phases and the individual activities of ions, *J. Phys. Chem. 33*:842.
15. Henderson, P. (1907). Zur Thermodynamik der Flüssigkeitsketten, *Z. Phys. Chem. 59*:118. (1908). *Z. Phys. Chem. 63*:325.
16. Rowlinson, J. S. (1988). The excess surface tensions of simple binary mixtures, *J. Chem. Soc. Faraday Trans. 1 84*:4125.
17. Helmholtz, H. (1879). Studien über electrische Grenzschichten, *Ann. Phys. 7*:337.
18. Billiter, J. (1903). Eine Theorie der Kolloide und Suspensionen, *Z. Phys. Chem. 45*:307.
19. Billiter, J. (1903). Eine Theorie der Kolloide und Suspensionen, *Z. Phys. Chem. 45*:307; see page 325.
20. Gouy, [G.] (1910). Sur la constitution de la charge électrique à la surface d'un électrolyte, *J. Phys. 9*:457.
21. Chapman, D. L. (1913). A contribution to the theory of electrocapillarity, *Philos. Mag. 25*:475.
22. Debye, P., and Hückel, E. (1923). Zur Theorie der Elektrolyte. I. Gefrierpunktserniedrigung und verwandte Erscheinungen; II. Das Grenzgesetz für die elektrische Leitfähigkeit, *Phys. Z. 24*:185, 305. English translation (1954). On the theory of electrolytes. I. Freezing-point depression and related phenomena; II. The limiting law for electrical conductivity. In *The Collected Papers of Peter J. W. Debye,* Interscience, New York, pp. 217, 264.
23. Rowlinson, J. S. (1989). The Yukawa potential, *Physica A 156*:15.
24. Fowler, R. H. (1929). *Statistical Mechanics,* University Press, Cambridge, Sections 8.7, 8.8, 13.6.
25. Onsager, L. (1933). Theories of concentrated electrolytes, *Chem. Rev. 13*:73.
26. Kirkwood, J. G. (1934). On the theory of strong electrolyte solutions, *J. Chem. Phys. 2*:767.
27. Stern, O. (1924). Zur Theorie der Elektrolytischen Doppelschicht, *Z. Elektrochem. 30*:508.
28. Langmuir, I. (1916). The constitution and fundamental properties of solids and liquids. 1. Solids, *J. Am. Chem. Soc. 38*:2221 (see p. 2267).
29. Bikerman, J. J. (1942). Structure and capacity of the electrical double layer, *Philos. Mag. 33*:384.
30. Dutta, M., and Baggchi, S. N. (1950). On the distribution of ions in solutions of strong electrolytes, *Indian J. Phys. 24*:61.
31. Eigen, M., and Wicke, E. (1951). Zur Theorie der starken Elekrolyte, *Naturwissenschaften 38*:453. Wicke, E., and Eigen, M. (1952). Zum Konzentrationsverlauf der Aktivitätskoeffizienten starker Elektrolyte, *Naturwissenschaften 39*:545.
32. Loeb, A. L. (1951). An interionic attractive theory applied to the diffuse layer around colloid particles, *J. Colloid Sci. 6*:75.

33. Levine, S., and Outhwaite, C. W. (1978). Comparison of theories of the aqueous electric double layer at a charged plane interface, *J. Chem. Soc. Faraday Trans. 2 74*:1670.
34. Langmuir, I. (1938). The role of attractive and repulsive forces in the formation of tactoids, thixotropic gels, protein crystals and coacervates, *J. Chem. Phys. 6*:873.
35. Derjaguin, B., and Kussakov, M. (1939). Anomalous properties of thin polymolecular films, *Acta Physicochim. URSS 10*:25, 153.
36. Levine, S. (1939). Problems of stability in hydrophobic colloidal solutions, *Proc. R. Soc. A 170*:145, 165.
37. Verwey, E. J. W., and Overbeek, J. Th. G. (1948). *Theory of the Stability of Lyophobic Colloids*, Elsevier, Amsterdam, p. 21.
38. Rideal, E. K., and others (1940). The electrical double layer, *Trans. Faraday Soc. 35*:1-322, 711-732.
39. Landau, L., and Lifshitz, E. (1935). On the theory of the dispersion of magnetic permeability in ferromagnetic bodies, *Phys. Z. Sowjetunion 8*:153. Reprint (1965). In *Collected Papers of L. D. Landau* (D. ter Haar, ed.). Pergamon Press, Oxford, p. 101.
40. Mitsui, T., and Furuichi, J. (1953). Domain structure of Rochelle salt and KH_2PO_4, *Phys. Rev. 90*:193.
41. Cahn, J. W., and Hilliard, J. E. (1958). Free energy of a nonuniform system. I. Interfacial free energy, *J. Chem. Phys. 28*:258. Cahn, J. W. (1959). II. Thermodynamic basis, *J. Chem. Phys. 30*:1121. Cahn, J. W., and Hilliard, J. E. (1959). III. Nucleation in a two-component incompressible fluid, *J. Chem. Phys. 31*:688.
42. Kirkwood, J. G., and Buff, F. P. (1949). The statistical mechanical theory of surface tension, *J. Chem. Phys. 17*:338.
43. Irving, J. H., and Kirkwood, J. G. (1950). The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics, *J. Chem. Phys. 18*:817 (Appendix).
44. Harasima, A. (1953). Statistical mechanics of surface tension, *J. Phys. Soc. Jpn. 8*:343.
45. Yvon, J. (1935). La théorie statistique des fluides et l'equation d'etat, in *Actualités scientifiques et industrielles*, No. 203, Hermann, Paris. Reprinted in Yvon, J. (1985). *Oeuvre Scientifique*, Comm. de l'Energie Atomique, Paris, Vol. 1, p. 37.
46. Kirkwood, J. G. (1935). Statistical mechanics of fluid mixtures, *J. Chem. Phys. 3*:300.
47. Born, M., and Green, H. S. (1946). A general kinetic theory of liquids. I. The molecular distribution functions, *Proc. R. Soc. A 188*:10. Reprinted with five further papers in Born, M., and Green, H. S. (1949). *A General Kinetic Theory of Liquids*, Cambridge University Press, Cambridge.
48. Bogoliubov, N. N. (1946). In Russian; English translation (1963). Problems of a dynamical theory in statistical physics, in *Studies in Statistical Mechanics* (J. de Boer and G. E. Uhlenbeck, eds.), North-Holland, Amsterdam, Vol. 1, p. 5.

49. Lee, T. D., and Young, C. N. (1960). Many-body problems in quantum statistical mechanics. IV. Formulation in terms of average occupation numbers in momentum spaces, *Phys. Rev. 117*:22.
50. Green, M. S. (1960). On the theory of the critical point of a simple fluid, *J Chem. Phys. 33*:1403.
51. Morita, T., and Hiroike, K. (1961). A new approach to the theory of classical fluids, III, *Prog. Theor. Phys. 25*:537.
52. De Dominicis, C. (1962). Variational formulations of equilibrium statistical mechanics, *J. Math. Phys. 3*:983.
53. Stillinger, F. H., and Buff, F. P. (1962). Equilibrium statistical mechanics of inhomogeneous fluids, *J. Chem. Phys. 37*:1.
54. Lebowitz, J. L., and Percus, J. K. (1963). Statistical mechanics of nonuniform fluids, *J. Math. Phys. 4*:116.
55. Mermin, N. D. (1965). Thermal properties of the inhomogeneous electron gas, *Phys. Rev. A 137*:1441.
56. Ebner, C., and Saam, W. F. (1975). Renormalized density functional theory of non-uniform superfluid ^{4}He at zero temperature, *Phys. Rev. B 12*:923. Ebner, C., Saam, W. F., and Stroud, D. (1976). Density-functional theory of classical fluids. I. Surfaces, *Phys. Rev. A 14*:2264.
57. Yvon, J. (1958). Note sur un calcul de perturbation en mécanique statistique, *Nuovo Cimento (Suppl.)* 9:144.
58. Evans, R. (1979). The nature of the liquid–vapour interface and other topics in the statistical mechanics of non-uniform classical fluids, *Adv. Phys. 28*:143.
59. Rowlinson, J. S., and Widom, B. (1982). *Molecular Theory of Capillarity*, Oxford University Press, Oxford.
60. Lebowitz, J. L., and Percus, J. K. (1966). Mean spherical model for lattice gases with extended hard cores and continuum fluids, *Phys. Rev. 144*:251.
61. Triezenberg, D. G., and Zwanzig, R. (1972). Fluctuation theory of surface tension, *Phys. Rev. Lett. 28*:1183.
62. Lovett, R., DeHaven, P. W., Vieceli, J. J., and Buff, F. P. (1973). Generalized van der Waals theories for surface tension and interfacial width, *J. Chem. Phys. 58*:1880.
63. Schofield, P. (1979). The statistical theory of surface tension, *Chem. Phys. Lett. 62*:413.
64. Waldor, M. H., and Wolf, D. E. (1986). Equality of two different formulas for the surface tension: a direct proof, *J. Chem. Phys. 85*:6082.
65. Tolman, R. C. (1948). Consideration of Gibbs theory of surface tension, *J. Chem. Phys. 16*:658. (1949). The superficial density of matter at a liquid–vapor boundary, *J. Chem. Phys. 17*:118. (1949). The effect of droplet size on surface tension, *J. Chem. Phys. 17*:333.
66. Ono, S., and Kondo, S. (1960). Molecular theory of surface tension in liquids, in *Encyclopedia of Physics* (S. Flügge, ed.), Springer-Verlag, Berlin, Vol. 10, p. 134.

2

Statistical Mechanical Sum Rules

J. R. Henderson
University of Leeds
Leeds, England

I. INTRODUCTION

The theory of inhomogeneous fluid phenomena has been transformed during the last decade by an explosive input from physicists, into what had previously been a long-established branch of physical chemistry (Derjaguin, 1940). Particularly striking are new conceptual frameworks obtained from rigorously based models (Nakanishi and Fisher, 1982) and a highlighting of the significance of fluctuations to interfacial phenomena (Lipowsky, 1987). For example, in just one subject known as wetting (Sullivan and Telo da Gama, 1986) we now possess a unified theory of all equilibrium behavior associated with adsorption at interfaces. We are likely to see an increasing emphasis on complex fluids (Chapters 11 and 12) and in dynamic phenomena such as interfacial motion (spreading and film growth) and transport at surfaces and in confined media (Chapter 13).

The interest by physicists was driven by the consequences of inhomogeneity rather than a desire to study fluids; in particular, excitement centered on newly understood phase transitions (Cahn, 1977; Ebner and Saam, 1977). In fact, as explained in Section II.B, the physics of interfacial phenomena has opened up a new world inhabited by a myriad of phase transitions, including new classes of critical phenomena. Here the pioneering work has come from phenomenological physics, such as Landau

dau or van der Waals theories (Sullivan, 1981) and interfacial Hamiltonians (Lipowsky, 1984), and also from rigorously based lattice models (Pandit et al., 1982), including some exact results (Abraham, 1980) and Monte Carlo simulations (Binder et al., 1986). However, liquid-state theory based on molecular Hamiltonians has much to offer that complements the foregoing work (van Swol and Henderson, 1984), thanks in particular to the availability of powerful techniques borrowed from functional calculus. In this chapter we are concerned with contributions to the fundamental understanding of equilibrium interfacial phenomena obtainable from liquid-state theory founded on realistic many-body Hamiltonians.

Much of what is discussed below can be traced back to seminal work in the 1950s and 1960s that was focused toward an understanding of liquid–vapor interfaces in simple atomic fluids. In the event, the so-called simple liquid–vapor interface turned out to be an extremely difficult problem from a fundamental point of view, due to the essential role that external fields play in broken symmetry systems (Section II.A). Reviews of this work can be found in Evans (1979), Percus (1982), Rowlinson and Widom (1982), and Henderson (1986c). Here it is convenient to avoid repetition of the painful gestation by presenting a unified discussion of fluids in the presence of external fields. In particular, let us seek to establish a general context appropriate to the many varied problems that have recently been tackled from within this single mathematical framework.

In this chapter the term *statistical mechanical sum rule*, hereafter shortened to *sum rule*, refers to an equilibrium identity between a statistical thermodynamic property and an integral over correlation functions. The former are macroscopic physical quantities such as a free energy, a generalized thermodynamic field, or derivatives of a free energy with respect to thermodynamic fields. To be a sum rule a relationship must be exact (i.e., derivable from a partition function defined by a microscopic Hamiltonian). The significance of sum rules lies in the direct links they create between microscopic correlations and physical properties. Thus sum rules provide a framework for interpreting molecular models of physical phenomena. Of special importance is the fact that all macroscopic phenomena appropriate to a given Hamiltonian must be compatible with any sum rule derived from that Hamiltonian. Hence dramatic collective behavior, such as associated with phase transitions or the appearance of structural order, is often highlighted by deceptively simple sum rules (see Section IV.A). Computer simulation studies benefit from sum rules, both as a framework with which to analyze masses of computer-generated data and as checks on equilibration and incorrect procedures. Density functional theories (Chapter 3), especially those constructed to possess full internal statistical mechanical consistency, are similarly enhanced by a

close attention paid to sum rules (see Section III.D). In contrast, approximate integral equation theories invariably violate key sum rules, pointing the way to future improvements.

It is not our purpose in this chapter to attempt a complete survey of specific applications of sum rule analyses applied to inhomogeneous fluids, but rather, to present a general introduction illustrated by selected examples; the reader might also like to note a recent review by Evans and Parry (1990). In Section II we present an overview of a collection of important concepts known as broken symmetry, as relevant to inhomogeneous fluids. It is important to review these concepts because they form the context in which theoreticians derive and use sum rules applied to interfacial systems. Perhaps the greatest thrill of theoretical science is to be able to see such deep physical understanding emerge miraculously, yet inevitably, from the dry logic of statistical mechanical sum rules. In Section III we present key sum rules of particular importance to the study of inhomogeneous fluids and review their derivation from statistical mechanical theory. In Sections IV and V we discuss examples of sum rule analyses of wall–fluid interfaces and confined fluids, reflecting the tastes of the present author. Such powerful methods are always available for the study of equilibrium properties of inhomogeneous molecular systems, thanks in particular to the utility of functional calculus.

II. BROKEN SYMMETRY

A. Importance of External Fields

Equilibrium states of a translationally invariant Hamiltonian must themselves be translationally invariant. This is particularly significant for fluid states because fluid interfaces lack rigidity and thus readily develop fluctuations that lead to instability in the absence of stabilizing boundaries or surface fields. Thus the presence of an external field term in the Hamiltonian is essential to theories of inhomogeneous fluids. A formal proof of this statement is provided by a theorem due to Mermin (1965), reviewed in Chapter 3, which implies that the one-body density profile $\rho(\mathbf{r})$ is uniquely determined by the one-body external field $v(\mathbf{r})$, and vice versa.

An external field can play up to three direct roles in controlling the behavior of inhomogeneous fluids. First, the symmetry of $v(\mathbf{r})$ will invariably define the interfacial geometry. This includes the important case of confined fluids, where a pore width or radius is defined by a parameter contained in the wall–fluid potential. Second, $v(\mathbf{r})$ will often be defined by a set of parameters that act as surface fields controliing the microscopic structure of wall–fluid interfaces and in special circumstances the mac-

roscopic behavior as well. The repulsive part of $v(\mathbf{r})$ will induce oscillatory layering structure in fluids packed against walls by moderate to high pressure. An attractive contribution to $v(\mathbf{r})$ will always contain at least two important parameters: a well depth and a range parameter, which play key roles in determining adsorption phenomena (wetting). Third, an external field can be used to stabilize an interface against instabilities arising from collective fluctuations. For example, a weak gravitational field is sufficient to stabilize a liquid–vapor interface from the wandering induced by capillary-wave fluctuations (see Section II.E).

An important class of indirect external field effects arises because boundary conditions cannot be ignored when considering phase transitions and when comparing results from different choices of ensemble. Often these macroscopic phenomena run counter to intuition obtained from work based on the strict thermodynamic limit. For example, consider the consequences of planar interfacial symmetry on wetting transitions at wall–vapor interfaces. Namely, if one induces a wetting transition (first order or continuous) by varying temperature or a surface field, the final outcome is a system with an infinitely thick film of liquid separating the wall from bulk vapor (which constitutes the far boundary condition). Clearly, this interfacial phase transition cannot be reversed; instead, to dry the wall one would now have to induce a drying transition, thereby creating an infinitely thick layer of vapor between the wall and the infinitely thick liquid film. Similar comments apply to finite systems as used in computer simulation studies. Related to these effects is the significance of boundary conditions to collective dynamics. Thus a planar liquid film does not have the same collective modes available to it as does a curved meniscus, which can spread out along a boundary wall. One should therefore bear in mind that a particular choice of geometry may induce extremely high metastable barriers preventing passage to true equilibrium states. Interesting examples of boundary effects have also been highlighted in studies of confined fluids. For example, models of capillary condensation based on pores of infinite length possess metastable barriers to pore filling (a drop must nucleate inside the pore) and to pore emptying (a bubble must nucleate inside the pore), whereas a finite open pore can empty from the pore ends without encountering significant metastability (provided that the system is not immersed in a bath of liquid) (Saam and Cole, 1975; Marini Bettolo Marconi and van Swol, 1989). Similarly, very strongly confined fluids exhibit extreme sensitivity to the choice of ensemble; for example, modeling a planar slit of width L in the canonical ensemble (fixed N, closed pore) will yield a true two-dimensional system as L tends to zero, whereas in this limit an open pore at finite chemical potential (grand canonical ensemble) can only approach a two-dimen-

sional weak gas, due to fluid being squeezed out from the pore ends (Henderson, 1986b).

B. Phase Diagrams and Surface Phase Transitions

Experimental science tends to restrict physical understanding to that available from the use of variables that are readily to hand in the laboratory. This leads to a limited view of phenomena that are better understood as part of a larger reality—like the ant confined to a two-dimensional table top. The most important collection of physical knowledge is summarized in phase diagrams, which plot the points, lines, and surfaces that separate distinct phases of matter distinguishable via their macroscopic properties. Passage across such a boundary in phase space is known as a *phase transition*. Describing and understanding phase transitions, a favorite pastime of physicists, is of special significance because here one controls large changes in physical behavior through small changes in thermodynamic variables.

A proper understanding of phase behavior is based on the identification of thermodynamic fields (Griffiths and Wheeler, 1970). Fields are those thermodynamic variables that always yield identical values in two or more phases that lie in thermodynamic equilibrium. Thus all-field phase diagrams are the most succinct way to represent phase behavior. Furthermore, each thermodynamic field acts as a degree of freedom relevant to the Gibbs phase rule. In general, thermodynamics proceeds by constructing a thermodynamic potential to be a concave function of all relevant fields (and thus the potential itself can be regarded as a field). Taking partial derivatives of the potential with respect to each field generates a set of conjugate densities (defined to be intensive variables by appropriate division with respect to volume or area, as necessary). One then chooses spaces from within this collection of fields and densities in which to describe physical phenomena under study. The topology of such a phase space is controlled by the number of densities used. In the following section an explicit version of such a theory is given for inhomogeneous fluids, and Section II.D introduces the development of statistical mechanical models that are mathematically exact realizations of this structure.

From the above it follows that a failure to identify thermodynamic fields relevant to a class of phenomena will severely cripple attempts at a rigorous understanding, particularly with regard to conceptual meaning. This lesson has been highlighted more than once during the recent expansion of theories of inhomogeneous fluids. Here, the problem can be broken

down into two questions. First, what bulk thermodynamic fields remain applicable in inhomogeneous systems? In general, the answer is temperature (T) and chemical potential (μ) but not pressure, because interfacial stress must be described by pressure tensor component profiles (Section II.D). Both T and μ can be defined at any point within an inhomogeneous system such that each take on a constant value everywhere throughout an equilibrium state; T is given by the average molecular kinetic energy (in classical systems) and μ by a statistical mechanical sum rule known as potential distribution theory (Section II.D). Second, do interfacial systems possess additional thermodynamic fields arising from the presence of surfaces? It was the latter question that remained largely unrecognized until the 1980s.

For example, consider wetting phenomena in the context of fluid adsorbed at a wall. Can a wall–fluid interface exhibit two or more states of adsorption at the same grand potential? If so, varying the temperature will be one route to induce interfacial phase transitions, but are there field variables present with more direct impact? In particular, note that a mathematical description of the wall–fluid interaction will introduce at least two new parameters; a well depth (ϵ_w) and a range (a_w) of the wall–fluid attractive potential energy. Since we are asking whether two equilibrium states of adsorption correspond to the same values of (T,μ,ϵ_w,a_w), it follows that ϵ_w and a_w act as generalized thermodynamic fields. One realizes immediately that if the fluid is held at bulk liquid–vapor coexistence at a given temperature, a sufficiently high value of ϵ_w will induce complete wetting of a wall–vapor interface (liquid prefers to lie next to the wall). At the other extreme, too little attraction will ensure that a wall–liquid interface can lower its free energy by interposing an infinite film of vapor between the wall and bulk liquid (a drying transition). Thus the key phase diagram for understanding wetting phenomena lies in the space (T,μ,ϵ_w); usually, one subtracts from μ its value at bulk liquid–vapor coexistence, $\mu_{sat}(T)$, so that wetting transitions are confined to the plane $\mu - \mu_{sat}(T) = 0$. In general, Nakanishi and Fisher (1982) argue that for wall–fluid systems one expects to find a tricritical point on the wetting transition line separating first-order transitions at lower T from second-order behavior (called *critical wetting*). Adding the field a_w will turn such a tricritical point into a tricritical line; thus the key phase diagram for recording the separatrix between first-order and critical wetting lies in the field space (ϵ_w,a_w). It is perhaps worth noting that ϵ_w has a direct magnetic analog: namely, ϵ_w translates to a surface magnetic field strength and its conjugate density to the surface magnetization [see (56)]. Experimental progress in studying wall–fluid wetting phenomena has been severely hampered by a lack of control over the fields (ϵ_w,a_w), in

complete contrast to statistical mechanical theory and computer simulation studies, although pioneering work by Durian and Franck (1987) has shown the way to the development of early approaches toward laboratory control. The future will surely see such work gaining enormous technological significance.

Another example of the significance of statistical physics to inhomogeneous fluid phenomena concerns the subject of fluids absorbed in porous media (confined fluids). The basic parameter of a pore is its width (L) or radius (R) (i.e., the extent along the confining direction). Two-phase coexistence of fluid within an open pore (equilibrated with an outside reservoir via exchange of particles through the ends of the pore) is known as capillary condensation or evaporation and has long been observed as a shift in the bulk liquid–vapor coexistence curve.* A full conceptual understanding of this phenomena requires the realization that L (or R) plays the role of yet another thermodynamic field. Thus capillary condensation takes place in the space (T,μ,L). The density conjugate to L is known as the *solvation force* (or *disjoining pressure*) (Evans and Marini Bettolo Marconi, 1987). The field ϵ_w is also important, since capillary condensation will interact strongly with any nearby instability toward wetting at a pore wall. Also, it is thermodynamically possible for layering transitions at the pore wall(s) to precede capillary condensation.

The foregoing considerations constitute a qualitative explanation of the striking richness of the phase behavior of inhomogeneous fluids. Each new relevant field increases phase space by an extra dimension and introduces additional potential phase transitions, following the Gibbs phase rule. Recent work on theories of single-component atomic fluids adsorbed at walls and in pores has identified examples of layering transition sequences, prewetting transitions between thin and thick films, various classes of wetting transitions, and capillary condensation transitions. Generalizations to molecular fluids or fluid mixtures would greatly enhance this complexity (Section II.C). Furthermore, the full mathematical technology of thermodynamics is immediately applicable to these extended phase spaces (i.e., Maxwell relations, Clapeyron equations, C_p–C_v relations, etc.). This includes standard approaches to phase transitions, such as mean-field van der Waals loops and renormalization group critical phenomena but now involving the newly identified thermodynamic variables. Surface critical phenomena have excited particular theoretical interest

*Strictly speaking, one should limit this use of the phrase two-phase coexistence to systems in which the fluid remains unconfined in at least two dimensions (slit pores). However, the effects of finite-size rounding of phase transitions are usually exponentially small and thus will rarely be seen (Evans, 1990).

because of the association with exponent relations, scaling theories, and dramatic collective fluctuations (soft modes). In some cases these modes are familiar from bulk statistical mechanics, but here they possess the character of a lower dimensionality (the dimensionality of the interface), while other surface critical phenomena concern interfacial modes such as capillary waves.

Explicit examples of the foregoing conclusions are discussed below; here, let us end by noting that phase diagrams of inhomogeneous fluid phenomena are not always restricted to spaces belonging to the set of fields and densities discussed above, where one can be assured of the usual rules. Instead, some workers have found it useful to introduce hybrid surface–bulk diagrams, in which surface thermodynamic variables are mixed with purely bulk variables. For example, in discussing wetting transitions it has been common to plot, at fixed (ϵ_w, a_w), the transition point and any associated prewetting transition line, superimposed on the bulk liquid–vapor coexistence curve drawn in (T,ρ) space. T is a valid surface field, but ρ is merely the bulk fluid density in equilibrium with the interface. Given the bulk equation of state, (T,ρ) can be transformed into (T,μ), which are both true surface fields. Similar comments can be made about adsorption isotherms for confined fluids plotted versus a bulk pressure p belonging to a reservoir with which the pore fluid is defined to be in thermodynamic equilibrium; in particular, p is not related to any pressure tensor component describing the inhomogeneous fluid.

C. Statistical Thermodynamics of Broken Symmetry

Gibbs (1906) has shown how to construct thermodynamic theories of interfacial properties that remain consistent with rigorous models without the need for explicit statistical mechanical solutions of inhomogeneous systems. The trick is to introduce mathematically defined dividing surfaces, equivalent to factoring the partition function into a bulk term (hereafter assumed to be known or defined) and surface terms; the latter are simply the parts left over. For model wall–fluid interfaces the natural choice of dividing surface lies on the equipotential of infinite wall–fluid repulsion. This choice has the benefit that at fixed geometry the dividing surface is independent of all relevant thermodynamic fields. The statistical thermodynamics of inhomogeneous fluids is invariably most conveniently based on Ω, the *grand-canonical potential* (often shortened to *grand potential*). This is because most fluid interfaces exchange molecules with neighboring bulk fluids; consider, for example, a wall–liquid interface

modeled in semi-infinite geometry, or liquid confined to an open pore immersed in a reservoir of gas.

Thus, following Gibbs, let us define the surface excess grand potential with respect to a choice of dividing surface as

$$\Omega^{ex} = \Omega - \Omega_b \tag{1}$$

where subscript b denotes the usual bulk term. Then the second law of thermodynamics applied to a planar wall–fluid interface of area A, such as discussed in Section II.B, will read

$$d\Omega^{ex} = -S^{ex}\,dT - A\Gamma\,d\mu + \left(\frac{\Omega^{ex}}{A}\right)dA - A\Theta\,d\epsilon_w - A\Psi\,da_w \tag{2a}$$

where S^{ex} is the surface excess entropy defined analogous to (1), and Γ denotes adsorption (i.e., $N^{ex} \equiv A\Gamma$, where N is the total number of fluid molecules). If the system was bounded by two planar walls of separation L (a slit pore), then, in addition to doubling the single wall terms appearing on the right side of (2a), one would need to add the term

$$-Af\,dL \tag{2b}$$

where f is known as the *solvation force*.

Underlying (2) is the fact that the thermodynamic potential can be expressed as a function of all relevant field variables [i.e., $\Omega^{ex}(T,\mu,A,\epsilon_w,a_w,L)$]. The coefficients of each term on the right side of Eq. (2) are the thermodynamic densities conjugate to each field:

$$\frac{\partial(\Omega^{ex}/A)}{\partial T} = \frac{-S^{ex}}{A} \tag{3a}$$

$$\frac{\partial(\Omega^{ex}/A)}{\partial \mu} = -\Gamma \tag{3b}$$

$$\frac{\partial \Omega^{ex}}{\partial A} = \frac{\Omega^{ex}}{A} \tag{3c}$$

$$\frac{1}{A}\frac{\partial \Omega}{\partial \epsilon_w} = -\Theta \tag{3d}$$

$$\frac{1}{A}\frac{\partial \Omega}{\partial a_w} = -\Psi \tag{3e}$$

$$\frac{\partial(\Omega^{ex}/A)}{\partial L} = -f \tag{3f}$$

where Ω has appeared in (3d) and (3e) to highlight the fact that ϵ_w and

a_w are strictly surface fields. In Section II.D we explain how statistical mechanics leads to sum rule expressions for the densities above; in particular, Γ, Θ, Ψ, and f are all defined by simple integrals over the one-body density profile and thus act as order parameters for interfacial structure. It follows that further differentiation of these densities with respect to the fields (μ,ϵ_w,a_w,L) generates surface compressibilities.

Equations (2) and (3) admit many generalizations. One example would be the addition of external electric field terms; see Chapters 5 and 6 for explicit results concerning inhomogeneous plasmas and electrolytes. Generalizations to molecular fluids and fluid mixtures are straightforward and will be touched on briefly later in this chapter. Here, let us confine our remarks to noting the rapid increase in complexity that such generalizations must bring (i.e., the large growth in the number of relevant field variables). For example, a molecular fluid interacts with a wall not just through center of mass fields (ϵ_w,a_w) but also through fields that act on orientational degrees of freedom. Thus molecular fluids exist within an extended space that allows for such phenomena as orientational wetting. The situation with mixtures is particularly complex in general, since each additional fluid component will involve not just one extra chemical potential field but will also add a new member to every class of surface field; the latter fields directly control phenomena such as selective adsorption. In this way, (2) enables us to truly appreciate the scale of the complexity of phase behavior belonging to the natural world.

An important paper by Evans and Marini Bettolo Marconi (1987) has emphasized that the standard mathematical procedures of bulk thermodynamics are equally applicable to inhomogeneous fluid problems, such as defined by (2). For example, a plethora of surface Maxwell relations follow directly from (2) and (3), such as

$$\frac{\partial\Gamma}{\partial\epsilon_w} = \frac{\partial\Theta}{\partial\mu} \tag{4a}$$

$$\frac{\partial\Gamma}{\partial L} = \frac{\partial f}{\partial\mu} \tag{4b}$$

and defining additional functions (e.g., $Z \equiv \Omega^{ex} + N^{ex}\mu$) yields many more of these relations:

$$\left(\frac{\partial\mu}{\partial\epsilon_w}\right)_\Gamma = -\left(\frac{\partial\Theta}{\partial\Gamma}\right)_{\epsilon_w} \tag{4c}$$

$$\left(\frac{\partial\mu}{\partial L}\right)_\Gamma = -\left(\frac{\partial f}{\partial\Gamma}\right)_L \tag{4d}$$

Further standard manipulations lead to results mathematically equivalent to the well-known C_p–C_v relation; for example, at fixed $\{T,A,a_w\}$ we obtain from $\Theta(\mu,\epsilon_w)$ and $\Theta(\Gamma,\epsilon_w)$ a result of particular significance to critical wetting transitions (Evans and Parry, 1989):

$$\left(\frac{\partial\Theta}{\partial\epsilon_w}\right)_\Gamma - \left(\frac{\partial\Theta}{\partial\epsilon_w}\right)_\mu = \left(\frac{\partial\Theta}{\partial\mu}\right)_{\epsilon_w}\left(\frac{\partial\mu}{\partial\epsilon_w}\right)_\Gamma \tag{5a}$$

$$= -\left(\frac{\partial\Gamma}{\partial\epsilon_w}\right)_\mu^2\left(\frac{\partial\Gamma}{\partial\mu}\right)_{\epsilon_w}^{-1} \tag{5b}$$

where the second version follows from (4a) and $\Gamma(\mu,\epsilon_w)$.

Phase coexistence in inhomogeneous fluids can be analyzed similarly. In particular, consider all the Clapeyron equations that follow from inserting (2) into

$$d\Omega_\alpha^{ex} = d\Omega_\beta^{ex}$$

Thus the slope of phase coexistence at fixed (A,ϵ_w,a_w,L) is determined by

$$\left(\frac{d\mu}{dT}\right)_{cc} = -\frac{(S_\alpha^{ex} - S_\beta^{ex})/A}{\Gamma_\alpha - \Gamma_\beta} \tag{6a}$$

where cc denotes coexistence curve. Similarly, at fixed (T,A,ϵ_w,a_w),

$$\left(\frac{d\mu}{dL}\right)_{cc} = -\frac{f_\alpha - f_\beta}{\Gamma_\alpha - \Gamma_\beta} \tag{6b}$$

and so on. Evans and Marini Bettolo Marconi (1987) point out that mean-field theories of first-order transitions will yield van der Waals loops in isotherms of order parameters such as Γ, Θ, Ψ, and f when plotted versus their conjugate fields. Furthermore, it follows from (3b) and (3d) to (3f), respectively, that the coexisting values of these order parameters are determined by applying equal-area constructions to the loops. The vanishing of such loops yields standard criteria for criticality, but now in unfamiliar variables.

D. Grand-Canonical Ensemble

Explicit statistical mechanical realizations of the thermodynamic structure discussed in Section II.C follow from introduction of the grand partition function Ξ:

$$\Omega \equiv -kT \ln \Xi \tag{7a}$$

$$\Xi = \sum_{N=0}^{\infty} \frac{\Lambda^{-3N}}{N!} \int \prod_{i=1}^{N} di \exp\left(\frac{-H_N + \mu N}{kT}\right) \tag{7b}$$

where $\Lambda \equiv (h^2/2m\pi kT)^{1/2}$ is the translational de Broglie wavelength (i.e., let us assume classical dynamics) and H_N denotes an N-body potential. For models of inhomogeneous fluids we write

$$H_N \equiv \Phi(1\cdots N) + \sum_{i=1}^{N} v(i) \tag{7c}$$

where Φ is a fluid–fluid many-body potential and v is the external field. Note that H_N does not depend on the fields T and μ, while all the remaining fields appearing on the right side of (2) can be introduced exclusively in terms of the one-body potential. For example, at fixed a_w a basic model of a planar wall–fluid interface is

$$v(z,\epsilon_w) = \epsilon_w v_{\epsilon_w}(z) \tag{8}$$

and a simple model of confined fluids is given by a symmetric planar slit of width L, at fixed (ϵ_w, a_w):

$$v(z,L) = v^{\infty}(z) + v^{\infty}(L - z) \qquad [v^{\infty}(z) \equiv \infty, \quad z < 0] \tag{9}$$

where the superscript ∞ denotes the potential of a single planar wall situated at $z = 0$. It follows that it is not necessary to specify the intermolecular potential Φ in order to take derivatives of the grand potential with respect to field variables. In fact, at fixed temperature the effects of Φ can be entirely subsumed within the distribution functions. Thus isothermal statistical mechanics generates theories of inhomogeneous fluids that are completely general with respect to fluid–fluid intermolecular forces.

A formal statement of this statistical mechanical approach is given by the following hierarchy of functional derivatives, which follow immediately from (7):

$$\left(\frac{\delta\Omega}{\delta[\mu - v(1)]}\right)_T = -\rho(1) \tag{10a}$$

$$\left(\frac{\delta^2\Omega}{\delta[\mu - v(1)]\delta[\mu - v(2)]}\right)_T = -\frac{1}{kT}[\rho^{(2)}(12) - \rho(1)\rho(2) + \rho(1)\delta(12)] \tag{10b}$$

and so on; that is, the sth functional derivative generates the s-body distribution function $\rho^{(s)}(1\cdots S)$. Note that (10a) constitutes a general sum rule expression for all of the field derivatives listed in (3), apart from (3a) and (3c) (van Swol and Henderson, 1986):

$$\left(\frac{\partial\Omega}{\partial\lambda_i}\right)_{T,\{\lambda_{j\neq i}\}} = -\int d1\ \rho(1)\frac{\partial}{\partial\lambda_i}[\mu - v(1)] \qquad \lambda_i \in \{\mu,\epsilon_w,a_w,L\} \tag{11}$$

Further differentiation and (10b) yields sum rules for a set of compressibilities:

$$\left(\frac{\partial^2\Omega}{\partial\lambda_i\,\partial\lambda_j}\right)_{T,\{\lambda_{k\neq i,j}\}} = -\frac{1}{kT}\int d1\,\rho(1)\frac{\partial}{\partial\lambda_i}[\mu - v(1)]\int d2\,\rho(2)\frac{\partial}{\partial\lambda_j}[\mu - v(2)]$$
$$\times\left[h(12) + \frac{\delta(12)}{\rho(1)}\right] - \int d1\,\rho(1)\frac{\partial^2}{\partial\lambda_i\,\partial\lambda_j}[\mu - v(1)]\quad \lambda_i,\lambda_j \in \{\mu,\epsilon_w,a_w,L\} \tag{12}$$

where we have introduced the total correlation function $h(12) \equiv [\rho^{(2)}(12) - \rho(1)\rho(2)]/\rho(1)\rho(2)$. In recognition of the importance of (12), the statistical mechanical theory above is referred to as the *compressibility route*. The compressibility route to the statistical mechanics of inhomogeneous fluids is of special significance because it focuses directly on the thermodynamic fields and densities of relevance, apart from the temperature and entropy. Thus (11) generates single integral sum rules for the densities Γ, Θ, Ψ, and f, and (12) highlights the microscopic nature of collective modes responsible for interfacial critical phenomena. The surface excess grand potential itself is a special case that is treated separately in Section III.C.

Other important results also follow from the hierarchy (10). In particular, consider the consequences of translational symmetry (i.e., applied to the entire system, including boundaries):

$$\delta\rho(\mathbf{r}) \equiv [\rho(\mathbf{r} + \boldsymbol{\Delta}) - \rho(\mathbf{r})] \to \boldsymbol{\Delta}\cdot\nabla\rho(\mathbf{r}) + \mathcal{O}\Delta^2$$
$$\delta v(\mathbf{r}) \equiv [v(\mathbf{r} + \boldsymbol{\Delta}) - v(\mathbf{r})] \to \boldsymbol{\Delta}\cdot\nabla v(\mathbf{r}) + \mathcal{O}\Delta^2 \tag{13}$$

Combining (13) with (10a) just confirms that finite pressure and finite volume go together [see (23b)]. However, (13) and (10b) yield a key integrodifferential equation for the density profile (Lovett et al., 1976; Wertheim, 1976):

$$\nabla^\alpha\rho(1) = -\frac{1}{kT}\left[\rho(1)\nabla^\alpha v(1) + \rho(1)\int d2\,\rho(2)h(12)\nabla^\alpha v(2)\right] \tag{14}$$

Note that integrating equations such as (14) across an interface generates potentially useful sum rules. Even in cases where $v(\mathbf{r})$ contains discontinuities it is straightforward to make use of the integral equations and sum rules above; one simply introduces the one-body y-function $n(\mathbf{r})$, defined by

$$\rho(\mathbf{r}) \equiv n(\mathbf{r})\exp\left[\frac{-v(\mathbf{r})}{kT}\right] \tag{15}$$

and uses the fact that graphical analysis will always prove that $n(\mathbf{r})$ is a continuous function, even across a hard wall boundary [see also (17) and (29)].

The statistical mechanical hierarchy (10) treats the grand potential as a functional of the one-body field, $\Omega[\mu - v]$. It is equally permissible to regard Ω as a functional of the one-body density, $\Omega[\rho]$, because at fixed (T,μ) we know that $\rho(\mathbf{r})$ is uniquely defined by $v(\mathbf{r})$ and vice versa (Mermin, 1965). This inversion of (10) generates a complementary hierarchy of correlation functions, called direct correlation functions. Since this approach forms the basis of density functional theory (Chapter 3), let me restrict the discussion here to listing a few key results:

$$\Omega[\rho] \equiv \mathcal{F}[\rho] - \int d1\ \rho(1)[\mu - v(1)] \tag{16a}$$

$$\mathcal{F}[\rho] \equiv \mathcal{F}_{ex}[\rho] + kT \int d1\ \rho(1)\ \{\ln[\Lambda^3\rho(1)] - 1\} \tag{16b}$$

$$\frac{\delta^s \mathcal{F}_{ex}[\rho]}{\delta\rho(1)\cdots\delta\rho(S)} \equiv -kTc^{(s)}(1\cdots S) \tag{16c}$$

The first two of these results decompose the grand potential into one-body terms plus the many-body free energy, $\mathcal{F}_{ex}$; here, the subscript refers to the excess over ideal free energy. The hierarchy (16c) is the analog of (10), with $c^{(s)}(1\cdots S)$ denoting the s-body direct correlation function. At equilibrium, the first member of this hierarchy can be rewritten as

$$\ln\ [\Lambda^3\rho(1)] = c^{(1)}(1) + \frac{\mu - v(1)}{kT} \tag{17}$$

that is, $-kTc^{(1)}(1)$ is the excess (over ideal) chemical potential. Applying (13) to the second member of (16c) and inserting this into the gradient of (17) yields the inverse of (14):

$$\nabla^\alpha\rho(1) = -\frac{1}{kT}\rho(1)\nabla^\alpha v(1) + \rho(1)\int d2\ c^{(2)}(12)\nabla^\alpha\rho(2) \tag{18}$$

The general statement of this functional inversion between density and external field, known as the Ornstein–Zernike equation, can be written compactly as

$$\int d3\ G(13)G^{-1}(32) = \delta(12) \tag{19a}$$

where I have used (10b) and (16c) and (17) to define

$$G(12) \equiv kT\frac{\delta\rho(1)}{\delta[\mu - v(2)]} = \rho^{(2)}(12) - \rho(1)\rho(2) + \rho(1)\delta(12) \tag{19b}$$

$$G^{-1}(12) \equiv \frac{1}{kT}\frac{\delta[\mu - v(1)]}{\delta\rho(2)} = \frac{\delta(12)}{\rho(1)} - c^{(2)}(12) \tag{19c}$$

Macroscopic symmetry will considerably simplify (19a); for example, in planar symmetry we have

$$\int_{-\infty}^{\infty} dz_3 G(z_1,z_3;Q)G^{-1}(z_3,z_2;Q) = \delta(z_{12}) \tag{20a}$$

where I have introduced the transverse Fourier transform

$$G(z_1,z_2;Q) \equiv \int d^{d-1}\mathbf{R}_{12} e^{i\mathbf{Q}\cdot\mathbf{R}_{12}} G(z_1,z_2,R_{12}) \tag{20b}$$

and similarly for its inverse; $\mathbf{R}_{12}$ lies in the plane of the surface and d denotes the overall dimensionality (e.g., in $d = 3$, $R_{12}^2 = x_{12}^2 + y_{12}^2$). Provided that G and G^{-1} can be expanded up to order Q^2,

$$G(z_1,z_2;Q) = G_0(z_1,z_2) + Q^2 G_2(z_1,z_2) + \cdots \tag{20c}$$

one can equate terms up to order Q^2 in (20a) to derive

$$G_2(z_1,z_2) = -\int_{-\infty}^{\infty} dz_3 \int_{-\infty}^{\infty} dz_4 G_0(z_1,z_4)G_2^{-1}(z_4,z_3)G_0(z_3,z_2) \tag{20d}$$

This result is important for an understanding of surface critical phenomena because it enables singular contributions to $G_2(z_1,z_2)$ to be deduced from divergences in G_0, given that G_2^{-1} is sufficiently well behaved (Evans and Parry, 1989). Of course, the above relies on the assumption that the decay of $G(z_1,z_2,R_{12})$ along the interface is not such as to render the integral

$$G_2(z_1,z_2) = -\frac{1}{2(d-1)} \int d^{d-1}\mathbf{R}_{12} R_{12}^2 G(z_1,z_2,R_{12}) \tag{20e}$$

ill defined. In Section III.C we shall see that $G_2(z_1,z_2)$ determines the interfacial tension; thus planar surface critical phenomena at which the interface remains intact (neither zero nor infinite surface tension) should always be described by (20) up to and including order Q^2. This is in contrast to Ornstein–Zernike theory of bulk critical phenomena, which breaks down in $d \leq 4$; that is, below some upper-critical dimension the disappearance of interfacial structure is associated with a nonzero value of the exponent η.

Compressibility route analyses of interfacial fluid phenomena are carried out without explicit reference to fluid–fluid intermolecular forces; instead, one proceeds via increasing familiarity with the behavior of the distribution functions $\rho(1)$ and $\rho^{(2)}(12)$. Alternatively, one can introduce specific molecular models of fluids

$$\Phi(1\cdots N) = \sum\sum_{i<j} \varphi^{(2)}(ij) + \sum\sum\sum_{i<j<k} \varphi^{(3)}(ijk) + \cdots \tag{21}$$

and generate the distribution function hierarchy by taking functional derivatives with respect to $\varphi^{(s)}$. This is the formal basis of the virial route to statistical mechanics; in particular, introducing a displacement field $\mathbf{e}(\mathbf{r})$, it follows that the leading-order change in free energy is given by

$$(\delta\Omega)_{T,\mu} = \langle -kT \sum_i \nabla\cdot\mathbf{e}_i + \sum_i \mathbf{e}_i\cdot\nabla v_i + \sum_i \mathbf{e}_i\cdot\nabla_i\Phi\rangle \tag{22}$$

which is the general expression for a virial equation of state; hereinafter $\langle\ \rangle$ denotes a statistical mechanical average defined by a partition function. It is convenient to introduce a pressure tensor $p^{\alpha\beta}$:

$$\nabla^\beta p^{\alpha\beta}(\mathbf{r}) \equiv kT\nabla^\alpha\rho(\mathbf{r}) + \langle\sum_i \delta(\mathbf{r} - \mathbf{r}_i)\nabla_i{}^\alpha\Phi\rangle \tag{23a}$$

$$= -\rho(\mathbf{r})\nabla^\alpha v(\mathbf{r}) \tag{23b}$$

(where the usual summation convention over Greek indices applies) so that (22) reduces to the usual stress–strain form [e.g., Henderson, 1986c)]:

$$(\delta\Omega)_{T,\mu} = -\int p^{\alpha\beta}\nabla^\beta e^\alpha + \int \rho e^\alpha\nabla^\alpha v \tag{24}$$

The step from the right side of (23a) to (23b) follows directly from the statistical mechanical definition of $\rho(\mathbf{r})$ [eq. (26c)]. The hierarchy of equations obtained from direct differentiation of the distribution functions is usually referred to as the YBG hierarchy (Yvon, 1935). Equation (23b) shows that formally the two terms on the right side of (24) cancel (i.e., the force exerted on the walls by the fluid balances the force exerted on the fluid by the walls). Thus (23b) expresses mechanical equilibrium. When applying the virial theorem to inhomogeneous fluids, one notes that it is the first term on the right side of (24) that concerns the work done on the fluid. For example, in planar symmetry it follows that [e.g., Henderson and van Swol (1984)]

$$\frac{\Omega}{A} = -\int_{-\infty}^{\infty} dz\, p_T(z) \tag{25a}$$

where p_T denotes the transverse component of the pressure tensor:

$$p^{\alpha\beta}(z) = \begin{pmatrix} p_T(z) & 0 & 0 \\ 0 & p_T(z) & 0 \\ 0 & 0 & p_N(z) \end{pmatrix} \tag{25b}$$

However, only the normal component is determined by (23) in planar symmetry:

$$p_N'(z) = -\rho(z)v'(z) \tag{25c}$$

More generally, (23) follows from consideration of the rate of change of momentum density:

$$j^{\alpha}(\mathbf{r},t) = -\nabla_{\mathbf{r}}{}^{\beta}\left[\sum_i \frac{p_i{}^{\alpha}p_i{}^{\beta}}{m_i}\delta(\mathbf{r}-\mathbf{r}_i)\right] - \sum_i \delta(\mathbf{r}-\mathbf{r}_i)(\nabla_i{}^{\alpha}\Phi + \nabla_i{}^{\alpha}v) \quad (26a)$$

$$\equiv \nabla^{\beta}\sigma^{\alpha\beta}(\mathbf{r},t) - \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla^{\alpha}v(\mathbf{r}) \quad (26b)$$

and the conservation of linear momentum (i.e., $\langle \dot{J}\alpha \rangle = 0$) implies (23), with

$$p^{\alpha\beta} \equiv -\langle \sigma^{\alpha\beta} \rangle \qquad \rho(\mathbf{r}) \equiv \langle \sum_i \delta(\mathbf{r}-\mathbf{r}_i) \rangle \quad (26c)$$

Equation (26) does not uniquely define the many-body contribution to the pressure tensor:

$$\nabla^{\beta}p_c{}^{\alpha\beta}(\mathbf{r}) \equiv \langle \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i{}^{\alpha}\Phi \rangle \quad (27)$$

In particular, for pair potential fluids, Schofield and Henderson (1982) have shown that (27) is consistent with an infinite class of pressure tensors, with the nonuniqueness expressed in terms of an arbitrary path integral representing lines of intermolecular stress. The result (25a) is invariant with respect to such a choice of pressure tensor, but not higher-order moments of $p_T(z)$. For computer simulation studies of pair potential models of inhomogeneous fluids a convenient choice is to adopt the pressure tensor of Irving and Kirkwood (1950):

$$p_{\mathrm{IK}}^{\alpha\beta}(\mathbf{r}) = kT\rho(\mathbf{r})\delta^{\alpha\beta} - \frac{1}{2}\int d\mathbf{r}_{12}r_{12}^{\alpha}r_{12}^{\beta}\frac{\varphi'(r_{12})}{r_{12}}\int_0^1 dl \times \rho^{(2)}[\mathbf{r} - l\mathbf{r}_{12},\mathbf{r} + (1-l)\mathbf{r}_{12}] \quad (28)$$

but one must beware of using this result to calculate expressions which, in contrast to (25a), are ill defined by (27).

In general, it is notoriously difficult to transform explicitly between complementary integral equations and sum rules generated by the compressibility route and the virial route, respectively [see, e.g., eqs. (14), (18), and (23)]. Thus it is of some interest to note that a specific correspondence between the two routes can be made (Henderson, 1983). The link is via a sum rule for the configurational chemical potential (i.e., the one-body direct correlation function), known as potential distribution theory (Widom, 1963):

$$c^{(1)}(\mathbf{r}) = \ln\left\langle \exp\left[-\frac{\Phi'(\mathbf{r})}{kT}\right]\right\rangle \quad (29a)$$

where $\Phi'(\mathbf{r})$ denotes the potential field due to a hypothetical test molecule fixed at position $\mathbf{r}$:

$$\Phi_N'(\mathbf{r}) = \Phi(1\cdots N+1)_{\mathbf{r}_{N+1}=\mathbf{r}} - \Phi(1\cdots N) \tag{29b}$$

The proof of (29) follows immediately from consideration of the partition function $\Xi'(\mathbf{r})$ defined by a molecular system (7c) in the presence of the many-body external field defined by (29b):

$$\Xi'(\mathbf{r}) = \exp\left[-\frac{\mu - v(\mathbf{r})}{kT}\right]\Lambda^3\rho(\mathbf{r})\Xi \tag{30a}$$

That is, the configurational chemical potential at position $\mathbf{r}$ is given by the work done, at constant (T,μ), to insert a test molecule at that position [see (17)]:

$$\begin{aligned}\Omega'(\mathbf{r},T,\mu) - \Omega(T,\mu) &\equiv -kT\ln\frac{\Xi'(\mathbf{r})}{\Xi} \\ &= -kTc^{(1)}(\mathbf{r})\end{aligned} \tag{30b}$$

Direct differentiation of (29a) yields (Henderson, 1983)

$$\nabla^\alpha c^{(1)}(\mathbf{r}) = -\frac{\left\langle \sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i^\alpha\Phi\right\rangle}{kT\rho(\mathbf{r})} \tag{31a}$$

or, using (17) together with $\nabla^\alpha\mu = 0$,

$$\nabla^\alpha\rho(\mathbf{r}) = -\frac{1}{kT}\left[\rho(\mathbf{r})\nabla^\alpha v(\mathbf{r}) + \left\langle\sum_i \delta(\mathbf{r}-\mathbf{r}_i)\nabla_i^\alpha\Phi\right\rangle\right] \tag{31b}$$

Thus, for a pure field, Eqs. (14), (18), and (23) are all equivalent expressions of mechanical equilibrium (hydrostatic stability). One can also use potential distribution theory to derive various statistical mechanical hierarchies based on alternative formulations of configurational chemical potential, such as Kirkwood–Hill scaling (Hill, 1959) and scaled particle theory (Reiss et al., 1960). Namely, if λ is any parameter that one chooses to introduce into the test particle field Φ', then $\partial c^{(1)}/\partial\lambda$ follows immediately from (29a), or alternatively, from the functional derivative $\delta\Omega'/\delta\Phi'$ (Henderson, 1983). Note that for pair potential fluids Φ' is a one-body field and $\rho'(\mathbf{r}')$, in the presence of a test molecule at $\mathbf{r}$, is just $\rho^{(2)}(\mathbf{r},\mathbf{r}')/\rho(\mathbf{r})$ in the real system (defined by unfreezing the test molecule); in this case by transforming to a frozen test-molecule system one can immediately make use of the compressibility route hierarchy (10). Thus, for pair potential fluids Kirkwood–Hill scaling and scaled particle theory are just specific versions of the general sum rule (11).

The test-molecule approach is easily generalized to encompass the entire distribution function hierarchy; in particular, (30a) generalizes to

$$\Xi^{t_1\cdots t_s}(\mathbf{r}_1\cdots\mathbf{r}_s) = (\Lambda^3 e^{-\mu/kT})^s \exp\left[\sum_{i=1}^{s}\frac{v(\mathbf{r}_i)}{kT}\right]\rho^{(s)}(\mathbf{r}_1\cdots\mathbf{r}_s)\Xi \quad (32a)$$

and (30b) generalizes to

$$\Omega^{t_1\cdots t_s}(\mathbf{r}_1\cdots\mathbf{r}_s,T,\mu) - \Omega(T,\mu) - \sum_{i=1}^{s}[\Omega^t(\mathbf{r}_i,T,\mu) - \Omega(T,\mu)] = -kT\ln[g^{(s)}(\mathbf{r}_1\cdots\mathbf{r}_s)] \quad (32b)$$

where $g^{(s)}$ denotes the s-body correlation function $\rho^{(s)}(\mathbf{r}_1\cdots\mathbf{r}_s)/\rho(\mathbf{r}_1)\cdots\rho(\mathbf{r}_s)$. The right side of (32b) is referred to as a potential of mean force, that is, the work done to insert s test molecules at positions $\mathbf{r}_1\cdots\mathbf{r}_s$ minus the work required to insert them singly at these same positions. In homogeneous fluids a potential of mean force is just the work required to bring test molecules together from infinity. Note that the partition function (32a) includes direct interactions between the test molecules; alternatively, one can leave this contribution out and then (32b) would yield the s-body y-function. Note also that the left sides of (30b) and (32b) are surface excess grand potentials, belonging to test-molecule systems. For homogeneous pair potential fluids (32) implies that

$$-kT\frac{\partial}{\partial r_{12}}\ln[g(r_{12})] = \frac{\partial}{\partial r_{12}}\Omega^{t_1t_2}(r_{12}) \quad (33)$$

Here the potential of mean force is a type of solvation force similar to (3f).

To conclude this survey of the grand ensemble applied to inhomogeneous fluids, let me indicate the nature of generalizations to mixtures and to molecular fluids. If the fluid has ν components, the partition function (7b) is extended to include a sum over μ_ν and in general, $v_\nu(\mathbf{r})$ (i.e., each type of molecule is associated with a particular one-body field). Functional derivatives of Ω with respect to these one-body fields generates $\rho_\nu(1)$ and $\rho^{(2)}_{\nu\nu'}(12)$, directly analogous to (10). Thus (11) to (15) are straightforwardly generalized to the case of mixtures. Similar remarks apply to the density functional formalism [Eqs. (16) to (20)] (i.e., one considers functional derivatives with respect to $\delta\rho_\nu$) (Lebowitz, 1964). The only point that requires some care is to note that whereas the correlation functions are defined by partial functional derivatives (all remaining variations set to zero), results such as the generalizations of (14), (18), and (19a) concern full variations and thus contain a sum over ν (i.e., each ρ_ν is a functional of all the external fields and each v_ν is a functional of all the

density profiles). The virial route is generalized to mixtures by extending (21) to include all the various classes of intermolecular potentials. The pressure tensor gradient, defined as (26), now involves a sum over components ν; however, the YBG hierarchy splits into separate equations (one for each component):

$$\nabla^\alpha \rho_\nu(\mathbf{r}) = -\frac{1}{kT}[\rho_\nu(\mathbf{r})\nabla^\alpha v_\nu(\mathbf{r}) + \langle \sum_{i_\nu} \delta(\mathbf{r} - \mathbf{r}_{i_\nu})\nabla^\alpha_{i_\nu}\Phi\rangle] \tag{34}$$

From the obvious generalization of potential distribution theory [i.e., (29a) with $\Phi_\nu'(\mathbf{r})$] it follows that for any given ν, (34) is equivalent to $\nabla\mu_\nu = 0$ (Henderson, 1983). Thus in mixtures the condition for diffusive equilibrium applies to each component separately, but not mechanical equilibrium [i.e., the latter requires all members of the set (34) to hold].

A conceptually straightforward generalization to molecular fluids follows from decomposing all the forces into center-of-mass interactions (involving $\mathbf{r}$) and angular interactions (involving a set of angles ω). Thus the external field is now $v(\mathbf{r},\omega)$. Note that this form applies even in simple cases such as planar boundary, since the center of mass of, say, a rod-shaped molecule can approach more closely to the wall when the molecule is oriented along the wall than when it points perpendicular to the wall. In this formalism all of the statistical mechanics above is immediately applicable, with positions and gradients referring to center-of-mass coordinates and integrations over angular variables included to remove dependences on ω. For example, since linear momentum is concerned with center-of-mass motion, (26) carrys over essentially as before, apart from the addition of a simple integral over angles; for example, (23b) now reads

$$\begin{aligned}\nabla^\beta P^{\alpha\beta}(\mathbf{r}) &\equiv -\langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i)\nabla^\alpha_{\mathbf{r}_i} v(\mathbf{r}_i,\omega_i)\rangle \\ &= -\int d\omega \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i)\,\delta(\omega - \omega_i)\nabla^\alpha_{\mathbf{r}_i} v(\mathbf{r}_i,\omega_i)\rangle \\ &= -\int d\omega\, \rho(\mathbf{r},\omega)\nabla_{\mathbf{r}}{}^\alpha v(\mathbf{r},\omega)\end{aligned} \tag{35}$$

Integral equations such as (14) and (18), sum rules for free-energy derivatives carried over from the atomic fluid case, sum rules for surface tension (Section III.C), and explicit formulas for pressure tensors such as (28) are similarly modified only by the inclusion of trivial integrals over angular degrees of freedom [see, e.g., Walton and Gubbins (1985)]. This is because up to now only center-of-mass coordinates have been directly involved (i.e., no new physics associated with the extension to molecular systems has so far been discussed). However, this situation changes as

soon as one focuses on the angular degrees of freedom. In particular, the hierarchy (10) is equally applicable to differentiation with respect to angular field variables. It follows that (11) and (12) generate a set of sum rules determining order parameters and interfacial compressibilities that are directly concerned with orientational order, in addition to sum rules concerned with positional order. Similarly, the presence of angular coordinates means that applying rotational invariance to the hierarchies (10) and (16c), analogous to the derivation of (14) and (18), will lead to new physics: namely, integral equations for $(\mathbf{r} \times \nabla_{\mathbf{r}} + \nabla_{\omega})\rho(\mathbf{r},\omega)$. For explicit derivations, see Percus (1980) and Tarazona and Evans (1983). In addition, one has a hierarchy of orientational YBG equations, involving angular derivatives of the many-body potential: generated by direct differentiation of the distribution functions and presumably also by potential distribution theory analogous to the derivation of (31). To date, little application has been made of this powerful statistical mechanical approach to orientational phenomena at fluid interfaces.

E. Fluctuations and Instabilities of Fluid Interfaces

Light-scattering experiments demonstrate that fluid–fluid interfaces are constantly in motion, due to the thermal excitation of long-wavelength capillary-wave (cw) modes [see, e.g., Katyl and Ingard (1967, 1968)]. In addition, computer simulation studies have confirmed that this picture remains down to microscopic wavelengths (Kalos et al., 1977). The longer the wavelength of a collective mode, the longer is its oscillatory period. Thus in the long-wavelength limit it should be correct to treat the capillary-wave modes of a fluid interface as hydrodynamic fluctuations (e.g., considering distortions from planar symmetry) (Buff et al., 1965):

$$\hat{l}(\mathbf{R},t) = l + \sum_{\mathbf{Q}} \zeta_{\mathbf{Q}}(t) e^{i\mathbf{Q}\cdot\mathbf{R}} \qquad l \equiv \langle \hat{l} \rangle \tag{36a}$$

$$H_{\mathrm{cw}} = \tfrac{1}{2} \sum_{\mathbf{Q}} m_Q \left(|\dot{\zeta}_{\mathbf{Q}}|^2 + \omega_Q^2 |\zeta_{\mathbf{Q}}|^2 \right) \qquad Q \to 0 \tag{36b}$$

where $\hat{l}$ denotes the position of the midpoint of the fluctuating interface. The probability $P(\hat{l})$ that any point in the interface lies at height $\hat{l}$ is determined by the free-energy contribution to (36b):

$$P(\hat{l}) = \mathcal{N} \exp\left(\frac{-F_{\mathrm{cw}}}{kT}\right) \qquad F_{\mathrm{cw}} \equiv \tfrac{1}{2} \sum_{\mathbf{Q}} m_Q \omega_Q^2 |\zeta_{\mathbf{Q}}|^2 \tag{36c}$$

where $\mathcal{N}$ is a normalization factor. F_{cw} is often referred to as an interface

Hamiltonian. In general, the interface Hamiltonian takes the form

$$F_{cw} \equiv \gamma \Delta A + \delta V(l) \tag{37a}$$

where γ is the interfacial tension and $\delta V(l)$ denotes the modification of the capillary-wave contribution due to damping from external fields. It is straightforward to show that up to order Q^2 the change in interfacial area due to the distortions (36a) is

$$\Delta A = \tfrac{1}{2} A \sum_{\mathbf{Q}} Q^2 \mid \zeta_{\mathbf{Q}} \mid^2 \tag{37b}$$

where A denotes the planar interfacial area. The leading-order contribution from the damping term is

$$\delta V(l) = \frac{1}{2} (\delta l)^2 \left[\frac{\partial^2 V(l)}{\partial l^2} \right]_{l = l_{eq}} \tag{37c}$$

$$(\delta l)^2 \equiv A^{-1} \int d^{d-1}\mathbf{R} (\hat{l} - l)^2 = \sum_{\mathbf{Q}} \mid \zeta_{\mathbf{Q}} \mid^2 \tag{37d}$$

where l_{eq} denotes the equilibrium position of the interface. Thus (37) is of the form (36), with

$$m_Q \omega_Q^2 = AkT \left(W + \frac{Q^2 \gamma}{kT} \right) \tag{38a}$$

and

$$W \equiv (AkT)^{-1} \left[\frac{\partial^2 V(l)}{\partial l^2} \right]_{l = l_{eq}} \tag{38b}$$

Since we have ignored any contribution from the damping at order Q^2 it follows that (38) is relevant to the weak damping regime. For a specific example, consider a liquid–vapor interface in the presence of the earth's gravitational field:

$$V(l) = \tfrac{1}{2} mg \, \Delta\rho \, A l^2 + c \tag{39a}$$

$$W = \frac{mg \Delta \rho}{kT} \tag{39b}$$

where $\Delta\rho$ is the number density difference between liquid and vapor. Applying equipartition to (36) and (38),

$$kT = m_Q \omega_Q^2 \langle \mid \zeta_{\mathbf{Q}} \mid^2 \rangle = AkT \left(W + \frac{Q^2 \gamma}{kT} \right) \langle \mid \zeta_{\mathbf{Q}} \mid^2 \rangle \tag{40a}$$

it follows that the capillary-wave correlation length for correlations perpendicular to the interface diverges as W tends to zero in dimensions d

≤ 3 [see, e.g., Bedeaux and Weeks (1985)]:

$$\xi_\perp{}^2 \equiv \sum_{\mathbf{Q}} \langle |\zeta_{\mathbf{Q}}|^2 \rangle = \frac{1}{A} \sum_{\mathbf{Q}} \left(W + \frac{Q^2 \gamma}{kT} \right)^{-1} \sim \begin{cases} W^{(d-3)/2} & d < 3 \\ \ln(W^{-1}) & d = 3 \\ \text{finite} & d > 3 \end{cases} \tag{40b}$$

So in $d \leq 3$ capillary-wave wandering of a fluid interface will increase progressively as the external damping is reduced. Note also that the three-dimensional world is the borderline dimension for this singular, or critical, capillary-wave phenomena.

Capillary-wave correlations will be strongest in the plane of the interface: $\xi_\parallel$. For the interface Hamiltonian (36) we can define

$$G(z_1,z_2,R_{12})_{\text{cw}} = \langle \delta\hat{\rho}(1)\delta\hat{\rho}(2) \rangle$$
$$\delta\hat{\rho}(\mathbf{r},t) \equiv -\rho'(z)[\hat{l}(\mathbf{r},t) - l] \qquad Q \to 0 \tag{41a}$$

where the subscript cw reminds us that only capillary-wave correlations are present in a system defined by an interface Hamiltonian. It therefore follows that

$$G(z_1,z_2;Q)_{\text{cw}} = A\langle |\zeta_{\mathbf{Q}}|^2 \rangle \rho'(z_1)\rho'(z_2) = \frac{\rho'(z_1)\rho'(z_2)}{W(1 + Q^2\xi_\parallel{}^2)} \qquad Q \to 0 \tag{41b}$$

where we have defined

$$\xi_\parallel{}^2 \equiv -\frac{G_2(z_1,z_2)_{\text{cw}}}{G_0(z_1,z_2)_{\text{cw}}} = \frac{\gamma}{kTW} \tag{41c}$$

Thus long-wavelength correlations in the plane of the interface are universally divergent in the zero damping limit. Furthermore, we see that interface Hamiltonians lead to Ornstein–Zernike behavior [i.e., Eq. (41b)]. This is a consequence of the fact that γ is defined in (37a) to be a constant and thus cannot introduce additional singular behavior. Combining (37) to (41), we can identify the form of the singular part of the free energy as

$$F^{\text{sing}} \sim \left(\frac{\xi_\perp}{\xi_\parallel} \right)^2 \tag{42a}$$

Substituting (40b), it follows that below the upper critical dimension capillary-wave phenomena obey a universal hyperscaling relation:

$$F^{\text{sing}} \xi_\parallel{}^{d-1} = \text{constant} \qquad d < 3 \tag{42b}$$

Mathematically rigorous versions of interface Hamiltonian theory may be based on column models, as introduced by Weeks (1977). Here, one divides the planar interface into columns and treats capillary-wave modes as inducing correlations between different columns. The probability that the instantaneous position of the interface lies at height $\hat{l}$ within any column is $P(\hat{l})$ defined by (36c). Evaluating the Fourier transform of (36c) [i.e., $\langle \exp(iQ\hat{l}) \rangle$] implies that $P(\hat{l})$ is a Gaussian [see, e.g., Evans (1979)]:

$$P(\hat{l}) = (2\pi\xi_\perp^2)^{-1/2} \exp\left(\frac{-\hat{l}^2}{2\xi_\perp^2}\right) \tag{43a}$$

where $\xi_\perp$ is defined by (40b). Similarly, the joint fluctuation probability $P(\hat{l}(1), \hat{l}(2))$, defining intercolumn correlations, follows from evaluating the quantity $\langle \exp[iQ\hat{l}(1) + iQ'\hat{l}(2)] \rangle$ given the normal mode probability (36c) [see, e.g., Rao et al. (1979)]. The long-wavelength expansion of this "transform" is consistent with the following "inversion" to order $|\zeta_Q|^2$ (Rao et al., 1979; Percus, 1981):

$$\begin{aligned} P(\hat{l}(1), \hat{l}(2)) &= P(\hat{l}(1))P(\hat{l}(2)) \\ &\quad + P'(\hat{l}(1))P'(\hat{l}(2)) \sum_{\mathbf{Q}} \langle |\zeta_{\mathbf{Q}}|^2 \rangle e^{i\mathbf{Q}\cdot\mathbf{R}_{12}} \qquad Q \to 0 \end{aligned} \tag{43b}$$

From (43) we identify the long-wavelength contribution to the surface structure factor as shown in (41b).

Let us now consider the relationship of the capillary-wave theory above to molecular models of inhomogeneous fluids. Following Weeks (1977), it is natural to inquire if one may treat the fluctuating interface (the so-called *bare profile*) as being qualitatively defined by including all correlations up to a length of order of the bulk correlation length, ξ_b (and hence ξ_b defines the column width). If so, the full profile is obtained by including capillary waves of wavelengths larger than ξ_b; in particular, from (43a) we have

$$\rho'(z) = (2\pi\xi_\perp^2)^{-1/2} \int_{-\infty}^{\infty} d\hat{l}\, \rho_0'(z - \hat{l}) \exp\left(\frac{-\hat{l}^2}{2\xi_\perp^2}\right) \tag{44}$$

where subscript 0 denotes the bare profile. The central portion of the full interfacial profile gradient receives contributions from both the bare profile gradient and the capillary-wave term, each of which is a strongly peaked factor. However, in the tails (or asymptotic wings) of the density profile the right side of (44) is dominated by only one of the factors appearing in the integrand, depending on the behavior of $\xi_\perp$ [i.e., (40b) implies a crossover in the nature of the profile tails, at the upper critical dimension ($d_>$) for interfacial fluctuations]. For $d > d_>$ the capillary-wave

fluctuations do not broaden the tails of the profile because in comparison with the rest of the integrand, the Gaussian factor on the right side of (44) acts as a delta function:

$$\rho'(z) \to \rho_0'(z) \qquad |z| \to \infty, \quad d > d_> \tag{45}$$

This is precisely what is found in mean-field theories (i.e., from van der Waals–like approximations to integral equation theories or density functional theories, in any dimension), where $\xi_\perp$ is always a finite microscopic length and the mean-field profile gradient decays no faster than an exponential. It follows that for the purpose of evaluating the asymptotic behavior of the full profile it is appropriate to calculate the bare profile from mean-field theory. For example, consider the case of strictly finite-range interactions, or exponentially decaying interactions, for which the mean-field liquid–vapor profile is readily shown to have the following asymptotic behavior [see, e.g., Henderson (1987b)]:

$$\rho_{MF}'(z) \to ae^{-\lambda|z|} + be^{-2\lambda|z|} + \cdots \qquad \lambda|z| \gg 1 \tag{46}$$

where in general the liquid and vapor tails will have different values of the temperature-dependent parameters λ, a, and b. Close to the mean-field critical point, (46) becomes symmetric and $1/\lambda$ reduces to the bulk correlation length (Fisk and Widom, 1969). Identifying (46) with $\rho_0'(z)$, as discussed above, we see that in $d > 3$ we obtain (45) because a Gaussian decays faster than an exponential. However, for $d \leq 3$, one must take into account the divergent behavior of $\xi_\perp$. In particular, substituting (46) into (44), as $\rho_0'(z)$, we see that the behavior of the tails of $\rho(z)$ is determined by the integral

$$\begin{aligned} I(\lambda,|z|) &= (2\pi\xi_\perp^2)^{-1/2} \int_z^\infty d\hat{l}\, e^{\lambda(z-\hat{l})} e^{-\hat{l}^2/2\xi_\perp^2} \qquad z \to -\infty \\ &= (2\pi\xi_\perp^2)^{-1/2} \int_0^\infty dx\, e^{-\lambda x} e^{-(x-|z|)^2/2\xi_\perp^2} \end{aligned} \tag{47a}$$

This integral is readily evaluated in terms of a standard error function, giving the following asymptotic behavior:

$$I(\lambda,|z|) \to \begin{cases} \exp\left(-\lambda|z| + \dfrac{\lambda^2\xi_\perp^2}{2}\right) & |z| \gg \lambda\xi_\perp^2, \ \text{large}\,|z| \\ \left[(2\pi)^{1/2}\left(\lambda\xi_\perp - \dfrac{|z|}{\xi_\perp}\right)\right]^{-1} \exp\left(\dfrac{-z^2}{2\xi_\perp^2}\right) & |z| \ll \lambda\xi_\perp^2, \ \text{large}\,\xi_\perp \end{cases} \tag{47b}$$

For large enough $|z|$ and at nonzero damping, the value of (47) is dominated by the $\hat{l} \simeq 0$ contribution (i.e., the profile eventually decays ex-

ponentially). However at $|z| < \lambda\xi_\perp^2$ the leading-order part of (47) arises from the region $\hat{l} \simeq z$. In the latter case, we must include the contribution to (44) arising from $\hat{l} < z$ [i.e., now it is the $\rho_0'(z - \hat{l})$ factor in the integrand of (44) that acts as a delta function]:

$$\rho'(z) \to \pm\Delta\rho(2\pi\xi_\perp^2)^{-1/2} \exp\left(\frac{-z^2}{2\xi_\perp^2}\right) \qquad 1 \ll \lambda|z| \leqslant \lambda^2\xi_\perp^2 \tag{48}$$

The physical interpretation of the origin of (48) is that it arises from capillary-wave fluctuations that move the central portion of the bare profile across the height z. Below the upper critical dimension, which for short-range forces is clearly $d = 3$, such fluctuations will always arise in the limit of infinitesimal damping ($W \to 0_+$), regardless of the value of z [i.e., (48) dominates the asymptotic decay of $\rho(z)$ as $\xi_\perp \to \infty$]. For this reason, $d \leqslant d_>$ is known as the fluctuation regime. In the borderline case for short-range interactions, $d = 3$, one should add a warning concerning the extremely slow divergence of $\xi_\perp$; namely, note the logarithmic behavior of (40b) in $d = 3$ and the fact that (40b) arises only if one is permitted to include arbitrarily small values of Q [this involves the requirement that one must be able to perform statistical averages over arbitrarily large times (Henderson, 1987b)].

With regard to capillary-wave correlations in the plane of the interface, the link between interface Hamiltonians and molecular-based theories is described in a seminal paper by Wertheim (1976). In particular, Wertheim (1976) noted that in the weak damping limit the single eigenfunction result (41) for $G_0(z_1,z_2)$ is more or less implied by the integral equation (14), giving (Rao et al., 1979):

$$W = -\frac{1}{kT}\int dz\, \rho'(z)v'(z) \qquad W \to 0 \tag{49}$$

and note, for example, that (49) rederives (39b) in the case $v(z) = mgz$. Furthermore, substituting $G_0(z_1,z_2)_{cw} = \rho'(z_1)\rho'(z_2)/W$ into (20d) and using sum rule (77b) implies that

$$G_2(z_1,z_2) = -\frac{\gamma}{kTW^2}\rho'(z_1)\rho'(z_2) \qquad W \to 0 \tag{50}$$

This result explains the mysterious surface tension sum rule (77a), given (49). Thus in a few short lines the statistical mechanical theory of inhomogeneous fluids is able to confirm the general significance of the Ornstein–Zernike behavior of interfacial fluctuations in the weak damping regime [Eq. (41b)]. Some workers have attempted to cast doubt on this conclusion in the special but physically relevant case of $d = 3$. In par-

ticular, Robert (1985), Ciach (1987), and Requardt and Wagner (1989) have suggested that capillary-wave correlations might be so divergent as to make the right sides of sum rules (77a) and (77b) ill defined, thus invalidating (50); room for such doubt is provided by the nonscaling nature of capillary-wave correlations in the borderline dimension $d = 3$. However, all such arguments to date have relied on explicit results concerning the strict $W = 0$ limit (i.e., the thermodynamic limit). Here, one does not have a well-defined interface ($\gamma = 0$), so one should not anticipate any link with the $W \to 0_+$ limit of a nondiffuse interface [see the remark following (20e)]. Weeks et al. (1989) have provided the strongest rebuff to date of doubts concerning the validity of Ornstein–Zernike interfacial behavior in $d = 3$; if such arguments are ultimately proved to be incorrect, interface Hamiltonians of the class (36) would not be applicable to $d = 3$ and none of the critical wetting phenomena that have been derived from such theories would be relevant to molecular Hamiltonians (Section IV).

To conclude this section, let me highlight the significance of soft interfacial modes to fluid interfacial phenomena such as wetting. As discussed above, all fluid interfaces are inherently unstable to capillary-wave contributions to $\xi_\parallel$, and in $d \leq 3$ this further involves a divergent $\xi_\perp$. For example, consider the growth of a liquid film at a substrate–vapor interface. At finite film thickness the capillary-wave modes are damped by the wall–fluid potential (e.g., capillary waves cannot penetrate the substrate). However, the thicker the film, the smaller the damping, and thus we see that the continuous growth of a liquid film is an example of the $W \to 0_+$ limit discussed above. Similar remarks apply to continuous melting of a solid–gas interface (i.e., surface melting) (Lipowsky, 1986). Soft modes also play significant roles in first-order interfacial transitions. For example, consider the growth of liquid films at the walls of a planar slit pore of width L. As soon as the film thicknesses are large enough to support capillary-wave modes, the size of $\xi_\perp$ will determine the likelihood of collisions between interfacial fluctuations spanning the gap L, thereby nucleating capillary condensation.

F. Intermolecular Forces and Nonuniversality in Interfacial Systems

The conclusions of the preceding section imply that below the upper critical dimension for interfacial fluctuations ($d < d_>$) the details of intermolecular forces are irrelevant to the qualitative nature of fluid interfacial phenomena in the weak damping regime. That is, in this situation capillary-wave fluctuations will dominate the properties of fluid interfaces and interfacial critical phenomena will belong to universal scaling regimes

described by (40b), (41c), and (42b). However, this universality cannot be expected to be generally applicable to the physical world because $d_> \leqslant 3$. Thus it is necessary to inquire whether or not the details of intermolecular forces play a qualitative role in interfacial critical phenomena. In particular, the ubiquitous presence of dispersion forces in molecular systems means that the asymptotic decay of intermolecular interactions is invariably a power law:

$$\varphi^{(2)}(r) \sim r^{-(3+m)} \tag{51a}$$

where $m = 3$ at intermediate range (i.e., until retardation effects take over and enforce the macroscopic limit $m = 4$). Thus in planar symmetry the wall–fluid potential arising from a semi-infinite solid will decay as

$$v_{\epsilon_w}(z) \sim z^{-(m+3-d)} + \mathcal{O}z^{-(m+4-d)} \qquad z \to \infty \tag{51b}$$

where (as usual) d denotes the overall dimensionality.

For an example of the significance of (51), consider a wall–gas interface with $v(z)$ defined by (51b). From (15) and (17) it follows that in the low-density limit of the bulk gas, the density profile falls off as

$$\rho(z) \to \rho_b\left(1 - \frac{v(z)}{kT}\right) \qquad \rho_b \to 0, \quad z \to \infty \tag{52}$$

That is, a power-law decay of $v(z)$ induces a power-law decay of $\rho(z)$ and so in molecular systems one cannot treat a wall–fluid interface as being of finite range. The crucial significance of this fact to fluid adsorption and wetting phenomena has long been apparent to the Soviet school [see, e.g., Derjaguin (1940)] but has only been generally appreciated in the West since the early 1980s (de Gennes, 1981).

The asymptotic wings of a liquid–vapor or a fluid–fluid profile are similarly affected by the presence of power-law interactions. Naively, one can construct an analogy with (52) by treating the liquid phase as the wall, and then on the liquid side of the interface it follows that the same effect must arise from the absence of intermolecular interactions; for an explicit theory, see Barker and Henderson (1982). Thus, in molecular systems we should replace (46) with

$$\rho'_{\mathrm{MF}}(z) \to a|\lambda z|^{-(m+4-d)} + b|\lambda z|^{-(m+5-d)} + \cdots \tag{53}$$

Here, the relevant integral replacing (47) is

$$(2\pi\xi_\perp{}^2)^{-1/2} \int_\sigma^\infty dx\, x^{-(m+4-d)} e^{-(x-|z|)^2/2\xi_\perp{}^2} \tag{54}$$

where σ is a microscopic length of order $1/\lambda$ or perhaps $10/\lambda$. The mean-field region of the profile is now given by the condition $|z|[\ln(|z|/\sigma)]^{-1/2}$

$\gg \xi_\perp$, and comparing with (47b) we see that at a fixed value of the capillary-wave damping, power-law interactions are much more dominant over the fluctuation contribution than are short-range interactions. For the case of a liquid–vapor interface in a gravitational field, these remarks are of little significance, except to experiments that specially probe the asymptotic region, because the amplitude of these profile tails is so insignificant in comparison with $\Delta\rho$ (Barker and Henderson, 1982). In contrast, the study of continuous wetting phenomena is concerned directly with the behavior of the profile wings, which are the means by which a liquid–vapor surface interacts with the short-range field of a distant substrate. In particular, Lipowsky (1984) noted that power-law interactions imply a lowering of the upper-critical dimension for interfacial critical phenomena; $2 < d_> < 3$. So in the absence of bulk critical phenomena, mean-field theory is qualitatively correct in $d = 3$ (but not $d = 2$), except for models restricted to finite-range or exponentially decaying interactions.

It follows from the above that interfacial critical phenomena is nonuniversal in $d = 3$, with free-energy scaling functions and associated exponents that depend explicitly on the power-law interaction parameter m (Dietrich and Schick, 1985; Ebner et al., 1985). In fact, even exponential forces cannot always be treated as short ranged, since an exponential wall field can compete with the exponential decay of a mean-field profile determined by strictly finite-ranged fluid–fluid interactions or by exponential interactions [i.e., Eq. (46)]. Furthermore, because $d = 3$ is the upper-critical dimension for capillary-wave fluctuations in the absence of power-law forces, these competing exponential effects can in turn compete with capillary-wave broadening (Hauge and Olaussen, 1985). It follows that models of inhomogeneous fluids without power-law forces possess particularly complex nonuniversal behavior in the special case of $d = 3$. Physically, one must get used to the fact that distant surfaces will always interact via the tails of their density profiles, even if the only effect present is exponential decay; for example, exponential decay is associated with logarithmic film growth. For inhomogeneous fluids the qualitative details of molecular models have qualitative consequences and one cannot get away with pretending that intermolecular forces are strictly short ranged.

III. SUM RULES

A. Compressibility Route

For the sake of definitiveness I restrict this discussion of explicit examples to the models defined by external fields (8) and (9). These two models

are simplified prototypes that highlight two important aspects of interfacial physics, associated with the fields ϵ_w and L, respectively. Obviously, many generalizations could readily be incorporated; for example, (8) could represent just part of the external field (say, the short-range repulsive part or a power-law attraction), or a range parameter a_w could be introduced into (8), and (9) could be modified to discuss nonplanar geometry (Section V.B) or generalized to nonsymmetric walls (e.g., antagonistic walls). Thus the potential breadth of physics that could be studied by the following methods is very large indeed.

Substituting (8) and (9) into the general sum rule (11) leads to explicit results for the densities conjugate to the field variables μ, ϵ_w, and L:

$$\Gamma \equiv -\frac{1}{A}\frac{\partial \Omega^{ex}}{\partial \mu} = \int_0^\infty dz[\rho(z) - \rho_b] \tag{55}$$

$$\Theta \equiv -\frac{1}{A}\frac{\partial \Omega}{\partial \epsilon_w} = -\frac{1}{\epsilon_w}\int_0^\infty dz\, \rho(z)v(z) \tag{56}$$

$$f \equiv -\frac{1}{A}\frac{\partial \Omega^{ex}}{\partial L} = -\int_{-\infty}^\infty dz\, \rho(z)v^{\infty\prime}(z) - p_b \tag{57}$$

where in sum rules (55) and (56) and hereafter (wherever specifically required) the equipotential of infinite wall–fluid repulsion is located at the plane $z = 0$. Note also that hereafter, unless otherwise stated, all partial derivatives denote fixed field variables.

Equation (12) enables us to introduce the following interfacial compressibilities, among others:

$$\chi \equiv \frac{\partial \Gamma}{\partial \mu} = \frac{1}{kT}\int_0^\infty dz_1 \int_{-\infty}^\infty dz_2\, [G_0(z_1,z_2) - G_0(|\, z_{12}\,|)_b] \tag{58}$$

(Evans et al., 1983)

$$\chi_1 \equiv \frac{\partial \Gamma}{\partial \epsilon_w} = \frac{\partial \Theta}{\partial \mu} = -\frac{1}{kT\epsilon_w}\int_{-\infty}^\infty dz_1 \int_{-\infty}^\infty dz_2\, v(z_2)G_0(z_1,z_2) \tag{59}$$

(Henderson, 1986a)

$$\chi_{11} \equiv \frac{\partial \Theta}{\partial \epsilon_w} = \frac{1}{kT\epsilon_w^2}\int_{-\infty}^\infty dz_1 \int_{-\infty}^\infty dz_2\, v(z_1)v(z_2)G_0(z_1,z_2) \tag{60}$$

(Henderson, 1986a)

$$\chi_L \equiv \frac{\partial f}{\partial L} = \frac{1}{kT}\int_{-\infty}^\infty dz_1 \int_{-\infty}^\infty dz_2\, v^{\infty\prime}(z_1)v^{\infty\prime}(L - z_2)G_0(z_1,z_2) \tag{61}$$

(Henderson, 1986b)

where in (61) and elsewhere a dash denotes derivative with respect to the

stated argument, and note that since (57) has made use of the symmetry of $\rho(z)$ and $v(z)$ about $z = L/2$, it is best to derive sum rule (61) via direct differentiation of (57).

It is also worth pointing out that sum rules (58) to (61) are associated with the following integral equations for partial derivatives of the one-body density:

$$kT\frac{\partial\rho(z_1)}{\partial\mu} = \int_{-\infty}^{\infty} dz_2\, G_0(z_1,z_2) \tag{62}$$

$$kT\frac{\partial\rho(z_1)}{\partial\epsilon_w} = -\frac{1}{\epsilon_w}\int_{-\infty}^{\infty} dz_2\, v(z_2)G_0(z_1,z_2) \tag{63}$$

$$kT\frac{\partial\rho(z_1)}{\partial L} = -\int_{-\infty}^{\infty} dz_2\, v^{\infty\prime}(L - z_2)G_0(z_1,z_2) \tag{64}$$

and to this collection one should add Eq. (14):

$$kT\rho'(z_1) = -\int_{-\infty}^{\infty} dz_2\, v'(z_2)G_0(z_1,z_2) \tag{65}$$

which is equally applicable to models (8) and (9) (i.e., single wall and slit) (Henderson, 1986b).

In cases where the external field is discontinuous it is important to introduce (15). Similarly, the connection between capillary-wave theory and molecular Hamiltonians is most rigorously handled by going to the trouble of removing the $\delta(z_{12})$ term in $G_0(z_1,z_2)$ via introduction of (15) (Henderson and van Swol, 1985; Henderson, 1986a). For example, one can generate an interesting sum rule by integration of Eq. (65) across a wall–fluid interface:

$$\rho_b - \frac{p_b}{kT} = \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2\, \rho(z_1)n(z_2)h_0(z_1,z_2)\frac{d}{dz_2}\exp\left[\frac{-v(z_2)}{kT}\right] \tag{66}$$

where $h_0(z_1,z_2)$ is defined analogous to (20) and p_b has been introduced via the virial route expression for mechanical stability [Eq. (25c)]. In the case of slit pores, one can integrate (62), (64), and any appropriately generalized version of (63) to obtain sum rules for the change in total absorption within the pore; for example, from (64) (Kjellander and Sarman, 1990):

$$\frac{kT}{A}\frac{\partial N(L)}{\partial L} = -\int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2\, v^{\infty\prime}(L - z_2)G_0(z_1,z_2) \tag{67}$$

In addition, one can invoke the Ornstein–Zernike equation (20a) at $Q = 0$ to invert the results (62) to (65), thereby generating a complementary

set of equations involving the two-body direct correlation function; thus (58) to (61) and (66) and (67) can all be reexpressed as direct correlation function sum rules. The Q^2 Ornstein–Zernike inversion, Eq. (20d), is particularly important in conjuncture with sum rules for the interfacial free energy (see Section III.C).

B. Virial Route

The significance of the virial route lies in its direct connection with mechanical stability. Thus integration of (25c) across a planar wall–fluid interface generates a sum rule for the bulk pressure in terms of a balance with the force exerted by the wall:

$$\frac{p_b}{kT} = \int_{-\infty}^{z_b} dz\, n(z) \frac{d}{dz} \exp\left[\frac{-v(z)}{kT}\right] \tag{68a}$$

and note that (68a) can be rewritten in the language of the compressibility route by introducing the one-body direct correlation function via (17):

$$\frac{p_b}{kT} - \rho_b = -\int_{-\infty}^{z_b} dz\, \rho(z) c^{(1)\prime}(z) \tag{68b}$$

In narrow slits, mechanical equilibrium is no longer associated with a bulk pressure (defined by T and μ) but with the normal component of the pressure tensor inside the slit: in particular, for a symmetric slit of width L:

$$\frac{p_N(L/2)}{kT} = \int_{-\infty}^{\infty} dz\, n(z) \exp\left[\frac{-v^{\infty}(L-z)}{kT}\right] \frac{d}{dz} \exp\left[\frac{-v^{\infty}(z)}{kT}\right] + \frac{2}{kT} \int_{L/2}^{\infty} dz\, \rho(z) v^{\infty\prime}(z) \tag{69}$$

In the absence of direct overlap between the two wall fields, the last term on the right side of (69) is zero and the solvation force (57) reduces to the pressure difference $[p_N(L/2) - p_b]$ (Henderson, 1986b). In molecular dynamics simulation studies of pair potential models of inhomogeneous fluids with planar symmetry, it is now routine practice to calculate pressure tensor profiles directly from (28).

An alternative but equivalent expression of mechanical force balance is provided by the first YBG equation (31b) [i.e., one now avoids explicit introduction of the pressure tensor (23a)]. Thus mechanical stability is also expressed by sum rules generated from integration of (31b). For example, Lekner and Henderson (1980) introduced the following set of equations expressing integrated force balance across a liquid–vapor interface

of a pair potential fluid:

$$(1 - s^{-1})(\rho_L{}^s - \rho_V{}^s) - \frac{p}{kT}(\rho_L{}^{s-1} - \rho_V{}^{s-1})$$
$$= \frac{\pi}{kT}\int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2\,[\rho(z_2)^{s-1} - \rho(z_1)^{s-1}] \qquad (70)$$
$$\times (z_1 - z_2)\int_{|z_{12}|}^{\infty} dr_{12}\rho^{(2)}(r_{12},z_1,z_2)\varphi'(r_{12})$$

where the first term on the left-hand side reduces to $-\ln(\rho_L/\rho_V)$ in the limit $s = 0$. Analogous sum rules for the compressibility route are generated via (18). Of course, to understand how sum rules obtained from (14) apply to liquid–vapor interfaces, one must make the explicit connection with capillary-wave theory, as described in Section II.E.

C. Surface Free Energy

Superficially, the straightforward route to sum rules for surface excess grand potential is via the virial theorem (24). In particular, in planar geometry we can write, from (25a),

$$\frac{\Omega^{ex}}{A} = \int_{-\infty}^{\infty} dz[p_b\Theta_G(z) - p_T(z)] \qquad (71)$$

where $\Theta_G(z)$ defines a choice of Gibbs dividing surface [i.e., $\Theta_G(z) = 0$ for all z lying outside any boundary surface and is unity otherwise]. Expressions such as (71) can be evaluated from pressure tensors like (28), or equivalently, one can work directly with specific examples of equation (22) (Kirkwood and Buff, 1949). However, surface free energy is dependent on the geometry of an interface and outside planar symmetry the analog of (25a) is in general ill defined by pressure tensors of the class (23a) [see, e.g., Henderson (1986c)]. Accordingly, beyond planar symmetry the virial route to surface free energy in molecular systems becomes problematical. The compressibility route to sum rules for Ω^{ex} is more technically demanding because it requires one to evaluate second-order changes in free energy, in contrast to the first-order virial result (22). However, the compressibility route is relatively free of boundary condition problems that plague the virial route applied to nonplanar geometry; thus, below and in Section V.B it is the compressibility route that is invoked to generate rigorous results for spherical or cylindrical symmetry.

In the case of a planar wall–fluid interface [model (8)], it is instructive to note that (71) splits into an external field term plus a many-body fluid–fluid term; the latter is conveniently termed the surface tension (γ). That

is, inserting $p_N(z)$ defined by (25c) into (71) yields

$$\frac{\Omega^{ex}}{A} = \gamma - \int_0^\infty dz\, z\rho(z)v'(z) \tag{72a}$$

$$\gamma \equiv \int_0^\infty dz[p_N(z) - p_T(z)] \tag{72b}$$

For a symmetric planar slit [model (9)] we can similarly rewrite (71) as

$$\frac{\Omega^{ex}}{A} \equiv \int_0^L dz[p_b - p_T(z)]$$

$$= 2\left[\gamma - \int_0^L dz\, z\rho(z)v^{\infty\prime}(z)\right] - Lf \tag{73a}$$

$$\gamma \equiv \int_0^{L/2} dz[p_N(z) - p_T(z)] \tag{73b}$$

where f is the solvation force (57) and to obtain (73a) I have made use of (69); note that in (73), γ, $\rho(z)$, and f are all dependent on the slit width L.

Compressibility route sum rules for the surface tension of a planar interface in a general external field $v(z)$ can be obtained from the virial route expression (72b) via the derivation of Schofield (1979); for a review that explains why Schofield's proof is limited neither to pair potential fluids nor to fluctuating liquid–vapor interfaces, see Appendix A of Henderson (1983). However, this approach involves lengthy algebraic manipulation, the use of complex correlation functions, and an integration by parts that requires a careful consideration of boundary conditions. Accordingly, I shall review a direct method introduced by Henderson (1983), based exclusively on a straightforward use of the compressibility route hierarchy (10).

The trick is first to transform to d-dimensional spherical symmetry and then take the planar limit. For the sake of definitiveness I shall consider radial external fields of the class

$$v(r;R_G) = v(r - R_G) \tag{74}$$

where R_G denotes a choice of dividing surface. For any planar $v(z)$ one can always introduce a radial external field of the class (74) that reduces to $v(z)$ in the limit $R_G \to \infty$; $z \equiv (r - R_G)$. I shall also require that $v'(r - R_G)$ vanish at $r > R_W$ for some $R_W \geq R_G$. The latter condition could be achieved in various ways; for example, if one is interested in wall–fluid systems, the fluid could be placed inside a cavity (say, of radius

R_G), or alternatively, one could consider fluid surrounding a spherical wall (at R_G), with R_W denoting a far boundary. Weak stabilizing fields (analogous to the earth's gravity) can equally well be considered, provided that one truncates or smoothly heals the potential at $r > R_W$, directly analogous to the planar case, where again one prefers to use models that possess true bulk phases (Weeks, 1984). The important point here is that by extracting the planar result from the limit $R_G \to \infty$ one avoids the difficulties of defining rigorous boundary conditions for infinite planar interfaces such as those associated with integrating by parts along a planar interface possessing long-range capillary-wave correlations. Provided only that the excess grand potential following subtraction of bulk terms [e.g., proportional to ${R_G}^d$ and $(R_W - R_G)^d$] varies strictly as ${R_G}^{d-1}$ at leading order, the limit $R_G \to \infty$ will yield a well-defined surface tension. Thus, in the absence of pathological behavior such as the presence of a ${R_G}^{d-1} \ln(R_G)$ term in $\Omega(R_G)$, it follows that planar interfacial correlations must be consistent with a finite surface tension in the limit $R_G \to \infty$.

For any choice of T, μ, R_G let us consider an infinitesimal variation of R_G at fixed (T,μ), as defined by the compressibility route (10):

$$\frac{\partial \Omega}{\partial R_G} = -\int d\mathbf{r}\, \rho(r;R_G) v'(r - R_G) \tag{75a}$$

$$\frac{\partial^2 \Omega}{\partial {R_G}^2} = -\frac{1}{kT} \int d1 \int d2\, v'(r_1 - R_G) v'(r_2 - R_G) \times (1 - \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) G(12) - (d-1) \int d\mathbf{r}\, r^{-1} \rho(r;R_G) v'(r - R_G) \tag{75b}$$

The leading-order area dependence can be extracted from the combination

$$\frac{1}{{R_G}^{d-2}} \left(\frac{\partial}{\partial R_G} - \frac{R_G}{d-1} \frac{\partial^2}{\partial {R_G}^2} \right) \Omega(R_G) = \frac{1}{(d-1) kT {R_G}^{d-3}} \times \int d1 \int d2\, v'(r_1 - R_G) v'(r_2 - R_G)(1 - \hat{\mathbf{r}}_1 \cdot \hat{\mathbf{r}}_2) G(12) - \frac{1}{{R_G}^{d-1}} \int d\mathbf{r}\, \frac{R_G}{r} (r - R_G) \rho(r;R_G) v'(r - R_G) \tag{76}$$

Taking the planar limit [$z \equiv (r - R_G)$; $R_G \to \infty$] of (76) we regain the form (72a) except that here γ is defined by a compressibility route sum

rule:

$$\gamma = \frac{1}{2(d-1)kT}\int_{-\infty}^{\infty} dz_1\, v'(z_1)\int_{-\infty}^{\infty} dz_2\, v'(z_2)\int d\mathbf{R}_{12}R_{12}^2 G(R_{12},z_1,z_2)$$
$$= -\frac{1}{kT}\int_{-\infty}^{\infty} dz_1\, v'(z_1)\int_{-\infty}^{\infty} dz_2\, v'(z_2)G_2(z_1,z_2) \tag{77a}$$

In addition, one can insert Eq. (20d) and then use of Eq. (65) immediately transforms sum rule (77a) into the inverted form

$$\gamma = kT\int_{-\infty}^{\infty} dz_1\, \rho'(z_1)\int_{-\infty}^{\infty} dz_2\, \rho'(z_2)G_2^{-1}(z_1,z_2) \tag{77b}$$

The results (77a) and (77b) were first derived from fluctuation theory of liquid–vapor interfaces [Wertheim (1976) and Triezenberg and Zwanzig (1972), respectively]. However, the derivation above holds regardless of the strength of the external field and is equally appropriate in the absence of capillary-wave modes. Note that the Schofield (1979) proof of equivalence between (77) and the virial route expression (72b) concerns the many-body contribution to Ω^{ex}; in general, the total surface excess grand potential contains a one-body term, the last term on the right sides of (72a) and (76). Finally, in the light of the derivation above, let me reconsider the doubts concerning the validity of sum rules (77) in the case of $d = 3$, as reviewed briefly toward the end of Section II.E. Namely, the derivation above would naturally associate any singular behavior in the planar limit (77) with a failure of the leading-order radius expansion of the free energy in spherical symmetry [e.g., the presence of an $R_G{}^{d-1}\ln(R_G)$ term in $\Omega(R_G)$]. Thus in any model in which the surface free energy per unit area of a spherical cluster can be proven to remain finite and non-zero as $R_G \to \infty$, it follows that sum rules (77) should be well defined.

D. Sum Rules and Approximate or Phenomenological Theories

Integral equation theories based on equations such as (62) to (65) will invariably violate the statistical mechanical consistency that links these equations to each other and to the rest of statistical mechanics. That is, only very special sets of correlation functions are capable of maintaining strict compatibility between one integral equation and another. For example, an approximate $G_0(z_1,z_2)$ inserted into, say, (65) will not usually yield a density profile that satisfies mechanical stability sum rules such as (66) and (68). However, if one enforces such compatibility, it is rea-

sonable to expect that the quality of the approximate solution should be enhanced. Thus an obvious approach to integral equation theory is to develop correlation function approximations that can be optimized by enforcing agreement with sum rules. This use of sum rules also has the benefit that the resulting integral equation theory can be said to satisfy the underlying physics of the sum rules, such as integrated mechanical stability or compressibility fluctuation phenomena. A related approach is to enforce agreement between corresponding pairs of virial route and compressibility route equations. In fact, Lovett (1988) has argued on formal grounds that approximate integral equation theories will probably possess no solutions at all unless additional consistency requirements are included.

In general, phenomenological density functional theories (see Chapter 3) yield correlation function hierarchies that are fully consistent with statistical mechanics. However, by approximating the grand potential functional (16) it is no longer possible to claim any clear link with a Hamiltonian of the class (7c). The significance of sum rules to density functional theory involves two aspects. First, the use of sum rules can enhance the numerical analysis of a density functional by providing criteria for numerical stability and through the direct correspondence between correlation functions and physical phenomena expressed by sum rules [see, e.g., van Swol and Henderson (1989)]. The second aspect is associated with a need to understand those special cases in which statistical mechanical consistency breaks down. In particular, careful consideration should always be given to ensuring that physically sensible realizations of (15) and (17) arise from the one-body direct correlation function that any density functional theory will define.

Two classes of violations of statistical mechanical consistency by density functional theory have been noted in the literature.

1. Those simple density functionals that do not make use of coarse-grained density profiles (i.e., attempt delta function smoothing) cannot satisfy the fundamental requirement that $c^{(1)}(\mathbf{r})$ be a continuous function across a hard-wall boundary [i.e., note that $c^{(1)}(\mathbf{r})$ determines the one-body y-function $n(\mathbf{r})$ via (15) and (17)]. Thus such crude density functional theories applied to wall–fluid interfaces in the presence of discontinuous external fields will violate any sum rule that involves integrated force across a discontinuous boundary. However, this effect is fairly benign because all that happens is that each such integration picks up a spurious factor associated with the discontinuity in $n(\mathbf{r})$ (Parry and Evans, 1988).

2. From a careful consideration of (17) it follows that one should avoid density functionals that include a logarithmic term in the excess chemical potential (i.e., the one-body direct correlation function). This is because,

formally, such a term can induce a violation of the positivity of $\rho(\mathbf{r})$ that would otherwise be guaranteed by (17) or its exponentiated form. Thus density functional theories using a one-dimensional hard-rod contribution to the equation of state can lead to spurious negative density profiles if (17) is not treated with special care (Vanderlick et al., 1986).

Given that the one-body direct correlation function defined by a density functional satisfies basic physical requirements, it follows that functional differentiation of the grand potential functional will generate all of the compressibility route sum rules above, in complete consistency with one another. Furthermore, mechanical stability and compatibility with the virial route is ensured via the links (31a), (68b), and so on. It follows that modern density functional theories, based on coarse-grained density profiles, are fully consistent with the structure of statistical mechanics (Tarazona, 1985; Curtin and Ashcroft, 1985). In particular, defining $c^{(1)}(\mathbf{r})$ in terms of a coarse-grained profile guarantees smooth behavior at all times and so avoids violations of class 1 above, as shown explicitly with regard to the mechanical stability sum rule (68) by van Swol and Henderson (1989). However, note that even this extreme internal consistency with statistical mechanics can still be associated with violations of physical behavior, such as failures inherent in all mean-field theories, presumably indicating an unphysical underlying Hamiltonian.

Inverting the emphasis of the discussion above, it should be stressed that density functional theory can contribute much to the interpretation of sum-rule analyses (Evans, 1981; Parry and Evans, 1988). Namely, density functional theories provide explicit realizations of correlation functions that satisfy statistical mechanical sum rules. It follows that sum-rule analyses, such as presented in Section IV, had better not be mathematically incompatible with density functional theory, even if one is suspicious of the physics underlying a given functional.

IV. SUM-RULE ANALYSES OF PLANAR WALL–FLUID INTERFACES

A. Fluids at a Hard Wall

For the sake of clarity it is convenient to begin a discussion of wall–fluid interfaces by focusing on the special case of a hard-wall boundary. Clearly, this model cannot display phenomena associated with the field ϵ_w; this topic is left for Section IV.B. But in other respects the hard-wall limit acts as a prototype of models of wall–fluid interfaces that do not involve power-law interactions (i.e., all the derivations presented in this section have extensions to models with finite-range wall–fluid forces). At

the end of this section I shall discuss how the inclusion of power-law interactions can lead to qualitatively new physics.

Some of the equations derived in Section III acquire a remarkable but occasionally deceptive simplicity when applied to the hard-wall model. In particular, sum rules (68a), (66) and (77a) reduce to, respectively:

$$\frac{p_b}{kT} = \rho_w \equiv \rho(0_+) \tag{78}$$

$$\rho_b - \frac{p_b}{kT} = \rho_w \int_0^\infty dz\, \rho(z) h_0(z,0) \tag{79}$$

$$\frac{\gamma}{kT} = -\rho_w^2 h_2(0,0) \tag{80}$$

and (65) yields

$$\frac{\rho'(z)}{\rho(z)} = \rho_w h_0(z,0) \qquad z > 0 \tag{81}$$

Multiplying (81) by $\rho(z)^s$ and integrating from $z = 0$ out into the bulk fluid generates an infinite set of sum rules, of which (79) is the $s = 1$ version:

$$\frac{\rho_b{}^s - \rho_w{}^s}{s} = \rho_w \int_0^\infty dz\, \rho(z)^s h_0(z,0) \tag{82}$$

where the left side reduces to $\ln(\rho_b/\rho_w)$ in the limit $s \to 0$.

Sum rule (78) links the structure at the wall with the bulk equation of state. Thus moderate to high pressure corresponds to $\rho_w \sigma^d \gg 1$ (where σ denotes a molecular diameter), which implies a wall–liquid profile with strong oscillatory layering structure. The prototype system for studying this packing structure is hard-sphere fluid at a hard wall; for example, Henderson and van Swol (1984) have used molecular dynamics simulation to confirm (81) in the limit $z \to 0$, and Henderson and Plischke (1987) have solved integral equation theories based on (81) and (62). It is equally clear from (78) and (81) that oscillatory interfacial structure will disappear at sufficiently low pressure (i.e., the hard wall will become dry at low p_b). In particular, consider an interface between a hard wall and a bulk liquid at values of (T,μ) close to saturation (i.e., close to bulk liquid–vapor coexistence). Here, p_b approaches the vapor pressure, which will always be lower than $kT\rho_V$ (where ρ_V denotes the vapor density) because the second virial coefficient has to be negative at temperatures below the liquid–vapor critical point (T_c). It follows from (78) that $\rho_w < \rho_V$ and from (81) that $\rho'(z)$ will be very small near the wall. So the equilibrium state of an interface between a hard wall and a saturated liquid is completely

dry (i.e., a macroscopic region of vapor lies between the wall and the liquid).

Since the previous argument applies at all $T < T_c$ it follows that we do not expect to find partial drying at any temperature along the bulk liquid–vapor coexistence curve and with no drying transition there should be no region of prewetting transitions in the approach to saturation, either. So for $T < T_c$ the adsorption isotherms of liquid at a hard wall should all show a continuous growth of adsorbed gas that becomes divergent in the approach to bulk liquid–vapor coexistence (van Swol and Henderson, 1984). Furthermore, this behavior is clearly associated with the divergence of at least one correlation length: the $\xi_\parallel$ capillary-wave correlation length belonging to the liquid–vapor interface that moves out from the wall as the layer of adsorbed gas increases in thickness. So an approach to complete drying is a type of critical phenomena, with the liquid–vapor coexistence curve acting as a line of critical points, associated with a divergent order parameter (the adsorption Γ or the film thickness t) and a divergent correlation length ($\xi_\parallel$):

$$| \Gamma | \equiv (\rho_L - \rho_V)t \sim | \delta\mu |^{-\beta} \tag{83a}$$

$$\xi_\parallel \sim | \delta\mu |^{-\nu_\parallel} \tag{83b}$$

where $\delta\mu$ denotes the distance from saturation in terms of the field chemical potential and I have introduced the complete drying (or wetting) exponents β, $\nu_\parallel$. Below the upper critical dimension for capillary-wave phenomena (Section II.E) we should also include a second divergent correlation length:

$$\xi_\perp \, \xi_\parallel{}^{\zeta} \qquad \zeta \equiv \frac{3 - d}{2} \qquad d < 3 \tag{83c}$$

$$(\lambda\xi_\perp)^2 = \omega \ln [(\lambda\xi_\parallel)^2] \qquad \omega \equiv \frac{kT\lambda^2}{4\pi\gamma_{LV}} \qquad d = 3 \tag{83d}$$

where the value of ζ follows from (40b) and (41c) and the inverse length λ has been introduced from (46) to define the high-Q capillary-wave cutoff, which plays an important role in $d = 3$. Lipowsky and Fisher (1986a) have argued that (83) is immediately generalizable to discussions of wetting phenomena in the presence of random fields; one simply alters the value of the exponent ζ to take account of the fact that the presence of random fields will increase the value of $d_>$. As in all critical phenomena, the divergences above will manifest themselves as a singular contribution to the interfacial free energy [see (42)]:

$$\tilde{\Omega}^{ex} \sim | \delta\mu |^{2-\alpha} \tag{83e}$$

where hereafter a tilde denotes the "singular part of" whenever necessary.

Let us now focus attention on the startling consequences that are implied by attempts to reconcile (83) with sum rule (80) and Eq. (81). For example, note that $h_2(0,0)$ contains the liquid–vapor contribution to the interfacial excess grand potential (γ_{LV}), even in the limit $t \to \infty$. Less mysteriously, (81) is highly suggestive of the capillary-wave correlation phenomena discussed in Section II.E. That is, inside the liquid–vapor part of the profile we anticipate

$$\rho(t)\rho(t)h_0(t,t) \to \frac{\rho'(t)\,\rho'(t)}{W} \qquad \{W,\delta\mu\} \to 0 \tag{84a}$$

Furthermore, substituting (84a) into (20d) implies that

$$G_2(t,t) \to \frac{-\xi_\parallel{}^2\rho'(t)\rho'(t)}{W} \qquad \{W,\delta\mu\} \to 0 \tag{84b}$$

with $\xi_\parallel$ identified from sum rule (77b) as

$$\xi_\parallel{}^2 = \frac{\gamma_{LV}}{kTW} \qquad \delta\mu \to 0 \tag{84c}$$

We can now use (20d) in conjunction with (81) and (84) to derive

$$G_2(t,0) \to -\xi_\parallel{}^2\rho'(t) \qquad \delta\mu \to 0 \tag{85}$$

That is, the wall and the liquid–vapor interface communicate with each other through the presence of long-wavelength capillary-wave correlations, even in the limit $t \to \infty$.

Since $G_2(t,0)$ is singular it is natural to conclude that $G_0(t,0)$ will also contain a singular capillary-wave contribution. But note from (81) that this contribution is finite; that is, the wall–liquid correlation length, $[-G_2(t,0)/G_0(t,0)]^{1/2}$, is identical to the fluctuating interface correlation length, $[-G_2(t,t)/G_0(t,t)]^{1/2}$.

Henderson (1986a) introduced a single-eigenfunction ansatz that contains the behavior (84) and (85):

$$\tilde{G}(z_1,z_2;Q) = \frac{\epsilon(z_1)\epsilon(z_2)}{W(1 + Q^2\xi_\parallel{}^2)} \qquad W \to 0 \tag{86}$$

with W and $\xi_\parallel$ linked by (84c). To solve for the exponents β and $\nu_\parallel$ one can try to identify the behavior of W. When z_1 and z_2 both lie inside the fluctuating part of the interface it follows from (84) that $\epsilon(z) \sim \rho'(z)$. Now, fix z_1 at t and let $z_2 \to 0$ in (86); if (81) or (85) are to hold, we identify

$$W \sim \tilde{\rho}'(0) \tag{87a}$$

From (87a), (84c), and the results of Section II.E, one concludes

$$\xi_{\parallel}^{-2} \sim \begin{cases} I(\lambda,t) & d \leq d_> \\ e^{-\lambda t} & d > d_> \end{cases} \tag{87b}$$

where λ is defined by (46) and $I(\lambda,t)$ by (47). At and below the upper critical dimension $d_>$, the argument leading to (87b) is simply to use (44) with capillary-wave fluctuations truncated at the wall (Percus, 1981; Henderson, 1987b); It is believed that this procedure is equivalent to a linear renormalization-group analysis of fluctuating surfaces in the presence of a hard-wall boundary (see Forgacs et al., 1991, Appendix A).

We are now in a position to present a full sum-rule analysis of the approach to complete drying at a hard wall. First, from sum rule (58) and Eq. (86) at ($z_1 \simeq t$, $z_2 \simeq t$),

$$-\frac{\partial t}{\partial|\delta\mu|} \to \frac{\xi_{\parallel}^{2}(\rho_L - \rho_V)}{\gamma_{LV}} \tag{88}$$

From this result and from sum rule (55) follow exponent relations linking the critical exponents defined in (83):

$$1 + \beta = 2\nu_{\parallel} \tag{89a}$$

$$1 - \alpha = -\beta \tag{89b}$$

and combining these results, we have

$$2 - \alpha = 2\nu_{\parallel} - 2\beta \tag{89c}$$

The rest of the analysis is provided by taking the logarithm of (87b) and combining with (88). Thus in mean field we have

$$\lambda t \to \ln[(\lambda\xi_{\parallel})^2] \qquad d > d_> \tag{90a}$$

with $\nu_{\parallel} = \frac{1}{2}$, $\beta = 0(\ln)$, and $\alpha = 1$. In $d < d_>$, (87b) reduces to

$$(\lambda\xi_{\parallel})^2 \sim \lambda\xi_{\perp} \exp\left(\frac{t^2}{2\xi_{\perp}^{2}}\right) \qquad d < d_> \tag{90b}$$

so provided that $\xi_{\parallel}$ varies as some power of t, it follows that $t \sim \xi_{\perp}$ to within a logarithmic factor; that is,

$$\nu_{\parallel} = \frac{1}{2-\zeta} \qquad \nu_{\perp} = \beta = \frac{\zeta}{2-\zeta} \qquad 2 - \alpha = \frac{2(1-\zeta)}{2-\zeta} \qquad d < d_> \tag{90c}$$

In the special case $d = d_>$, (47b) and (87b) imply the existence of two

regimes:

$$(\lambda\xi_{\parallel})^2 \sim \begin{cases} \exp\left(\lambda t - \dfrac{\lambda^2\xi_{\perp}^{\,2}}{2}\right) & \lambda\xi_{\perp}^{\,2} < t \\ \lambda\xi_{\perp} \exp\left(\dfrac{t^2}{2\xi_{\perp}^{\,2}}\right) & \lambda\xi_{\perp}^{\,2} > t \end{cases} \tag{90d}$$

In the absence of random fields, we can combine the logarithm of this result with (83d) (Henderson, 1987b):

$$\lambda t \to \begin{cases} \left(1 + \dfrac{\omega}{2}\right) \ln\left[(\lambda\xi_{\parallel})^2\right] & \omega < 2 \\ & d = 3 \\ (2\omega)^{1/2}(\ln[(\lambda\xi_{\parallel})^2] - \frac{1}{4}\ln\{\ln[(\lambda\xi_{\parallel})^2]\}) & \omega > 2 \end{cases} \tag{90e}$$

The results above have all been confirmed from specific mean-field models and from renormalization-group analyses of interface Hamiltonians. It is therefore tempting to conclude that ansatz (86) combined with (87) represents the true many-body correlations present in an approach to complete drying. It is interesting, however, that none of the analysis above has relied on sum rule (80); namely, the physics of complete drying does not require (86) to hold at ($z_1 = 0$, $z_2 = 0$), although we can readily see that such an ansatz is a mathematical solution to (80) [i.e., $-\tilde{G}_2(0,0) \sim (\bar{\rho}'(0)/W)^2 \to \gamma_{LV}/kT$]. Despite this, Parry and Evans (1988) have used a specific mean-field theory to provide a counterexample showing that the simplest ansatz for an approach to complete drying (or wetting) is not correct in mean field. That is, they find that (84) and (85) can be associated with the absence of any singular contribution to $G_2(0,0)$.* It is not known if this applies to fluctuation regimes; if so, it could suggest a breakdown of interface Hamiltonian theory applied to complete wetting correlations along the surface of a wall. Note, however, from (80) and (81) that a finite correlation length along the wall will still be dramatically large in the case of complete drying:

$$(\xi_{\parallel}{}^{w})^2 \equiv \frac{-h_2(0,0)}{h_0(0,0)} = \frac{\gamma}{kT\rho_w'} \qquad \rho_w' \equiv \rho'(0_+) \tag{91}$$

where ρ_w' belongs to a wall–vapor interface, while γ is dominated by the liquid–vapor contribution. In fact, Henderson and van Swol (1985) have

* Note in proof: this statement has recently been withdrawn; namely, a reanalysis of their original results has yielded $\xi_{\parallel}{}^{w} \sim \xi_{\parallel}$ (A. O. Parry, 1991, private communication). See also, Mikheev and Weeks (1991), Parry (1992).

used molecular dynamics simulation to verify the presence of dramatic long-range correlations along a hard-wall system close to complete drying; note that these correlations stand out against the virtually weak-gas contribution appropriate to a wall–vapor profile.

Henderson (1986a) extended the foregoing treatment to encompass complete drying or complete wetting at a general wall–fluid interface by invoking (65) in place of the hard-wall result (81) [i.e., now $W \sim \xi_{\parallel}^{-2}$ is given by the result (49)]. For finite-range interactions (87) continues to follow from (49). However, this is not necessarily the case when power-law interactions are included. In particular, if $v(z)$ varies as (51b), then in mean field (49) will be dominated by a power-law term coming from the liquid–vapor interface:

$$\xi_{\parallel}^{-2} \sim t^{-(m+4-d)} \qquad d > d_{>} \tag{92a}$$

The value of $d_{>}$ follows from equating (92a) with (83c) at $\xi_{\perp} \sim t$ (Lipowsky, 1984): $\zeta_{>} = 2/(m + 4 - d_{>})$ and note that in the absence of random fields $d = 3$ is a mean-field regime. Combining (92a) with (89) implies that

$$\nu_{\parallel} = \frac{m + 4 - d}{2(m + 3 - d)} \qquad \beta = \frac{1}{m + 3 - d}$$

$$2 - \alpha = \frac{m + 2 - d}{m + 3 - d} \tag{92b}$$

This result is equally relevant to mean-field regimes of models that include power-law contributions to the fluid–fluid interactions, even if $v(z)$ is finite range. That is, now (53) applies in the tails of the liquid–vapor profile, so (87a) will also imply (92a), for $d > d_{>}$. Interestingly, wall–fluid and fluid–fluid power-law contributions could conceivably cancel at leading order (at a particular temperature), thereby inducing higher-order critical regimes (as in Section IV.B).

B. Critical Wetting

Introduction of the field ϵ_w, as in model (8), enables us to control wetting phenomena at bulk liquid–vapor coexistence without having to vary temperature, as discussed in Section II.B. The most dramatic consequences of this degree of freedom is the appearance of wetting transitions at bulk liquid–vapor coexistence, between partial wetting and complete wetting (increasing ϵ_w) or complete drying (decreasing ϵ_w). If a wetting (or drying) transition is second order, it is termed critical wetting (drying). Here, one has a set of exponents directly analogous to (83) but not necessarily taking the same values, with $\delta\mu$ replaced by $\delta\epsilon_w$. In fact, complete wetting and

critical wetting may be combined within a single scaling ansatz [except at $d = d_>$, which can involve complications leading to a breakdown of simple scaling (Parry and Evans, 1989)]:

$$\tilde{\Omega}^{ex} = |\delta\epsilon_w|^{2-\alpha^1} f(|\delta\mu| \, |\delta\epsilon_w|^{-\Delta}) \tag{93a}$$

where hereafter superscript 1 denotes critical wetting, and the crossover exponent Δ links the two sets of exponents:

$$\Delta = \frac{2-\alpha^1}{2-\alpha^c} = \frac{\nu_\parallel{}^1}{\nu_\parallel{}^c} = \frac{\beta^1}{\beta^c} \tag{93b}$$

with superscript c denoting an approach to the critical wetting point from off two-phase coexistence (a special case of complete wetting). The scaling form (93) is equally applicable to fixed ϵ_w with $\delta\epsilon_w$ replaced by δT; that is, away from any tricritical point associated with a change from first-order wetting to critical wetting, the fields ϵ_w and T play equivalent roles. This point is significant to proposed comparisons with experiment (Henderson, 1987a).

The sum rules of Section III have much to say concerning critical wetting, whenever it occurs (Henderson, 1986a). For example, if the external field $v(z)$ is strictly finite range [$v(z) = 0$, $z > a_w$], then (59) and (60) imply, introducing a length $a < a_w$,

$$\chi_1 \sim G_0(t,a) \qquad \text{more precisely, } G_0(t,a) \sim \rho'(t)\chi_1 \tag{94a}$$

$$\tilde{\chi}_{11} \sim \tilde{G}_0(a,a) \tag{94b}$$

and substituting (94a) into (20d) gives

$$\chi_1{}^2 \sim G_2(a,a) \tag{94c}$$

Thus critical wetting in finite-range models must be associated with a divergent correlation length along the wall:

$$(\xi_\parallel{}^w)^2 \equiv \frac{-G_2(a,a)}{\tilde{G}_0(a,a)} \sim \frac{\chi_1{}^2}{\tilde{\chi}_{11}} \sim |\delta\epsilon_w|^{-(2-\alpha^1+2\beta^1)} \sim (\xi_\parallel{}^1)^2 \tag{95}$$

where the very last proportionality follows from the scaling ansatz (93b) together with the generalized hyperscaling relation (89c) [i.e., the exponent relation (89c) is equally valid for critical wetting]. Thus at critical wetting not only do the capillary-wave correlations manifest themselves at the wall, they do so with a correlation length that diverges just as strongly along the wall as it does inside the detaching liquid–vapor interface. Parry and Evans (1990a) have made an important connection between the conclusion above and the exact "C_p–C_v" relation (5), which

can be rewritten in terms of the compressibilities (58) to (60) as

$$\chi\left[\chi_{11} - \left(\frac{\partial\Theta}{\partial\epsilon_w}\right)_t\right] = \chi_1^2 \tag{96}$$

In particular, note that if one can neglect any singular contribution to $(\partial\Theta/\partial\epsilon_w)_t$, it follows that

$$(\xi_\parallel^{\,1})^2 \to \frac{\gamma_{LV}}{(\rho_L - \rho_V)^2}\frac{\chi_1^2}{\bar{\chi}_{11}} \tag{97a}$$

where $\xi_\parallel^{\,1}$ has been defined from (88); that is, (88) holds below complete wetting as well, given only that (86) be appropriate inside a fluctuating interface ($z_1 \simeq t$, $z_2 \simeq t$). So (95), and hence scaling, arises because $(\partial\Theta/\partial\epsilon_w)_t$ is no more singular than χ_{11}, as would seem eminently reasonable.

The obvious solution of (95) is ansatz (86) (i.e., to assume that there is only one capillary-wave correlation length $\xi_\parallel$ at all values of z_1 and z_2). If the single-eigenfunction ansatz (86) is inserted into sum rules (58) to (60), one finds that (Henderson, 1986a)

$$\chi\bar{\chi}_{11} \to \chi_1^2 \tag{97b}$$

Comparing (96) and (97) it is natural to associate any singular contribution to $(\partial\Theta/\partial\epsilon_w)_t$ with a renormalized capillary-wave correlation length along the wall (Parry and Evans, 1990a):

$$K^{-1} \equiv \left(\frac{\xi_\parallel^{\,w}}{\xi_\parallel^{\,1}}\right)^2 \equiv 1 - \frac{(\partial\bar{\Theta}/\partial\epsilon_w)_t}{\bar{\chi}_{11}} \tag{98a}$$

$$\chi\bar{\chi}_{11} \to K\chi_1^2 \tag{98b}$$

In fact, Parry and Evans (1990a) show that when $d < d_>$ hyperscaling implies that $K \neq 1$. Thus, at the least, ansatz (86) cannot be completely correct in the strong fluctuation regime of critical wetting.

In the presence of a power-law external field, (51b), sum rule (56) implies that Θ contains a power-law singular contribution:

$$\bar{\Theta} \sim t^{-(m+2-d)} \tag{99a}$$

In regimes in which (99a) dominates over the capillary-wave contribution to $\bar{\Theta}$, it follows immediately that $(\partial\Theta/\partial\epsilon_w)_t \to 0$ at leading order. Thus, in such power-law regimes, $K = 1$ [i.e., (97b) is the relevant solution]. In fact, it is trivial to show that $\bar{\Theta} \to \bar{\Theta}(t)$ is a general solution to (97b), as required by the C_p–C_v relation (96). This is important because (99a) is a first-order partial differential equation for the singular part of the free energy and is thus readily solved by standard methods. In particular, by substituting the scaling form (93) into (99a), Henderson (1987a) obtained

a parametric solution to the scaling function $f(|\,\delta\mu\,|\,|\,\delta\epsilon_w\,|^{-\Delta})$; more revealing is the resulting equation of state:

$$|\,\Gamma\,|^{1/\beta^c}|\,\delta\mu\,| = b(T)|\,\Gamma\,|^{1/\beta^1}\,\delta\epsilon_w + c(T) \tag{99b}$$

A full sum-rule analysis shows that (99) is appropriate to two regimes of critical wetting (Henderson, 1986a). First, one has, from (99a),

$$1 - \alpha^1 = (m + 2 - d)\beta^1 \tag{99c}$$

and from (97),

$$2 - \alpha^1 = 2\nu_{\parallel}{}^1 - 2\beta^1 \tag{99d}$$

So it only remains to determine one of the exponents, say β^1. If ansatz (86), which should apply when $K = 1$, is inserted into sum rule (59), χ_1 reduces to the ratio of two terms of the form $v(t)/(49)$, assuming that $\bar{\epsilon}(z) \sim \bar{\rho}'(z)$ at all z. Thus ansatz (86) forces us to associate the critical wetting point with $W = 0$ at leading order; in fact, mean-field theories associate critical wetting with a leading-order cancellation of terms in the free energy (Dietrich and Schick, 1985; Ebner et al. 1985) and this effect should also appear in W because of the general relation (38b). It follows that the mean-field regime of noncritical complete wetting, (92), is associated with two regimes of critical wetting: that is, a mean-field regime (MF) determined by a higher-order power-law term in W, and a so-called weak fluctuation regime (WFL) when the term that takes over after $W = 0$ at leading order is the capillary-wave term (90b). That is,

$$W \sim a\,\delta\epsilon_w\,t^{-(m+4-d)} + bt^{-(m+5-d)} + cI(\lambda,t) \tag{99e}$$

where the analysis of Henderson (1986a) holds because the first term on the right side of (99e) is never more singular than the dominant remaining term (A. O. Parry, 1989, private communication). In both cases one finds that $\Delta = 1 + (m + 3 - d)\beta^1$, so the two regimes are distinguished by their values of β^1 (Henderson, 1986a):

$$\begin{aligned} &\text{MF:}\quad \beta^1 = 1 \qquad \zeta < \frac{2}{m+5-d} \\ &\text{WFL:}\quad \beta^1 = \frac{\zeta}{2-(m+4-d)\zeta} \qquad \frac{2}{m+5-d} < \zeta < \frac{2}{m+4-d} \end{aligned} \tag{99f}$$

Note that the equation of state (99b) is equally applicable to an approach to complete wetting (or drying) states lying above the critical wetting (drying) point (i.e., $[(1/\beta^c) - (1/\beta^1)] = (\Delta - 1)/\beta^1 = m + 3 - d$ as appropriate to MF complete wetting). Furthermore, this remarkably complete solution contains the regime directly relevant to experiments on critical wetting in $d = 3$. Oddly, no experimental tests of the results

summarized in Henderson (1987a) have been announced; note, for example, that the theoretical equation of state (99b) is equally applicable to fixed ϵ_w with $\delta\epsilon_w \to \delta T$ and thus (99b) could be used to locate a critical wetting point from adsorption data collected along sets of complete wetting isotherms.

A similar level of success has been achieved with a sum rule analysis of critical wetting in models without power-law forces (Henderson, 1987b). Following Aukrust and Hauge (1985), it is convenient to consider an exponential external field model:

$$v(z) = \begin{cases} \infty & z < 0 \\ -\epsilon_w \exp(-\lambda_w z) & z > 0, \quad \lambda < \lambda_w \leq 2\lambda \end{cases} \tag{100}$$

where $\lambda_w = 2\lambda$ would be mathematically equivalent to finite range forces [see (46)] and $\lambda_w < \lambda$ always corresponds to first-order wetting transitions. The sum rules of Section III are especially revealing for model (100); namely, combining with (65) it is straightforward to derive the following exact results:

$$\Theta = \frac{kT}{\epsilon_w \lambda_w}\left(\rho_w - \frac{p_b}{kT}\right) \tag{101a}$$

$$\chi_1 = \frac{1}{\epsilon_w \lambda_w}\left[\int_0^\infty dz\, G_0(z,0_+) - \rho_b\right] \tag{101b}$$

$$\begin{aligned} \chi_{11} &= \frac{-kT}{\epsilon_w{}^2 \lambda_w}\left[\frac{1}{kT}\int_0^\infty dz\, v(z) G_0(z,0_+) + \rho_w - \frac{p_b}{kT}\right] \\ &= \frac{-kT}{\epsilon_w{}^2 \lambda_w{}^2}\left[\rho_w' - \rho_w^2 h_0(0,0) + \lambda_w\left(\rho_w - \frac{p_b}{kT}\right)\right] \end{aligned} \tag{101c}$$

Note in particular that $\tilde{\chi}_{11} \sim [\tilde{\rho}_w' - \tilde{G}_0(0,0)]$ is inescapable from (101c). Combining (101a) and (101c), we have

$$\frac{\epsilon_w}{kT}\frac{\partial}{\partial \epsilon_w}(\epsilon_w \Theta) = \frac{\rho_w{}^2 h_0(0,0) - \rho_w'}{\lambda_w{}^2} \tag{101d}$$

which is perhaps the most precise way of highlighting the fact that in models with short-range interactions critical wetting correlations contributing to the singular part of the free energy must show up at the wall. The only source of such correlations is presumably capillary-wave phenomena. The analysis of Henderson (1987b) proceeds as before via (101b):

$$\xi_\parallel{}^{-2} \sim W \sim \begin{cases} a\,\delta\epsilon_w e^{-\lambda t} + b e^{-\lambda_w t} & d > d_> \\ a\,\delta\epsilon_w I(\lambda,t) + b I(\lambda_w,t) & d = 3 \end{cases} \tag{102}$$

In $d = 3$ one again has two regimes (renormalized mean-field (RMF) and WFL) depending on which form of (47b) is determining the value of

$I(\lambda_w,t)$; in both regimes the parameter ω appears in the exponents, as with complete wetting in $d = 3$ [see (90e)]. In MF and RMF regimes the ratio λ_w/λ enters the exponents; this nonuniversality is directly analogous to the power-law case because once again it results from a term dominated by $v'(t)$. So, in certain circumstances even an exponential force need not always be a short-range effect.

The sum-rule analysis described above generates all the singular behavior appropriate to complete wetting and critical wetting phenomena, apart from strong fluctuation regimes (SFL) of critical wetting. Thus many of the results known from phenomenological mean field and interface Hamiltonian studies have been shown to apply to molecular Hamiltonians and explicit connections with microscopic correlations have been made via ansatz (86). This success is especially striking in those cases of direct relevance to experiment. SFL regimes of critical wetting constitute an especially difficult case because the cancellation argument employed by Henderson (1986a) is no longer helpful and may not even be correct. That is, in a SFL regime all terms in W are now dominated by capillary-wave singularities and further analysis of critical wetting is impossible without knowing how all these terms combine at the critical wetting point; the sentence following (101c) is possibly the key here. Furthermore, the C_p–C_v analysis of Parry and Evans (1990a) that leads to $K \neq 1$ in (98b) for $d < d_>$, shows that ansatz (86) cannot be the full story behind SFL critical wetting. Nevertheless, Evans and Parry (1989) have gone some way to providing a sum-rule analysis of the SFL regime in $d = 3$, based on identifying $\bar{\Theta}$ from capillary-wave theory; for this case they even suggest that (86) could be applicable provided that the singular behavior of the eigenfunction is generalized beyond $\tilde{\epsilon}(z) \sim \tilde{\rho}'(z)$ as $z \to 0$. However, in $d < d_>$ even (47)–(48) is in doubt, since it apparently leads to incorrect exponents previously associated with a linear renormalization group analysis of interface Hamiltonian theory (Parry and Evans, 1990a). Here it would be helpful to know what the structure of $\tilde{G}(z_1,z_2;Q)$ is in interface Hamiltonian theory, since this should be directly relevant to molecular models of a SFL regime; in particular, see the numerical nonlinear renormalization-group technique introduced by Lipowsky and Fisher (1986b) and the recent analytical treatment given by David and Leibler (1990), and for explicit results in $d = 2$ see Parry (1992).

V. SUM-RULE ANALYSES OF THE EFFECTS OF GEOMETRY

A. Confinement

Let me turn attention to model (9), the basic model of confinement, to discuss how sum rules lead to an understanding of the significance of the

field L. Once more, it is revealing to begin with a discussion of the limit of hard-wall boundary conditions. In particular, (64) and (65) reduce to (Henderson, 1986b):

$$\frac{\partial \rho(z)}{\partial L} = G_0(z,L_-) \qquad 0 < z < L_- \tag{103a}$$

$$\rho'(z) = G_0(z,0_+) - G_0(z,L_-) \qquad 0 < z < L \tag{103b}$$

and sum rules (57), (61), (67), (73), and (77a) give

$$f = kT(\rho_w - \rho_w^{\infty}) \tag{103c}$$

(Henderson, 1986b)

$$\chi_L = kT\frac{\partial \rho_w}{\partial L} = kTG_0(0_+,L_-) \tag{103d}$$

(Henderson, 1986b)

$$\frac{1}{A}\frac{\partial N(L)}{\partial L} = \int_0^{\infty} dz\, G_0(z,L_-) = kT\frac{\partial \rho_w}{\partial \mu} = \frac{\partial f}{\partial \mu} + \rho_b \tag{103e}$$

$$\frac{\Omega^{ex}}{A} = 2\gamma - Lf \tag{103f}$$

$$\frac{\gamma}{kT} = -\rho_w^2[h_2(0,0) - h_2(0,L)] \tag{103g}$$

The final result requires only a trivial generalization of the proof leading to (77). Namely, replace (74) by

$$v(r;R_G) = v(r - R_G) + v(R_G + L - r)$$

and repeat the derivation finishing with $R_G \to \infty$ at fixed L; that is, the $\sim R_G{}^{d-1}Lp_b$ volume term and the two external field contributions to Ω yield the last two terms on the right side of (73a), while the many-body contribution shows that (77) is equally applicable to the slit problem, except that as $L \to \infty$ it is now the sum of two many-body interfacial free energies.

Sum rule (103c) explains most of the physics behind oscillatory solvation force structure, as arising from a competition between repulsive wall–fluid and repulsive fluid–fluid interactions. Here, one should also note that for hard-wall slits $kT\rho_w$ is just the normal component of the pressure tensor in the slit; p_N is constant throughout the slit because the external field gradient is zero except at the walls, and see, for example, (69). Sum rule (103d) is particularly striking in the presence of two-dimensional critical phenomena (Evans and Parry, 1990); I shall return to this below. In addition, it is worth pointing out that further compressibility

sum rules are readily added to follow (61), by combining isothermal derivatives among the full set (μ,ϵ_w,a_w,L); for example,

$$\frac{\partial f}{\partial \mu} = \frac{1}{A}\frac{\partial N(L)}{\partial L} - \rho_b \tag{104a}$$

where I have substituted (67). For a hard-wall slit (104a) reduces to (103e) (Evans and Parry, 1990):

$$\frac{\partial f}{\partial \mu} = kT\frac{\partial \rho_w}{\partial \mu} - \rho_b \tag{104b}$$

Furthermore, as pointed out at the end of Section III.A, all of the results above are readily transformed into direct correlation function integral equations and sum rules [see also (68b)].

Henderson (1986b) has used (65) to argue that a good approximation to the density profile of the general symmetric slit model (9) is a product of single-wall solutions ($L \to \infty$) multiplied by an exponential amplitude factor that oscillates with the solvation force. Interestingly, it has recently been demonstrated that in the absence of phase transitions equations (64) and (65) provide numerically stable routes to integral equation theories of confined fluids; both in the case of hard-wall slits [with (103b) combined with (103a) at $z = 0$ (Kjellander and Sarman, 1988)] and in continuous force models [combining (65) with (67) (Kjellander and Sarman, 1990)]. Here the idea is to integrate in from the large L regime. Henderson (1986b) obtained a limiting solution to (103b) as $L \to 0$. Recently, this result has been generalized to arbitrary potentials of the class (9), including confined fluid mixtures (Adams et al., 1989). The solution reduces to a universal form for the one-body y-function:

$$\rho(z) \to n\left(\frac{L}{2}\right)\exp\left[\alpha\left(z - \frac{L}{2}\right)^2\right]\exp\left[\frac{-v(z)}{kT}\right] \tag{105}$$

valid to a remarkably high degree of accuracy throughout the quasi two-dimensional regime; that is, for L small enough such that $\rho^{(2)}(|\, z_{12}\,|,z_1,z_2)$ is everywhere negligible (molecules cannot pass directly over one another when moving parallel to the slit walls).

Henderson (1986b) also discussed the significance of potential distribution theory, (29), to highly confined fluids. In particular, in open pores (the grand ensemble) fluid is squeezed out the ends of the pore as $L \to 0$, so that at finite chemical potential the system must reduce to a two-dimensional weak gas at $L = 0$:

$$\rho(z) \to \Lambda^{-3}\exp\left[\frac{\mu - v(z)}{kT}\right] \qquad L \to 0 \tag{106a}$$

For example, note the explicit solution of Robledo and Rowlinson (1986) for a grand ensemble of hard rods on a line of finite length, which confirms that the density profile approaches a finite limit as $L \to 0$; in a three-dimensional system it thus follows that the effective two-dimensional density $\int dz\, \rho(z)$ tends to zero as $L \to 0$. The potential distribution theorem can also be used to extract the leading-order small L dependence of the density profile and the solvation force (Henderson, 1986b). Not surprisingly, it is determined by the two-dimensional second virial coefficient. However, these results can be obtained more directly from (64) and sum rule (61) [e.g., Henderson (1986b)]:

$$kT\chi_L \to \int_{-\infty}^{\infty} dz\, \rho(z)v^{\infty\prime}(z)v^{\infty\prime}(L - z) - 2B_2\left[\int_{-\infty}^{\infty} dz\, \rho(z)v^{\infty\prime}(z)\right]^2 \qquad L \to 0 \quad (106b)$$

where B_2 denotes the two-dimensional second virial coefficient [see also Wertheim et al. (1989)].

The discussion above concerning the limit $L \to 0$ has ignored one significant point. Namely, as three-dimensional systems pass through a quasi two-dimensional regime, one will often expect to observe phase transitions of a two-dimensional character. Thus at temperatures below the two-dimensional critical temperature, a slit filled with liquid must undergo capillary evaporation at some sufficiently small value of L. At any T, sufficiently high μ will lead to capillary freezing, so freezing transitions will sometimes be induced by a decrease in L at fixed (T,μ). At temperatures lying between the two- and three-dimensional critical points, liquid–vapor coexistence will end at a capillary critical point somewhere before $L \to 0$. A capillary critical point is a shifted bulk critical point, as obtained from the scaling theory of Fisher and Nakanishi (1981). At a capillary critical point in slit geometry the correlation length diverges along the slit and the system displays critical phenomena of a two-dimensional universality class. As pointed out by Evans and Parry (1990), a sum rule analysis of capillary criticality would show striking behavior in the compressibility route correlation functions. For example, since $f_L - f_V$ is an order parameter, both χ_L and $\partial f/\partial\mu$ are divergent at a capillary critical point; once again, hard-wall boundary conditions are particularly revealing [see sum rules (103d) and (103e)].

Finally, let me emphasize that sum-rule analyses should be equally revealing of confined fluid phenomena outside the realm of model (9). For example, by introducing left- and right-hand attractive field strengths $(\epsilon_W{}^L, \epsilon_W{}^R)$, one could investigate the consequences of antagonistic walls.

Here one should note the seminal work of Parry and Evans (1990b), who have shown that liquid–vapor coexistence in an asymmetric slit ($\epsilon_W{}^L = -\epsilon_W{}^R$) is governed by the wetting transition and drying transition temperatures of the two walls (i.e., the two semi-infinite systems), not by a bulk critical temperature as in the symmetric slit case discussed above.

B. Nonplanar Geometry

The symmetry of inhomogeneous fluid systems can be controlled through an external field and hence through the compressibility route hierarchy (10). In nonplanar geometry one will often expect to be able to write the grand potential as a function of curvature variables. For example, consider fluid confined to a cylindrical pore or a spherical cavity and introduce a dividing surface at radius R_G:

$$\Omega^{ex} \equiv \begin{cases} 2\pi R_G L\sigma(R_G) \equiv \Omega + \pi R_G{}^2 L p_b & \text{cylinder} \quad (107a) \\ 4\pi R_G{}^2\sigma(R_G) \equiv \Omega + \frac{4\pi}{3} R_G{}^3 p_b & \text{sphere} \quad (107b) \end{cases}$$

where in the limit $R_G \to \infty$ we expect $\sigma(R_G) \to \sigma^\infty = \Omega^{ex}/A$ belonging to a planar wall–fluid system. Analogous to the planar solvation force (57), we can introduce a disjoining pressure:

$$f \equiv -\frac{1}{A}\frac{\partial\Omega^{ex}}{\partial R_G} = -\begin{cases} \frac{\sigma}{R_G} + \frac{\partial\sigma}{\partial R_G} & \text{cylinder} \\ \frac{2\sigma}{R_G} + \frac{\partial\sigma}{\partial R_G} & \text{sphere} \end{cases} \quad (107c)$$

Introducing one-body external fields of the form

$$v(R - R_G) \quad \text{cylinder} \quad (108a)$$

$$v(r - R_G) \quad \text{sphere} \quad (108b)$$

and using (10a) gives sum rules analogous to (57):

$$f = \begin{cases} \int_0^\infty dR\, \frac{R}{R_G}\, \rho(R) v'(R - R_G) - p_b & \text{cylinder} \quad (109a) \\ \int_0^\infty dr \left(\frac{r}{R_G}\right)^2 \rho(r) v'(r - R_G) - p_b & \text{sphere} \quad (109b) \end{cases}$$

The choice (108) has the property that it possesses a well-defined planar limit, as can be seen by comparison of (109) with (57). The virial route

can be expressed in terms of pressure tensors defined by (23) in the appropriate symmetry:

$$p_N'(R) + \frac{1}{R}[p_N(R) - p_T(R)] = -\rho(R)v'(R - R_G) \quad \text{cylinder} \quad (110a)$$

$$p_N'(r) + \frac{2}{r}[p_N(r) - p_T(r)] = -\rho(r)v'(r - R_G) \quad \text{sphere} \quad (110b)$$

Thus we can rewrite (109) in the form

$$f = -\begin{cases} \dfrac{1}{R_G}\displaystyle\int_0^\infty dR[p_b\Theta(R_G - R) - p_T(R)] & \text{cylinder} \quad (111a) \\ \dfrac{2}{R_G}\displaystyle\int_0^\infty dr\,\frac{r}{R_G}[p_b\Theta(R_G - r) - p_T(r)] & \text{sphere} \quad (111b) \end{cases}$$

where $\Theta(x) \equiv \{0, x < 0; 1, x > 0\}$ is Heaviside's step function. Note, for example in the latter case, that it is the first moment of $p_T(r)$ that is well defined by (23b) in spherical symmetry, not the second moment.

We can equally well use (108) to discuss fluids surrounding a cylinder or a sphere. A special case of the latter is provided by the test-particle route to the statistical mechanics of pair potential fluids, discussed previously in conjunction with potential distribution theory (Section II.D). When the frozen test particle acts as a hard sphere, the statistical mechanics above is known as scaled particle theory; a detailed discussion of this connection is given by Henderson (1983). In Section III.C, I introduced the compressibility route to spherical geometry as a means of deriving σ^∞ from the $R_G \to \infty$ limit. Given suitably behaved interfacial correlations, one expects to be able to define a radius term in the grand potential for three-dimensional spherical systems:

$$\sigma(R_G) \to \sigma^\infty\left(1 - 2\frac{\delta}{R_G}\right) \quad \text{sphere} \quad (112)$$

where from the case of liquid drops δ is known as Tolman's length. Let us now consider the compressibility route to Tolman's length. Beyond (112) it is not realistic to expect a curvature expansion of $\sigma(R_G)$ to be of much physical relevance; for example, there could well be a $\ln R_G$ term in Ω and anyway, a constant term in the grand potential cannot be studied by the compressibility route. For these reasons one does not introduce a Tolman length in two dimensions or in cylindrical symmetry. Thus I shall restrict the following discussion to spherical symmetry in $d = 3$; the cylindrical geometry version of the derivation leading to (76) has been given by Henderson and Rowlinson (1984).

From (76) it is apparent that δ is defined in general by a third-order derivative of Ω (i.e., derivatives up to and including $\partial^3\Omega/R_G{}^3$) and thus will involve the three-body density–density distribution function:

$$\left(\frac{\delta^3\Omega}{\delta[\mu - v(1)]\delta[\mu - v(2)]\delta[\mu - v(3)]}\right)_T \equiv -\left(\frac{1}{kT}\right)^2 G(123) \quad (113)$$

which follows on from (10b). Again, let us control R_G through the external field (74). Then we need to evaluate

$$\delta \equiv \frac{1}{2\sigma^\infty} R_G{}^2 \frac{\partial}{\partial R_G}\sigma(R_G) \qquad R_G \to \infty \quad (114a)$$

$$= \frac{1}{8\pi\sigma^\infty} R_G{}^2 \frac{\partial}{\partial R_G}(76)_{d=3} \qquad R_G \to \infty \quad (114b)$$

Inserting (76) into (114b) and making use of (113), one finds, after a certain amount of algebra, the following general expression for Tolman's length:

$$\begin{aligned}\delta = -\frac{kT}{32\sigma^\infty}\Bigg[&\left(\frac{1}{kT}\right)^3 \int_{-\infty}^{\infty} dz_1 \int d2 \int d3 v'(z_1)v'(z_2)v'(z_3) \\ &\times R_{12}^2(R_{12}^2 - R_{13}^2 - R_{23}^2)G(123) \\ &+ 8\left(\frac{1}{kT}\right)^2 \int_{-\infty}^{\infty} dz_1(z_1 - z_G)\int d2 v'(z_1)v'(z_2)R_{12}^2 G(12) \\ &- 16\frac{1}{kT}\int_{-\infty}^{\infty} dz(z - z_G)^2\rho(z)v'(z)\Bigg] \end{aligned} \quad (114c)$$

where σ^∞ is given by (77a):

$$\begin{aligned}\frac{\sigma^\infty}{kT} = \frac{1}{4}\left(\frac{1}{kT}\right)^2 &\int_{-\infty}^{\infty} dz_1 \int d2 v'(z_1)v'(z_2)R_{12}^2 G(12) \\ &- \frac{1}{kT}\int_{-\infty}^{\infty} dz\, z\rho(z)v'(z) \end{aligned} \quad (114d)$$

and all the distribution functions now refer to the planar limit with fluid occupying the space $z > 0$. The last two terms on the right side of (114c) depend explicitly on the choice of R_G and are thus external field terms. For the case of a hard-wall boundary [i.e., a hard-sphere (HS) cavity] both these terms disappear:

$$\delta_{HS} = \left(\frac{1}{8}\right)\frac{\int d\mathbf{R}_2 \int d\mathbf{R}_3 R_2{}^2(R_2{}^2 - R_3{}^2 - R_{23}^2)G_w(\mathbf{R}_2,\mathbf{R}_3;\mathbf{R}_1 \equiv \mathbf{0})}{\int d\mathbf{R}_{12}R_{12}^2 G_w(R_{12})} \quad (115)$$

where subscript w denotes all coordinates lying on the wall. Stecki and Toxvaerd (1990) have used molecular dynamics simulation in an attempt to extract δ_{HS} from a curvature expansion of the grand potential of Lennard–Jones liquid surrounding a hard-sphere cavity, via (109b). Perhaps not surprisingly, given the expression (115), Stecki and Toxvaerd (1990) find that the magnitude of δ_{HS} is very small.

To apply (114) to a liquid–vapor drop, one needs to negotiate the tricky question of how to take the limit of infinitesimal external field in the second term appearing on the right side of (114c). One is immediately suspicious of this term because it depends on the choice of R_G and yet will not obviously vanish in the weak-field limit, unlike the final terms on the right sides of (114c) and (114d). In fact, if we substitute the capillary-wave expression (49)–(50) for $G_2(12)$, it follows that this two-body contribution to Tolman's length is simply

$$\begin{aligned} \delta_{2\text{-body}} &= \frac{\int_{-\infty}^{\infty} dz(z_G - z)\rho'(z)v'(z)}{\int_{-\infty}^{\infty} \rho'(z)v'(z)} \qquad W \to 0 \\ &\equiv z_G - z_0 \end{aligned} \tag{116}$$

If $v(z)$ was taken to be a step function at R_G, (116) would be zero. Or, if one uses $v(z) = mgz$ and argues that a gravitational field would have to be able to act on the equimolar surface, then again (116) would be zero. The conclusion here is that in the limit of vanishing external field, the $\partial/\partial R_G$ derivatives used to derive (114) to (116) only lead to a well-defined planar limit at fixed T,μ in the case $R_G \to z_0$, defined by (116). Thus we conclude that Tolman's length for a free liquid–vapor drop (δ_{LV}) is, like δ_{HS}, given by a pure three-body correlation function sum rule [i.e., the first term on the right side of (114c)]. Of course, δ_{LV} only exists if the particular moment of $G(123)$ defined in (114c) is well defined. That is, a noninfinite value for Tolman's length places strong restrictions on the capillary-wave contribution to the three-body density–density distribution function [see (117)]. It is also worth noting the penetrable sphere model result of Rowlinson (1984), who found that δ_{LV} diverges much less strongly than the bulk correlation length in the approach to the critical point.

Finally, it should be mentioned that it would be expected to be possible to invert (114c) to obtain a corresponding direct correlation function expression; for example, from the result proposed by Phillips and Mohanty (1985), one might hope to be able to show that $(kT)^{-3}v'(z_1)v'(z_2)v'(z_3)G(123)$ can be replaced by $-\rho'(z_1)\rho'(z_2)\rho'(z_3)$

$G^{-1}(123)$, directly analogous to (77). Here we would presumably need the second order Ornstein–Zernike equation [see, e.g., Henderson (1987c)]:

$$G(123) = -\int d4 \int d5 \int d6\, G^{-1}(456)G(14)G(25)G(36) \qquad (117)$$

However, using this result, the present author did not find the simple inversion suggested by analogy with (77). The direct correlation function expression obtained by Phillips and Mohanty (1985) is somewhat dubious because the derivation relies on extracting a second-order curvature term from fluctuation theory of a planar interface; nevertheless, the basic form of their solution is highly suggestive of the structure of (114c). Phillips and Mohanty (1985) placed great emphasis on a term analogous to $\delta_{2\text{-body}}$, but the analogy with (116) suggests that their interpretation and the consequences claimed for the critical point divergence of δ_{LV} are most unlikely.

ACKNOWLEDGMENTS

I thank F. van Swol for many years of fascinating collaboration on the physics of inhomogeneous fluids. The conceptual framework that I have presented in this chapter owes much to the stimulus and encouragement of R. Evans. A. O. Parry is thanked for detailed critical comments that have been used in formulating parts of the discussion of Section IV.

REFERENCES

Abraham, D. B. (1980). Solvable model with a roughening transition for a planar Ising ferromagnet, *Phys. Rev. Lett. 44*:1165.

Adams, P., Henderson, J. R., and Walton, J. P. R. B. (1989). The quasi-two-dimensional regime of fluids absorbed in porous media, *J. Chem. Phys. 91*: 7173.

Aukrust, T., and Hauge, E. H. (1985). Nonuniversal ν from a van der Waals theory of the wetting transition, *Phys. Rev. Lett. 54*: 1814.

Barker, J. A., and Henderson, J. R. (1982). Generalized van der Waals theories and the asymptotic form of the density profile of a liquid–vapor interface, *J. Chem. Phys. 76*: 6303.

Bedeaux, D., and Weeks, J. D. (1985). Correlation functions in the capillary wave model of the liquid–vapor interface, *J. Chem. Phys. 82*: 972.

Binder, K., Landau, D. P., and Kroll, D. M. (1986). Critical wetting with short-range forces: is mean-field theory valid? *Phys. Rev. Lett. 56*: 2272.

Buff, F. P., Lovett, R. A., and Stillinger, F. H. (1965). Interfacial density profile for fluids in the critical region, *Phys. Rev. Lett. 15*: 621.

Cahn, J. W. (1977). Critical point wetting, *J. Chem. Phys. 66*: 3667.

Ciach, A. (1987). Slow decay of the direct correlation function in the three-dimensional capillary wave model of the interface, *Phys. Rev. A 36*: 3990.

Curtin, W. A., and Ashcroft, N. W. (1985). Weighted-density-functional theory of inhomogeneous liquids and the freezing transition, *Phys. Rev. A 32*: 2909.

David, F., and Leibler, S. (1990). Multicritical unbinding phenomena and non-linear functional renormalization group, *Phys. Rev. B 41*: 12926.

de Gennes, P. G. (1981). Some effects of long range forces on interfacial phenomena, *J. Phys. Lett. 42*: L377.

Derjaguin, B. (1940). A theory of capillary condensation in the pores of sorbents and of other capillary phenomena taking into account the disjoining action of polymolecular liquid films, *Acta Physicochim. URSS 12*: 181.

Dietrich, S., and Schick, M. (1985). Critical wetting of surfaces in systems with long-range forces, *Phys. Rev. B 31*: 4718.

Durian, D. J., and Franck, C. (1987). Wetting phenomena of binary liquid mixtures on chemically altered substrates, *Phys. Rev. Lett. 59*: 555.

Ebner, C., and Saam, W. F. (1977). New phase-transition phenomena in thin Argon films, *Phys. Rev. Lett. 38*: 1486.

Ebner, C., Saam, W. F., and Sen, A. K. (1985). Wetting transitions in systems with van der Waals forces, *Phys. Rev. B 32*: 1558.

Evans, R. (1979). The nature of the liquid–vapor interface and other topics in the statistical mechanics of non-uniform classical fluids, *Adv. Phys. 28*: 143.

Evans, R. (1981). The role of capillary wave fluctuations in determining the liquid–vapor interface: analysis of the van der Waals model, *Mol. Phys. 42*: 1169.

Evans, R. (1990). Fluids adsorbed in narrow pores: phase equilibria and structure, *J. Phys. Condens. Matter 2*:8989.

Evans, R., and Marini Bettolo Marconi, U. (1987). Phase equilibria and solvation forces for fluids confined between parallel walls, *J. Chem. Phys. 86*: 7138.

Evans, R., and Parry, A. O. (1989). Wetting transitions in fluids with short-ranged forces: correlation functions and criticality, *J. Phys. Condens. Matter 1*: 7202.

Evans, R., and Parry, A. O. (1990). Liquids at interfaces: what can a theorist contribute?, *J. Phys. Condens. Matter 2*:SA15.

Evans, R., Tarazona, P., and Marini Bettolo Marconi, U. (1983). On the failure of certain integral equation theories to account for complete wetting at solid fluid interfaces, *Mol. Phys. 50*: 993.

Fisher, M. E., and Nakanishi, H. (1981). Scaling theory for the criticality of fluids between plates, *J. Chem. Phys. 75*: 5857.

Fisk, S., and Widom, B. (1969). Structure and free energy of the interface between fluid phases in equilibrium near the critical point, *J. Chem. Phys. 50*: 3219.

Forgacs, G., Lipowsky, R., and Nieuwenhuizen, Th.M. (1991). The behavior of interfaces in ordered and disordered systems, in *Phase Transitions and Critical Phenomena* (C. Domb and J. Lebowitz, eds.), in press.

Gibbs, J. W. (1906). On the equilibrium of heterogeneous substances, in *The Scientific Papers of J. Willard Gibbs*, Longmans, Green, London, Vol. I, p. 55.

Griffiths, R. B., and Wheeler, J. C. (1970). Critical points in multicomponent systems, *Phys. Rev. A 2*: 1047.

Hauge, E. H., and Olaussen, K. (1985). Renormalization of a nonuniversal wetting exponent, *Phys. Rev. B 32*: 4766.

Henderson, D., and Plischke, M. (1987). Sum rules for the pair-correlation functions of inhomogeneous fluids: results for the hard-sphere hard-wall system, *Proc. R. Soc. London A 410*: 409.

Henderson, J. R. (1983). Statistical mechanics of fluids at spherical structureless walls, *Mol. Phys. 50*: 741.

Henderson, J. R. (1986a). Sum rule analysis of continuous wetting phenomena, *Mol. Phys. 59*: 1049.

Henderson, J. R. (1986b). Compressibility route to solvation structure, *Mol. Phys. 59*: 89.

Henderson, J. R. (1986c). Statistical mechanics of spherical surfaces, in *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, Chichester, West Sussex, England, p. 555.

Henderson, J. R. (1987a). Critical wetting in the presence of power-law surface fields, *Phys. Rev. B 35*: 7303.

Henderson, J. R. (1987b). Three-dimensional critical wetting and the statistical mechanics of fluids with short-range forces, *Mol. Phys. 62*: 829.

Henderson, J. R. (1987c). Three-particle direct correlation function of inhomogeneous fluids, *Phys. Rev. A 36*: 4527.

Henderson, J. R., and Rowlinson, J. S. (1984). Statistical mechanics of fluid interfaces in cylindrical symmetry, *J. Phys. Chem. 88*: 6484.

Henderson, J. R., and van Swol, F. (1984). On the interface between a fluid and a planar wall: theory and simulations of hard sphere fluid at a hard wall, *Mol. Phys. 51*: 991.

Henderson, J. R., and van Swol, F. (1985). On the approach to complete wetting by gas at a liquid–wall interface: exact sum rules, fluctuation theory and the verification by computer simulation of the presence of long-range pair correlations at the wall, *Mol. Phys. 56*: 1313.

Hill, T. L. (1959). Exact definition of quasi-thermodynamic point functions in statistical mechanics, *J. Chem. Phys. 30*: 1521.

Irving, J. H., and Kirkwood, J. G. (1950). The statistical mechanical theory of transport processes. IV. The equations of hydrodynamics, *J. Chem. Phys. 18*: 817.

Kalos, M. H., Percus, J. K., and Rao, M. (1977). Structure of a liquid–vapor interface, *J. Stat. Phys. 17*: 111.

Katyl, R. H., and Ingard, U. (1967). Line broadening of light scattered from a liquid surface, *Phys. Rev. Lett. 19*: 64.

Katyl, R. H., and Ingard, U. (1968). Scattering of light by thermal ripplons, *Phys. Rev. Lett. 20*: 248.

Kirkwood, J. G., and Buff, F. P. (1949). The statistical mechanical theory of surface tension, *J. Chem. Phys. 17*: 338.

Kjellander, R., and Sarman, S. (1988). On the statistical mechanics of inhomogeneous fluids in narrow slits: an application to a hard-sphere fluid between hard walls, *Chem. Phys. Lett. 149*: 102.

Kjellander, R., and Sarman, S. (1990). A study of anisotropic pair distribution theories for Lennard–Jones fluids in narrow slits. I. Density profiles and surface interactions, *Mol. Phys. 70*: 215.

Lebowitz, J. L. (1964). Exact solution of generalized Percus–Yevick equation for a mixture of hard spheres, *Phys. Rev. A133*: 895.

Lekner, J., and Henderson, J. R. (1980). Liquid–vapor coexistence and correlations in the interface, *Mol. Phys. 39*: 1437.

Lipowsky, R. (1984). Upper critical dimension for wetting in systems with long-range forces, *Phys. Rev. Lett. 52*: 1429.

Lipowsky, R. (1986). Melting at grain boundaries and surfaces, *Phys. Rev. Lett. 57*: 2876.

Lipowsky, R. (1987). Critical behavior of interfaces: wetting, surface melting and related phenomena, Habilitationsschrift, Ludwig-Maximilians-Universitat, Munchen. Published as Juel-Spez-438 report by KFA Jülich (1988).

Lipowsky, R., and Fisher, M. E. (1986a). Wetting in random systems, *Phys. Rev. Lett. 56*: 472.

Lipowsky, R., and Fisher, M. E. (1986b). Unusual bifurcation of renormalization-group fixed points for interfacial transitions, *Phys. Rev. Lett. 57*: 2411.

Lovett, R. (1988). The constraint on the integral kernels of density functional theories which results from insisting that there be a unique solution for the density function, *J. Chem. Phys. 88*: 7739.

Lovett, R., Mou, C. Y., and Buff, F. P. (1976). The structure of the liquid–vapor interface, *J. Chem. Phys. 65*: 570.

Marini Bettolo Marconi, U., and van Swol, F. (1989). A model of hysteresis in narrow pores, *Eur. Phys. Lett. 8*: 531.

Mermin, N. D. (1965). Thermal properties of the inhomogeneous electron gas, *Phys. Rev. A 137*: 1441.

Mikheev, L.V., and Weeks, J.D. (1991). Sum rules for interface hamiltonians, *Physica A 177*:495.

Nakanishi, H., and Fisher, M. E. (1982). Multicriticality of wetting, prewetting, and surface transitions, *Phys. Rev. Lett. 49*: 1565.

Pandit, R., Schick, M., and Wortis, M. (1982). Systematics of multilayer adsorption phenomena on attractive substrates, *Phys. Rev. B 26*: 5112.

Parry, A. O. (1992). Anomalous decay of pair correlations for two dimensional critical wetting, preprint.

Parry, A. O., and Evans, R. (1988). Correlation functions in the approach to complete drying at a wall–liquid interface, *Mol. Phys. 65*: 455.

Parry, A. O., and Evans, R. (1989). Comment on simple scaling theory for three-dimensional critical wetting with short-range forces, *Phys. Rev. B 39*: 12336.

Parry, A. O., and Evans, R. (1990a). Critical amplitudes for critical wetting with short-ranged forces: the approach to $d = 3^-$, *J. Phys. Condens. Matter 2*: 7687.

Parry, A. O., and Evans, R. (1990b). Influence of wetting on phase equilibria: a novel mechanism for critical-point shifts in films, *Phys. Rev. Lett. 64*: 439.

Percus, J. K. (1980). Shielded distribution approximation for a wall-bounded classical fluid, *J. Stat. Phys. 23*: 657.

Percus, J. K. (1981). The intrinsic liquid–vapor interface, *Faraday Symp. 16*: 23.
Percus, J. K. (1982). Nonuniform fluids, in *The Liquid State of Matter* (E. W. Montroll and J. L. Lebowitz, eds.), North-Holland, Amsterdam, p. 31.
Phillips, P., and Mohanty, U. (1985). The correlation length divergence of Tolman's length, *J. Chem. Phys. 83*: 6392.
Rao, M., Berne, B. J., Percus, J. K., and Kalos, M. H. (1979). Structure of a liquid–vapor interface in the presence of a hard wall in the transition region, *J. Chem. Phys. 71*: 3802.
Reiss, H., Frisch, H. L., Helfand, E., and Lebowitz, J. L. (1960). Aspects of the statistical thermodynamics of real fluids, *J. Chem. Phys. 32*: 119.
Requardt, M., and Wagner, H. J. (1989). Are liquid–vapor interfaces really "rough" in three dimensions? Going beyond the capillary wave model, preprint.
Robert, M. (1985). Direct correlation function in fluid interfaces, *Phys. Rev. A 32*: 500.
Robledo, A., and Rowlinson, J. S. (1986). The distribution of hard rods on a line of finite length, *Mol. Phys. 58*: 711.
Rowlinson, J. S. (1984). The critical exponent of Tolman's length, *J. Phys. A 17*: L357.
Rowlinson, J. S., and Widom, B. (1982). *Molecular Theory of Capillarity*, Clarendon Press, Oxford.
Saam, W. F., and Cole, M. W. (1975). Excitations and thermodynamics for liquid-helium films, *Phys. Rev. B 11*: 1086.
Schofield, P. (1979). The statistical theory of surface tension, *Chem. Phys. Lett. 62*: 413.
Schofield, P., and Henderson, J. R. (1982). Statistical mechanics of inhomogeneous fluids, *Proc. R. Soc. London A 379*: 231.
Stecki, J., and Toxvaerd, S. (1990). Hard sphere cavity in a Lennard–Jones liquid, *J. Chem. Phys. 93*:7342.
Sullivan, D. E. (1981). Wetting transitions at fluid–solid interfaces, *Faraday Symp. 16*: 191.
Sullivan, D. E., and Telo da Gama, M. M. (1986). Wetting transitions and multilayer adsorption at fluid interfaces, in *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, Chichester, West Sussex, England, p. 45.
Tarazona, P. (1985). Free-energy density functional for hard spheres, *Phys. Rev. A 31*: 2672.
Tarazona, P., and Evans, R. (1983). On the validity of certain integrodifferential equations for the density-orientation profile of molecular fluid interfaces, *Chem. Phys. Lett. 97*: 279.
Triezenberg, D. G., and Zwanzig, R. (1972). Fluctuation theory of surface tension, *Phys. Rev. Lett. 28*: 1183.
Vanderlick, T. K., Scriven, L. E., and Davis, H. T. (1986). Statistical mechanics of rigid particles in an external field, *J. Chem. Phys. 85*: 6699.
van Swol, F., and Henderson, J. R. (1984). Complete wetting in systems with short-range forces, *Phys. Rev. Lett. 53*: 1376.
van Swol, F., and Henderson, J. R. (1986). Wetting at a fluid–wall interface.

Computer simulation and exact statistical sum rules, *J. Chem. Soc. Faraday Trans. 2 82*: 1685.

van Swol, F., and Henderson, J. R. (1989). Wetting and drying transitions at a fluid–wall interface: density-functional theory versus computer simulation, *Phys. Rev A 40*: 2567.

Walton, J. P. R. B., and Gubbins, K. E. (1985). The pressure tensor in an inhomogeneous fluid of non-spherical molecules, *Mol. Phys. 55*: 679. (1986). Erratum, *Mol. Phys. 58*: 1013.

Weeks, J. D. (1977). Structure and thermodynamics of the liquid–vapor interface, *J. Chem. Phys. 67*: 3106.

Weeks, J. D. (1984). Scaling relations between correlations in the liquid–vapor interface and the interface width, *Phys. Rev. Lett. 52*: 2160.

Weeks, J. D., van Saarloos, W., Bedeaux, D., and Blokhuis, E. (1989). Consistency of capillary wave theory in three dimensions: divergence of the interface width and agreement with density functional theory, *J. Chem. Phys. 91*: 6494.

Wertheim, M. S. (1976). Correlations in the liquid–vapor interface, *J. Chem. Phys. 65*: 2377.

Wertheim, M. S., Blum, L., and Bratko, D. (1989). Statistical mechanics of confined systems, in *Statistical Thermodynamics of Micellar and Microemulsion Systems* (S. H. Chen, ed.), Springer-Verlag, Berlin.

Widom, B. (1963). Some topics in the theory of fluids, *J. Chem. Phys. 39*: 2808.

Yvon, J. (1935). La théorie statistique des fluides et l'equation d'etat, in *Actualitiés scientifiques et industrielles*, No. 203, Hermann, Paris, p. 49.

3

Density Functionals in the Theory of Nonuniform Fluids

R. Evans
University of Bristol
Bristol, England

I. INTRODUCTION

A. Why Density Functionals?

The last 15 years or so have seen an enormous growth in the use of density functional methods applied to inhomogeneous classical fluids. Such systems are characterized by spatial variation of the average one-body density $\rho(\mathbf{r})$. Density functional methods are based on the idea that the free energy of the inhomogeneous fluid can be expressed as a functional of $\rho(\mathbf{r})$. From a knowledge of this functional all the relevant thermodynamic functions can be calculated so that tensions can be computed for interface problems, solvation forces can be determined for confined fluids, and phase transitions can be investigated for various types of inhomogeneity. Moreover, derivatives of the functional determine the equilibrium distribution (or correlation) functions that describe the microscopic structure of the inhomogeneous fluid. Determining the *exact* free-energy functional is, of course, equivalent to solving exactly the statistical mechanics for the particular fluid under investigation (i.e., calculating the partition function). So what is to be gained by adopting the density functional approach? First, we can employ the methods of functional differentiation and integration to obtain important formally exact results for correlation functions and thermodynamic functions (see Chapter 2), which are less easily or

less elegantly derived by methods that focus on the partition function directly. Second, we can seek explicit *approximations* for the free-energy functional that will allow calculations to be made for a wide variety of inhomogeneous systems. One of the key results of density functional theory is that the intrinsic Helmholtz free energy is a unique functional $\mathcal{F}[\rho]$ for a given interatomic or intermolecular potential energy Φ. That part of the free energy, which is not associated directly with the external potential $V(\mathbf{r})$ producing the inhomogeneity, has the same dependence on $\rho(\mathbf{r})$ for all $V(\mathbf{r})$. Thus the same $\mathcal{F}[\rho]$ should be valid for the liquid–gas interface stabilized by a gravitational field, for the same fluid adsorbed by a substrate or confined in a capillary or for the bulk crystalline solid whose periodic density can be regarded as that of the highly inhomogeneous, symmetry-broken fluid. A major goal of the theory is to find suitable approximations for $\mathcal{F}[\rho]$ (for a given type of fluid) that are computationally tractable and that are sufficiently accurate for application in a wide range of problems [i.e., remain accurate for various choices of $V(\mathbf{r})$]. For simple (atomic) fluids described by a Lennard–Jones (or cruder) pairwise potential there exists a plethora of approximations. Some of the more sophisticated versions have been applied successfully to a wide variety of interfacial, confinement, and freezing problems. Others have been designed specifically for one particular type of inhomogeneity (this has often been the case in theories of freezing), so their usefulness in other problems is not determined. Relationships between the different approximations, when they exist, are often obscure. It is the purpose of this chapter to review briefly the formal aspects of density functional theory, to discuss some of the approximations that have been used, and to describe some of the applications to interfacial and adsorption problems. In other chapters of the book applications of density functional theory to freezing, nucleation, and liquid–crystalline ordering are described.

B. Origins

Rowlinson will have mentioned some of the history of the use of functionals in Chapter 1. Bogoliubov [1] is often credited with introducing functional techniques into statistical physics. Morita and Hiroike [2], de Dominicis [3], Stillinger and Buff [4], and Lebowitz and Percus [5] developed much of the formalism of inhomogeneous fluids in the early 1960s. The influential articles of Percus [6] and Stell [7] described functional methods and their applications to the theory of the structure of bulk fluids. At about the same time, Hohenberg and Kohn [8] and Kohn and Sham [9] developed a powerful density functional treatment for the ground state of the inhomogeneous interacting electron liquid. Their work provided an

important synthesis of the approach used by Thomas, Fermi, Dirac, and others and the self-consistent field approach of Hartree, Slater, and so on. Hohenberg and Kohn used a fundamental variational principle for the ground-state energy to show that the intrinsic part of this energy is a unique functional of the electron density $n(\mathbf{r})$. By recasting the many-electron problem in terms of a variational principle based on an energy functional of the electron density they showed that Thomas–Fermi theory consists simply of a local density approximation for the kinetic energy functional plus complete neglect of all exchange and correlation contributions—only the electrostatic energy is treated properly. Hartree theory, on the other hand, constitutes an exact (Schrödinger equation) treatment of the kinetic energy but also neglects all exchange and correlation contributions. More significantly, the density functional approach suggested possible ways, different from the Hartree–Fock, or configuration interaction approaches, of including exchange and correlation. Kohn and Sham used their variational principle to rewrite the interacting many-electron Schrödinger equation as a formally exact one-electron equation with an additional effective potential arising from the (unknown) exchange and correlation energy functional $E_{xc}[n]$. The crudest (local density) approximation to the latter already constitutes a significant improvement on Hartree theory for the electronic structure and ground-state properties of atoms, molecules, and solids. Indeed, almost all modern band structure calculations are based on the density functional theory of Kohn and Sham, with some simple approximation for the functional $E_{xc}[n]$. We refer the reader to the recent comprehensive review [10] and book [11] for detailed accounts of density functional theory applied to electronic systems. Mermin [12] extended Hohenberg and Kohn's treatment to nonzero temperature, writing down a variational principle for the grand potential functional of the interacting electron liquid. The realization that such methods were readily applicable to classical fluids, where they could also generate useful approximation schemes, came in 1976, or thereabouts, in the work of Ebner and Saam [13,14] and Yang et al. [15]. By then, several groups (e.g., F. F. Abraham, H. T. Davis, and S. Toxvaerd) had been using approximate free-energy functionals to calculate the density profile and surface tension of the liquid–gas interface. (See Refs. 16 and 17 for reviews.) In the late 1970s it was recognized that the square-gradient approximation of van der Waals and some of its generalizations could be derived in a systematic fashion using density functional theory. This work [13–17] was influenced strongly by that of Hohenberg and Kohn on the gradient expansion for the electron liquid. A certain amount of cross-fertilization between classical and quantum liquids has occurred subse-

quently, but sometimes approximations have been (re)discovered by authors unaware of the corresponding results in the other field.

The late 1970s also saw important applications of density functional techniques in deriving [18,19] exact integrodifferential equations for the density profile of a fluid in an external potential (see Chapter 2) and in analyzing the nature of correlations at the liquid–gas interface [20,21]. The development of approximation schemes and their application was a major activity during the 1980s. Weighted-density approximations to $\mathcal{F}[\rho]$ have become especially popular and have met with widespread success. These have their origins in a variety of sources that we describe.

C. Scope

The subject of this chapter has a large and rapidly growing literature. It would be inappropriate to attempt a comprehensive review here. We specialize in one-component, argonlike fluids, mentioning work on mixtures of simple fluids where this is considered relevant. Density functional methods have been applied to more complex fluids, and some of these applications are described elsewhere in this book. *Representative* papers on extensions of density functional theory to the surface properties of ionic fluids can be found in [22,23], to electrical double layers in [24], to inhomogeneous dipolar and quadrupolar fluids in [25,26], and to liquid metal surfaces in [27].

In Section II we give a self-contained account of the density functional formalism and a discussion of the exact free-energy functional for the case of hard rods in one dimension. Approximations for $\mathcal{F}[\rho]$ are described in Section III. We have attempted to include most of the schemes that have been employed, pointing out their strengths and weaknesses and listing the problems to which they have been applied. In Section IV we describe some recent applications of density functional theory to the determination of the structure of *uniform* fluids [i.e., the radial distribution function $g(r)$ and the three-body direct correlation function $c^{(3)}$ of bulk liquids]. By examining the consequences of approximate theories for bulk structure it is possible to learn something about their intrinsic limitations. In Section V we describe some applications of approximate density functional theories to fluid interfacial phenomena. Since this is an enormous field we cannot do it justice and we restrict ourselves to a few topics, emphasizing the successes and failings of approximation schemes in accounting for phase transitions and criticality at various types of interface. Section VI contains some concluding remarks.

Parts of the present chapter draw heavily on a set of lecture notes by the author entitled "Microscopic Theories of Simple Fluids and Their

Interfaces" [28]. Several sections of those notes have been amplified and brought up to date here.

II. FORMALISM

In this section we summarize the formal structure of density functional theory, specializing to one-component atomic fluids. We work in the grand canonical ensemble, which is the most convenient for most problems involving inhomogeneous fluids. The presentation is similar to that given in [28]. Details of proofs are omitted; these can be found in an earlier review [16] or in Ref. 12. More rigorous treatments of the foundations of classical density functional theory can be found in Ref. 29.

A. Generating Functionals and Hierarchies of Correlation Functions

The Hamiltonian for the fluid of N atoms, each of mass m, is

$$\begin{aligned} H_N &= \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m} + \Phi(\mathbf{r}_1, \ldots, \mathbf{r}_N) + \sum_{i=1}^{N} V(\mathbf{r}_i) \\ &\equiv \text{K.E.} \quad + \Phi \qquad\qquad + V \end{aligned} \tag{1}$$

where $\mathbf{p}_i$ is the momentum of atom i and Φ is the total interatomic potential energy (not necessarily pairwise additive). The one-body external potential $V(\mathbf{r})$ is, as yet, arbitrary. The grand potential Ω is a function of chemical potential μ, inverse temperature $\beta = (k_B T)^{-1}$, and the available volume; it is also a functional of $V(\mathbf{r})$ and therefore of the combination

$$u(\mathbf{r}) \equiv \mu - V(\mathbf{r}) \tag{2}$$

A hierarchy of correlation functions is obtained by functional differentiation of Ω with respect to $u(\mathbf{r})$. The first derivative is simply the average one-body density

$$\rho(\mathbf{r}) \equiv \rho^{(1)}(\mathbf{r}) \equiv \langle \hat{\rho}(\mathbf{r}) \rangle = -\frac{\delta \Omega}{\delta u(r)} \tag{3}$$

where $\hat{\rho}(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i)$ and $\langle\ \rangle$ denotes the ensemble average. For a fluid $\rho(\mathbf{r})$ must have the symmetry of the external potential $V(\mathbf{r})$. A second derivative yields the density–density correlation function

$$\begin{aligned} G(\mathbf{r}_1, \mathbf{r}_2) &\equiv \langle (\hat{\rho}(\mathbf{r}_1) - \langle \hat{\rho}(\mathbf{r}_1) \rangle)(\hat{\rho}(\mathbf{r}_2) - \langle \hat{\rho}(\mathbf{r}_2) \rangle) \rangle \\ &= \beta^{-1} \frac{\delta \rho(\mathbf{r}_1)}{\delta u(\mathbf{r}_2)} = -\beta^{-1} \frac{\delta^2 \Omega}{\delta u(\mathbf{r}_2)\, \delta u(\mathbf{r}_1)} \end{aligned} \tag{4}$$

which is related to the two-body distribution function $\rho^{(2)}$ via

$$G(\mathbf{r}_1,\mathbf{r}_2) = \rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) + \rho(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{5}$$

Further differentiation yields three-body, four-body, and so on, correlation functions. This procedure is the standard one in equilibrium statistical mechanics (see Chapter 2 and Ref. 30). Note that for a bulk fluid of uniform density ρ, $\rho(\mathbf{r}) \to \rho$ and translational invariance demands that $G(\mathbf{r}_1,\mathbf{r}_2) \to G(r_{12}) = \rho^2(g(r_{12}) - 1) + \rho\delta(r_{12})$, where $r_{12} \equiv |\mathbf{r}_1 - \mathbf{r}_2|$ and $g(r)$ is the usual radial distribution function. It follows that the Fourier transform $G(k) = \rho S(k)$, where $S(k)$ is the static structure factor of the bulk liquid [30].

The density functional approach focuses on functionals of $\rho(\mathbf{r})$ rather than $u(\mathbf{r})$. While it is clear that $\rho(\mathbf{r})$ is a functional of $u(\mathbf{r})$, one can prove [12,16,29] the less obvious result that for given Φ, μ, and T only one external potential $V(\mathbf{r})$ can determine a specified equilibrium one-body density $\rho(\mathbf{r})$. From this result it follows that the probability density

$$f_N = \Xi^{-1} \exp[-\beta(H_N - \mu N)] \tag{6}$$

where Ξ is the partition function, is uniquely determined by $\rho(\mathbf{r})$—the latter fixes $V(\mathbf{r})$, which then determines f_N. Since f_N is a unique functional of $\rho(\mathbf{r})$, so is the quantity

$$\begin{aligned} \mathcal{F}[\rho] &\equiv \langle \text{K.E.} + \Phi + \beta^{-1} \ln f_N \rangle \\ &= \text{Tr}_{\text{cl}}[f_N(\text{K.E.} + \Phi + \beta^{-1} \ln f_N)] \end{aligned} \tag{7}$$

where Tr_{cl} denotes the classical trace. The same form of $\mathcal{F}[\rho]$ will be valid for any external potential. A second functional is constructed from a Legendre transform of $\mathcal{F}$:

$$\Omega_V[\tilde{\rho}] = \mathcal{F}[\tilde{\rho}] - \int d\mathbf{r}\, u(\mathbf{r})\tilde{\rho}(\mathbf{r}) \tag{8}$$

When $\tilde{\rho} = \rho$, the equilibrium density, $\Omega_V[\tilde{\rho}]$, reduces to the grand potential Ω. It can also be shown [16] that Ω is the minimum value of $\Omega_V[\tilde{\rho}]$, so that we have a variational principle

$$\left.\frac{\delta\Omega_V[\tilde{\rho}]}{\delta\tilde{\rho}(\mathbf{r})}\right|_{\rho} = 0 \qquad \Omega_V[\rho] = \Omega \tag{9}$$

for determining the equilibrium density of a fluid in an external potential. $\mathcal{F}[\rho]$ is the intrinsic Helmholtz free-energy functional, since the total Helmholtz free energy $F = \Omega + \mu \int d\mathbf{r}\, \rho(\mathbf{r}) = \mathcal{F}[\rho] + \int d\mathbf{r}\, \rho(\mathbf{r})V(\mathbf{r})$. Combining (8) and (9) we have

$$\mu = V(\mathbf{r}) + \frac{\delta\mathcal{F}[\rho]}{\delta\rho(\mathbf{r})} \tag{10}$$

which expresses the constancy of the chemical potential through the inhomogeneous fluid. $\delta\mathcal{F}[\rho]/\delta\rho(\mathbf{r})$ can be regarded as the intrinsic chemical potential; in general this will not be a local function of $\rho(\mathbf{r})$.

By virtue of (9), use of $\Omega_V[\rho]$ as a generating functional is identical to the standard procedure. A second hierarchy of correlation functions is generated by differentiating $\mathcal{F}[\rho]$. The latter contains an ideal gas (noninteracting) contribution

$$\mathcal{F}_{id}[\rho] = \int d\mathbf{r}\, f_{id}(\rho(\mathbf{r})) \tag{11}$$

with $f_{id}(\rho) = \beta^{-1}\rho(\ln \Lambda^3\rho - 1)$, the free-energy density of a uniform ideal gas. Λ is the thermal de Broglie wavelength. Subtracting the ideal contribution we generate the *direct correlation* function hierarchy:

$$c^{(1)}(\mathbf{r}) = -\frac{\delta(\beta\mathcal{F}_{ex}[\rho])}{\delta\rho(\mathbf{r})} \equiv -\frac{\delta(\beta\mathcal{F}[\rho] - \beta\mathcal{F}_{id}[\rho])}{\delta\rho(\mathbf{r})} \tag{12}$$

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\delta c^{(1)}(\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} = -\frac{\delta^2(\beta\mathcal{F}_{ex}[\rho])}{\delta\rho(\mathbf{r}_2)\,\delta\rho(\mathbf{r}_1)} = c^{(2)}(\mathbf{r}_2,\mathbf{r}_1) \tag{13}$$

that is,

$$c^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{\delta c^{(n-1)}(\mathbf{r}_1, \ldots, \mathbf{r}_{n-1})}{\delta\rho(\mathbf{r}_n)}$$

where the tilde in the density variable is omitted. Using (11) and (12), (10) can be reexpressed as

$$\Lambda^3\rho(\mathbf{r}) = \exp[\beta u(\mathbf{r}) + c^{(1)}(\mathbf{r})] \tag{14}$$

where $c^{(1)}$ is itself a functional of $\rho(\mathbf{r})$. For an ideal gas $c^{(1)} \equiv 0$ and (14) reduces to the familiar barometric law for the density distribution in the presence of an external field. Thus (14) implies that $-\beta^{-1}c^{(1)}(\mathbf{r})$ acts as an additional effective one-body potential in determining self-consistently the equilibrium density. This quantity is the classical analog of the effective one-body potential $\int d\mathbf{r}'\, n(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'| + \delta E_{xc}[n]/\delta n(\mathbf{r})$ entering the theory of Kohn and Sham [9] for the electron density $n(\mathbf{r})$. That potential enters a one-electron Schrödinger equation appropriate to a noninteracting electron liquid. The presence of the exponential in (14) reflects the corresponding classical behavior.

In a uniform fluid with $V(\mathbf{r}) \equiv 0$, (14) reduces to

$$\mu(\rho) = \mu_{id}(\rho) - \beta^{-1}c^{(1)}(\rho) \tag{15}$$

with $\mu_{id}(\rho) = df_{id}/d\rho$, so that $c^{(1)}(\rho)$ is proportional to the excess (over ideal) chemical potential. Equation (14) is also equivalent to Widom's potential distribution formula (see Chapter 2). From (12), (13), and (10)

we find that

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\delta(\mathbf{r}_1 - \mathbf{r}_2)}{\rho(\mathbf{r}_1)} - \beta \frac{\delta u(\mathbf{r}_1)}{\delta \rho(\mathbf{r}_2)} \tag{16}$$

where the second term is [via (4)] $- G^{-1}(\mathbf{r}_1,\mathbf{r}_2)$, the functional inverse being defined as

$$\int d\mathbf{r}_3 G^{-1}(\mathbf{r}_1,\mathbf{r}_3)G(\mathbf{r}_3,\mathbf{r}_2) = \delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{17}$$

Thus $c^{(2)}$ is (essentially) the inverse of the density–density correlation function G. With (5), (16) and (17) together imply the integral equation

$$h(\mathbf{r}_1,\mathbf{r}_2) = c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) + \int d\mathbf{r}_3 h(\mathbf{r}_1,\mathbf{r}_3)\rho(\mathbf{r}_3)c^{(2)}(\mathbf{r}_3,\mathbf{r}_2) \tag{18}$$

relating the two-body direct correlation function $c^{(2)}$ to the *total* correlation function h defined by

$$\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)h(\mathbf{r}_1,\mathbf{r}_2) \equiv \rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) - \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) \tag{19}$$

Equation (18) is the familiar Ornstein–Zernike equation for an inhomogeneous fluid. Often this is used to define $c^{(2)}$, however, we see that this equation follows as a natural consequence of having two generating functionals Ω_V and $\mathcal{F}$ linked by the Legendre transform (8); that is, (17) is equivalent to

$$\int d\mathbf{r}_3 \frac{\delta^2 \mathcal{F}}{\delta\rho(\mathbf{r}_1)\,\delta\rho(\mathbf{r}_3)} \frac{\delta^2 \Omega_V}{\delta u(\mathbf{r}_3)\,\delta u(\mathbf{r}_2)} = -\delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{20}$$

Thus the direct correlation function hierarchy has equal status with the standard distribution function hierarchy. Indeed, the existence of two hierarchies, generated by two generating functionals, is a common procedure in many-body theory. In field theoretical treatments [31] of statistical mechanics, the analog of $\mathcal{F}[\rho]$ is $\Gamma[\bar{\Phi}]$, the generating functional for the vertex functions $\Gamma^{(N)}$, with $\bar{\Phi}$, the *averaged* order parameter, being the analog of the average density $\rho(\mathbf{r})$.

Henderson has described, in Chapter 2, how the formalism above can be used to derive exact sum rules for inhomogeneous fluids. Here we apply the formalism to the determination of thermodynamic functions.

B. Thermodynamic Functions via Functional Integration

There are several routes to the calculation of the free energy of an inhomogeneous fluid. We describe two of these. Consider an initial fluid state with density $\rho_i(\mathbf{r})$ and a final state with density $\rho(\mathbf{r})$ at the same

temperature T and suppose that these can be linked by a linear path in the space of density functions characterized by a single coupling parameter α:

$$\begin{aligned}\rho_\alpha \equiv \rho(\mathbf{r};\alpha) &= \rho_i(\mathbf{r}) + \alpha(\rho(\mathbf{r}) - \rho_i(\mathbf{r})) \\ &\equiv \rho_i(\mathbf{r}) + \alpha\Delta\rho(\mathbf{r})\end{aligned} \tag{21}$$

with $0 \leq \alpha \leq 1$. Integration of (12) yields

$$\beta\mathcal{F}_{ex}[\rho] = \beta\mathcal{F}_{ex}[\rho_i] - \int_0^1 d\alpha \int d\mathbf{r}\, \Delta\rho(\mathbf{r}) c^{(1)}([\rho_\alpha];\mathbf{r}) \tag{22}$$

where the functional dependence of $c^{(1)}$ is made explicit. A second integration [see (13)] gives

$$c^{(1)}([\rho_\alpha];\mathbf{r}_1) = c^{(1)}([\rho_i];\mathbf{r}_1) + \int_0^\alpha d\alpha' \int d\mathbf{r}_2\, \Delta\rho(\mathbf{r}_2) c^{(2)}([\rho_{\alpha'}];\mathbf{r}_1,\mathbf{r}_2) \tag{23}$$

For a uniform fluid (23) simplifies, with $\rho_i \equiv 0$, to

$$c^{(1)}(\rho) = \int_0^\rho d\rho' \int d\mathbf{r}_2\, c^{(2)}(\rho';\mathbf{r}_1,\mathbf{r}_2)$$

so that

$$\frac{\partial c^{(1)}(\rho)}{\partial\rho} = \int d\mathbf{r}\, c^{(2)}(\rho;r) \tag{24a}$$

which, by virtue of (15), is equivalent to

$$\beta\rho\left(\frac{\partial\mu}{\partial\rho}\right)_T = 1 - \rho\int d\mathbf{r}\, c^{(2)}(\rho;r) \tag{24b}$$

the compressibility sum rule. $c^{(2)}(\rho;\mathbf{r}_1,\mathbf{r}_2) \equiv c^{(2)}(\rho;|\,\mathbf{r}_1 - \mathbf{r}_2\,|)$ is the (two-body) direct correlation function of the bulk fluid. Equations (22) and (23) can be combined to give

$$\begin{aligned}\beta\mathcal{F}_{ex}[\rho] = \beta\mathcal{F}_{ex}[\rho_i] &- \int d\mathbf{r}\, \Delta\rho(\mathbf{r}) c^{(1)}([\rho_i];\mathbf{r}) \\ &- \int_0^1 d\alpha \int d\mathbf{r}_1\, \Delta\rho(\mathbf{r}_1) \int_0^\alpha d\alpha' \int d\mathbf{r}_2\, \Delta\rho(\mathbf{r}_2) c^{(2)}([\rho_{\alpha'}];\mathbf{r}_1,\mathbf{r}_2)\end{aligned} \tag{25}$$

Using the identity

$$\int_0^1 d\alpha \int_0^\alpha d\alpha'\, q(\alpha') = \int_0^1 d\alpha(1-\alpha)q(\alpha)$$

valid for any function $q(\alpha)$, (25) simplifies to

$$\beta\mathcal{F}_{ex}[\rho] = \beta\mathcal{F}_{ex}[\rho_i] - \int d\mathbf{r}\, \Delta\rho(\mathbf{r})c^{(1)}([\rho_i];\mathbf{r}) + \int_0^1 d\alpha(\alpha - 1) \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \Delta\rho(\mathbf{r}_1)\, \Delta\rho(\mathbf{r}_2)c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2) \quad (26)$$

As emphasized by Saam and Ebner [14] this result should be independent of the choice (21) of integration path; recall that $\mathcal{F}_{ex}[\rho]$ is a unique functional of $\rho(\mathbf{r})$. A more familiar version of (26) emerges when $\rho_i \equiv 0$ and the final state is a uniform fluid of density ρ. The total Helmholtz free energy density is

$$f(\rho) = f_{id}(\rho) + \beta^{-1}\rho^2 \int_0^1 d\alpha(\alpha - 1) \int d\mathbf{r}\, c^{(2)}(\alpha\rho;r) \quad (27)$$

The thermodynamic functions $\mathcal{F}$ and Ω can be obtained from (26) by adding the external potential contribution and the resulting formulas are equivalent [16] to those obtained by Stillinger and Buff [4] and Lebowitz and Percus [5] using different methods. We shall see that (26) forms the basis for several approximate theories of inhomogeneous fluids. It also provides the starting point for the modern theory of freezing, where the reference state refers to that of a bulk liquid and the final state refers to the bulk crystal, viewed as an inhomogeneous fluid (see Chapter 9). Note that while (26) is formally exact its evaluation requires $c^{(2)}$ as a functional of $\rho(\mathbf{r})$. This is a tall order for any theory and we shall find that gross approximations must be made for this quantity. Note also that use of the corresponding bulk expression (27) is restricted to integration paths (at fixed T) in a single-phase region, where $c^{(2)}(\rho;r)$ is single valued.

A second formula for the free energy can be obtained for the particular case when the potential function Φ is pairwise additive, that is,

$$\begin{aligned}\Phi(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N) &= \frac{1}{2}\sum_{i\neq j}\sum_{j=1}^{N}\phi(\mathbf{r}_i,\mathbf{r}_j) \\ &= \frac{1}{2}\iint d\mathbf{r}\, d\mathbf{r}'\, \phi(\mathbf{r},\mathbf{r}')\hat{\rho}(\mathbf{r})\,(\hat{\rho}(\mathbf{r}') - \delta(\mathbf{r} - \mathbf{r}'))\end{aligned} \quad (28)$$

The grand potential is a functional of $\phi(\mathbf{r},\mathbf{r}') \equiv \phi(|\mathbf{r} - \mathbf{r}'|)$ and has the property, for fixed T and $u(\mathbf{r})$,

$$\frac{\delta\Omega}{\delta\phi(\mathbf{r}_1,\mathbf{r}_2)} = \tfrac{1}{2}\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) \quad (29)$$

since, from (5) and (4),

$$\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \langle\hat{\rho}(\mathbf{r}_1)\hat{\rho}(\mathbf{r}_2)\rangle - \langle\hat{\rho}(\mathbf{r}_1)\rangle\delta(\mathbf{r}_1 - \mathbf{r}_2) \quad (30)$$

Equation (29) implies that

$$\frac{\delta \mathcal{F}[\rho]}{\delta \phi(\mathbf{r}_1,\mathbf{r}_2)} = \frac{1}{2}\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) \tag{31}$$

We now consider a reference fluid at the same temperature and density $\rho(\mathbf{r})$, for which the pairwise potential is ϕ_r and integrate (31) using a "charging" parameter α:

$$\phi_\alpha(\mathbf{r}_1,\mathbf{r}_2) = \phi_r(\mathbf{r}_1,\mathbf{r}_2) + \alpha\phi_p(\mathbf{r}_1,\mathbf{r}_2) \qquad 0 \le \alpha \le 1 \tag{32}$$

The result is

$$\mathcal{F}[\rho] = \mathcal{F}_r[\rho] + \frac{1}{2}\int_0^1 d\alpha \int d\mathbf{r}_1 \int d\mathbf{r}_2\, \rho^{(2)}(\phi_\alpha;\mathbf{r}_1,\mathbf{r}_2)\phi_p(\mathbf{r}_1,\mathbf{r}_2) \tag{33}$$

where $\mathcal{F}_r[\rho]$ is the intrinsic free energy of the reference fluid ($\alpha = 0$), and $\phi_p = \phi - \phi_r$ is the perturbation potential that is turned on via (32). As α increases between 0 and 1 the density $\rho(\mathbf{r})$ must not alter, so it is necessary to envisage [32] an effective external potential, depending on α, that imposes such a density for any intermediate system with pairwise potential ϕ_α. This external potential will only reduce to the actual one $V(\mathbf{r})$ in the final system $\alpha = 1$.

Equation (33) is the basis of all *perturbation theories* [30] of bulk and inhomogeneous fluids. Usually, ϕ_r is taken to be the repulsive part of the full pairwise potential and the remaining, attractive part is treated as the perturbation ϕ_p. By making suitable approximations for $\mathcal{F}_r[\rho]$ and for the two body distribution function $\rho^{(2)}(\phi_\alpha;\mathbf{r}_1,\mathbf{r}_2)$ of the inhomogeneous fluid with potential ϕ_α, approximate free-energy density functionals can be derived (see Section III.D).

C. Hard Rods in One Dimension: An Exactly Solvable Model

There is no continuum model for which the statistical mechanics can be solved exactly (at liquid densities) in three dimensions. Thus there is no model for which the functional $\mathcal{F}[\rho]$ is known exactly in three dimensions. In one dimension, however, where correlations are of a much simpler nature, exact results do exist for particles with nearest-neighbor interactions. Percus [33] derived an integral equation for the density profile $\rho(z)$ of a one-dimensional fluid of hard rods (length σ) in an arbitrary external potential $V(z)$ via functional differentiation of the grand partition function with respect to $u(z) \equiv \mu - V(z)$. Robledo [34] obtained the same equation from potential distribution theory, while Robledo and Varea [35]

and Percus [36] constructed the functional $\Omega_V[\rho]$. This has the usual structure

$$\Omega_V[\rho] = \mathcal{F}_{id}[\rho] + \mathcal{F}_{ex}[\rho] - \int dz\, u(z)\rho(z)$$

with the excess free-energy functional given by

$$\mathcal{F}_{ex}[\rho] \equiv \mathcal{F}_{ex}^{hr}[\rho] = -\beta^{-1} \int dz\, \rho(z) \ln[1 - t(z)] \tag{34}$$

where $t(z) = \int_{z-\sigma}^{z} dy\, \rho(y)$. Requiring $\Omega_V[\rho]$ to be a minimum yields Percus's equation for the profile:

$$\beta u(z) = \ln \frac{\Lambda\rho(z)}{1 - t(z)} + \int_{z}^{z+\sigma} dy \frac{\rho(y)}{1 - t(y)} \tag{35}$$

$c^{(2)}$, $c^{(3)}$, and so on, can be obtained by further functional differentiation and can be shown [33] to be of finite range in all pairs of variables [e.g., $c^{(2)}(z, z') = 0$ unless $| z - z' | \leq \sigma$]. Percus [36] also considered the case of sticky hard rods, deriving $\Omega_V[\rho]$ and showing that $c^{(2)}$ vanishes beyond the range of the core. For the special case of hard rods confined by two hard walls, Robledo and Rowlinson [37] have obtained a complete set of results, including the n-body distribution functions and the solvation force (the excess pressure brought about by confinement). (See also Davis [38] for a treatment of the confined Tonks–Takahashi one-dimensional fluid.) Vanderlick et al. [39] have extended the work of Percus [33,36] to mixtures of hard rods in an arbitrary external field, and these authors provide many results for density profiles, selective adsorption, and solvation force of a binary mixture confined between two walls. Monson [40] has determined the partition function and density profiles for inhomogeneous one-dimensional square-well mixtures.

Can we learn anything from these exact solutions that will guide us toward an effective approximation for real fluids? This question was posed by Percus [36,41] and Robledo and Varea [35], who suggested possible approximation schemes for higher dimensions based on the *form* of (34). Rewriting this equation as

$$\mathcal{F}_{ex}^{hr}[\rho] = \frac{1}{2} \int dz \left[\rho\left(z + \frac{\sigma}{2}\right) + \rho\left(z - \frac{\sigma}{2}\right) \right] \psi_{ex}(\rho_\tau(z)) \tag{36}$$

with

$$\psi_{ex}(n) \equiv -\beta^{-1} \ln(1 - n\sigma) \tag{37}$$

and

$$\rho_\tau(z) \equiv \sigma^{-1} \int_{-\sigma/2}^{\sigma/2} dy\, \rho(z + y) \tag{38}$$

it is clear that the excess free-energy functional is an integral over the excess free energy of a uniform fluid: (37) is the excess free energy per particle of the uniform hard-rod fluid of density n. [This result can be obtained from the thermodynamic identity $(\partial\psi_{ex}/\partial n)_T = n^{-2}(p - n\beta^{-1})$ applied to the hard-rod equation of state $\beta p = n(1 - n\sigma)^{-1}$.] However, (36) is not merely a local density formula—the excess free energy is that of a bulk fluid at a density $\rho_\tau(z)$, which is the *average* over the particle in question. It is tempting to assume that (36) will generalize to hard disks and spheres, provided that suitable average densities can be constructed and the ψ_{ex} appropriate to the higher dimension is utilized. This was the strategy proposed in Refs. 35 and 36. A skeptic might argue that there is no reason to suppose that correlations between hard particles will have the same form in higher dimensions as in one. Nevertheless, such generalizations have proved very instructive and we discuss them further in Section III.E.5, where we make connection with other, closely related approaches that also involve a local average of the density.

III. APPROXIMATIONS FOR FREE ENERGY FUNCTIONALS

A. Basic Strategy and Its Pitfalls

Any practical implementation of density functional theory for a particular physical problem almost always requires some explicit approximation for the functional $\mathcal{F}[\rho]$. Once this is given, the equilibrium $\rho(\mathbf{r})$ and grand potential Ω are determined, via (8) and (9), for specified T, μ, and $V(\mathbf{r})$. (There is the tacit assumption that $\Omega_V[\rho]$ still obeys the minimum principle when $\mathcal{F}[\rho]$ is an approximation to the exact functional.) Direct correlation functions, if required, are obtained by further functional differentiation, distribution functions $G(\mathbf{r}_1,\mathbf{r}_2)$, and so on, follow via the Ornstein–Zernike equation (17) or (18). From an engineering viewpoint this strategy is very appealing. The reliability and accuracy of the results should reflect the skill with which $\mathcal{F}[\rho]$ is constructed for the particular model Hamiltonian. For certain problems one might be able to extract understanding of the essential phenomena using very crude approximations. For others, which demand detailed information about the microscopic structure, say, sophisticated approximations will be required. From a statistical mechanics viewpoint the strategy is less satisfactory. There is always great danger of losing sight of the Hamiltonian.* Once $\mathcal{F}[\rho]$ is specified, all equilibrium properties are determined, so there is a temptation to regard $\mathcal{F}[\rho]$ as defining some model fluid. If $\mathcal{F}[\rho]$ corresponds to some exactly

* Henderson (Chapter 2) also discusses these issues.

solved model, as in the case of hard rods, there is a one-to-one relationship between the functional and the Hamiltonian. If, however, $\mathcal{F}[\rho]$ is merely some (intelligently chosen) functional, there is no reason to expect the resulting properties to be those which would correspond to exact solution of *any* Hamiltonian, let alone the original. There is a model due to Percus [42] for which

$$\beta\mathcal{F}_{ex}[\rho] = -\frac{1}{2}\iint d\mathbf{r}_1\, d\mathbf{r}_2\, c(\mathbf{r}_1 - \mathbf{r}_2)\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)$$

with c the basic model function. Functional differentiation implies that the direct correlation function $c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = c(\mathbf{r}_1 - \mathbf{r}_2)$, independent of any external field, for this model. Percus showed that provided that $c(\mathbf{r}_1 - \mathbf{r}_2) \geq 0$, such a model could be represented by a Hamiltonian consisting of an infinite sequence of many-body interaction potentials whose forms are determined by c. This model fluid exhibits very peculiar phase transitions, however [42].

Phase transitions warrant special mention. We cannot expect any approximate density functional treatment to answer *subtle* questions regarding phase transitions. Most approximate functionals are mean field in character, so that certain (but not all) effects of fluctuations will necessarily be omitted. An important case is that of the liquid–gas interface in a weak gravitational field where capillary-wave-like fluctuations play a role. Weeks and co-workers [43,44] (see also Bedeaux [45]) have emphasized the pitfalls that are encountered when one attempts [46] to apply simple approximations to $\mathcal{F}[\rho]$ in this problem. They have also devised [43] a method for constructing the exact functional for an "anisotropic" Kac–van der Waals fluid. We will return to the failings of the approximate treatments of the liquid–gas interface in Section V.A.

Perhaps it is worth pointing out the difference between the density functional strategy and the more conventional field-theoretical approach. There one does not normally make *direct* approximations to the generating functional $\Gamma[\bar{\Phi}]$; rather, one uses the machinery of loop expansions, and so on, to generate systematic approximations for thermodynamic functions and correlation functions [31]. This usually allows one to keep track of fluctuation effects. However, field-theoretical treatments of realistic models of inhomogeneous fluids are not easy! By approximating $\Gamma[\bar{\Phi}]$ (or $\mathcal{F}[\rho]$) directly it is sometimes difficult to ascertain what, if any, fluctuation effects are being incorporated into the theory. The Fisk–Widom [47] theory of the liquid–gas interface near the bulk critical point is in this spirit (see Section V.A). Although their functional omits interfacial (capillary-wave induced) broadening of the density profile, it is constructed so as to incorporate the effects of bulk critical fluctuations (i.e., the correct

bulk critical exponents). Setting aside the issue of fluctuations at phase transitions, it should be clear that the classical density functional strategy is very much in keeping with that used for electronic properties. In the Hohenberg–Kohn–Sham scheme the analog of $\mathcal{F}_{ex}[\rho]$ is the energy functional

$$\frac{1}{2}\iint d\mathbf{r}\, d\mathbf{r}' \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n]$$

and approximations are sought for $E_{xc}[n]$. As in simple Hartree theory, the direct Coulomb repulsion between electrons is treated exactly. For classical atomic fluids where the pairwise potential $\phi(r)$ has both repulsive and (shorter-range) attractive contributions a corresponding division of $\mathcal{F}_{ex}[\rho]$ is not so obvious. For ionic liquids, however, it is natural to separate out the total electrostatic energy and seek approximations for the remaining part of $\mathcal{F}_{ex}$, which is then the classical analog of the exchange and correlation functional (see, e.g., Refs. 22 and 23).

In the remainder of this section we describe various approximations for $\mathcal{F}[\rho]$ that have been developed for simple, atomic fluids.

B. Square-Gradient Approximation and a Descendent

The best known approximation is probably that arising from truncating the gradient expansion of $\mathcal{F}[\rho]$. This is derived by supposing that the density $\rho(\mathbf{r}) \equiv \Psi(\mathbf{r}/r_0)$, where the scale parameter $r_0 \to \infty$. Then the density may vary by large amounts but over a long distance scale. Following Hohenberg and Kohn [8], one finds that [16]

$$\mathcal{F}[\rho] = \int d\mathbf{r}[f(\rho(\mathbf{r})) + f_2(\rho(\mathbf{r}))(\nabla\rho(\mathbf{r}))^2 + O(\nabla\rho)^4] \tag{39}$$

Successive terms correspond to successive powers of r_0^{-1} and symmetry arguments, equivalent to those used in Landau theory, eliminate certain terms. $f(\rho)$ is the Helmholtz free-energy density of a uniform fluid of density ρ, so that truncating the expansion after the first term constitutes the local density approximation. This first term also contains all ideal gas contributions, since these are given exactly by the local density form [see (11)]. The other coefficients, $f_2(\rho)$, and so on, can only be determined by imposing additional requirements on $\mathcal{F}[\rho]$. A natural choice is that (39) should be consistent with linear response theory [i.e., with the change of free energy obtained by creating an infinitesimal perturbation of the density $\delta\rho(\mathbf{r})$ away from that of a uniform fluid]. That change can be obtained by functional Taylor expansion and involves the direct correlation func-

tions of the uniform fluid [8,15,16]. It is found that

$$f_2(\rho) = (12\beta)^{-1} \int d\mathbf{r}\, r^2 c^{(2)}(\rho;r) \tag{40}$$

The coefficient of the square-gradient term in the expansion of $\mathcal{F}[\rho]$ is $-(2\beta)^{-1}$ times the coefficient of k^2 in the Fourier transform of the two-body direct correlation function of the uniform fluid. Higher-order coefficients depend on integrals of $c^{(n)}$ with $n > 2$ and are much less amenable to calculation. For this reason and for overall simplicity (39) is usually truncated after the second term. With some means (e.g., via integral equation or perturbation theory) of calculating $c^{(2)}(\rho;r)$, and hence $f(\rho)$ and $f_2(\rho)$, for bulk fluids (39) constitutes a very simple but fully microscopic theory for an inhomogeneous fluid.

Formally, the square-gradient approximation should be valid only for the case of very slowly varying density profiles, such as would pertain for the liquid–gas interface near the bulk critical point [47,17] or for a single phase fluid in a gravitational field. However, the theory is often applied to situations were the profile varies rapidly—by orders of magnitude over a distance of a few atomic diameters in the case of the liquid–gas interface near the bulk triple point [16,17]. There is no reason to expect the theory to be accurate in these circumstances.

The neglect of higher-order terms in the expansion has severe repercussions when the theory is employed for fluids with power-law (algebraically decaying) pairwise potentials $\phi(r)$. A formal gradient expansion does not exist for such potentials. Since $c^{(2)}(\rho;r) \sim -\beta\phi(r)$ as $r \to \infty$, higher moments of $c^{(2)}$ will diverge if $\phi(r) \sim -r^{-n}$ as $r \to \infty$. The most relevant case is the Lennard–Jones 12–6 potential, which has $n = 6$. $f_2(\rho)$ exists but not higher-order coefficients. This is reflected in the small k behavior of the Fourier transform $c^{(2)}(\rho;k)$, where the presence of a k^3 term reflects the $-r^{-6}$ decay of $\phi(r)$ [48,49]. The square-gradient approximation fails to describe the z^{-3} algebraic decay of the tails of $\rho(z)$ at the liquid–gas interface of the Lennard–Jones fluid [50,51]; rather, it predicts exponential decay. More important, it cannot account for the proper growth law for wetting films in systems that exhibit van der Waals forces [52]. Such difficulties are best surmounted by treating attractive forces in a nonlocal fashion that avoids the gradient expansion. In the extreme case of an ionic liquid ($n = 1$) a gradient expansion is meaningful only for the residual, non-Coulombic, part of $\mathcal{F}_{ex}[\rho]$ [22].

There is another difficulty associated with the implementation of (39) with (40). If the fluid exhibits attractive, as well as repulsive interatomic forces, bulk liquid–gas coexistence will occur. The density of the inhomogeneous fluid may take values locally that lie within the bulk two-phase

region. This is certainly the case for many interfacial problems. What values should be used for $f(\rho)$, and how is one to calculate $c^{(2)}(\rho;r)$? In practice, $f(\rho)$ always has some (generalized) van der Waals form, so that $\mu(\rho) = \partial f/\partial\rho$ has a loop for subcritical temperatures [17], and some extrapolation of $c^{(2)}(\rho;r)$ [or $f_2(\rho)$] into the two-phase region is made. For the Lennard–Jones fluid, $f_2(\rho)$ is calculated [53,54] to be weakly density and temperature dependent. The integral in (40) is dominated by the larger r portion, where $c^{(2)}(\rho;r) \sim -\beta\phi(r)$ and f_2 can be regarded as a positive constant determined primarily by the attractive part of $\phi(r)$. The resulting square-gradient theory is then essentially the same as that of van der Waals (see Chapter 1 and Ref. 17).

Despite the shortcomings described above, the square-gradient approximation has proved extremely valuable for a wide variety of interfacial problems. Its generalization to fluid mixtures is straightforward [54,55] and has proved useful for understanding surface tension and relative adsorption at the liquid–gas [55,56] and liquid–liquid [56] interfaces of binary mixtures of simple fluids. It has also been used to investigate a wetting transition at a fluid–fluid interface and surface tension near a critical endpoint [56]. In these applications some prescription for calculating the coefficients $f_{2ij}(\rho_i,\rho_j)$ with i,j referring to species 1 and 2 is required. A simple perturbation theory, using a mixture of hard spheres as a reference system, has been used to calculate $c_{ij}^{(2)}$ for the bulk mixtures [55,56].

The square-gradient approximation formed the crucial ingredient for Cahn's [57] seminal paper on wetting transitions. Cahn employed (for a planar interface) the grand potential functional per unit area

$$\frac{\Omega_V[\rho]}{A} = \int_0^\infty dz \left[f(\rho(z)) + f_2 \left(\frac{d\rho}{dz}\right)^2 - \mu\rho(z) \right] + \Phi_w(\rho_w) \quad (41)$$

where the substrate (or spectator phase)–fluid contribution is taken to be

$$\Phi_w(\rho_w) = 2f_2 \left(\frac{C}{2} \rho_w^2 - \bar{\epsilon}\rho_w \right) \quad (42)$$

with $\rho_w \equiv \rho(0^+)$, the density of the fluid at the substrate (wall). The quantity $2f_2\bar{\epsilon} > 0$ measures the strength of the attractive substrate potential, while the term $f_2C\rho_w^2$ represents a modification of pair interactions between fluid particles at the substrate. For pure fluids it is natural to take $C > 0$ since the absence of fluid for $z < 0$ means that there is less net attraction from fluid atoms, causing an increase in surface energy. $C < 0$ implies some enhancement of the attractive pair attraction at the substrate, lowering the surface energy. Although (42) is not a particularly

realistic representation of the substrate–fluid free energy (it can be derived as the continuum limit of a lattice-gas model, where $\tilde{\epsilon}$ and C are more meaningful parameters [58]), it does have the advantage that the Euler–Lagrange equation obtained by minimizing (41) has a first integral so that the profile $\rho(z)$ is easily determined and the grand potential has a simple form. This makes (41) very well suited for *global* investigations of interfacial properties, including surface phase transitions such as wetting. Different classes of behavior can occur depending on the values of the parameters C (surface enhancement) and $\tilde{\epsilon}$ (surface field). Following Cahn's pioneering work there has been an enormous number of studies of (Landau) theories of this type (see the reviews [58,59]). Such a functional does not incorporate the short-ranged correlations, which lead to the oscillations in the density profile that usually occur for liquids near substrates. Any liquidlike film that develops near the substrate will be represented by a density that is high but almost constant. $\rho(z)$ can only vary monotonically. While this defect is serious for the microscopic structure of adsorbed films, it should not be crucial as regards predictions of global features of wetting transitions (at the mean-field level). The choice (42) is appropriate to short-range wall–fluid forces. In reality van der Waals forces between the atoms of the rigid substrate and those of the fluid give rise to an effective one-body potential $V(z) \sim -z^{-3}$, large z. The presence of such power-law potentials, $V(z) \sim z^{-m}$, leads to critical exponents for film growth at complete wetting (from off bulk coexistence; see Chapter 2 and Refs. 58 and 59) that depend directly on the value of m. The Cahn functional with (42) does not capture these effects; any critical exponents will take the mean-field values appropriate to a system where *all* forces are strictly short range. This is discussed further in Section V.B.

We complete this subsection by mentioning a close descendent of the square-gradient approximation. This is the functional

$$\mathcal{F}[\rho] = \int d\mathbf{r}\, f(\rho(\mathbf{r})) + (4\beta)^{-1} \int d\mathbf{r}_1 \int d\mathbf{r}_2\, c^{(2)}(\rho^m; r_{12})[\rho(\mathbf{r}_1) - \rho(\mathbf{r}_2)]^2 \tag{43}$$

introduced by Ebner et al. [13] in a study of a liquid–gas interface and used subsequently by Ebner and Saam [60,61] in their important work on gas adsorption at substrates, where they discovered (independently of Cahn) a first-order wetting transition. Equation (43) can be derived from a partial resummation of the gradient expansion, as suggested by Hohenberg–Kohn–Sham for the electronic case. It may also be justified in terms of an approximation [14,16] for $c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2)$ in (25) for the exact

functional. The important feature is that (43) is consistent with linear response (even for rapidly varying pertubations) and reduces to the gradient expansion for slowly varying densities. Practical implementation requires some prescription for the "mean" density ρ^m (e.g., $\rho^m = [\rho(\mathbf{r}_1) + \rho(\mathbf{r}_2)]/2$) and some means of calculating $c^{(2)}(\rho;r)$. Ebner et al. [13] solved the bulk Percus–Yevick equation numerically at a large set of state points for a Lennard–Jones fluid and extrapolated $c(\rho;r)$ into the two-phase region. They also make a complicated parametrization of the density profile in order to effect the minimization of $\Omega_V[\rho]$.

It is not clear that (43) will provide an accurate description of the highly structured density profiles that one finds for fluids near substrates. The results of Ebner and Saam [60,61] for gaseous "argon" adsorbed at a planar "carbon dioxide" substrate caused some controversy and stimulated many subsequent calculations, using various techniques, for the same model.* Recent Monte Carlo results [63] (see Section V.C) for the density profiles and adsorption provide convincing evidence for the existence of the wetting and prewetting transitions, with a wetting temperature similar to that calculated by Ebner and Saam. However, the Monte Carlo density profiles are considerably different from the Ebner–Saam density functional results as regards the shape of the oscillatory part. Whether this reflects some fundamental shortcoming of the theory or is an artifact of the parametrization [60] of the profile is not obvious. One disturbing feature of (43) is the presence of the local density term $f(\rho(\mathbf{r}))$. Packing effects induced by both the repulsive and attractive parts of the substrate potential $V(z)$ can cause the local density $\rho(z)$ to be very high—sometimes greater than the close-packing density. Under these circumstances $f(\rho(z))$ is divergent. Since the theory simply cannot tolerate such high densities, any minimization procedure would force the maxima in $\rho(z)$ to be smaller. In this respect (43) is similar to (41). Because of such difficulties and considerable numerical complexity, (43) has fallen out of fashion and appears to have been superseded by other approximations to be described later.

C. Density Expansions

Let us return to the exact expression (26) for the excess free-energy functional $\mathcal{F}_{ex}[\rho]$. We are free to choose the initial density $\rho_i(\mathbf{r})$ to be ρ_b, that of a uniform (bulk) reference fluid at the same chemical potential. Then

* Some of the history is given in footnote 79(b) of Ref. 62.

$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_b$ and the grand potential functional becomes

$$\Omega_V[\rho] = \Omega[\rho_b] + \int d\mathbf{r}\, V(\mathbf{r})\rho(\mathbf{r}) + \beta^{-1}\int d\mathbf{r}\left[\rho(\mathbf{r})\ln\frac{\rho(\mathbf{r})}{\rho_b} - \rho(\mathbf{r}) + \rho_b\right] + \beta^{-1}\int_0^1 d\alpha(\alpha - 1)\int d\mathbf{r}_1\int d\mathbf{r}_2\, c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2)\,\Delta\rho(\mathbf{r}_1)\Delta\rho(\mathbf{r}_2) \tag{44}$$

having used (15). Suppose now that we neglect the dependence of $c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2)$ on the coupling parameter α and, for simplicity, set this function equal to the two-body direct correlation function of the uniform fluid, so that

$$c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2) \approx c^{(2)}(\rho_b;|\,\mathbf{r}_1 - \mathbf{r}_2\,|) \tag{45}$$

Then (44) simplifies to [64]:

$$\Omega_V[\rho] = \Omega[\rho_b] + \int d\mathbf{r}\, V(\mathbf{r})\rho(\mathbf{r}) + \beta^{-1}\int d\mathbf{r}\left[\rho(\mathbf{r})\ln\frac{\rho(\mathbf{r})}{\rho_b} - \rho(\mathbf{r}) + \rho_b\right] - (2\beta)^{-1}\int d\mathbf{r}_1\int d\mathbf{r}_2\, c^{(2)}(\rho_b;r_{12})(\rho(\mathbf{r}_1) - \rho_b)(\rho(\mathbf{r}_2) - \rho_b) \tag{46}$$

This functional can now be minimized, according to (9), and yields the following integral equation for the density profile:

$$\rho(\mathbf{r}_1) = \rho_b \exp\left[-\beta V(\mathbf{r}_1) + \int d\mathbf{r}_2\, c^{(2)}(\rho_b;r_{12})(\rho(\mathbf{r}_2) - \rho_b)\right] \tag{47}$$

which is the same as that obtained by making the HNC closure of the wall–particle Ornstein–Zernike equation (see Chapter 4). The Percus–Yevick (PY) version is a simple linearization of part of the exponent in (47), that is,

$$\rho(\mathbf{r}_1) = \rho_b \exp[-\beta V(\mathbf{r}_1)]\left[1 + \int d\mathbf{r}_2\, c^{(2)}(\rho_b;r_{12})(\rho(\mathbf{r}_2) - \rho_b)\right] \tag{48}$$

Both equations, along with closely related approximations based on alternative closures of the wall–particle Ornstein–Zernike equation, have been used in many studies of the density profile of liquids and gases near walls. Some of this work is reviewed in Chapter 4. While these theories are quite successful at describing the oscillatory profiles of hard spheres near hard walls, they are less successful when the fluid possesses an attractive, as well as a repulsive, component in the interatomic potential. One severe drawback of this type of integral equation theory is their inability to account for the presence of macroscopically thick wetting (or drying) films at a wall–fluid interface [65] or for the phenomenon of critical

adsorption, which arises from the slow, algebraic decay of the density profile associated with the diverging bulk correlation length [66]. Such theories are not capable of describing phase transitions at fluid interfaces.

Their deficiencies can best be understood by reconsidering their generating functional (46). Clearly, this approximation retains only terms quadratic in the "perturbation" $\rho(\mathbf{r}) - \rho_b$; the logarithm merely reflects the ideal gas contribution. Note that $c^{(2)}(\rho_b;r)$ is fixed once μ and T are specified. If one considers (46) for bulk densities ρ, other than the initial one, one finds that the grand potential density

$$\omega(\rho) = \omega(\rho_b) + \beta^{-1}\left(\rho \ln \frac{\rho}{\rho_b} - \rho + \rho_b\right) - \frac{\beta^{-1}}{2}\int d\mathbf{r}\, c^{(2)}(\rho_b;r)(\rho - \rho_b)^2$$

cannot account for liquid–gas coexistence [65]. A quadratic is insufficient to describe two minima, which is a necessary requirement for coexistence in the uniform fluid. If the functional does not exhibit two minima (corresponding to bulk liquid and bulk gas), it will not be able to describe the development of a macroscopic wetting film, nor will it be able to describe correctly the interface near the critical point, whose properties reflect (within the mean field approach) the coalescence of two minima [66].

How does one improve on (46)? It is straightforward to show that this approximation is equivalent to retaining only the quadratic term in the functional Taylor expansion of $\mathcal{F}_{ex}[\rho]$ about the uniform reference value ρ_b. The next term in the expansion is

$$\frac{1}{3!}\int d\mathbf{r}_1 \int d\mathbf{r}_2 \int d\mathbf{r}_3 \left.\frac{\delta^3 \mathcal{F}_{ex}[\rho]}{\delta\rho(\mathbf{r}_1)\,\delta\rho(\mathbf{r}_2)\,\delta\rho(\mathbf{r}_3)}\right|_{\rho_b} \times (\rho(\mathbf{r}_1) - \rho_b)(\rho(\mathbf{r}_2) - \rho_b)(\rho(\mathbf{r}_3) - \rho_b)$$

where the derivative can be identified with $-\beta^{-1}c^{(3)}(\rho_b;\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$, the three-body direct correlation function of the uniform fluid. Including such a term is sufficient to ensure that $\omega(\rho)$ does have two minima, so that coexistence and wetting are possible [67]. Rickayzen and Augousti [68] have included the third-order term by making a simple ansatz for $c^{(3)}$, characterized by a single parameter. The resulting theory, obtained by minimizing a new functional, does improve on (47) and (48) for the density profile of hard spheres at a hard wall [68] and has proved very successful for describing the density profile of certain (supercritical) states of a Lennard–Jones fluid confined in a slit [69]. The parameter in the third-order theory is *chosen* in such a way that the exact sum rule ($\beta p = \rho_w$) for the density at a single hard wall is satisfied automatically. [It is well known that the sum rule is not satisfied by either the HNC (47) or PY (48) equations.] A generalization of the third-order theory to binary hard-sphere

mixtures has been made and tested against simulations for a mixture confined in hard-wall slit [70]. Note that the input to both the quadratic and the third-order theories is the direct correlation function $c^{(2)}(\rho_b;r)$ at some specified bulk density ρ_b and temperature. For fluid absorption at a single wall it is natural to choose ρ_b equal to the density of the bulk fluid far from the wall. For confined fluid, however, the choice of ρ_b is not quite so obvious. If the fluid has access to a reservoir, as for an open slit or cylinder, one might choose ρ_b to be the density of the bulk fluid at the chemical potential fixed by the reservoir. Then the comparison with grand canonical Monte Carlo can be made. Such a comparison is shown in Fig. 1 for a supercritical Lennard–Jones fluid between two "10–4–3" walls. For a subcritical Lennard–Jones fluid capillary condensation of an undersaturated gas to a "liquid" will occur in pores with attractive walls (see Section V.D; a brief review of the phase equilibria of fluids confined in pores is given in Ref. 71). If the walls are hard capillary evaporation of the liquid to a "gas" will occur for μ slightly greater than its value at saturation [72]. While it is clear that (47) or (48) cannot account for such phase transitions, it is not clear how the third-order theory should fare. Reference 69 reports comparison of theory with canonical Monte Carlo results—the simulations and theory referring to the same average density of fluid in the hard-wall slit. Such a comparison is problematical at subcritical temperatures, especially when phase transitions are possible.

Other improvements on the HNC and Percus–Yevick approximations have been developed. Some of these are known to correct the fundamental deficiencies described above. In particular, Zhou and Stell [73] have developed powerful nonlocal integral-equation approximations, which they term hydrostatic HNC and PY, based on a density functional expansion of direct correlation functions developed earlier by Stell and co-workers. Zhou and Stell show that the HHNC approximation is much superior to HNC and PY for Lennard–Jones fluids near walls and confined in slits. They also demonstrate that HHNC can account for wetting and drying transitions at a single wall and for capillary condensation in a slit pore. A second class of improvements makes use of the inhomogeneous Ornstein–Zernike equation (18) (see Chapter 4 and the paper by Kjellander and Sarman [74] for a comprehensive list of references and a thorough discussion of some of the most recent results). Such theories take both the singlet density and the pair distribution functions as unknown (i.e., the closure is effected at the pair level, which demands a very large computational effort). We do not discuss this type of approach further.

Rather we return to the question of the usefulness of the functional Taylor expansion. This has received much attention from the practitioners of the density functional theory of bulk freezing. Indeed, the quadratic

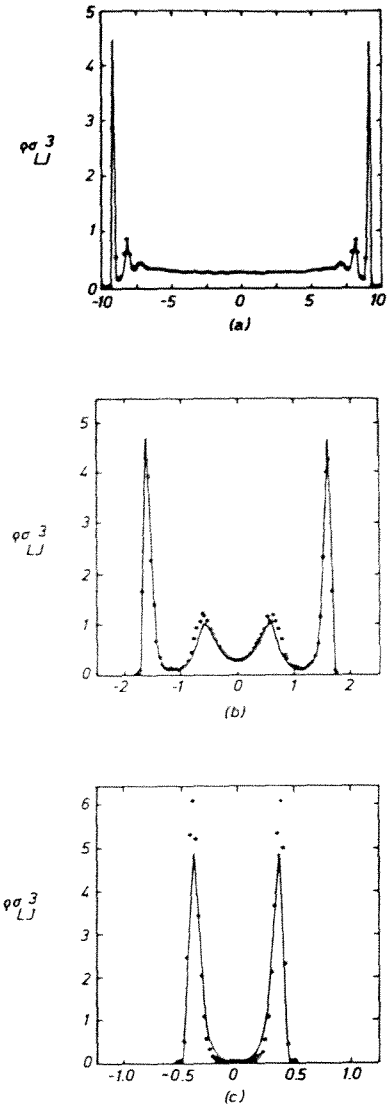


Fig. 1 Density profiles of a (truncated) Lennard–Jones fluid confined between two parallel walls exerting 10–4–3 potentials mimicking ethylene on graphite: (a) wall separation $L = 20\sigma_{LJ}$; (b) $L = 5\sigma_{LJ}$; (c) $L = 2.5\sigma_{LJ}$. In each case $k_BT/\epsilon = 1.35$, which lies above the critical temperature, and $\mu = -3\epsilon$, which fixes the density $\rho_b\sigma^3_{LJ} = 0.28$. The solid curves are the results of Rickayzen–Augousti theory and the points are the grand-canonical Monte Carlo results of Walton and Quirke. (Redrawn from Ref. 69.)

approximation (46) formed the starting point for modern developments of this subject (see Chapter 9). Within the context of freezing ρ_b in (46) refers to the density of a uniform liquid and $\rho(\mathbf{r})$ to that of the crystal; $V(\mathbf{r}) \equiv 0$. Considered as a generating functional (46) asserts that the direct correlation function for any inhomogeneous fluid, including the crystal, is $c^{(2)}(\rho_b;r_{12})$. [This follows directly using (13).] That correlations in a crystal should be identical to those in a uniform liquid is an assertion that is somewhat difficult to sustain! Nevertheless, results for freezing properties obtained from (46) have been found to be rather good in certain cases (Chapter 9). It appears that this quadratic approximation is capable of describing a solid–liquid transition—provided that the spontaneous breaking of translational symmetry is inserted by hand (i.e., a specific periodic crystalline density is imposed). Those of us raised on a diet of fluid interfaces cannot fail to be amazed by the apparent success of this (lowest-order) theory for freezing. Higher-order terms in the functional expansion have been considered and their effects are often substantial (a very recent discussion of the convergence of the density expansion for hard-sphere fluids is given in Ref. 75; see also Refs. 76 and 90) but this subject is reviewed in Chapter 9.

D. Perturbation About a Hard-Sphere Reference Fluid: A van der Waals Approximation

In this section we consider approximate free-energy functionals obtained from (33), the exact expression for the intrinsic free energy for a fluid with a pairwise interatomic potential $\phi(r)$. In the perturbation theory of *uniform* liquids $\rho^{(2)}(\phi_\alpha;\mathbf{r}_1,\mathbf{r}_2)$ is replaced by $\rho^2 g_r(\rho;r)$—the pair correlation function of the uniform reference fluid ($\alpha = 0$) with the potential $\phi_r(r)$ corresponding to the repulsive force part of the full potential [30]. This is equivalent to assuming that pairwise correlations are determined primarily by excluded volume effects resulting from the repulsive forces. For a dense uniform Lennard–Jones liquid this is known to be correct [30,77] and the resulting approximation for the Helmholtz free energy is rather accurate. The construction of an analogous perturbation theory for the inhomogeneous fluid is less straightforward because much less is known about the nature of pair correlations in such fluids. There are many interfacial situations when one *does* expect $\rho^{(2)}(\phi;\mathbf{r}_1,\mathbf{r}_2)$ to be substantially different from $\rho^{(2)}(\phi_r;\mathbf{r}_1,\mathbf{r}_2)$, where ϕ_r again refers to the repulsive part of ϕ. Notwithstanding, a simple choice, suggested by the success of the corresponding bulk theory, is to set

$$\rho^{(2)}(\phi_\alpha;\mathbf{r}_1,\mathbf{r}_2) \approx \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)g_r(\rho^m;r_{12}) \tag{49}$$

where ρ^m is some mean of the local densities at $\mathbf{r}_1$ and $\mathbf{r}_2$. Toxvaerd [78] pioneered such an approach using hard spheres as the reference fluid. For the free-energy functional of the latter he made the local density approximation (LDA):

$$\mathcal{F}_r[\rho] = \mathcal{F}_{hs}[\rho] \approx \int d\mathbf{r}\, f_{hs}(\rho(\mathbf{r})) \tag{50}$$

where $f_{hs}(\rho)$ is the Helmholtz free-energy density of a uniform hard-sphere fluid. By parametrizing $\rho(z)$ he minimized the resulting $\Omega_V[\rho]$ for the planar liquid–gas interface of a Lennard–Jones fluid to obtain estimates of the surface tension and equilibrium density profile. Other work of this type is reviewed in Refs. 17 and 54.

An even simpler approximation,

$$\rho^{(2)}(\phi_\alpha;\mathbf{r}_1,\mathbf{r}_2) \approx \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) \tag{51}$$

which ignores correlations completely, leads to what is commonly termed a van der Waals theory of nonuniform fluids. With (51) the free-energy functional (33) reduces to

$$\mathcal{F}[\rho] = \mathcal{F}_r[\rho] + \frac{1}{2}\int d\mathbf{r}_1 \int d\mathbf{r}_2\, \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\phi_p(r_{12}) \tag{52}$$

which after making the LDA (50) becomes

$$\mathcal{F}[\rho] = \int d\mathbf{r}\, f_{hs}(\rho(\mathbf{r})) + \frac{1}{2}\int d\mathbf{r}_1 \int d\mathbf{r}_2\, \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\phi_{att}(r_{12}) \tag{53}$$

ϕ_p is taken to be the attractive part ϕ_{att} of the pairwise potential ϕ, but there is ambiguity in defining ϕ_{att} inside the core region, reflecting the absence of correlations in (51). For a Lennard–Jones fluid ϕ_{att} is often set equal to ϕ_{min}, the value of ϕ at its minimum, for $r < \sigma$, the effective hard-sphere diameter—the latter being determined by standard bulk procedures [30].

Using (53) the variational principle (9) yields

$$\mu = V(\mathbf{r}_1) + \mu_{hs}(\rho(\mathbf{r}_1)) + \int d\mathbf{r}_2\, \rho(\mathbf{r}_2)\phi_{att}(r_{12}) \tag{54}$$

for the equilibrium profile $\rho(\mathbf{r}_1)$. The intrinsic chemical potential at $\mathbf{r}_1$ consists of two distinct parts: a hard-core piece determined by local thermodynamics ($\mu_{hs} \equiv \partial f_{hs}/\partial\rho$) and a nonlocal piece, arising from the attractive tails of surrounding atoms. Such a division lies at the heart of van der Waals' own approach to nonuniform fluids (see Chapter 1). Our present "derivation" of the van der Waals functional (53) emphasizes the assumptions and approximations that must be made if the resulting theory

is to be applied to a realistic model fluid. There are other ways of arriving at (53). These are summarized by Sullivan [32] (see also Ref. 43), who discusses a γ ordering procedure for obtaining systematic corrections to (53) in the limit $\gamma \to 0$. (γ measures the strength and inverse range of ϕ_{att}.)

There has been an enormous number of applications of (53) to interfacial problems. This simple theory has provided insight into the free liquid–gas interface, adsorption and wetting phenomena of all types [58,59], and phase transitions of confined fluids [71]. But what is the status of (53)? For a uniform fluid the free energy per atom reduces to

$$\frac{F}{N} = \rho^{-1} f_{hs}(\rho) - \frac{\alpha}{2}\rho \tag{55}$$

with $\alpha = -\int d\mathbf{r}\ \phi_{att}(r) > 0$, the integrated strength of the attractive potential. With the Carnahan–Starling [79] prescription for f_{hs} (55) is a respectable but not an accurate approximation for a simple fluid; its shortcomings have been discussed [80]. The direct correlation function obtained from the second functional derivative of (53) is

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \left[-\beta f''_{hs}(\rho(\mathbf{r}_1)) + \frac{1}{\rho(\mathbf{r}_1)} \right] \delta(\mathbf{r}_1 - \mathbf{r}_2) - \beta\phi_{att}(r_{12}) \tag{56}$$

where the prime denotes differentiation with respect to density. The unphysical delta function merely reflects the LDA (50), which is tantamount to assuming the correlation length for hard-sphere repulsion is zero. No information about the detailed short-range correlations that characterize a real fluid are included in (56). For a uniform fluid (56) predicts the correct asymptotic behavior $c^{(2)}(\rho;r) \to -\beta\phi_{att}(r)$ as $r \to \infty$ [81]. More significantly, it implies the same asymptotics for *any* inhomogeneous fluid. This must reflect the underlying mean-field character of the approximation; it does not incorporate any fluctuation corrections that may arise. One good feature of (56) is that for uniform fluid the compressibility sum rule (24) is satisfied and one might hope that certain sum rules for inhomogeneous fluids are also satisfied by this functional.

Let us list some of the advantages of (53) over other approximations:

1. It is not necessary to parametrize the profile; in most cases the integral equation (54) is easily solved by Picard iteration. When convergence is slow, as is the case with thick wetting films, the grand potential as a function of an imposed film thickness can be calculated [82], which is particularly useful for investigating phase transitions.
2. The theory is versatile and can be extended straightforwardly to mixtures [58,59,83].

3. The division into hard-sphere and attractive portions avoids the difficulties associated with specifying $f(\rho)$ and $c(\rho;r)$ in the two-phase region.
4. Treating attractive forces in a nonlocal fashion avoids the shortcoming of the square-gradient approximation in not accounting properly for algebraically decaying $\phi(r)$. Within a mean-field situation, the van der Waals theory will treat long-range forces correctly. In the square-gradient approximation (39) the whole of the direct correlation function is delta function-like:

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \left[-\beta f''(\rho(\mathbf{r}_1)) + \frac{1}{\rho(\mathbf{r}_1)} - 2\beta f_2 \nabla^2 \right] \delta(\mathbf{r}_1 - \mathbf{r}_2) \quad (57)$$

   where we have neglected any density dependence of f_2. Note that Taylor expansion of (53) or use of (40) imply that $f_2 = -\frac{1}{12} \int d\mathbf{r}\, r^2 \phi_{\text{att}}(r)$ for the van der Waals theory.
5. The relative simplicity of (56) means that it is sometimes possible to invert the Ornstein–Zernike equation and obtain (an approximation for) the pairwise distribution function of the inhomogeneous liquid without enormous computational effort. This is especially useful for phase transitions, where singular contributions can often be ascertained by analytical means [84,85] (see Section V.B).

The crucial disadvantages of (53) are:

1. The form of $c^{(2)}$ is such that the theory is necessarily mean-field like. This has significant repercussions for phase transitions at interfaces (see Sections V.A and V.B). However, systematic improvements, incorporating fluctuation corrections, are notoriously difficult to make (e.g., Ref. 32).
2. Short-range correlations are absent, so that the theory cannot describe oscillatory density profiles, nor will it give a correct description of the density at a sharp boundary such as a hard wall; the relevant sum rule will not be satisfied.

Most schemes for improvement have focused on item 2, often specializing to a hard-sphere fluid from the outset (i.e., the division $\mathcal{F}[\rho] = \mathcal{F}_{\text{hs}}[\rho] + \mathcal{F}_{\text{att}}[\rho]$ is accepted and an approximation, better than LDA, is sought for $\mathcal{F}_{\text{hs}}[\rho]$). The attractive force part $\mathcal{F}_{\text{att}}[\rho]$ is still treated at the mean-field level. Some schemes claim, explicitly or implicitly, to apply to the full functional. In Section III.E we describe weighted-density approximations for $\mathcal{F}_{\text{hs}}[\rho]$, pointing out, where necessary, difficulties that arise when these are applied to $\mathcal{F}[\rho]$.

E. Weighted-Density Approximations

The initial idea behind weighted (or smoothed)-density approximations (WDAs) was the introduction of a coarse-graining procedure whereby a smoothed density $\bar{\rho}(\mathbf{r})$ is constructed as an average of the true density profile $\rho(\mathbf{r})$ over a local volume that is determined by the range of interatomic forces. The pronounced peaks that occur in an oscillatory profile, where the local density may exceed that for close packing, are smoothed out in the coarse-grained $\bar{\rho}(\mathbf{r})$ so that the (excess) free energy should be well approximated by a *local* function of $\bar{\rho}(\mathbf{r})$:

$$\mathcal{F}_{ex}[\rho] = \mathcal{F}_{hs}[\rho] - \mathcal{F}_{id}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\psi_{ex}(\bar{\rho}(\mathbf{r})) \tag{58}$$

where $\psi_{ex}(\rho) = [f_{hs}(\rho) - f_{id}(\rho)]/\rho$ is the excess, over ideal, free energy per atom. The different versions of WDA correspond to different recipes for $\bar{\rho}(\mathbf{r})$.

1. *Recipes of Nordholm et al. and Tarazona (Mark I)*

Perhaps the earliest recipe is that of Nordholm and co-workers [86], for which

$$\bar{\rho}(\mathbf{r}) = \int d\mathbf{r}'\, w_0(|\,\mathbf{r} - \mathbf{r}'\,|)\rho(\mathbf{r}') \tag{59}$$

with a *weight function* w_0 proportional to the Heaviside step function

$$w_0(r) \equiv \frac{3}{4\pi\sigma^3}\,\Theta(\sigma - r) \tag{60}$$

This choice of weight function was considered earlier by Stell [7] in a different application of effective density techniques. It is equivalent to his lowest-order result $w_0(r) = f(r)/\int d\mathbf{r}\, f(r)$, with the Mayer function $f(r) = \exp[-\beta\phi(r)] - 1$, applied to the hard-sphere fluid. The normalization ensures that $\bar{\rho}(\mathbf{r})$ reduces to the constant value ρ in the limit of a uniform fluid. Nordholm et al. used the simple approximation $\beta\psi_{ex}(\rho) = -\ln(1 - \rho v_0)$ with the excluded volume $v_0 = \sigma^3$ [see (37)]. Their results for hard spheres near a hard wall gave sensible oscillatory profiles but were not in good agreement with simulation. Two subsequent applications of the recipes, with the mean-field approximation for $\mathcal{F}_{att}$, to the calculation of solvation forces for a Lennard–Jones fluid confined between two substrates [87] and to the study of wetting transitions [88] demonstrated its usefulness for phase transitions. Since the approximation was formulated in terms of a coarse-graining argument for a nonlocal entropy functional (in the same spirit as van Kampen [89]), any connections between what was termed a fine-grained generalized van der Waals theory

and density functional approaches remained somewhat obscure until Tarazona [90,91] examined the consequences of (58) with (59) as a theory for $c_{hs}^{(2)}(\rho;r)$, the direct correlation function of a uniform hard-sphere fluid. Replacing the excluded volume approximation by the accurate Carnahan and Starling [79] result

$$\beta\psi_{ex}(\rho) = \eta(4 - 3\eta)(1 - \eta)^{-2} \qquad \eta = \frac{\pi\sigma^3\rho}{6}$$

Tarazona compared the $c_{hs}^{(2)}$ resulting from functional differentiation of (58) with the PY results for various densities [91]. The WDA defined by (60) grossly underestimates the magnitude of $c_{hs}^{(2)}(\rho;r)$ for $r \ll \sigma$ and yields a sizable decaying tail for $r > \sigma$. While this still represents a poor theory of pair correlations in a bulk hard-sphere liquid, the resulting $c_{hs}^{(2)}$ is a major improvement on a delta function! Results [91] for hard spheres near a hard wall are still not accurate but represent an improvement on those of Ref. 86, which can be traced to the use of the Carnahan–Starling equation of state. The WDA does satisfy the hard-wall sum rule $\beta p = \rho_w$. Thus the more accurate the bulk equation of state, the better is the agreement between simulation and theory for the contact density ρ_w. That a WDA should give a proper description of the density discontinuity at a hard wall arises from the fact that the smoothed density $\bar{\rho}(z)$, and hence $c^{(1)}(z)$, remain continuous [92]; this feature constitutes a significant improvement over the LDA.

Tarazona and Evans [91] showed that the WDA, with the mean-field approximation for $\mathcal{F}_{att}$, provides a realistic description of the phenomenon of complete drying at the hard wall–liquid interface. The density profiles of a Lennard–Jones liquid exhibit oscillations that become much less pronounced as the bulk chemical potential is reduced toward its value at saturation. In the limit $\mu \to \mu_{sat}^{+}$ a thick layer of gas develops at the wall. Theirs was probably the first successful theoretical treatment of this phenomenon. Tarazona [90] used the same functional in his first study of the freezing of hard spheres and disks. He also calculated the phase diagram for the bulk Lennard–Jones system, showing that a *single* theory can predict reasonable gas–liquid and solid–liquid phase boundaries, including the triple point. Later, Mederos et al. [93] applied Tarazona Mark I to the investigation of phase transitions in submonolayer films adsorbed at attractive substrates. By modeling adsorbate–adsorbate interactions in terms of hard disks, treated within the WDA, they obtained rich phase diagrams, which include a fluid modulated by the (periodic) substrate potential $V(\mathbf{r})$, a commensurate crystal phase, occupying a sublattice of the substrate and an "intrinsic" crystal phase, associated with the intrinsic crystalline order of the hard-disk system, that is incommensurate

with the substrate. Possible defect and domain structures were also considered. This study appears to be the first based on a realistic *continuum* theory. It avoids the inherent difficulties that lattice-gas models have in describing the intrinsic ordering of the adsorbate.

It should be evident that the WDA does not correspond to any finite-order density expansion of $\mathcal{F}_{ex}[\rho]$. In this respect it is superior to the approximations described in Section III.C and constitutes an alternative nonperturbative approach.

2. *Tarazona (Mark II)*

In a subsequent paper Tarazona [94] developed a more sophisticated version of his original WDA. The basic strategy was to construct a smoothed density, for use in (58), which would lead to an accurate $c_{hs}^{(2)}(\rho;r)$ for the uniform fluid over a wide range of (bulk) densities ρ. By ensuring that a functional gives a good account of bulk *pair* correlations, one might expect the same functional to provide an accurate description of the singlet density $\rho(\mathbf{r})$ and the free energy of an inhomogeneous fluid. To this end the weight function is allowed to depend on the smoothed density so that the latter is determined by the implicit equation

$$\bar{\rho}(\mathbf{r}) = \int d\mathbf{r}'\, w(|\,\mathbf{r} - \mathbf{r}'\,|;\bar{\rho}(\mathbf{r}))\rho(\mathbf{r}') \tag{61}$$

which replaces (59). w is specified by requiring (58) with (61) to produce, upon differentiation, $c_{hs}^{(2)}(\rho;r)$ close to the corresponding PY result. The analysis is simplified by adopting a (truncated) power-series expansion

$$w(r;\rho) = w_0(r) + w_1(r)\rho + w_2(r)\rho^2 \tag{62}$$

and the first two coefficients, w_0 and w_1, are calculated by comparison with the virial expansion of $c_{hs}^{(2)}(\rho;r)$, while the third, w_2, is obtained from a fit to the PY result [94]. w_0 is identical to the zeroth-order result (60) used in the Mark I version of the theory. Explicit formulas for w_1 and w_2 are given in Ref. 72; beware of typographical errors in earlier papers [94].

Then (61) reduces to a quadratic equation for $\bar{\rho}(\mathbf{r})$:

$$\bar{\rho}(\mathbf{r}) = \bar{\rho}_0(\mathbf{r}) + \bar{\rho}_1(\mathbf{r})\bar{\rho}(\mathbf{r}) + \bar{\rho}_2(\mathbf{r})[\bar{\rho}(\mathbf{r})]^2 \tag{63}$$

where

$$\bar{\rho}_j(\mathbf{r}) = \int d\mathbf{r}'\, w_j(|\,\mathbf{r} - \mathbf{r}'\,|)\rho(\mathbf{r}') \qquad j = 0, 1, 2$$

are coefficients depending on $\bar{\rho}(\mathbf{r})$. The physical root of (63) is

$$\bar{\rho}(\mathbf{r}) = \frac{1 - \bar{\rho}_1(\mathbf{r}) - [(1 - \bar{\rho}_1(\mathbf{r}))^2 - 4\bar{\rho}_0(\mathbf{r})\bar{\rho}_2(\mathbf{r})]^{1/2}}{2\bar{\rho}_2(\mathbf{r})}$$

$$= \frac{2\bar{\rho}_0(\mathbf{r})}{1 - \bar{\rho}_1(\mathbf{r}) + [(1 - \bar{\rho}_1(\mathbf{r}))^2 - 4\bar{\rho}_0(\mathbf{r})\bar{\rho}_2(\mathbf{r})]^{1/2}}$$

where the second form is more convenient for calculations since in the limit of a uniform fluid $\bar{\rho}_0(\mathbf{r}) = \rho$, while $\bar{\rho}_1(\mathbf{r}) = \bar{\rho}_2(\mathbf{r}) = 0$. Note the requirement that $\bar{\rho}(\mathbf{r})$ reduce to ρ in the case of a uniform fluid implies the normalization condition

$$\int d\mathbf{r}\, w(r;\rho) = 1$$

for all ρ, which [with (62)] implies that

$$\int d\mathbf{r}\, w_j(r) = \begin{cases} 1 & j = 0 \\ 0 & j = 1, 2 \end{cases}$$

Like Mark I this Mark II version satisfies the bulk compressibility sum rule (24). As pointed out by Tarazona [94], this implies that use of the Carnahan–Starling result for ψ_{ex} already fixes the value of $\int d\mathbf{r}\, c_{hs}^{(2)}(\rho;r)$, for each density. There is then a slight inconsistency involved in fitting w_1 and w_2 to PY results, but this is not significant since all that is being required is a realistic description of $c_{hs}^{(2)}(\rho;r)$ throughout the range of fluid densities.

Tarazona [94] showed that his Mark II version gave substantially better agreement with simulation for the density profile and surface tension of hard spheres near a hard wall than did Mark I. Moreover, the improved version gave a better account of hard-sphere freezing. There have been numerous subsequent applications of Mark II. Tarazona et al. [72] compared the results of this WDA with those of the LDA (53) for an example of a wetting transition, for capillary evaporation of a liquid confined between two hard walls, and for capillary condensation of a fluid confined between two adsorbing walls and in a cylinder. Unlike the LDA, the WDA yields realistic oscillatory density profiles, which in the case of a fluid confined in a narrow slit or cylinder, reflect packing constraints on the arrangements of atoms. The WDA also yields a realistic equation of state in the limit of a two-dimensional fluid [72]. Other applications to wetting transitions can be found in Refs. 95 and 96 and to prewetting in Ref. 97 (see Section V.C.1). A careful comparison between the results of WDA and simulation for the wetting and drying transitions of a square-well fluid at the square-well wall has been carried out by van Swol and Henderson

[92,98]. Ball and Evans [99] have investigated layering transitions for gases adsorbed at strongly attractive planar substrates and in pores (see also Ref. 100 and Section V.C.2).

Several groups have applied the theory successfully to the structure and phase equilibria of confined fluids. References 101 and 102 describe results for capillary condensation in cylinders, showing that the theory yields density profiles in good agreement with simulation and gives an adequate estimate of the location (in chemical potential) of the phase transition for a wide range of pore radii. Vanderlick et al. [103] compare the results of Tarazona Mark II for the density profile of a hard-sphere fluid confined between two hard walls and of a Lennard–Jones fluid confined by Lennard–Jones 10–4–3 walls with those of simulation and other approximation schemes. The comparison is made at fixed mean pore density. Mark II is considerably more accurate than Mark I or the original Nordholm version (with the simplified bulk equation of state). In the case of the hard-sphere fluid (see Fig. 2) at high values of the mean pore density, Mark I predicts erroneous peak positions and peak heights reminiscent of crystalline densities for the layered fluid, while the Nordholm version leads to negative densities in some of the troughs. The applications listed here all employ the mean-field treatment for $\mathcal{F}_{\text{att}}[\rho]$ when attractive fluid–fluid forces are present. Some modifications to $\phi_{\text{att}}(r)$ are sometimes made [95–98] to ensure a better bulk equation of state. Solving the integral equation for $\rho(\mathbf{r})$, given by minimizing $\Omega_V[\rho]$, requires more computational effort than solving (54) for the van der Waals approximation, since $\bar{\rho}(\mathbf{r})$ must be determined self-consistently. Nevertheless, it is tractable without recourse to parametrization—provided that there are no solid phases present. For the freezing problem Tarazona [94] introduced a set of identical gaussians centered on lattice sites $\mathbf{R}_l$ to represent the density of the crystal:

$$\rho_{\text{crys}}(\mathbf{r}) = \left(\frac{\alpha}{\pi}\right)^{3/2} \sum_l \exp[-\alpha(\mathbf{r} - \mathbf{R}_l)^2] \tag{64}$$

The width parameter α is the only variational parameter once the lattice type has been specified; the lattice spacing is fixed by the average density. The elastic constants of the hard-sphere crystal have also been calculated within this approximation scheme [104].

There is no doubt that Tarazona Mark II has proved a versatile and, in many cases, an accurate approximation scheme for inhomogeneous fluids, and its success lead the Cornell group [105,106] to consider its extension to mixtures. For a binary fluid the appropriate generalization

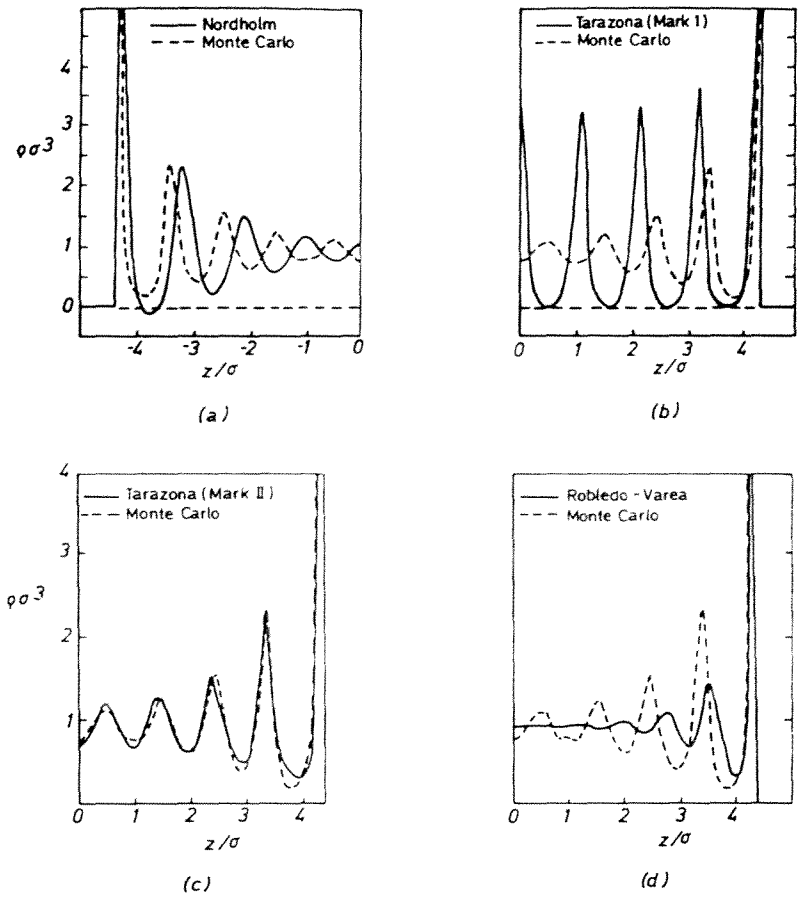


Fig. 2 Density profiles for a hard-sphere fluid confined between two parallel hard walls separated by a distance $L = 8.74\sigma$. The mean pore density is fixed at $0.897\sigma^{-3}$ in all cases, but the chemical potential μ is different in the various theories. (a) The Nordholm et al. recipe, using the simple excluded volume equation of state; (b) Tarazona (Mark I), with the Carnahan–Starling equation of state; (c) Tarazona (Mark II); (d) Robledo–Varea recipe with the Carnahan–Starling equation of state. The dashed lines are the Monte Carlo results for the same mean pore density. (Note that the profiles are drawn for half the slit.) (Redrawn from Ref. 103.)

of the grand potential functional is

$$\Omega_V(\rho_1,\rho_2] = \mathcal{F}_{id}[\rho_1,\rho_2] + \mathcal{F}_{ex}[\rho_1,\rho_2] - \sum_{i=1}^{2} \int d\mathbf{r}\,(\mu_i - V_i(\mathbf{r}))\rho_i(\mathbf{r}) \qquad (65)$$

where $\rho_i(\mathbf{r})$ and μ_i are the number density and chemical potential of species i, respectively and $V_i(\mathbf{r})$ is the external potential exerted on species i. The ideal gas term is simply

$$\beta\mathcal{F}_{id}[\rho_1,\rho_2] = \sum_{i=1}^{2} \int d\mathbf{r}\ \rho_i(\mathbf{r})[\ln \Lambda_i^3\rho_i(\mathbf{r}) - 1] \qquad (66)$$

The excess free-energy functional is approximated by

$$\mathcal{F}_{ex}[\rho_1,\rho_2] = \sum_{i=1}^{2} \int d\mathbf{r}\ \rho_i(\mathbf{r})\psi_{ex}^i(\bar{\rho}_1(\mathbf{r}),\bar{\rho}_2(\mathbf{r})) + \frac{1}{2}\sum_{i,j} \iint d\mathbf{r}\,d\mathbf{r}'\ \phi_{att}^{ij}(|\mathbf{r} - \mathbf{r}'|)\rho_i(\mathbf{r})\rho_j(\mathbf{r}') \qquad (67)$$

where $\psi_{ex}^i(\rho_1,\rho_2)$ is the excess Helmholtz free energy per atom of species i of the uniform hard-sphere mixture and the smoothed densities are defined [106] by analogy with (61):

$$\bar{\rho}_i(\mathbf{r}) = \int d\mathbf{r}'\ w_i(|\mathbf{r} - \mathbf{r}'|;\bar{\rho}_1(\mathbf{r}),\bar{\rho}_2(\mathbf{r}))\rho_i(\mathbf{r}') \qquad i = 1, 2 \qquad (68)$$

The *two* weight functions w_1 and w_2 are assumed to have the form

$$w_i(r;\bar{\rho}_1,\bar{\rho}_2) = w_{i0}(r) + \sum_{j=1}^{2} w_{ij}(r)\bar{\rho}_j \qquad i = 1, 2 \qquad (69)$$

with w_{10}, and so on, density-independent coefficients. These are determined by requiring the density expansions of the direct correlation functions of the homogeneous hard-sphere mixture, as obtained by differentiating (67) (with the attractive pair potentials $\phi_{att}^{ij} = 0$) to be close to those obtained from the PY approximation. Since the expansion (69) is truncated at the term linear in density, this prescription does not yield an accurate fit to the PY results at liquid densities [106]. Nevertheless, comparison with Monte Carlo simulations shows that the resulting theory gives a good account of the density profiles of hard-sphere mixtures near a hard wall for size ratios $\sigma_2/\sigma_1 \lesssim 3$ [106].

3. *Curtin–Ashcroft Recipe*

Despite the indisputable success of Tarazona Mark II for a vast range of problems, the purist must regard the scheme as utilitarian. (It was once

described, by an eminent theorist, as a chemical engineer's presciption. Whether this constitutes a compliment or an insult is not obvious.) The density expansion (62) of the weight function is a little inelegant for a pure fluid and any implementation becomes problematical for a mixture. The WDA introduced by Curtin and Ashcroft (CA) [107] avoids this expansion at the expense of increased computational complexity. Their strategy is the same as that of the Tarazona Mark II in that the excess free-energy functional is given by (58), the weighted density by (61), and w is determined by requiring the uniform fluid limit of $-\beta\delta^2\mathcal{F}_{ex}[\rho]/\delta\rho(\mathbf{r}_1)\,\delta\rho(\mathbf{r}_2)$ to generate known $c^{(2)}(\rho;r_{12})$ at all fluid densities. Thus

$$-\beta c^{(1)}([\rho];\mathbf{r}_1) = \frac{\delta\mathcal{F}_{ex}[\rho]}{\delta\rho(\mathbf{r}_1)} = \int d\mathbf{r}_2\, \rho(\mathbf{r}_2)\psi'_{ex}(\bar{\rho}(\mathbf{r}_2))\frac{\delta\bar{\rho}(\mathbf{r}_2)}{\delta\rho(\mathbf{r}_1)} + \psi_{ex}(\bar{\rho}(\mathbf{r}_1))$$

with

$$\frac{\delta\bar{\rho}(\mathbf{r}_2)}{\delta\rho(\mathbf{r}_1)} = \frac{w(\mathbf{r}_2 - \mathbf{r}_1;\bar{\rho}(\mathbf{r}_2))}{1 - \int d\mathbf{r}_3\, w'(\mathbf{r}_2 - \mathbf{r}_3;\bar{\rho}(\mathbf{r}_2))\rho(\mathbf{r}_3)}$$

Note that in the uniform limit the integral in the denominator vanishes by virtue of the normalization requirement on w. A second differentiation yields

$$-\beta^{-1}c^{(2)}(\rho;r_{12}) = 2\psi'_{ex}(\rho)w(r_{12};\rho) + \rho\psi''_{ex}(\rho)\int d\mathbf{r}_3\, w(r_{13};\rho)w(r_{32};\rho) + \rho\psi'_{ex}(\rho)\int d\mathbf{r}_3\,[w'(r_{13};\rho)w(r_{32};\rho) + w(r_{13};\rho)w'(r_{32};\rho)]$$

in the uniform limit. The uniform fluid direct correlation functions, along with ψ_{ex}, constitute the input for the theory. It follows that the Fourier transform of the weight function $w_k(\rho) \equiv \int d\mathbf{r}\, \exp(i\mathbf{k}\cdot\mathbf{r})w(r;\rho)$ satisfies the nonlinear differential equation

$$-\beta^{-1}c_k^{(2)}(\rho) = 2\psi'_{ex}(\rho)w_k(\rho) + \rho\frac{\partial}{\partial\rho}[\psi'_{ex}(\rho)w_k^2(\rho)] \tag{70}$$

where $c_k^{(2)}(\rho)$ is the Fourier transform of $c^{(2)}(\rho;r)$ and the prime denotes differentiation with respect to density ρ. Since $w(r;\rho)$ must be normalized $w_{k=0}(\rho) = 1$ and (70) is consistent with the compressibility sum rule (24). For the hard-sphere fluid, $c_k^{(2)}(\rho)$ and $\psi_{ex}(\rho)$ are given analytically within the PY approximation and (70) can be solved (numerically) to determine $w(r;\rho)$ for all ρ, without recourse to any density expansion.

This version of the WDA was applied successfully to hard-sphere freezing [107] and later to the determination of the structure and surface free energy of the hard-sphere crystal–liquid interface [108,109]. The para-

metrization (64) is employed for the bulk crystal density, while a flexible, two-parameter scheme is used to describe $\rho(\mathbf{r})$ at the interface; this allows for variations of the width of the interface and of the rate of broadening of the peaks in density through the interface region. Curtin's procedure improves upon earlier square-gradient treatments of this interfacial problem [109]. Kroll and Laird [110] have employed the Curtin–Ashcroft WDA for hard spheres at a hard wall and the resulting density profiles are very similar to those from Tarazona Mark II [94].

CA argue that their WDA contains all information obtainable from the two-point function $c^{(2)}$ of the uniform fluid and emphasize [111] that the $c^{(n)}_{\mathrm{WDA}}$ satisfy the exact sum rule

$$\int d\mathbf{r}_n \, c^{(n)}(\rho;\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_n) = \frac{\partial c^{(n-1)}}{\partial \rho}(\rho;\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_{n-1}) \quad (71)$$

appropriate to higher-order direct correlation functions of the uniform fluid, implying that a subset of higher-order correlations are treated correctly in the WDA. However, requiring an approximation to satisfy (71) is not particularly stringent. Equation (71) follows as a direct consequence of the definition of the hierarchy of direct correlation functions [see the derivation of (24a), the first member of (71)]. Thus we might expect that *any* free-energy functional approach that does not correspond to a truncated density expansion should satisfy (71) automatically. The higher-order $c^{(n)}$ need not be especially realistic in order that (71) be obeyed (see Section IV.A).

In their original treatment CA [107] presented their theory in terms of the full free-energy functional rather than the hard-sphere part $\mathcal{F}_{\mathrm{hs}}[\rho]$. However, it is clear that making the WDA on the full functional (including $\mathcal{F}_{\mathrm{att}}[\rho]$) would require some prescription for $c_k^{(2)}(\rho)$ and $\psi_{\mathrm{ex}}(\rho)$ for densities ρ corresponding to two phase states of the bulk phase diagram—a difficulty that we have encountered for other approximate theories. In a later paper [112] CA calculated the bulk freezing properties and the liquid–gas coexistence curve for a Lennard–Jones system, treating repulsive forces by means of the WDA for $\mathcal{F}_{\mathrm{hs}}[\rho]$ and attractive forces by means of a perturbation theory based on a hard-sphere reference fluid rather than the simple mean-field approximation given in (53). Their approximation reduces to the standard hard-sphere perturbation theory result for the free energy in the limit of a uniform fluid. The resulting temperature–density diagram is in remarkably good agreement with simulation [112]. Curtin [109] applied the theory to the crystal–liquid interface of a Lennard–Jones system using a further simplification for $\mathcal{F}_{\mathrm{att}}[\rho]$.

Denton and Ashcroft [113] have proposed an extension of the CA WDA

to mixtures. Their starting approximation for $\mathcal{F}_{ex}[\rho_1,\rho_2]$ is the same as that of Tan et al. [106] [see (67)] but *three* independent weight functions $\bar{w}_{ij}$ are involved, rather than two, so that

$$\bar{\rho}_i(\mathbf{r}) = \sum_{j=1}^{2} \int d\mathbf{r}'\, \bar{w}_{ij}(|\,\mathbf{r} - \mathbf{r}'\,|;\bar{\rho}_1(\mathbf{r}),\bar{\rho}_2(\mathbf{r}))\rho_j(\mathbf{r}') \qquad i = 1, 2 \quad (72)$$

The $\bar{w}_{ij}$ are specified by requiring the approximate $\mathcal{F}_{ex}$ to generate the known direct correlation functions of a uniform mixture $c_{ij}^{(2)}(\rho_1,\rho_2;r)$. The appropriate generalization of (70) is a set of coupled nonlinear differential equations for the Fourier components of $\bar{w}_{ij}$ in terms of those of $c_{ij}^{(2)}$. In principle, these could be solved for a binary hard-sphere mixture using the PY results as input, but this was not attempted in Ref. 113. Rather, a modified version of the WDA (see Section III.F.1) that uses position-independent weighted densities was developed and applied to the freezing of hard-sphere mixtures.

4. *Recipes of Meister and Kroll and of Groot and van der Eerden*

An alternative method of construcing a WDA was introduced by Meister and Kroll (MK) [114] based on the idea that $\mathcal{F}_{ex}$ could be expressed as a functional of some slowly varying reference density $\rho_0(\mathbf{r})$, as well as of $\rho(\mathbf{r})$. By minimizing the dependence on $\rho_0(\mathbf{r})$ an expression for $\rho_0(\mathbf{r})$ can be determined that has the same *form* as (61) for a weighted density.

MK's starting point is the exact result (22), which can be reexpressed as

$$\beta\mathcal{F}_{ex}[\rho] = -\int_0^1 d\alpha \int d\mathbf{r}\, \rho(\mathbf{r})c^{(1)}([\alpha\rho];\mathbf{r}) \quad (73)$$

by setting $\rho_i \equiv 0$. They then expand $c^{(1)}$ about a uniform reference system at a (coarse-grained) density $\rho_0(\mathbf{r})$:

$$c^{(1)}([\alpha\rho];\mathbf{r}) = c^{(1)}(\alpha\rho_0(\mathbf{r})) + \sum_{n=1}^{\infty} \frac{1}{n!} \int d\mathbf{r}_1 \ldots d\mathbf{r}_n\, c^{(n+1)}(\alpha\rho_0(\mathbf{r});\mathbf{r},\mathbf{r}_1,\ldots,\mathbf{r}_n) \times \prod_{k=1}^{n} [\alpha\rho(\mathbf{r}_k) - \alpha\rho_0(\mathbf{r})] \quad (74)$$

where $c^{(1)}(\rho) \equiv -\beta\mu_{ex}(\rho)$ and the $c^{(n+1)}$ refer to the direct correlation functions of the uniform fluid. If (74) is truncated at the $n = 1$ term (73)

reduces to

$$\mathcal{F}_{ex}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\psi_{ex}(\rho_0(\mathbf{r})) - \beta^{-1} \int d\mathbf{r} \int d\mathbf{r}'\, \rho(\mathbf{r}) L(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |)(\rho(\mathbf{r}') - \rho_0(\mathbf{r})) \tag{75}$$

where

$$L(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |) = \int_0^1 d\alpha\, \alpha c^{(2)}(\alpha\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |) \tag{76}$$

In order to specify the reference density MK extremize the functional $\Omega_V[\rho, \rho_0]$ with respect to $\rho_0(\mathbf{r})$. From the requirement

$$\frac{\delta \Omega_V}{\delta \rho_0(\mathbf{r})} = \frac{\delta \mathcal{F}_{ex}}{\delta \rho_0(\mathbf{r})} = 0 \tag{77}$$

it follows that

$$\rho(\mathbf{r}) \left[\beta\psi'_{ex}(\rho_0(\mathbf{r})) + \int d\mathbf{r}'\, L(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |) - \int d\mathbf{r}'\, L'(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |)(\rho(\mathbf{r}') - \rho_0(\mathbf{r})) \right] = 0 \tag{78}$$

By virtue of the compressibility sum rule (24), along with the definition (76), the sum of the first two terms vanishes for all ρ_0 so that (78) implies that

$$\rho_0(\mathbf{r}) = \frac{\int d\mathbf{r}'\, L'(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |)\rho(\mathbf{r}')}{\int d\mathbf{r}'\, L'(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |)} \tag{79}$$

which is of the desired form (61) with a weight function

$$w(|\, \mathbf{r} - \mathbf{r}'\, |; \rho_0(\mathbf{r})) \equiv \frac{L'(\rho_0(\mathbf{r}); |\, \mathbf{r} - \mathbf{r}'\, |)}{L'_0(\rho_0(\mathbf{r}))} \tag{80}$$

where $L'_0(\rho) \equiv \int d\mathbf{r}\, L'(\rho; r)$. Clearly, w is normalized.

This MK recipe is certainly appealing. It appears to avoid the rather ad hoc nature of previous recipes. Using (79) in (75) one has a functional of $\rho(\mathbf{r})$ only, which can be differentiated in the usual way to determine the equilibrium density profile and the thermodynamics. MK used this recipe in successful investigations of drying a hard wall and of a first-order wetting transition at a Lennard–Jones 9–3 wall. Only $\mathcal{F}_{hs}$ was treated within their WDA, and L and ψ_{ex} were calculated from PY approximation. The contribution from attractive forces $\mathcal{F}_{att}$ was treated in

mean-field approximation. MK's results [114] indicated that including short-range correlations, via their WDA, significantly reduce the temperature of the wetting transition below the value obtained when using the LDA (50) for $\mathcal{F}_{hs}$. This important result was verified by several subsequent calculations using other versions of WDA [72,88,95,97].

Groot [115]* pointed out an important deficiency of the MK recipe. $\mathcal{F}_{ex}[\rho]$ resulting from (75), with (79), does *not* satisfy $-\beta\delta^2\mathcal{F}_{ex}/\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2) = c^{(2)}(\rho;|\mathbf{r}_1 - \mathbf{r}_2|)$ in the limit of a uniform fluid. [This is a consequence of the truncation of (74) at the $c^{(2)}$ term.] The $c^{(2)}$ generated by differentiation can be quite different from the bulk $c^{(2)}$, which is the input for the theory. Groot and van der Eerden (GvE) [118] modified the theory to remedy this deficiency. They consider a functional $\mathcal{F}_{ex}[\rho,\rho_0]$ that is identical in *form* to (75) but allow L to be an arbitrary two-point function, *not* (76) as given by truncation. It is assumed that the form of (75) remains a good approximation when higher-order ($n > 1$) contributions in (74) are included. Applying the extremum condition (77) yields (78) once more. Requiring the same condition to hold for a uniform fluid implies that

$$L_0(\rho) \equiv \int d\mathbf{r}\, L(\rho;r) = -\beta\psi'_{ex}(\rho) \tag{81}$$

The same prescription (79), with (80), follows for $\rho_0(\mathbf{r})$. However, L is now determined by requiring consistency between $c^{(2)}$ generated from $\mathcal{F}_{ex}$ and the uniform fluid result. This requirement leads to the following differential equation for the Fourier transform $L_k(\rho)$:

$$c_k^{(2)}(\rho) = 2L_k(\rho) + \frac{\rho}{L'_0(\rho)}(L'_k(\rho))^2 \tag{82}$$

[Note the similarity with the corresponding CA equation (70).] In the limit $k = 0$, (82), with (81), is identical to the compressibility sum rule (24). GvE describe the implementation of their theory for hard spheres at a hard wall, using as input $c^{(2)}(\rho;r)$ obtained from simulation. Kroll and Laird [110] investigated the same model system using PY results as input. They found (see Fig. 3) that the density profiles from the GvE recipe are almost identical to those from the CA recipe; both are markedly better than MK for large bulk densities. They conclude that the requirement

* Groot also discusses the extension of MK theory to mixtures. Sokolowski and Fischer [116] report results of MK theory applied to binary mixtures of hard spheres at a hard wall. The results are in good agreement with simulation, even for large size ratios (see Ref. 106). Sokolowski and Fischer [117] describe applications of the same theory to adsorption of binary Lennard–Jones mixtures in narrow pores.

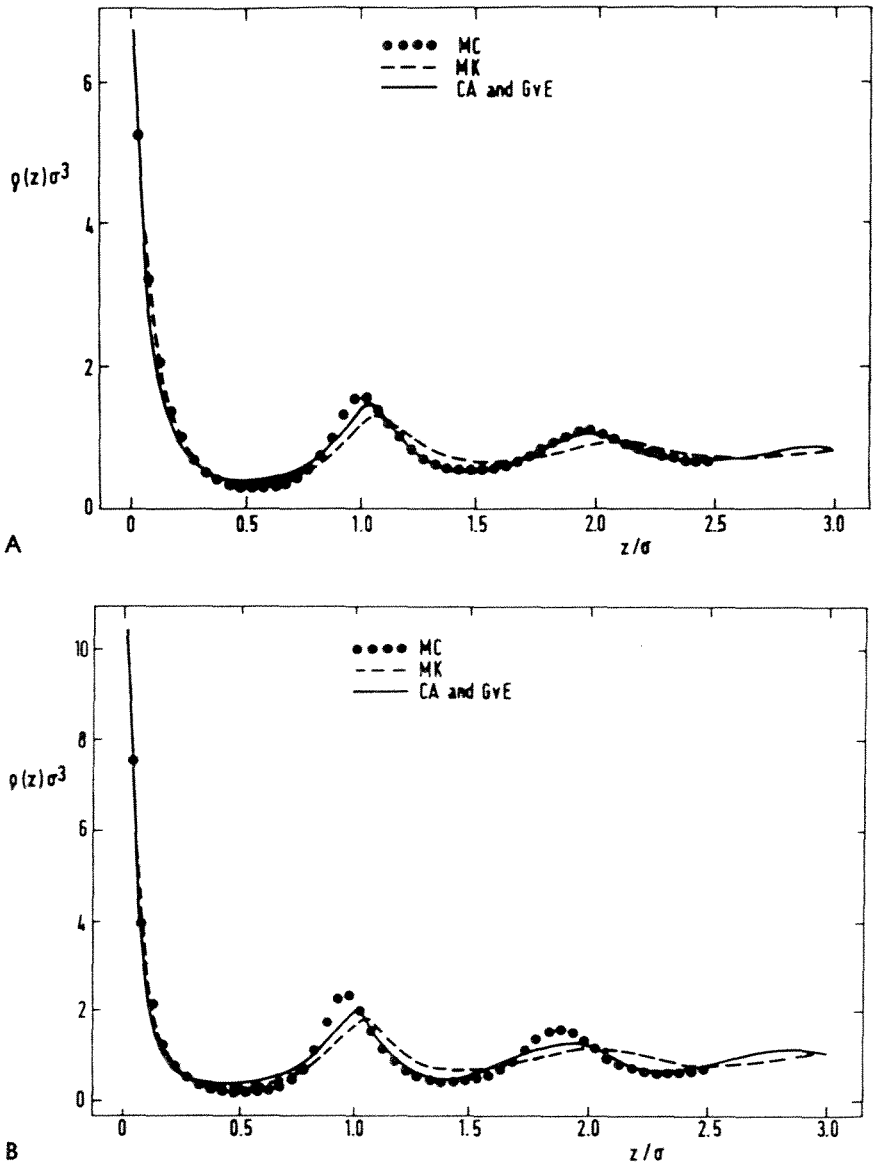


Fig. 3 Density profiles for hard spheres near a single hard wall: (A) bulk density $\rho_b\sigma^3 = 0.813$; (B) $\rho_b\sigma^3 = 0.9135$. The points are the Monte Carlo results, the dashed lines are the results of the original MK theory, and the solid lines are those of the CA and GVE recipes; the differences between the results of the latter theories cannot be resolved in the figure. All the calculations make use of the PY compressibility equation of state and direct correlation function, so that while the contact density satisfies the sum rule $\beta P = \rho(0^+)$ for all three theories, $\rho(0^+)$ is different from the simulation result. (Redrawn from Ref. 110.)

that $\mathcal{F}_{ex}$ yield the correct $c^{(2)}$ for a uniform fluid is a crucial ingredient of any theory.

There is a close relationship between the CA and GvE recipes. The former can be derived [110] by demanding that the final term in (75) vanish so that $\mathcal{F}_{ex}$ reduces to the purely "local" form $\int d\mathbf{r}\ \rho(\mathbf{r})\psi_{ex}(\rho_0(\mathbf{r}))$. This condition then implies that

$$\rho_0(\mathbf{r}) = \frac{\int d\mathbf{r}'\ L(\rho_0(\mathbf{r});|\ \mathbf{r} - \mathbf{r}'\ |)\rho(\mathbf{r}')}{\int d\mathbf{r}'\ L(\rho_0(\mathbf{r});|\ \mathbf{r} - \mathbf{r}'\ |)} \tag{83}$$

which is exactly of the form $\rho_0(\mathbf{r}) = \int d\mathbf{r}'\ w(\ |\ \mathbf{r} - \mathbf{r}'\ |; \rho_0(\mathbf{r}))\rho(\mathbf{r}')$ assumed in the CA or Tarazona (Mark II) recipes. By requiring the resulting $\mathcal{F}_{ex}$ to generate $c^{(2)}$ in the uniform fluid we recover the CA recipe. Which of the two is to be preferred? This last derivation suggests that CA corresponds to choosing the reference density $\rho_0(\mathbf{r})$ [$\equiv \bar{\rho}(\mathbf{r})$] so that all contributions to $\mathcal{F}_{ex}$, other than the local one, vanish, whereas GvE aims to minimize the dependence of $\mathcal{F}_{ex}$ on $\rho_0(\mathbf{r})$. [Both theories assume that (75), with arbitrary L, is an accurate approximation to the exact functional.] We speculate that the CA procedure might reduce the errors associated with approximating higher-order terms. One disturbing feature of GvE and of the original MK, which does not appear to have been pointed out in the literature [this was noticed by P. Tarazona (private communication, 1986) for the MK theory], is that in the limit of a uniform fluid the second derivative $\delta^2\mathcal{F}_{ex}/\delta\rho_0(\mathbf{r})\ \delta\rho_0(\mathbf{r}')$ reduces to $-\beta\rho\psi''_{ex}(\rho)\ \delta(\mathbf{r} - \mathbf{r}')$, which is *negative* for the case of a hard-sphere fluid, since $\psi''_{ex}(\rho)$ is positive for all fluid densities. This suggests that the extremal condition (77) actually determines a maximum; any perturbation $\delta\rho_0(\mathbf{r})$ will lower the free energy $\mathcal{F}_{ex}$ from its value at the extremum. Kroll and Laird [110] have found another rather surprising feature. They state that when GvE is reformulated for a position-independent weighted density the equation for the weighted density [analogous to (79)] has no solution for a crystalline phase. In contrast, the position-independent version of CA is known to yield accurate results for hard-sphere freezing [119].

Sokolowski and Fischer [120] have extended the MK recipe by including three-particle correlations. The functional (75) is supplemented by the term

$$-(2\beta)^{-1}\int d\mathbf{r}\ \rho(\mathbf{r})\iint d\mathbf{r}'\ d\mathbf{r}''\int_0^1 d\alpha\ \alpha^2 c^{(3)}(\alpha\rho_0(\mathbf{r});\mathbf{r},\mathbf{r}',\mathbf{r}'')(\rho(\mathbf{r}') - \rho_0(\mathbf{r}))(\rho(\mathbf{r}'') - \rho_0(\mathbf{r}))$$

and the triplet direct correlation function $c^{(3)}$ is approximated according

to the factorization scheme of Barrat et al. [121]. By including such a term, $\mathcal{F}_{ex}$ does generate the correct $c^{(2)}$ in the uniform limit; inconsistencies develop only at next order. The density profiles calculated for hard spheres at a hard wall and at a Lennard–Jones 9–3 wall represent a marked improvement [120] on the MK results, but for the hard wall they appear to be slightly less accurate than those from CA or GvE [110]. This procedure can be regarded as a generalization of the third-order density expansion used by Rickayzen and co-workers [68–70] (see Section III.C).

5. *Recipes of Percus and of Robledo and Varea*

We return now to the question raised in Section II.C as to whether the exact results for hard rods in one dimension can prove useful in higher dimensions. Percus [36,41,122] writes the excess free-energy functional of the hard-rod fluid (36) in the form

$$\mathcal{F}_{ex}[\rho] = \int dx\, \rho_\sigma(x)\psi_{ex}(\rho_\tau(x)) \tag{84}$$

where $\rho_\tau(x)$ is the average defined in (38) and

$$\rho_0(x) \equiv \frac{1}{2}\left[\rho\left(x + \frac{\sigma}{2}\right) + \rho\left(x - \frac{\sigma}{2}\right)\right] \tag{85}$$

is a second average of the density profile $\rho(x)$. [Note the resemblance between (84) and (58), the usual starting point for a WDA.] ρ_τ and ρ_σ are linear volume and surface averages, respectively. A natural generalization of (84) to a three-dimensional fluid is

$$\mathcal{F}_{ex}[\rho] = \int d\mathbf{r}\, \rho_\sigma(\mathbf{r})\psi_{ex}(\rho_\tau(\mathbf{r})) \tag{86}$$

where ψ_{ex} is again the excess free energy per atom and

$$\begin{aligned} \rho_\tau(\mathbf{r}) &= \int d\mathbf{r}'\, \tau(\mathbf{r} - \mathbf{r}')\rho(\mathbf{r}') \\ \rho_\sigma(\mathbf{r}) &= \int d\mathbf{r}'\, \bar{\sigma}(\mathbf{r} - \mathbf{r}')\rho(\mathbf{r}') \end{aligned} \tag{87}$$

are suitable averages. The weight functions τ and $\bar{\sigma}$ are assumed not to depend on the average densities. Percus [36] states that (86) "is an outrageous extrapolation based on minimal information" and suggests [41] that τ and $\bar{\sigma}$ can be obtained by requiring $\mathcal{F}_{ex}$ to generate known $c^{(2)}(\rho; r)$ in the uniform limit. For the hard-rod case the exact weight functions are $\tau(x) = \Theta(\sigma/2 - |x|)/\sigma$ and $\bar{\sigma}(x) = \delta(\sigma/2 - |x|)/2$. Robledo and Varea

[35] introduced direct generalizations for a hard-sphere fluid:

$$\tau(\mathbf{r}) = \frac{\Theta(\sigma/2 - |\mathbf{r}|)}{\pi\sigma^3/6}$$
$$\tilde{\sigma}(\mathbf{r}) = \frac{\delta(\sigma/2 - |\mathbf{r}|)}{\pi\sigma^2} \tag{88}$$

Both are normalized, as is required by (87) in the uniform limit. Significantly, the range of τ is half that of w_0, the weight function introduced by Nordholm et al. (see Section III.E.1). By substituting (88) into (86) and differentiating, Robledo and Varea showed that the resulting $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ has the range of the hard-core interaction [i.e., $c^{(2)}$ vanishes when $|\mathbf{r}_1 - \mathbf{r}_2| > \sigma$]. In the limit of a uniform fluid they obtained a cubic in r (with no r^2 term) for $c^{(2)}(\rho;r)$ inside the core. This form is the same as that obtained from PY or scaled particle theory for hard spheres. In contrast, the Nordholm et al. or Tarazona (Mark I) recipe yields a substantial (unrealistic) nonvanishing tail in $c^{(2)}(\rho;r)$ for $\sigma < r < 2\sigma$ [91], reflecting the greater range of the weight function. Fischer and Heinbuch [123] compare $c^{(2)}(\rho;r)$ as obtained from the Tarazona (Mark I) and the Robledo–Varea recipes using the same Carnahan–Starling ψ_{ex} as input. As expected, both approximations give a very poor representation of the exact result. Note that Tarazona (Mark I) and all subsequent WDAs are equivalent to setting $\tilde{\sigma}(\mathbf{r}) = \delta(\mathbf{r})$ (i.e., the surface average has zero range). The interesting paper of Fischer and Heinbuch [123] also points out the close connections between WDA and theories, such as that of Fischer and Methfessel [124], based on approximations to the Born–Green–Yvon equation for the density profile. In one dimension the exact relationship has been established [125].

Vanderlick et al. [103] compared the results of the Robledo–Varea recipe and those of the Fischer–Methfessel approximation with Monte Carlo results for a hard-sphere fluid confined between two hard walls. Unlike those from Tarazona (Mark II) (see Fig. 2), the density profiles from both approximation schemes are in very poor agreement with simulation. From Vanderlick et al.'s valuable study we conclude, in keeping with our earlier statements, that for a WDA to give an accurate description of the density profile of a highly inhomogeneous fluid it must generate a fairly accurate $c^{(2)}(\rho;r)$ for uniform fluids over a wide range of densities ρ. This, in turn, seems to demand that the weight function be density dependent.

6. *Recipes of Rosenfeld and of Kierlik and Rosinberg*

A different version of WDA was introduced by Rosenfeld [126] specifically for hard-sphere mixtures. Rosenfeld's free-energy functional is mo-

tivated by his earlier work [127] on uniform hard-particle fluids, where he argues that the pair direct correlation functions $c_{ij}^{(2)}$, in any number of dimensions, should be dominated by terms corresponding to the overlap volume ΔV_{ij} and overlap surface area ΔS_{ij} of two individual spheres i and j—quantities arising from convolutions of functions that describe the volume and surface of the spheres. His results [126,127] in three dimensions can be expressed (in his notation) as

$$-c_{ij}^{(2)}(r) = \chi^{(3)}\Delta V_{ij}(r) + \chi^{(2)}\Delta S_{ij}(r) + \chi^{(1)}\Delta R_{ij}(r) + \chi^{(0)}\Theta(R_i + R_j - r) \quad (89)$$

where R_i ($\equiv \sigma_i/2$) is the *radius* of the sphere of species i, Θ the Heaviside function, ΔR_{ij} a further geometrical measure, and the coefficients $\chi^{(k)}$ depend only on the four scaled particle [128] variables:

$$\xi^{(k)} = \sum_{i=1}^{\nu} \rho_i R_i^{(k)} \qquad k = 0, 1, 2, 3 \quad (90)$$

with $R_i^{(0)} = 1$, $R_i^{(1)} = R_i$, $R_i^{(2)} = 4\pi R_i^2$, and $R_i^{(3)} = 4\pi R_i^3/3$ in the case of hard spheres. ν is the number of species in the mixture. In a remarkable paper Rosenfeld derives (89) via a new graphical expansion of the direct correlation functions involving overlap volumes as basic functions and pair excluded volumes as basic variables [127]. His lowest-order "scaled field particle approximation" takes the form (89) for hard spheres and Rosenfeld shows that (89) is equivalent to the standard PY result for this case. Subsequently, he postulated [126] an excess free-energy functional $\mathcal{F}_{ex}$, motivated by the exact low-density expression (incorporating pair exclusion) that generates (89) in the uniform limit. We refer the reader to the original articles [126,129] for details of Rosenfeld's derivation, preferring to follow the subsequent, but somewhat more transparent arguments of Kierlik and Rosinberg [130].

The special structure of (89) and the experience gained from earlier sections (especially Sections III.C and III.E.5) suggest the form

$$\beta\mathcal{F}_{ex}[\{\rho_i(\mathbf{r})\}] = \int d\mathbf{r}\, \Phi(\{n_\alpha(\mathbf{r})\}) \quad (91)$$

where Φ is a function of linear averages (the weighted densities)

$$n_\alpha(\mathbf{r}) = \sum_{i=1}^{\nu} \int d\mathbf{r}'\, \rho_i(\mathbf{r}')\omega_i^{(\alpha)}(\mathbf{r} - \mathbf{r}') \quad (92)$$

and $\omega_i^{(\alpha)}(\mathbf{r})$, with $\alpha = 1, 2, \ldots, m$, are unknown but density-independent weight functions. (Percus [122] argued that the latter, and hence the weighted densities n_α, could be scalars or vectors.) Inserting (92) in (91)

and differentiating, we find that

$$c^{(n)}_{i_1 i_2 \ldots i_n}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{-\beta\delta^n \mathcal{F}_{ex}[\{\rho_i(\mathbf{r})\}]}{\delta\rho_{i_1}(\mathbf{r}_1)\delta\rho_{i_2}(\mathbf{r}_2)\cdots\delta\rho_{i_n}(\mathbf{r}_n)}$$
$$= -\int d\mathbf{r} \sum_{\alpha_1 \ldots \alpha_n} \frac{\partial^n \Phi}{\partial n_{\alpha_1} \partial n_{\alpha_2} \cdots \partial n_{\alpha_n}} \times \omega^{(\alpha_1)}_{i_1}(\mathbf{r}_1 - \mathbf{r})\omega^{(\alpha_2)}_{i_2}(\mathbf{r}_2 - \mathbf{r}) \cdots \omega^{(\alpha_n)}_{i_n}(\mathbf{r}_n - \mathbf{r}) \tag{93}$$

In particular, the Fourier transform of the pair function ($n = 2$) in the uniform fluid reduces to

$$c^{(2)}_{ij}(k) = -\sum_{\alpha,\beta} \frac{\partial^2 \Phi}{\partial n_\alpha \, \partial n_\beta} \omega_i^{(\alpha)}(k)\omega_j^{(\beta)}(k) \tag{94}$$

since the derivative is position independent in this limit. Equation (94) has the same form as the Fourier transform of (89). Of course, it remains to determine the weight functions and the excess free-energy density $\beta^{-1}\Phi[\{n_\alpha\}]$. Unlike Rosenfeld [126,129], who argues these quantities are, in some sense, uniquely predetermined by geometric considerations, Kierlik and Rosinberg [130] make a more heuristic, and probably more reasonable, specification. For the first four weight functions they require the normalization

$$\omega_i^{(\alpha)}(k = 0) \equiv \int d\mathbf{r} \; \omega_i^{(\alpha)}(r) = R_i^{(\alpha)} \qquad \alpha = 0, 1, 2, 3 \tag{95}$$

so that $n_\alpha(\mathbf{r}) \to \xi^{(\alpha)}$ ($\alpha = 0, 1, 2, 3$) in the limit of a uniform fluid. For the remaining weight functions they require that

$$\omega_i^{(\alpha)}(k = 0) = 0 \qquad \alpha > 3 \tag{96}$$

so that the n_α ($\alpha > 3$) vanish in the uniform limit. But in this limit Φ should be identical to the PY (compressibility) result, which expressed in terms of scaled particle variables is

$$\Phi(n_0, n_1, n_2, n_3) = \Phi_{PY} = -n_0 \ln(1 - n_3) + \frac{n_1 n_2}{1 - n_3} + \frac{{n_2}^3}{24\pi(1 - n_3)^2} \tag{97}$$

where $n_\alpha \equiv \xi^{(\alpha)}$. [Note that $n_3 = \sum_{i=1}^{\nu} \rho_i(4\pi/3)R_i^3$ is the packing fraction.] By imposing a natural scaling condition,

$$\frac{\omega_i^{(\alpha)}(k)}{\omega_i^{(\alpha)}(k = 0)} = \omega^{(\alpha)}(t_i) \tag{98}$$

with $t_i \equiv kR_i$, Kierlik and Rosinberg [130] show that four independent scalar functions $\omega^{(\alpha)}(t)$ are sufficient to recover (89) from (94) and they obtain unique solutions

$$\begin{aligned}
\omega_i^{(3)}(r) &= \Theta(R_i - r) \\
\omega_i^{(2)}(r) &= \delta(R_i - r) \\
\omega_i^{(1)}(r) &= \frac{1}{8\pi}\delta'(R_i - r) \\
\omega_i^{(0)}(r) &= -\frac{1}{8\pi}\delta''(R_i - r) + \frac{1}{2\pi r}\delta'(R_i - r)
\end{aligned} \tag{99}$$

No vector weight functions are introduced. In contrast, Rosenfeld [126,129] employs the first two of (99) but introduces vector functions to avoid the derivatives of delta functions that enter the third and fourth equations. However, the method by which he introduces a vector part of Φ is not completely clear [130]. There does not appear to be any fundamental objection to the appearance of derivatives of the delta function since these always appear inside integrals so that the weighted densities $n_\alpha(\mathbf{r})$ are not pathological [130]. In one dimension Rosenfeld et al. [129] show that their approach reproduces the exact hard-rod free-energy functional of Percus [i.e., (36)]. The relevant weight functions then reduce to those described in Section III.E.5 for hard rods.

These approaches are certainly appealing. They constitute a fundamental improvement on the recipes of Section III.E.5 in that they guarantee (by construction) that the PY direct correlation functions are reproduced in the uniform fluid. Morevoer, since they are specifically designed for hard-sphere *mixtures* that automatically avoid the difficulties encountered in extending some earlier recipes to multicomponent fluids (see Section III.E.2). Rosenfeld's approach makes optimal use of geometric considerations. Indeed, his determination of the weight functions is based purely on a decomposition of the pair exclusion step function. The simplified version of Kierlik and Rosinberg, which is removed slightly further from purely geometrical considerations, requires only four scalar $n_\alpha(\mathbf{r})$, irrespective of the number of species ν. These quantities can be regarded [130] as the generalization to nonuniform fluids of the four independent variables of scaled particle theory [e.g., $n_3(\mathbf{r})$ is the *local* packing fraction]. Unlike the uniform case, however, some of the $n_\alpha(\mathbf{r})$ can, and in practice [130] do, take on negative values, so that their physical interpretation is no longer obvious. By introducing four (density-independent) weight functions the recipe would seem to contain essentially the same ingredients as those of Tarazona (Mark II) and CA, who allow their single weight function to be density dependent. Whether the recipe

is better founded than these earlier WDAs is a matter for debate. Rosenfeld et al. [129] argue that their version is systematic because it has its origins in a diagramatic description which can, in principle, provide a systematic means of improving on the weight functions.

How well do the existing recipes fare? Kierlik and Rosinberg [130] applied their version to the canonical problem of hard spheres near a hard wall. The resulting density profile (see Fig. 4) for a high-density bulk ($\rho_b\sigma^3 = 0.878$) is very close to the simulation results; the oscillations coincide almost exactly—something not achieved by other approximate theories at this density. The only significant disagreement is very close to contact. As with other WDAs the theory satisfies the sum rule $\rho(0^+) = \beta p$, but the pressure p is that corresponding to the PY compressibility equation rather than the exact, or Carnahan–Starling, result. Adsorption at a Lennard–Jones 9–3 wall is also considered in [130]. As shown in Fig. 5a, the theory does remarkably well for a very strongly adsorbing wall. It only begins to fail (Fig. 5b) when the temperature is low *and* the bulk takes on a liquidlike density ($\rho_b\sigma^3 = 0.611$). Then the theory gives a poor description of the first peak of $\rho(z)$ (corresponding to the first adsorbed layer). In this region the local packing fraction $n_3(z) \sim 0.65$, so one cannot expect the PY theory for Φ or for $c_{ij}^{(2)}(\rho;r)$ to be accurate. Moreover, $n_0(z)$

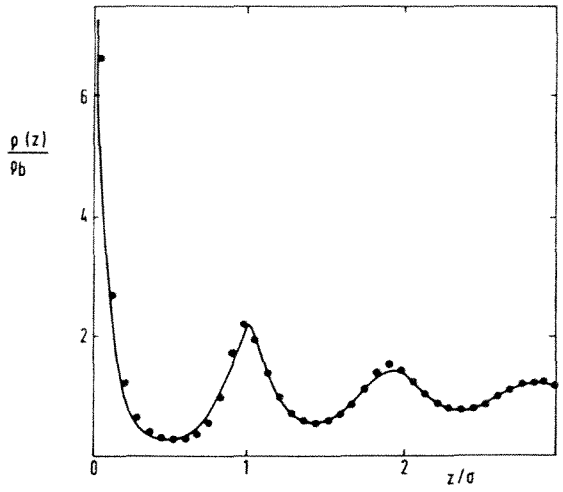


Fig. 4 Density profile of hard spheres near a single hard wall at bulk density $\rho_b\sigma^3 = 0.878$. The points are Monte Carlo results and the solid line denotes the density functional results from the Kierlik–Rosinberg recipe. (Redrawn from Ref. 130.)

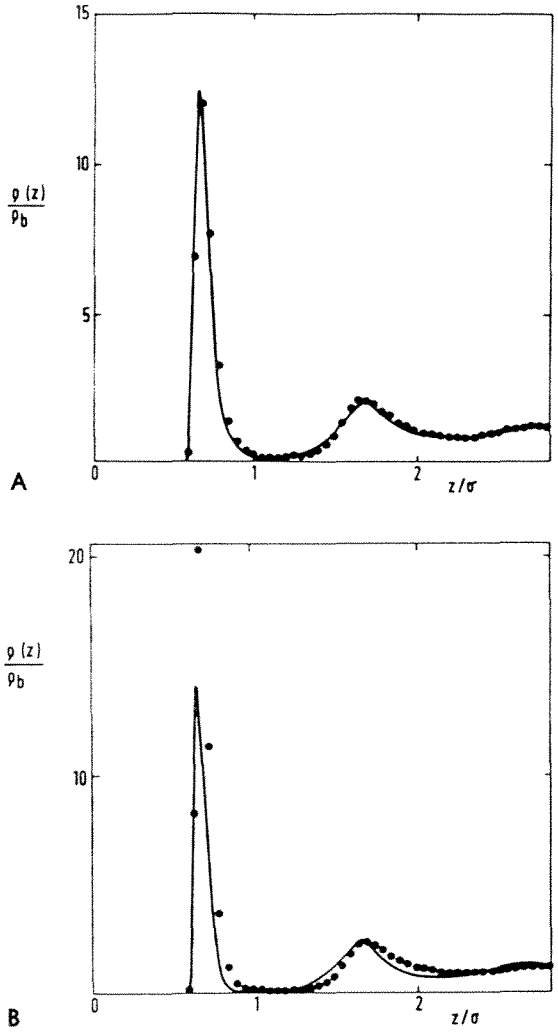


Fig. 5 Density profile of hard spheres at a Lennard–Jones 9–3 wall with ϵ/k_B = 2876 K and $z_0 = 0.562\sigma$, mimicking argon on graphite: (A) bulk density $\rho_b\sigma^3$ = 0.467 and temperature T = 150 K; (B) $\rho_b\sigma^3 = 0.611$ and T = 100 K. The points are the Monte Carlo results of Ref. 120 and the solid lines are the results of Kierlik and Rosinberg. (Redrawn from Ref. 130.)

takes on negative values in the same region. By contrast, the density functional results of Sokolowski and Fischer [120] (see Section III.E.4) for the same model remain rather accurate for the high-density state of Fig. 5b.

Rosenfeld [126,129] and Kierlik and Rosinberg [130] apply their recipes to the calculation of hard-sphere freezing, using (64) for the crystal density. The latter authors find that the free energy of the crystal decreases extremely rapidly for Gaussian widths $\alpha^{-1/2}$ that are unrealistically large. (According to simulation, freezing should occur near $\alpha\sigma^2 \approx 80$.) $n_0(\mathbf{r})$ and $n_1(\mathbf{r})$ become negative in regions where $n_3(\mathbf{r})$ is close to unity, so that the terms $-n_0 \ln(1 - n_3)$ and $n_1 n_2/(1 - n_3)$ in Φ_{PY} (97) become large and negative [130]. Once more the crucial problem is that associated with large values of the local packing fraction $n_3(\mathbf{r})$. In the limit $\alpha \to \infty$, when the crystal density becomes a sum of delta functions, $n_3(\mathbf{r}) = \sum_l \Theta(R - |\mathbf{r} - \mathbf{R}_l|)$ and the packing fraction is unity in a sphere of radius R around each lattice site $\mathbf{R}_l$. Evidently, the short range of the weight function $\omega^{(3)}(r)$ [equal to the hard-sphere radius R ($\equiv \sigma/2$)] is the source of the problem [130]. As mentioned in Section III.E.5, other WDAs have a range equal to σ, and these do not run into the difficulties encountered in the present case [i.e., the local packing fraction $\bar{\rho}(\mathbf{r})\pi\sigma^3/6$ remains $\lesssim 0.5$]. We are forced to conclude that the Rosenfeld or Kierlik–Rosinberg recipes are inherently unsuited for applications to freezing or to adsorption of liquids at strongly attractive substrates. Rosenfeld et al. [126,129] find, contrary to Ref. 130, that the free energy of the crystal is never lower than that of the liquid ($\alpha = 0$), so there is no freezing transition. They also attribute their result to the short range of their weight functions but defend their recipe by arguing that it is just this feature that makes their approach exact in one dimension. We see no reason why the range of the weight function(s) should not increase with increasing dimension.

The main thrust of Rosenfeld et al.'s work, however, lies in the calculation of triplet direct correlation functions $c^{(3)}$ for uniform hard-sphere liquids. $c^{(3)}$ is obtained easily from (93). The results are in good agreement with extensive Monte Carlo simulations—as are those from the simplified version in Ref. 130. We return to this aspect of density functional theory in Section IV.

F. Position-Independent WDAs

The approximations described in the preceding subsections all employ weighted densities that depend on position $\mathbf{r}$. Another class of WDA has been developed, primarily for the study of freezing, which makes use of a single, position-independent weighted density. These approximations

are often computationally simpler than many of the earlier versions. Here we discuss some that might be relevant for other types of inhomogeneity, as well as for the freezing problem.

1. *MWDA of Denton and Ashcroft*

Denton and Ashcroft [119] proposed a modified WDA (MWDA) based on the following relation for the excess free energy per atom:

$$\frac{\mathcal{F}_{ex}[\rho]}{N} = \psi_{ex}(\hat{\rho}) \tag{100}$$

where, by analogy with the earlier definition (61) of the weight function, the weighted density is postulated to be

$$\hat{\rho} = \frac{1}{N}\int d\mathbf{r}\, \rho(\mathbf{r}) \int d\mathbf{r}'\, \rho(\mathbf{r}')\hat{w}(|\,\mathbf{r} - \mathbf{r}'\,|;\hat{\rho}) \tag{101}$$

with $N = \int d\mathbf{r}\, \rho(\mathbf{r})$. The new weight function $\hat{w}$ is specified by requiring (100), with (101), to generate the (given) uniform fluid $c^{(2)}(\rho;r)$. It is found [119] that

$$\hat{w}(r;\rho) = -\frac{1}{2\psi'_{ex}(\rho)}\left[\beta^{-1}c^{(2)}(\rho;r) + \frac{1}{V}\rho\psi''_{ex}(\rho)\right] \tag{102}$$

where V is the total volume of the fluid. The second term, which ensures that the compressibility sum rule (24) is obeyed, is somewhat curious since it implies that $\hat{w}$ has a constant "tail" $O(1/N)$. If we ignore the tail, $\hat{w}$ is simply proportional to, and therefore has the same range as, $c^{(2)}(\rho;r)$. Note that (100) and (101) require that $\hat{w}$ is normalized in the uniform limit [i.e., $\int d\mathbf{r}\, \hat{w}(r;\rho) = 1$]. In the MWDA the weight function is given immediately, whereas in the CA recipe the differential equation (70) must be solved. As the MWDA also satisfies the sum rule (71), it retains many of the good features of the CA version of WDA.

Denton and Ashcroft [119] applied the MWDA to the freezing of the hard-sphere fluid, using as input $c^{(2)}(\rho;r)$ and $\psi_{ex}(\rho)$ from PY approximation. Their results for the freezing characteristics were very close to CA and the simplified theory appears to be equivalent (essentially) to the more sophisticated and computationally demanding WDA—at least as regards freezing. But how does it perform for interfacial problems, where the extent of the density inhomogeneity is *finite* rather than infinite, as is the case for the crystalline solid? For fluids adsorbed at walls the MWDA turns out to be completely equivalent [131] to (47), the HNC closure of the wall–particle Ornstein–Zernike equation. The structure of the MWDA is such that it corresponds to a truncation of the density

expansion at the quadratic term, as in (46), when it is applied to *adsorption*. Although the higher-order direct correlation functions $c^{(3)}$, and so on, generated from (100) are nonzero, their form is such [see (115)] that they make vanishingly small contributions when the integrals are performed in the higher-order terms of the density expansion of (44) [131]. It follows that the MWDA shares the strengths and the weaknesses of (47). In particular it cannot describe the growth of wetting films or surface phase transitions. Perhaps this is unsurprising. Since $\hat{\rho}$ is position independent it is forced to be equal to $\rho_b(1 + O(N^{-1}))$, where ρ_b is the density of the bulk fluid far from the wall; the bulk must dominate the integration in (101)—so there is no possibility of $\hat{\rho}$ taking on a value characteristic of the incipient wetting phase. When applied to the calculation of the structure of a uniform fluid, where the inhomogeneous density is caused by fixing one atom of the fluid, the MWDA again turns out to be equivalent to the (bulk) HNC approximation [131–133]. We postpone further discussion of this result to Section IV.

The MWDA has been extended to mixtures, following the procedure developed for the CA WDA (see Section III.E.3). Analytical results are obtained for the weight functions and the theory has been applied successfully to the freezing of binary hard-sphere mixtures [113,134]. It has also been applied, as one ingredient, in an investigation of surface melting at a crystal–gas interface [135] and in an extensive investigation [136] of the freezing of bulk fluids with soft repulsive (inverse-power and Yukawa) interatomic potentials. The latter uses the modified HNC theory to obtain the requisite bulk fluid input data $c^{(2)}(\rho;r)$ and $\psi_{ex}(\rho)$. The MWDA fails to predict freezing to a body-centered cubic (bcc) structure but does provide a good account of the liquid–face-centered cubic (fcc) transition.

Reference 136 also provides an instructive alternative derivation of the MWDA. This follows the derivation of the CA WDA given in Section III.E.4 but now $c^{(1)}([\alpha\rho];\mathbf{r})$ is expanded about a uniform fluid of constant density $\hat{\rho}$ so that we consider a functional

$$\mathcal{F}_{ex}[\rho,\hat{\rho}] = N\psi_{ex}(\hat{\rho}) - \beta^{-1}\int d\mathbf{r} \int d\mathbf{r}'\, \rho(\mathbf{r})L(\hat{\rho};|\,\mathbf{r}-\mathbf{r}'\,|)(\rho(\mathbf{r}') - \hat{\rho}) \tag{103}$$

where L is (another) arbitrary two-point function. $\hat{\rho}$ is treated as the density of a reference fluid to be chosen so that the second term in (103) is zero—then the free energy of the inhomogeneous system is equal to that of the uniform reference fluid of density $\hat{\rho}$. It follows that

$$\hat{\rho} = \frac{1}{N}\int d\mathbf{r}\, \rho(\mathbf{r}) \int d\mathbf{r}'\, \rho(\mathbf{r}')\, \frac{L(\hat{\rho};|\,\mathbf{r}-\mathbf{r}'\,|)}{L_0(\hat{\rho})} \tag{104}$$

with $L_0(\hat{\rho}) \equiv \int d\mathbf{r}\, L(\hat{\rho};r)$. Clearly, (104) is of the form postulated in (101), with $\hat{w} \equiv L(\hat{\rho};|\mathbf{r}-\mathbf{r}'|)/L_0(\hat{\rho})$ now normalized for all $\hat{\rho}$. When the function L is determined by requiring the resulting free-energy functional to yield the (given) $c^{(2)}$ in the uniform fluid, we arrive back at (102).

2. *GELA of Lutsko and Baus*

Baus and co-workers have developed a variety of "effective liquid theories" for use in density functional treatments of freezing. These involve different prescriptions for the effective density of a uniform liquid that is being used to describe the crystal [137]. One the latest versions, termed the generalized effective liquid approximation (GELA) [138], warrants special mention.

We begin with the exact result (25) for the excess free-energy functional and choose the initial density $\rho_i \equiv 0$. Then (25) simplifies to

$$\beta\mathcal{F}_{ex}[\rho] = -\int_0^1 d\alpha \int d\mathbf{r}\, \rho(\mathbf{r}) \int_0^\alpha d\alpha' \int d\mathbf{r}'\, \rho(\mathbf{r}')c^{(2)}([\alpha'\rho];\mathbf{r},\mathbf{r}') \tag{105}$$

For a uniform fluid of density ρ (105) reads

$$\beta\mathcal{F}_{ex}(\rho) = -N\rho \int d\mathbf{r} \int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}(\alpha'\rho;r) \tag{106}$$

One can map [138] the (unknown) excess free energy per atom of the nonuniform system onto that of some effective uniform liquid of densiy $\hat{\rho}_1$ by requiring that (105) and (106) be equal. Then

$$\hat{\rho}_1 \equiv \hat{\rho}_1[\rho] = \frac{1}{N} \int d\mathbf{r}\, \rho(\mathbf{r}) \int d\mathbf{r}'\, \rho(\mathbf{r}')w(\mathbf{r},\mathbf{r}';[\rho]) \tag{107}$$

with a weight function

$$w(\mathbf{r},\mathbf{r}';[\rho]) \equiv \frac{\int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}([\alpha'\rho];\mathbf{r},\mathbf{r}')}{\int d\mathbf{r} \int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}(\alpha'\hat{\rho}_1;r)} \tag{108}$$

Note that w is normalized in the limit of a uniform fluid $\rho(\mathbf{r}) \to \rho_1$, where $\hat{\rho}_1 \to \rho_1$, but not in the nonuniform system [see (104) for the MWDA).

A structural mapping can also be constructed [138] by requiring the density-averaged direct correlation function of the nonuniform fluid with density $\rho(\mathbf{r})$ [that enters (105)] to be identical to that of an effective uniform liquid of density $\hat{\rho}_2$ ($\equiv \hat{\rho}_2[\rho]$):

$$\int d\mathbf{r} \int d\mathbf{r}'\, \rho(\mathbf{r})\rho(\mathbf{r}')c^{(2)}([\rho];\mathbf{r},\mathbf{r}') = \int d\mathbf{r} \int d\mathbf{r}'\, \rho(\mathbf{r})\rho(\mathbf{r}')c^{(2)}(\hat{\rho}_2[\rho];|\mathbf{r}-\mathbf{r}'|) \tag{109}$$

Then (105) can be reexpressed as

$$\beta\mathcal{F}_{ex}[\rho] = -\int_0^1 d\alpha \int d\mathbf{r}\, \rho(\mathbf{r}) \int_0^\alpha d\alpha' \int d\mathbf{r}'\, \rho(\mathbf{r}') c^{(2)}(\hat{\rho}_2[\alpha'\rho];|\mathbf{r} - \mathbf{r}'|) \tag{110}$$

which, by comparison with (106), implies that $\hat{\rho}_2[\rho] \to \rho_1$ in the limit $\rho(\mathbf{r}) \to \rho_1$. If we now combine the first (thermodynamic) mapping with (110), we obtain

$$\hat{\rho}_1 = \frac{1}{N} \frac{\int d\mathbf{r}\, \rho(\mathbf{r}) \int d\mathbf{r}' \rho(\mathbf{r}') \int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}(\hat{\rho}_2[\alpha'\rho];|\mathbf{r} - \mathbf{r}'|)}{\int d\mathbf{r} \int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}(\alpha'\hat{\rho}_1; r)} \tag{111}$$

which, unlike (107), is expressed entirely in terms of $c^{(2)}$ for the uniform fluid. So far all is purely formal. Provided that both mappings exist, everything remains exact.

The GELA proceeds by making the assumption that

$$\hat{\rho}_1[\rho] = \hat{\rho}_2[\rho] \equiv \hat{\rho}[\rho]$$

The two mappings are taken to be equal so that the same effective liquid that is used to mimic the density-averaged correlation function of the nonuniform fluid is also required to reproduce its excess free energy. Equation (111) then takes the form of a self-consistency equation for $\hat{\rho}$ in terms of $\rho(\mathbf{r})$, given $c^{(2)}$ of the uniform fluid. It can be written, by analogy with the exact expression (107), as a doubly weighted density with w replaced by

$$w^{\mathrm{GELA}}(|\mathbf{r} - \mathbf{r}'|;[\rho]) = -\hat{\rho} \frac{\int_0^1 d\alpha \int_0^\alpha d\alpha'\, c^{(2)}(\hat{\rho}[\alpha'\rho];|\mathbf{r} - \mathbf{r}'|)}{\beta\psi_{ex}(\hat{\rho})} \tag{112}$$

Because of the functional dependence of $\hat{\rho}$ in the numerator, this weight function is normalized only in the limit of a uniform fluid.

To summarize, the GELA excess free-energy functional is defined by

$$\frac{\mathcal{F}_{ex}^{\mathrm{GELA}}[\rho]}{N} = \psi_{ex}(\hat{\rho})$$

as in the MWDA, or alternatively, by (110) with $\hat{\rho}_2 = \hat{\rho}$, where the weighted density is given by (111) with $\hat{\rho}_2 = \hat{\rho}_1 = \hat{\rho}$. The striking feature of the GELA is that the direct correlation functions of the nonuniform fluid can be expressed directly in terms of $c^{(2)}(\hat{\rho}; |\mathbf{r} - \mathbf{r}'|)$. Noticing

[138] that (110) retains the formal structure of the exact result (105) (the integrations are identical) it follows, via differentiation, that

$$c^{(1)}_{\mathrm{GELA}}([\rho];\mathbf{r}) = \frac{-\delta}{\delta\rho(\mathbf{r})}(\beta\mathcal{F}_{\mathrm{ex}}^{\mathrm{GELA}}[\rho]) \tag{113a}$$

$$= \int d\mathbf{r}' \int_0^1 d\alpha\, c^{(2)}(\hat{\rho}[\alpha\rho];|\,\mathbf{r}-\mathbf{r}'\,|)$$

$$c^{(2)}_{\mathrm{GELA}}([\rho];\mathbf{r}_1,\mathbf{r}_2) = \frac{-\delta^2(\beta\mathcal{F}_{\mathrm{ex}}^{\mathrm{GELA}}[\rho])}{\delta\rho(\mathbf{r}_1)\,\delta\rho(\mathbf{r}_2)} \tag{113b}$$

$$= c^{(2)}(\hat{\rho}[\rho];|\,\mathbf{r}-\mathbf{r}'\,|)$$

and

$$c^{(n)}_{\mathrm{GELA}}([\rho];\mathbf{r}_1,\ldots,\mathbf{r}_n) = \frac{\delta^{n-2}c^{(2)}(\hat{\rho}[\rho];|\,\mathbf{r}_1-\mathbf{r}_2\,|)}{\delta\rho(\mathbf{r}_3)\cdots\delta\rho(\mathbf{r}_n)} \qquad n>2 \tag{113c}$$

The key result is (113b), which states that $c^{(2)}$ of the nonuniform fluid is identical to that of the uniform fluid at the self-consistently determined density $\hat{\rho}_2[\rho]$. It is this feature that makes the GELA different from other theories, where the equivalence holds only in the uniform limit. Clearly, one cannot study the *structure* of inhomogeneous fluids using the GELA; it is a theory designed to provide the optimal way of calculating *thermodynamic* properties within an effective liquid picture. While (113c) does offer a way of obtaining the higher-order $c^{(n)}$ of the *uniform* fluid, the results are unrealistic.

Lutsko and Baus [138] describe the determination of $\hat{\rho}$ [i.e., the solution of (111)] for crystals where the density is parametrized, as in (64). Their results for hard-sphere freezing are remarkably close to simulation and rather better than those from other density functional approximations. Somewhat surprisingly, the GELA appears to fail to predict freezing into either bcc or fcc crystals for inverse-power law repulsive potentials, whereas the MWDA does account for the transition to fcc [136].

It is not obvious that the GELA is well suited for applications other than freezing. The self-consistent solution of (111) for densities that are not parametrized might prove intractable. Moreover, the status of such an approach, with a position-independent weighted density, for adsorption problems remains to be investigated. Does it suffer from the same defects as the MWDA when applied in this context? The GELA has much the same structure as the MWDA. For example, $c^{(3)}_{\mathrm{GELA}}$ for the uniform fluid [see (116)] is proportional to V^{-1}, where V is the total volume occupied by fluid. This suggests, as with MWDA [131], that contributions with $n > 2$ in the density expansion of $\mathcal{F}_{\mathrm{ex}}^{\mathrm{GELA}}$ are vanishing small in the thermodynamic limit for adsorption at a wall, therefore making zero contri-

bution to the surface tension and having no affect on the profile. The GELA would then also reduce to the HNC closure (47) for such problems. This contention remains to be proved. Note that for freezing, where the inhomogeneity extends throughout the system, the high-order contributions are nonvanishing and constitute very significant corrections to the quadratic ($n = 2$) approximation. For fluids at interfaces, however, a versatile approach would seem to demand a position-dependent weighted density.

Finally, we mention that Lutsko and Baus [138] show how the MWDA and an earlier self-consistent ELA (SCELA) of Baus [139] can be viewed as mathematical approximations to the solution of GELA. The SCELA replaces $\hat{\rho}[\alpha'\rho]$ in (111) by $\alpha'\hat{\rho}[\rho]$, which simplifies the calculation of $\hat{\rho}$. Although this changes the character of the theory [(113) no longer hold] it does yield the correct $c^{(2)}$ in the uniform limit and its numerical performance is close to that of the GELA—at least for hard-sphere freezing [138]. Zeng and Oxtoby [140] have also introduced an effective liquid mapping that yields another prescription for a position-independent $\hat{\rho}$ that is slightly less successful than the GELA for this problem.

IV. FREE-ENERGY FUNCTIONALS AND THE STRUCTURE OF UNIFORM FLUIDS

Before discussing applications to intrinsically nonuniform fluids it is instructive to inquire about the predictions of the various approximation schemes for the correlation functions of *uniform* fluids. There are two (distinct) procedures that can be applied. The first has already been mentioned several times. Given some approximate $\mathcal{F}_{ex}[\rho]$ one can generate the hierarchy $c^{(n)}$ via (12) and (13) and compare the results, in the uniform limit, with those from simulation or from other (integral equation) theories. In the crudest versions comparison for $c^{(2)}(\rho;r)$ ($n = 2$) is already informative. In the more sophisticated versions, where $c^{(2)}(\rho;r)$ is often an input for the theory, the predictions for $n \geq 3$ are relevant. Recently simulation data for $c^{(3)}$ has become available for a uniform soft-sphere fluid [121] and for the hard-sphere fluid [129]. The latter data are rather extensive.

A. Three-Body Direct Correlation Function

In Section III.E we indicated that any free-energy functional, which is not merely a truncation of the density expansion, should satisfy the sum rule (71). Thus, for $n = 3$ the Fourier transform of $c^{(3)}(\rho;\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) \equiv$

$c^{(3)}(\rho;\mathbf{r}_1 - \mathbf{r}_2,\mathbf{r}_1 - \mathbf{r}_3)$ should satisfy

$$c^{(3)}(\rho;k, 0) = c^{(3)}(\rho;0,k) = c^{(3)}(\rho;\mathbf{k}, -\mathbf{k}) = \frac{\partial}{\partial\rho} c^{(2)}(\rho;k) \qquad (114)$$

in an obvious notation; $c^{(2)}(\rho; k) \equiv c_k^{(2)}(\rho)$, introduced in (70). Even the crudest of approximations for $\mathcal{F}_{ex}$ will yield $c^{(3)}$ consistent with (114). Satisfying this sum rule does not constitute a severe test.

CA [111] calculated $c^{(3)}$ for a hard-sphere fluid using their WDA and compared their results with the soft-sphere results of Ref. 121. The agreement was reasonable. Rosenfeld et al. [129] carried out a detailed comparison of their simulation results for $c^{(3)}(\rho;\mathbf{k},\mathbf{k}')$ with the predictions of Rosenfeld's WDA; the latter yields analytical results for $c^{(3)}$ and the higher-order functions. These authors showed that the theoretical results agree quite well with the Monte Carlo data, apart from certain discrepancies in $c^{(3)}(\rho;\mathbf{k}, -\mathbf{k})$ at small $\mathbf{k}$, which they ascribe [see (114)] to the fact that the inputted PY $c^{(2)}(\rho;k)$ is itself inaccurate for $k\sigma \lesssim 3.5$. They also argued that the (seemingly successful) factorization ansatz in r space introduced in Ref. 121 may be "less conclusive than originally reported," noting that such a factorization is not in keeping with that suggested by their own approach, where $c^{(3)}$ is factorized in k space. The simplified version of Kierlik and Rosinberg [130] yields even an simpler expression for $c^{(3)}$ which appears to do slightly better at accounting for the Monte Carlo data.

The MWDA of Denton and Ashcroft [119] yields a simple analytical expression for the three-body direct correlation function [119,131]:

$$\begin{aligned} -\beta^{-1}c^{(3)}_{\text{MWDA}}(\rho;\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) &= 2V^{-1}\psi''_{ex}(\rho)[\hat{w}(\mathbf{r}_1 - \mathbf{r}_2;\rho) + \\ \hat{w}(\mathbf{r}_1 - \mathbf{r}_3;\rho) &+ \hat{w}(\mathbf{r}_2 - \mathbf{r}_3;\rho)] + 2V^{-1}\psi'_{ex}(\rho)[\hat{w}'(\mathbf{r}_1 - \mathbf{r}_2;\rho) + \\ \hat{w}'(\mathbf{r}_1 - \mathbf{r}_3;\rho) &+ \hat{w}'(\mathbf{r}_2 - \mathbf{r}_3;\rho)] + \rho V^{-2}\psi'''_{ex}(\rho) - 3V^{-2}\psi''_{ex}(\rho) \end{aligned} \qquad (115)$$

where V is the total volume of the fluid. While it is easy to check that (115) satisfies (114), the factors of V^{-1} ensure that this result is totally unrealistic otherwise. $c^{(3)}_{\text{MWDA}}(\rho;\mathbf{k},\mathbf{k}')$ is identially zero for $\mathbf{k},\mathbf{k}' \neq 0$ unless $\mathbf{k}' = -\mathbf{k}$, where it reduces to $-2\beta(\psi''_{ex}\hat{w}_1(k;\rho) + \psi'_{ex}\hat{w}'_1(k;\rho))$, consistent with (102) in (114). Since the GELA of Lutsko and Baus [138] has the same structure as the MWDA, it should yield similar $c^{(3)}$. Differentiating (113b) we find that

$$c^{(3)}_{\text{GELA}}(\rho;\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \frac{\partial}{\partial\rho} c^{(2)}(\rho;|\mathbf{r}_1 - \mathbf{r}_2|) \frac{\delta\hat{\rho}}{\delta\rho(\mathbf{r}_3)}\bigg|_\rho \qquad (116)$$

with $\delta\hat{\rho}/\delta\rho(\mathbf{r}_3) = V^{-1}$ in the uniform limit [132]. Even if (116) were to be

symmetrized, as mentioned in Ref. 138, it would remain just as unsatisfactory as (115).

To overcome this defect of the MWDA, Denton and Ashcroft [141] introduced a weighted density approximation for the one-body direct correlation function; that is, they set

$$c^{(1)}([\rho];\mathbf{r}) = c^{1)}(\bar{\rho}(\mathbf{r})) \tag{117}$$

where $c^{(1)}(\bar{\rho}(\mathbf{r}))$ refers to a uniform liquid of density $\bar{\rho}(\mathbf{r})$ with the weighted average defined via (61). Then

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \frac{\delta c^{(1)}([\rho];\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} = c^{(1)\prime}(\bar{\rho}(\mathbf{r}_1))\frac{\delta\bar{\rho}(\mathbf{r}_1)}{\delta\rho(\mathbf{r}_2)} \tag{118}$$

which is *not* equal to $c^{(2)}(\mathbf{r}_2,\mathbf{r}_1)$ for all densities [i.e., the last of the equalities in (13) is not satisfied since this particular approximation does not follow from an excess free-energy functional]. However, in the uniform limit, where $\rho(\mathbf{r}) = \bar{\rho}(\mathbf{r}) = \rho$, $\delta\bar{\rho}(\mathbf{r}_1)/\delta\rho(\mathbf{r}_2)$ reduces to $w(|\mathbf{r}_2 - \mathbf{r}_1|;\rho)$, and requiring (118) to reduce to $c^{(2)}(\rho;r_{12})$ (assumed known) gives the unique form

$$w(r;\rho) = \frac{c^{(2)}(\rho;r)}{c^{(1)\prime}(\rho)} \tag{119}$$

for the weight function. Differentiating again, Denton and Ashcroft obtained, in the uniform limit, an explicit formula for $c^{(3)}$ in terms of $c^{(2)}$:

$$c^{(3)}(k,k') = \frac{1}{c^{(1)\prime}}[c^{(2)}(k)c^{(2)\prime}(k') + c^{(2)}(k')c^{(2)\prime}(k)] - \frac{c^{(1)\prime\prime}c^{(2)}(k)c^{(2)}(k')}{(c^{(1)\prime})^2} \tag{120}$$

As before, a prime denotes differentiation with respect to ρ. While (120) satisfies the first two of (114), it does not satisfy the sum rule with $\mathbf{k}' = -\mathbf{k}$. Presumably, this is a consequence of beginning with an approximation for $c^{(1)}$ rather than for $\mathcal{F}_{ex}$. Denton and Ashcroft [141] argue that this is also the reason for the lack of any dependence of $c^{(3)}$ on the angle between wave vectors $\mathbf{k}$ and $\mathbf{k}'$, and propose to remedy this deficiency by imposing a symmetry requirement. Although a detailed comparison of (120) with simulation has not yet been reported, Rosenfeld et al. [129] state that the results are "able to account for the gross features in the simulation data." They also state that Curtin's WDA results for $c^{(3)}$ are close to those from the Denton–Ashcroft scheme.

Thus, at present, it is difficult to choose between the Rosenfeld recipe,

the CA recipe, and probably Tarazona (Mark II), on the basis of their predictions for $c^{(3)}$. This is a little disappointing. Given that the range of the relevant weight function is quite different in these recipes, we might have expected some propagation of the difference into $c^{(3)}(\rho;\mathbf{k},\mathbf{k}')$. Further comparisons might be more revealing.

B. Two-Body Correlations and the Percus Identity

The second procedure that can be applied to the determination of the structure of uniform fluids makes use of Percus' [6] observation. Consider a fluid of N atoms, with uniform density ρ_b, single out one atom, and measure the positions of the remaining $N - 1$ with respect to the center of that atom. Each of these $N - 1$ atoms will experience an "external" potential that is identical to the interatomic pair potential $\phi(\mathbf{r}) \equiv \phi(r)$ exerted by the atom fixed at the origin. Percus observed that the fluid then has a nonuniform density profile

$$\rho(\mathbf{r}) \equiv \rho(r) \equiv \rho_b g(r) \tag{121}$$

where $g(r)$ is the radial distribution function; recall the definitions of both functions. It is clear that an approximate density functional theory, when applied to this particular type of inhomogeneity, will yield an approximation for $g(r)$; minimization of the approximate $\Omega_V[\rho]$, with $V(\mathbf{r}) = \phi(\mathbf{r})$, and solution of the resulting Euler–Lagrange equation is all that is required. The status of results obtained from such a procedure requires some examination, however. To appreciate the issues that are involved, it is first necessary to recall standard integral equation theories of liquids.

The density profile can be expressed [see (14) and (15)] in the form

$$\rho(r) = \rho_b \exp[-\beta\phi(r) + c^{(1)}([\rho];r) - c^{(1)}(\rho_b)] \tag{122}$$

which leads, via (121), to a self-consistency equation for the radial distribution function

$$g(r) = \exp[-\beta\phi(r) + c^{(1)}([\rho_b g];r) - c^{(1)}(\rho_b)] \tag{123}$$

The exponent can be reexpressed, using (23), with $\rho_i = \rho_b$, as

$$c^{(1)}([\rho];r_1) - c^{(1)}(\rho_b) = \int_0^1 d\alpha \int d\mathbf{r}_2\, (\rho(\mathbf{r}_2) - \rho_b) c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2)$$

so that (123) becomes

$$\ln g(r_1) = -\beta\phi(r_1) + \int_0^1 d\alpha \int d\mathbf{r}_2\, \rho_b h(r_2) c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2) \tag{124}$$

where $h(r) \equiv g(r) - 1 = \rho(r)/\rho_b - 1$ is the total correlation function of the uniform fluid. Equation (124) is still exact but of no practical use until some approximation is made for $c^{(2)}([\rho_\alpha];\mathbf{r}_1,\mathbf{r}_2)$, the quantity that refers to the nonuniform fluid of density $\rho_\alpha(\mathbf{r})$. The HNC approximation simply replaces this quantity by $c^{(2)}(\rho_b;r_{12})$, its value in the initial (bulk) state with $\alpha = 0$ [see (45)]. Thus the HNC closure is

$$\ln g(r_1) = -\beta\phi(r_1) + \rho_b \int d\mathbf{r}_2\, h(r_2)c^{(2)}(\rho_b;r_{12}) \tag{125}$$

Using the uniform fluid OZ equation

$$h(r_1) = c^{(2)}(\rho_b;r_1) + \rho_b \int d\mathbf{r}_2\, h(r_2)c^{(2)}(\rho_b;r_{12}) \tag{126}$$

(125) reduces to the familiar HNC form

$$g(r) = \exp[-\beta\phi(r) + h(r) - c^{(2)}(\rho_b;r)] \qquad \text{HNC} \tag{127}$$

The bridge or elemental diagrams that are missing in the HNC can be reinstated formally via the inclusion of the bridge function $B(r)$:

$$g(r) = \exp[-\beta\phi(r) + h(r) - c^{(2)}(\rho_b;r) + B(r)] \tag{128}$$

Approximate integral equation theories correspond to different prescriptions for $B(r)$ and some of the more sophisticated versions yield results for both structure and thermodynamic functions that are in excellent agreement with simulation (e.g., Refs. 28 and 30).

Comparison of (123) and (128) shows that $B(r)$ may be expressed as

$$B(r) = c^{(1)}([\rho_b g];r) - c^{(1)}(\rho_b) - h(r) + c^{(2)}(\rho_b;r) \tag{129}$$

By expanding the first term about the density of the bulk fluid, we obtain

$$c^{(1)}([\rho_b g];r_1) = c^{(1)}(\rho_b) + \rho_b \int d\mathbf{r}_2\, h(r_2)c^{(2)}(\rho_b;r_{12}) + \sum_{n=2}^{\infty} \frac{{\rho_b}^n}{n!} \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_{n+1}\, h(r_2) \cdots h(r_{n+1})c^{(n+1)}(\rho_b;\mathbf{r}_1, \ldots, \mathbf{r}_{n+1}) \tag{130}$$

Use of the OZ equation (126) then leads to the exact expansion [121] for the bridge function

$$B(r_1) = \sum_{n=2}^{\infty} \frac{{\rho_b}^n}{n!} \int d\mathbf{r}_2 \cdots \int d\mathbf{r}_{n+1}\, h(r_2) \cdots h(r_{n+1})c^{(n+1)}(\rho_b;\mathbf{r}_1, \ldots, \mathbf{r}_{n+1}) \tag{131}$$

in terms of the high-order direct correlation functions of the uniform fluid.

Note that truncation of (130) at the first-order term yields $B(r) = 0$, which is the HNC approximation (127). Given some theory for $c^{(3)}$, and so on, $B(r)$ may be estimated from (131). Barrat et al. [121] used a real-space factorization ansatz for $c^{(3)}$ and ignored terms with $n > 2$ to construct an approximation for $B(r)$. The resulting theory represents a significant improvement on the HNC for soft spheres, the Lennard–Jones liquid and the OCP. Denton and Ashcroft [133] suggested that $c^{(3)}$ as derived from explicit density functional approximations (see Section IV.A) could also be used to investigate the form of the bridge function.

Given an explicit approximation for $\mathcal{F}_{ex}[\rho]$ it is, of course, straightforward to calculate $c^{(1)}([\rho_b g];r)$ appropriate to this type of inhomogeneity. In the MWDA of Denton and Ashcroft [119] (see Section III.F.1) it turns out [131,132] that $\hat{\rho} = \rho_b(1 + O(1/N))$ and

$$c^{(1)}([\rho_b g];r_1) = c^{(1)}(\rho_b) + \rho_b \int d\mathbf{r}_2\, h(r_2)c^{(2)}(\rho_b;r_{12}) + O\left(\frac{1}{N}\right)$$

Thus, in the thermodynamic limit, MDWA corresponds to neglecting all terms with $n \geq 2$ in (130). The resulting closure is simply (125), the HNC approximation. At this stage $\psi_{ex}(\rho)$ and $c^{(2)}(\rho;r)$ remain unspecified. If we require $c^{(2)}$ and g to satisfy the exact OZ equation (126), we recover the usual HNC integral equation theory for a uniform liquid [131,132]. Note that ψ_{ex} need not be specified; it would be natural to calculate the free energy via the compressibility equation of state. Other theories that employ a position-independent weighted density should also be equivalent to HNC for uniform fluids [142]. If, however, the weighted density is position dependent we should expect the resulting theory to be different from HNC and yield a nonzero $B(r)$.

Jones and Kim [143] have developed this strategy further, addressing the fundamental issue of how the free-energy functional is related to the interatomic pair potential $\phi(r)$. They consider the WDA as implemented by CA (see Section III.E.3) so that

$$\mathcal{F}_{ex}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})\psi_{ex}(\bar{\rho}(\mathbf{r})) \tag{132}$$

with $\bar{\rho}(\mathbf{r})$ given by (61) and the weight function $w(r;\rho)$ is related to $c^{(2)}(\rho;r)$ via (70). The functions ψ_{ex} and w are (initially) unspecified. Jones and Kim argue that use of the Percus identity (121) is sufficient to determine these functions, at least for a certain class of interatomic potentials and thermodynamic states. Their procedure is shown schematically in Fig. 6. Given (initial) ψ_{ex} and w, $g(r)$ is determined for a given potential $\phi(r)$, but then $c^{(2)}$, as obtained from the OZ equation, does not satisfy the requirement that $c^{(2)}$ is the second functional derivative of (132). The

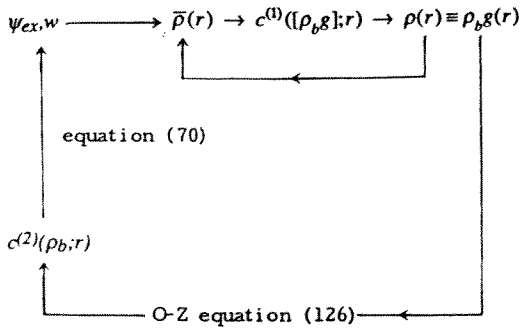


Fig. 6 Implementation of the WDA to determine the structure and thermodynamics of a uniform fluid.

iteration procedure in Fig. 6 may or may not converge for arbitary bulk densities and potentials. Following Tarazona [94], Jones and Kim assume that $w(r;\rho)$ has the density expansion (62) and that g, $c^{(2)}$, and ψ_{ex} have similar density expansions. They then proceed to obtain the low-order coefficients, arguing that the theory has a unique solution within the region of convergence of the density expansion, provided that $w_0(k) \equiv \int d\mathbf{r}\ e^{i\mathbf{k}\cdot\mathbf{r}} w_0(r) \neq 1$ for real wave vectors k. But the zeroth coefficient $w_0(r) = f(r)/\int d\mathbf{r}\ f(r)$, where $f(r) = e^{-\beta\phi(r)} - 1$, is the Mayer function (see Section III.E.1), so that for purely repulsive potentials ($\phi > 0$ everywhere), $| w_0(k) | < 1$ for all real $k \neq 0$. When ϕ has an attractive portion, $w_0(k)$ will be equal to 1 for some real k and sufficiently large β. Then the iteration procedure breaks down [143], suggesting that the basic assumptions are not valid for such pair potentials. Jones and Kim compute the coefficients $g_i(r)$ ($i = 0, 1, 2$) in the density expansion of $g(r)$, as well as the (pressure) virial coefficients up to B_4 for hard spheres and for the Gaussian model. B_4 can be determined from $g_2(r)$ using either the compressibility (C) equation or the virial equation (V). For hard spheres $B_4(V)$ is 0.6% larger than the exact value, while $B_4(C)$ is 1.6% greater than $B_4(V)$. For the Gaussian model, where $f(r) \equiv \exp(-\alpha r^2)$, corresponding to soft repulsive forces, $B_4(V)$ and $B_4(C)$ agree to within 0.3% and $B_4(V)$ lies within 0.3% of the exact result. $g_0(r)$ and $g_1(r)$ agree with the exact virial results, but $g_2(r)$ does not. Nevertheless, for both models $g_2(r)$ lies very close to the exact results and is much better than the corresponding results from PY and HNC, suggesting that the theory is very promising. Kim [142] has also investigated the GvE recipe (see Section III.E.4) for the free-energy functional [i.e., (75) with (79) for $\rho_0(\mathbf{r})$ and the two-point function L determined by (82)]. Making the same density expansions as

previously, Kim finds that B_4 and $g_2(r)$ are given much less accurately than in the CA version of WDA. This would appear to support our assertion that there is nothing to be gained in extremizing the functional (75). Of course, it remains to be seen whether the iteration procedure can be effected for dense liquids, where the expansion is no longer useful. Jones and Kim point out that the failure of the WDA theory at low temperatures for fluids with attractive potentials probably reflects some fundamental structural flaw in the form (132) assumed for the excess free-energy functional in the case where $\phi(r)$ can be negative. This observation reinforces our earlier contention that attractive and repulsive forces require separate treatments; (132) may be accurate for the repulsive force contribution to $\mathcal{F}_{ex}$ but not for the attractive force contribution.

Denton and Ashcroft [144] have reported some calculations of $g(r)$ for hard spheres based on their earlier ansatz (117) for $c^{(1)}[\rho];r)$. Their input are the PY results for $c^{(1)}(\rho)$ and $c^{(2)}(\rho;r)$, so that the weight function (119) has a range exactly equal to the hard-sphere diameter. Equation (122) can then be solved iteratively for $\bar{\rho}(r)$ and $\rho(r)$ and hence $g(r)$. Although the results are in reasonable agreement with simulation, they are not as accurate as the standard PY results. The bridge function $B(r)$ is estimated using (129) with the PY expressions for $c^{(2)}(\rho_b;r)$ and $h(r)$. Again the results are in reasonable agreement with simulation and appear to be more accurate than those obtained from the truncated expansion of $B(r)$ [i.e., taking the first term in (131) with $c^{(3)}$ given by (120)]. No attempt was made to recalculate $c^{(2)}(\rho_b;r)$ and iterate to self-consistency, in the sense of Fig. 6. A similar calculation has been performed by Brenan and Evans [145] using as input the Carnahan and Starling result for ψ_{ex} and the Tarazona (Mark II) recipe for the weight function. The resulting $g(r)$ is now very close to simulation even for high densities. However, the hard-sphere virial equation result, $\beta p = \rho_b[1 + \frac{2}{3}\pi\rho_b\sigma^3 g(\sigma^+)]$, is still not satisfied exactly. Freasier et al. [146] performed the equivalent calculation but with the zeroth-order weight function w_0 of Section III.E.1. When the Carnahan and Starling result for ψ_{ex} is employed (Tarazona Mark I) the peaks and troughs in $g(r)$ are greatly exaggerated. This trend is consistent with the results shown in Fig. 2 for the density profile of a confined hard-sphere fluid. Takamiya and Nakanishi [147] have investigated the radial distribution function of two-dimensional fluids (hard disks and Lennard–Jones) using a modification of Tarazona (Mark II) to construct $\mathcal{F}_{ex}$ for the hard-disk system. Their results are also encouraging. No attempt was made to iterate to self-consistency in [145–147].

For completeness we should mention the work of Zhou and Stell [73], referred to in Section III.C, who have developed a nonlocal density functional expansion of the direct-correlation function which allows calcula-

tion of an approximate $B(r)$. They regard their new approximations as adding "hydrostatic" correction terms to the standard PY and HNC approximations. Their approach differs from others described in this chapter in that they do not consider an explicit free-energy functional.

V. APPLICATIONS TO FLUID INTERFACES

As indicated in Section I, there have been numerous applications of approximate density functionals to the calculation of interfacial properties. Rather than attempting to review the entire field we select a few topics that illustrate the usefulness and the shortcomings of the density functional approach. The choice reflects the author's own interests.

A. Liquid–Gas Interface

We consider an atomic fluid in an external potential $V(z)$, which we shall set equal to the gravitational potential mgz. The simplest choice of the free-energy functional is the square-gradient approximation (39). If, for simplicity, we ignore the density dependence of the coefficient f_2, the Euler–Lagrange equation for $\rho(z)$, the density profile of the planar interface, is

$$\mu - V(z) = \mu(\rho(z)) - 2f_2\rho''(z) \tag{133}$$

where $\mu(\rho) = (\partial f/\partial\rho)_T$ exhibits a van der Waals loop for $T < T_c$. We are interested in the solution of (133) in the limit where the potential vanishes. It is well known [17,148] that a nontrivial solution characteristic of a two-phase system [i.e., with $\rho'(z) \equiv d\rho/dz \not\equiv 0$] exists in this limit when $\mu = \mu_{sat}$, the chemical potential at bulk liquid–gas coexistence. There are an infinite number of solutions, all with the same shape, and a unique solution is specified by fixing the position of the Gibbs dividing surface at $z = 0$, say. When $V(z) \equiv 0$ (133) has a first integral

$$\Psi(\rho(z)) \equiv \omega(\rho(z)) + p = f_2(\rho'(z))^2 \tag{134}$$

with $\omega(\rho) \equiv f(\rho) - \mu\rho$, the grand potential density, and $p = p_{sat}$, the pressure. $\Psi(\rho(z))$ vanishes deep in the liquid phase where $\rho(z) \to \rho_l$ and deep in the gas phase where $\rho(z) \to \rho_g$. Note that at coexistence $f(\rho_l) - \mu\rho_l = -p = f(\rho_g) - \mu\rho_g$. In fact, $\Psi(\rho)$ has two minima at ρ_g and ρ_l and a maximum at an intermediate density [17]. Once $f(\rho)$ and f_2 are specified, $\rho(z)$ is easily obtained by quadrature. The surface excess grand potential is

$$\Omega^{ex} = \int d\mathbf{r}\,(\omega(\rho(z)) + f_2(\rho'(z))^2 + p) \tag{135}$$

and it follows that in zero field the surface tension is

$$\gamma = \omega^{ex} \equiv \frac{\Omega^{ex}}{A} = 2f_2 \int_{-\infty}^{\infty} dz\, (\rho'(z))^2 \tag{136a}$$

or

$$\gamma = 2\int_{-\infty}^{\infty} d\rho [f_2 \Psi(\rho)]^{1/2} \tag{136b}$$

The second version remains valid when f_2 depends on ρ. Numerical work based on (134) and (136), with f_2 estimated from (40) (see Section III.B), yields sensible, monotonic density profiles whose width increases with increasing temperature, eventually diverging as $T \to T_c$ with the same (mean-field) exponent $\nu = \frac{1}{2}$ as the bulk correlation length ξ_b. The surface tension vanishes as $(T_c - T)^{\tilde{\mu}}$ with $\tilde{\mu} = \frac{3}{2}$. This is van der Waals' original result [17].

The internal consistency of this simple density functional approach can be tested by examination of the exact Triezenberg–Zwangzig [149] formula for the surface tension,

$$\gamma = -\beta^{-1}\int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2\, \rho'(z_1)\rho'(z_2)c_2(z_1,z_2) \tag{137}$$

where $c_2(z_1,z_2)$ is the second (transverse) moment of the two-body direct-correlation function of the nonuniform fluid. For planar geometry

$$c^{(2)}(\mathbf{r}_1,\mathbf{r}_2) \equiv c^{(2)}(z_1,z_2;R) \tag{138}$$

and the Fourier transform has the expansion, in powers of Q^2:

$$\begin{aligned} c^{(2)}(z_1,z_2;Q) &\equiv \int d\mathbf{R}\, e^{i\mathbf{Q}\cdot\mathbf{R}} c^{(2)}(z_1,z_2;R) \\ &= c_0(z_1,z_2) + c_2(z_1,z_2)Q^2 + \cdots \end{aligned} \tag{139}$$

$\mathbf{Q}$ is a wave vector parallel to the surface. Within the square-gradient approximation $c^{(2)}$ is very simple. From (57) we find that

$$c_2(z_1,z_2) = -2\beta f_2 \delta(z_1 - z_2) \tag{140}$$

which, when inserted into (137), reproduces (136a). Note that (137) applies in the presence of the external field (see Chapter 2) and that (136a) is also valid (for the *tension* γ) in a nonzero field. Moreover, the zeroth moment $c_0(z_1,z_2)$ obtained from (57) and inserted in the exact equation for $\rho'(z)$,

$$\frac{\rho'(z_1)}{\rho(z_1)} + \beta V'(z_1) = \int_{-\infty}^{\infty} dz_2\, c_0(z_1,z_2)\rho'(z_2) \tag{141}$$

yields the derivative of (133). In other words, the square-gradient ap-

proximation is consistent with the formally exact equations for interfacial properties. This is not surprising. The exact equations are obtained by minimizing the exact grand potential; minimizing an approximate functional should lead to equations of the same form.

What physics does the square-gradient approximation omit? As mentioned in Section III.B, this approximation predicts exponential decay of the profile $\rho(z)$ into the bulk liquid or gas with a decay length equal to that of the appropriate bulk correlation length. For a Lennard–Jones fluid a nonlocal treatment of attractive forces, via a functional such as that in (53), leads to z^{-3} decay of the tails of $\rho(z)$ [50,51]. Such behavior reflects the algebraic decay of the direct-correlation function, [i.e., $c^{(2)} \sim -\beta\phi_{\text{att}}(r_{12})$ for large separations r_{12}; see (56)]. Clearly, a nonlocal theory such as that based on (53), or on a more sophisticated WDA for repulsive forces, will correct this particular defect of the square-gradient approximation. However, such approximations are still strictly mean field in character; they will not account for the correct surface tension critical exponent $\tilde{\mu}$. Fisk and Widom [17,47] extended the classical van der Waals (square-gradient) treatment to include nonclassical (bulk) critical exponents. Recognizing that the square-gradient approximation should become increasingly accurate as $T \to T_c$, where $\rho_l - \rho_g$ decreases and the interfacial width increases, leading to small density gradients, they retained the functional (39) but replaced the mean-field expansion of the grand potential density by

$$\omega(\rho) = \omega(\rho_c) + \frac{a}{2}(\rho - \rho_c)^2 + \frac{b}{\delta + 1}(\rho - \rho_c)^{\delta+1} - (\rho - \rho_c)(\mu - \mu_{\text{sat}}) \quad (142)$$

where ρ_c is the critical density, b is a constant, δ is the exponent describing the $\mu(\rho)$ critical isotherm, and a is a quantity that vanishes as the inverse compressibility [i.e., $a \sim \kappa_T^{-1} \sim (T_c - T)^{\gamma}$]. This form guarantees that *bulk* thermodynamic exponents are properly incorporated into the theory. The second modification involves the coefficient f_2 of the square-gradient term, which can no longer be treated as a positive constant when $T \to T_c$. It is no longer appropriate to take $c^{(2)}(\rho;r) \sim -\beta\phi_{\text{att}}(r)$ when the bulk correlation length ξ_b is diverging; this result is valid only for $r > \xi_b$ [17]. In reality the second moment of $c^{(2)}(\rho;r)$ [see (40)] is very weakly divergent: $f_2 \sim (T_c - T)^{-\nu\eta}$, when η is the exponent that governs the decay of the radial distribution function $g(r)$ at $T = T_c$. From (136b) it follows that the surface tension vanishes as $f_2^{1/2}a^{1/2}(\rho_l - \rho_g)^2$, so its critical exponent is $\tilde{\mu} = -\nu\eta/2 + \gamma/2 + 2\beta$, where β is the order-parameter exponent: $\rho_l - \rho_g \sim (T_c - T)^{\beta}$. Standard (bulk) exponent relations permit $\tilde{\mu}$ to be expressed in several ways [17]. Perhaps the most concise one is $\tilde{\mu} = 2 - \alpha - \nu$, where α is the heat capacity exponent. Using the mean-

field values $\eta = 0$, $\gamma = 1$, $\beta = \frac{1}{2}$, $\alpha = 0$, we recover the mean-field result $\bar{\mu} = \frac{3}{2}$. For dimensions $d \leq 4$ we expect hyperscaling to be valid so that $d\nu = 2 - \alpha$ and $\bar{\mu} = (d - 1)\nu$. It is straightforward to show that the interfacial width must diverge as $f_2^{1/2}a^{-1/2}$ or $(T_c - T)^{-\nu\eta/2-\gamma/2}$. Use of the Fisher relation $(2 - \eta)\nu = \gamma$ then shows that the width diverges in the same fashion as the bulk correlation length but with the proper nonclassical value of the exponent ν. The predictions of this Fisk–Widom version of density functional theory are satisfied exactly for the two-dimensional Ising model ($\bar{\mu} = \nu = 1$) and experimental results for *real* fluids, including binary liquid mixtures near their consolute points, are in excellent agreement with the prediction $\bar{\mu} = 2\nu$ [17].

While the Fisk–Widom approach incorporates the effects of bulk fluctuations, on a length scale set by ξ_b, it does not incorporate the effects of interfacial (capillary-wave-like) fluctuations. In this respect it is still a mean-field theory and is no more satisfactory than the original square-gradient or van der Waals theory based on (53). The long-wavelength interfacial fluctuations are expected [150] to lead to thermal broadening of the average density profile so that the total width W can be regarded (effectively) as the sum of an "intrinsic" width $\sim\xi_b$ and a fluctuation or "wandering" contribution $\xi_\perp$. If the wandering is on a similar scale to the intrinsic width and thus remains finite in the thermodynamic limit of infinite interfacial area and vanishing external field, the interface is said to be *smooth*. If, however, $\xi_\perp$ and hence W diverge in this limit, the interface is described as *rough*. Analysis of effective interfacial Hamiltonians* [20,150] indicates that the liquid–gas interface is rough for all temperatures at which two-phase coexistence occurs, for dimension $d = 3$ or 2. (For $d > 3$ this interface is predicted to be smooth for all temperatures.) In the marginal dimension $d = 3$ $\xi_\perp \sim (\ln L)^{1/2}$ when the interfacial area $L^2 \to \infty$ in zero external field, or $\xi_\perp \sim (-\ln g)^{1/2}$ as the gravitational acceleration $g \to 0$ with $L^2 = \infty$. For laboratory-sized samples or for infinite samples in the earth's gravity the capillary-wave broadening of the density profile is only a few atomic diameters (i.e., it is on the same scale as the intrinsic width). It is therefore difficult to disentangle wandering from intrinsic contributions. In $d = 2$, where the interface is a line rather than a surface, the broadening is much stronger and $\xi_\perp \sim L^{1/2}$ or $g^{-1/4}$. Extensive molecular dynamics simulations [151] (using a special-purpose hard-ware processor and up to 16,000 particles) for a two-dimensional Lennard–Jones fluid have confirmed the prediction that the total width $W \sim L^{1/2}$. By contrast, theories based on approximate free-energy functionals yield density profiles $\rho(z)$ that are essentially g inde-

* Reviewed by Henderson in Chapter 2.

pendent. Taking the limit of vanishing external field does not lead to the broadening of the profile that is described above; $\rho'(z)$ remains nonzero for $g = 0$ and infinite area L^2. Moreover, the same form of profile is obtained for all bulk dimensions d, reflecting the mean-field character of the theories. However, this does not mean that such theories are completely useless for investigating the liquid–gas interface. Although they fail to account for a diverging interfacial width (roughening), they do predict the long-range *transverse* correlations that are the key signature of the capillary-wave fluctuations.

Given an explicit form, such as (56) or (57), for $c^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$, one can effect the OZ inversion (17) to obtain the density–density correlation function G. Provided that z_1 and z_2 lie in the interface and the external potential is weak, one finds [46,84]

$$G(z_1,z_2;Q) \approx \beta^{-1} \frac{\rho'(z_1)\rho'(z_2)}{-\int_{-\infty}^{\infty} dz\, \rho'(z)V'(z) + \gamma Q^2} + \text{H.O.T.} \quad (143)$$

for the Fourier transform at small wave vectors Q. When $V(z) = mgz$ this result can be reexpressed in Ornstein–Zernike form as

$$G(z_1,z_2;Q) \approx G_0(z_1,z_2)(1 + \xi_\parallel{}^2 Q^2)^{-1} \quad (144)$$

as small Q, where the zeroth moment [see (139)]

$$G_0(z_1,z_2) = \frac{\rho'(z_1)\rho'(z_2)\xi_\parallel{}^2}{\beta\gamma} \quad (145)$$

and the transverse correlation length

$$\xi_\parallel = \left[\frac{\gamma}{mg(\rho_l - \rho_g)}\right]^{1/2} \quad (146)$$

We emphasize that (144) and (145) are valid when z_1 and z_2 lie in the interface where $\rho'(z_1)$ and $\rho'(z_2)$ are large. $\xi_\parallel$ is the classical capillary length; it is of macroscopic extent ($\sim 10^{-3}$ m) for argon at its triple point in the earth's gravity. The physical picture behind (143) is one of capillary-wave-like fluctuations developing in the interface where $\rho'(z)$ is largest. These fluctuations, which act to increase the surface area, are restored by the tension γ and are damped by gravity g. In the limit $g \to 0$, $\xi_\parallel$ diverges as $g^{-1/2}$, for any $T < T_c$, suggesting that the bulk coexistence curve acts as a line of critical points for transverse interfacial correlations.

It is important to recognize that the *form* of (143) to (146) can be derived from analysis of the exact equations for the density profile and surface

tension. These read (see Chapter 2)

$$\rho'(z_1) = -\beta \int_{-\infty}^{\infty} dz_2\, G_0(z_1,z_2)V'(z_2) \tag{147}$$

which is equivalent to (141) and

$$\gamma = -\beta \int_{-\infty}^{\infty} dz_1 \int_{-\infty}^{\infty} dz_2\, V'(z_1)V'(z_2)G_2(z_1,z_2) \tag{148}$$

which is equivalent to (137). $G_2(z_1,z_2)$ is the second transverse moment of G [see (139)]. If the tension γ is to remain nonzero as $g \to 0$, the integral $\iint dz_1\, dz_2\, G_2(z_1,z_2)$ must diverge as g^{-2}. Similarly, if $\rho'(z_1)$ is not to vanish too quickly (the precise details are left unspecified at this stage) $\int dz_2\, G_0(z_1,z_2)$ must diverge no faster than g^{-1}. It is clear that (144), with (145), satisfies the sum rules (147) and (148) identically; any higher-order terms must make vanishingly small contributions to the integrals. Wertheim [19] (see also Refs. 20 and 21) was the first to propose (143), solely on the basis of an eigenfunction analysis of correlations. His analysis does not determine the dependence of $\rho'(z)$ on g; it is valid whether $\rho'(z)$ is nonzero in the limit $g \to 0$ (the prediction of mean-field density functional approaches) or is slowly vanishing (the prediction from theories based on effective interfacial Hamiltonians). Weeks [20] first argued that the presence of long-range transverse correlations must drive the interface rough, $\xi_\perp \to \infty$, whereas the mean-field treatment implies that $\xi_\perp$ remains finite as $g \to 0$, leaving the interface smooth. The mean-field treatment would appear to allow insufficient feedback from the transverse correlations to the density profile for the latter to be smeared out. This issue has aroused much interest and continues to attract a certain amount of controversy. (References 16 and 46 by the present author contributed to the confusion!) Here we give a simple argument that illustrates the deficiency of the mean-field treatments and supports the predictions of those theories based on capillary-wave models (effective interfacial Hamiltonians).

We note first that the simplest density functional approximations yield a second moment $c_2(z_1,z_2)$ that is independent of the density profile and therefore of the external field. This conclusion is apparent in the square-gradient result (140) and is readily obtained from the van der Waals result (56) (see also Ref. 28). If such a $c_2(z_1,z_2)$ is employed in the Triezenberg–Zwanzig formula (137), the tension γ will vanish unless $\rho'(z)$ remains nonzero in the limit $g = 0$. This observation attests once more to the internal consistency of the density functional approach. The tension obtained from (137) or (148) will be identical to that obtained from the excess grand potential. For certain choices of functional this can be provided explicitly [152]. Improving the treatment of repulsive forces via a WDA

does not alter the essential mean-field character of $c^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$. In reality we expect the direct-correlation function and, in particular, its second moment $c_2(z_1,z_2)$ to depend on $\rho'(z)$. For the capillary-wave model in $d = 2$ this quantity is known explicitly [44]:

$$c_2(z_1,z_2) = \frac{\beta\gamma\delta(z_1 - z_2)}{(\rho_l - \rho_g)\rho'(z_1)} \tag{149}$$

which satisfies (137) identically. Weeks et al. [153] have recently attempted to calculate the dominant contribution to $c_2(z_1,z_2;Q)$ for the capillary-wave model in $d = 3$. They state that

$$c^{(2)}(z_1,z_2;Q) \approx \frac{\beta\gamma\delta(z_1 - z_2)Q^2}{(\rho_l - \rho_g)\rho'(z_1)} \tag{150}$$

for $Q \gg \xi_\parallel^{-1}$, a quantity that can be made arbitrarily small. Clearly, (150) has the same form as the $d = 2$ result, (149). There is no reason to expect (150) to provide an accurate description of $c^{(2)}$ in the interface of a real fluid since effective interfacial Hamiltonians can only describe the very long wavelength fluctuations associated with thermal wandering, not the intrinsic structure of the interface, which is on a short length scale $O(\xi_b)$, or the long-wavelength bulk density fluctuations [20,44,153]. It is clear that this result is of a totally different character from the corresponding mean-field result. Recall that $\rho'(z)^{-1} \sim W$ *diverges* in the capillary-wave model. Rather than focusing on $c^{(2)}$, however, we return to the mean-field prediction (143) for G and demonstrate that this implies an unbounded correlation function in real space.

Assuming that (143) is the dominant contribution for $Q < Q_{\max}$ with $Q_{\max} = \pi\xi_b^{-1}$, a short-wavelength cutoff, it follows that

$$G(z_1,z_2;R) \sim \frac{\rho'(z_1)\rho'(z_2)}{\beta\gamma(2\pi)^{d-1}} \int_0^{Q_{\max}} dQ\, \frac{e^{-i\mathbf{Q}\cdot\mathbf{R}}}{\xi_\parallel^{-2} + Q^2} \tag{151}$$

for z_1 and z_2 in the interface. For transverse separations $R \gg \xi_\parallel$ (151) predicts the familiar (Ornstein–Zernike) exponential decay of (transverse) correlations. For $\xi_b \ll R \ll \xi_\parallel$ the integral in (151) predicts $R^{-(d-3)}$ power-law decay for $d > 3$ and $\ln R$ behavior in $d = 3$. The latter is characteristic of two-dimensional Ornstein–Zernike behavior. It is well known that in *bulk,* such behavior leads to unbounded correlations near the two-dimensional critical point. A proper treatment of fluctuations leads to correlations decaying as $g(r) - 1 \sim r^{-(d-2+\eta)}$ with $\eta = \frac{1}{4}$ in $d = 2$. The situation is different in the interface. There it is possible to retain OZ behavior (143) for transverse correlations, that is, to maintain the analog of η equal to zero but still have finite values of G provided that the fluc-

tuations drive $\rho'(z)$ to zero in the limit $\xi_{\|} \to \infty$. More explicitly, we set $R = 0$; then

$$G(z_1,z_2;0) \sim \rho'(z_1)\rho'(z_2)\xi_\perp{}^2 \tag{152}$$

where the interfacial roughness $\xi_\perp$ is given by

$$\xi_\perp{}^2 \equiv \frac{1}{\beta\gamma(2\pi)^{d-1}} \int_0^{Q_{\max}} d\mathbf{Q}\, \frac{1}{\xi_\perp{}^{-2} + Q^2} \tag{153}$$

[In the capillary-wave model $\xi_\perp$ is defined as the interfacial width [44] since $\xi_\perp{}^2 = \langle h(\mathbf{R})h(\mathbf{R})\rangle_{\text{cw}}$ is the height–height autocorrelation function for this model.] The integral in (153) is well known:

$$\xi_\perp{}^2 \sim \begin{cases} \text{const. } \xi_{\|}{}^{3-d} & d < 3 \\ \omega\xi_b{}^2 \ln(\xi_{\|}/\xi_b)^2 & d = 3 \\ \text{const.} & d > 3 \end{cases} \tag{154}$$

with $\omega \equiv (4\pi\beta\gamma\xi_b{}^2)^{-1}$, a dimensionless constant. Thus in mean-field approximation, where $\rho'(z)$ is nonzero in the limit of vanishing field, (152) implies an unbounded correlation function in the interface for $d \leq 3$. If we retain the Wertheim form (143) and G is to remain bounded, consistent with its definition (4) as a density–density correlation function, the product $\rho'(z)\xi_\perp$ must remain finite. For the case where the height–height autocorrelation function diverges, this divergence must be compensated by a vanishing $\rho'(z)$; that is, in a vanishing field,

$$\rho'(z) \sim \xi_\perp{}^{-1} \to 0 \qquad d \leq 3 \tag{155}$$

for z in the interface. But this is precisely the result of the capillary-wave model, which invokes a relationship between fluctuations of height $h(\mathbf{R})$ of the dividing surface and density fluctuations in order to obtain $\rho(z)$ [44]. Our present argument predicts $\rho'(z) \sim (-\ln g)^{-1/2}$ in $d = 3$ without direct appeal to the capillary-wave model. In other words, the Wertheim form for G (143), which satisfies the exact sum rules and which corresponds to $\eta = 0$, necessarily implies interfacial broadening (155) in the limit $g \to 0$. Only for $d > 3$ does $\xi_\perp$ remain finite and then the mean-field prediction $\rho'(z) \neq 0$ should be correct.

The argument given above [154] complements the elegant argument of Weeks [155], based on a scaling hypothesis for G, which showed that the total interfacial width must diverge for $d < 3$ but was inconclusive in $d = 3$. Although our argument is in the same spirit as work of Robert [156], it differs in important detail; Robert sets the external field to zero at the outset and argues that $\rho'(z) \neq 0$ leads to a contradiction—a divergent surface tension. He is not able to derive (155), which is, of course, mean-

ingful only in nonzero but small field. Our present argument will not satisfy those seeking rigorous proofs but does have the merit of focusing on the failing of the mean-field theory. The latter has the correct OZ-like pair correlation function but omits the effects of feedback of long-range transverse correlations into the density profile that produce broadening. We might expect a similar scenario for other interfacial problems, such as the growth of wetting films, which also involve capillary-wave fluctuations. Indeed, we turn to this in Section V.B.

Finally, we wish to emphasize that the present discussion of fluctuations at the liquid–gas interface is in no sense complete. Several authors, starting with Weeks [20], have attempted to "marry" the treatments of interfacial and ordinary, bulk density fluctuations. Huse et al. [157] argued that incorporation of the wandering of the interface should not change the value of the critical exponent $\tilde{\mu}$ from the Fisk–Widom (scaling) value. As treated by effective interfacial Hamiltonians, interfacial fluctuations should correspond to wavelengths larger than ξ_b, which is itself diverging near T_c. Fluctuations on longer length scales should not be important for gross features of the critical behavior, such as exponents. Sengers and van Leeuwen [158] treat the near-critical interface as a fluctuating Fisk–Widom interface and propose a method for determining capillary-wave contributions to the surface tension and density profile in order to effect comparison with experimental reflectivity measurements and with renormalization-group results. We recommend this paper to readers interested in the (subtle) issues involved in the marriage of the intrinsic interface, which should be described well via density functional approaches, with treatments that focus on the interfacial fluctuations. The effects of gravity also become very important very close to T_c, where the compressibility is large. These have been analyzed in some detail [159].

B. Criticality at Wetting Transitions

In Section III we mentioned that there have been many applications of density functional methods to problems involving wetting phenomena. The excellent reviews [58,59] describe much of the work based on the local density approximation (50) for $\mathcal{F}_{hs}[\rho]$ and on the square-gradient approximation (39). In this volume Franck (Chapter 7) has reviewed experimental work and Henderson (Chapter 2) has discussed sum-rule analyses for continuous wetting transitions at planar wall–fluid interfaces. Here we merely add some remarks regarding the efficacy of current density functional approximations in accounting for criticality at wetting transitions. Given the mean-field character of such approximations it is tempting to suppose that they will be totally inadequate for describing the

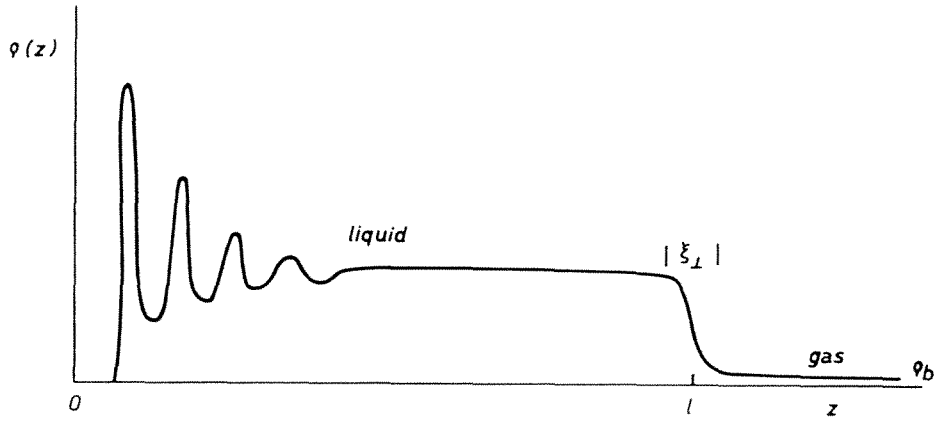


Fig. 7 Density profile of a wetting film (schematic). In the approach to complete or critical wetting, the film thickness l diverges and this is accompanied by a divergence in the transverse correlation length $\xi_{\parallel}$. The interfacial roughness $\xi_{\perp}$, and hence $\rho'(l)$, should also diverge for dimension $d \leq 3$. Note that a local density approximation will not describe the oscillations near the wall.

divergence of the wetting film thickness* l (see Fig. 7) and of the transverse correlation length $\xi_{\parallel}$, which characterize both complete wetting (from off-bulk coexistence) and critical wetting that occurs as $T \to T_w^-$, the wetting transition temperature, when the fluid is in bulk coexistence. That this is *not* necessarily the case was recognized first by Lipowsky [160] (see also Ref. 161). When fluid–fluid or substrate–fluid attractive forces are long range (i.e., exhibit power-law decay), the upper critical dimension d_c for both types of transition depends on the particular power law but is less than 3 (Chapter 2). Thus for real systems, governed by van der Waals forces, the critical exponents that describe the divergence of l and $\xi_{\parallel}$ should be given correctly by a mean-field theory, provided that the theory treats the attractive forces in a nonlocal fashion and not in the fashion of Cahn (see Section III.B). For example, the well-known result $l \sim (\mu_{sat} - \mu)^{-\beta_s}$, with $\beta_s = \frac{1}{3}$, for the growth of a thick wetting film of liquid from an undersaturated gas should remain valid beyond mean-field approximation. The effects of capillary-wave-like fluctuations in the liquid–gas edge of the growing film are insufficient, in dimension $d = 3$, to alter the critical exponents from their mean-field values when the relevant potentials are long range. This is an important bonus for the density func-

* Note that this thickness is denoted by t in Section IV of Chapter 2.

tional approach. If fluid–fluid and substrate–fluid forces are exponentially decaying or are of finite range, as one might employ in simulations, $d_c = 3$ for both types of wetting transition. Renormalization-group studies of effective interfacial Hamiltonians point to significant effects of fluctuations (see Chapter 2 and Refs. 58 and 59) at the marginal dimension $d = 3$ and there is no reason to suppose that mean-field (density functional) theory will yield the correct critical behavior. Although the critical exponents for complete wetting in $d = 3$ are predicted to retain their mean-field values, the amplitudes of l and $\xi_\parallel$ should be renormalized. More significantly, for critical wetting with short-range forces, the exponents are predicted to be nonuniversal and dependent on the liquid–gas surface tension.

Despite the shortcomings described above, simple density functional approximations *can* yield insight into the nature of criticality at wetting transitions. One of the fundamental issues concerns the form of the density–density correlation function $G(\mathbf{r}_1,\mathbf{r}_2)$ in a wetting film. This has been analyzed in some detail [84,85,152] for Sullivan's [162] model of a fluid adsorbed at a planar substrate. Using the local density approximation for $\mathcal{F}_{hs}[\rho]$ (50) and the mean-field treatment of attractive fluid–fluid forces (53), minimization of the grand potential functional yields (54), which reduces to

$$\mu = V(z_1) + \mu_{hs}(\rho(z_1)) + \int d\mathbf{r}_2\, \rho(z_2)\phi_{att}(|\mathbf{r}_1 - \mathbf{r}_2|) \tag{156}$$

when the substrate–fluid potential $V(\mathbf{r}) = V(z)$, so that the density profile $\rho(\mathbf{r}) \equiv \rho(z)$. Sullivan introduced a simple Yukawa form for the attractive fluid potential

$$\phi_{att}(r) = -\frac{\alpha\lambda^3 e^{-\lambda r}}{4\pi\lambda r} \tag{157}$$

where $\alpha = -\int d\mathbf{r}\, \phi_{att}(r)$ is the integrated strength and he assumed that

$$V(z) = \begin{cases} \infty & z < 0 \\ -\epsilon_w e^{-\lambda z} & z > 0 \end{cases} \tag{158}$$

with ϵ_w (>0) a strength parameter. The decay length λ^{-1} of both potentials is the same. Sullivan showed the model defined by (156) to (158) and the free-energy functional (53) has a critical wetting transition at a temperature T_w given by $\alpha\rho_l(T_w) = 2\epsilon_w$ and a critical drying transition at T_D given by $\alpha\rho_g(T_D) = 2\epsilon_w$. As $T \to T_w^-$, for $\mu = \mu_{sat}$, the thickness of the fluid film $l \sim -\ln(T_w - T)$ and $\xi_\parallel \sim (T_w - T)^{-1}$. Equivalent results pertain for the drying transition. For $T > T_w$ the wetting film thickens as $-\ln(\mu_{sat} - \mu)$ and $\xi_\parallel \sim (\mu_{sat} - \mu)^{-1/2}$. These exponents are the usual

mean-field results for short-range forces. The direct-correlation function $c^{(2)}$ of the inhomogeneous fluid has the simple form given in (56). Although it is not possible to effect a full inversion of the Ornstein–Zernike relation and obtain $G(z_1,z_2;Q)$ analytically, it is possible to derive results for the transverse moments $G_0(z_1,z_2)$ and $G_2(z_1,z_2)$ introduced in Section V.A [152]. When both particles lie in the edge of the film, $z_1,z_2 \sim l$, we find the same Ornstein–Zernike behavior (144) and (145) as that exhibited by the liquid–gas interface in a gravitational field, but now the transverse correlation length $\xi_{\|}$ diverges in the manner indicated above. The interfacial fluctuations are controlled by undersaturation $\mu_{sat} - \mu$ in the case of complete wetting or by the field $T_w - T$ for critical wetting. The mean-field character of the theory manifests itself in two ways: the exponents take their mean-field values, and $\rho'(l)$, the density derivative in the edge of the film (see Fig. 7), remains nonzero as $l \to \infty$ and the interface depins completely from the substrate. If we were to Fourier transform back to real space $G(l,l;R = 0)$ would be proportional to $(\rho'(l))^2\xi_\perp^2$ [see (151) to (153)], which is unbounded; the interfacial roughness $\xi_\perp$ still diverges as in (154). Once more the mean-field treatment implies an unbounded pair correlation function for $d \leq 3$. Nevertheless, as was the case for the free liquid–gas interface, we expect the Ornstein–Zernike *form* (144) and (145) to remain valid in an exact treatment of correlations (the analog of η should be zero) but with $\rho'(l)$ vanishing as $\xi_\perp^{-1}$. Similarly, we expect the density functional predictions for the *form* of the transverse moments, with z_1 or z_2 removed from the edge of the film, to be valid beyond mean field. The predictions for the moments are rather rich and have been shown [85,152] to be consistent with sum-rule analyses of both types of continuous wetting transition (see Chapter 2). For example, analysis of the Sullivan model yields [85] for the zeroth moment

$$G_0(z,0) = \frac{\rho'(z)\rho_w\chi_h(\rho_w)}{\rho_w - 2\alpha^{-1}\epsilon_w} \qquad z > 0 \tag{159}$$

where $\rho_w \equiv \rho(0^+)$ and $\chi_h(\rho) \equiv (\beta\rho\, d\mu_{hs}/d\rho)^{-1}$ is the susceptibility of a uniform hard-sphere fluid. In the approach to the critical wetting transition $\rho_w \to \rho_l$ and $\rho'(0) \to 0$ as the profile becomes flat. Then $G_0(z,0)$ diverges as $\rho'(z)(T_w - T)^{-1}$, when $z \sim l$ (i.e., when one particle is at the wall and the other is in the edge of the film). The sum-rule analysis [85] predicts that, in general, $G_0(l,0) \sim \rho'(l)(\partial l/\partial T)$ in the same limit. Inserting the mean-field results $\rho'(l) \neq 0$ and $l \sim -\ln(T_w - T)$ leads to the same divergence as the explicit density functional result.

In the special case of a hard wall ($\epsilon_w = 0$), (159) reduces to

$$G_0(z,0^+) = \rho'(z)\chi_h(\rho_w) \qquad z > 0 \tag{160}$$

which should be compared with the exact sum rule (Chapter 2) $G_0(z,0^+) = \rho'(z)$. The only discrepancy is the factor $\chi_h(\rho_w)$, which can be attributed [152] to the failure of the local density approximation to account for the exact sum rule $\beta p = \rho_w$. Within the Sullivan treatment the latter is replaced by $p_{hs}(\rho_w) = p$ when $\epsilon_w = 0$. For the hard wall it is also possible to prove [152] that when both particles are at the wall,

$$G_2(0^+,0^+) = -\beta\gamma\chi_h^2(\rho_w) \tag{161}$$

where γ is now the total wall–fluid interfacial tension. This result differs from the exact sum rule $G_2(0^+,0^+) = -\beta\gamma$ for the reason given above. A nonlocal WDA treatment of repulsive forces will remedy these defects and ensure that both sum rules are obeyed. The behavior of correlation functions for complete drying at a hard wall has been analyzed in some detail using the results from the Sullivan model [152]. Note that in complete wetting or drying $G_0(l,0^+)$ and $G_2(0^+,0^+)$ remain finite, while in critical wetting both quantities diverge. Analysis of the Sullivan model showed [85] that fluctuations manifest themselves in a much more pronounced fashion for critical wetting. Indeed, it was possible to calculate the transverse correlation length for two particles at the substrate, as defined by

$$\xi_\parallel^w \equiv \left[-\frac{G_2(0^+,0^+)}{G_0^{\text{sing}}(0^+,0^+)} \right]^{1/2} \tag{162}$$

where G_0^{sing} refers to the singular contribution due to the critical interface, and show that $\xi_\parallel^w = a_l \xi_\parallel$, with a_l a constant. Recall that $\xi_\parallel$ refers to two particles in the edge of the film:

$$\xi_\parallel = \left[-\frac{G_2(z_1,z_2)}{G_0(z_1,z_2)} \right]^{1/2} \qquad z_1,z_2 \sim l \tag{163}$$

That $\xi_\parallel^w$ diverges in the same fashion as $\xi_\parallel$ implies that transverse correlations are not localized in the depinning liquid–gas interface but extend all the way to the wall in the case of critical wetting. The local susceptibility $\chi(z) \equiv \beta^{-1}(\partial \ln \rho(z)/\partial\mu)_T$ provides another measure of the strength of the fluctuations. While $\chi(z)$ diverges as $\rho'(z)\xi_\parallel^2$, for $z \sim l$, at both types of wetting transition $\chi(0^+)$ remains finite at complete wetting but diverges as $(T_w - T)^{-1}$ or $(\partial l/\partial T)$ at critical wetting. We reiterate that all of these results are consistent with sum-rule analyses (Chapter 2).

Perhaps we should add a cautionary note here lest the reader becomes overexcited by tales of dramatic fluctuation behavior. A critical wetting transition has not yet been found in real experiments or in computer simulations for continuum fluids. However, new results, based on massive

molecular dynamics simulations, for a (truncated) Lennard–Jones fluid at a truncated Lennard–Jones substrate do appear to find some evidence for a continuous drying transition [163,164].

C. First-Order Phase Transitions in Adsorption at Substrate–Fluid Interfaces

Density functional techniques might seem better suited to the investigation of first-order phase transitions in inhomogeneous fluids, where the effects of fluctuations should not be of crucial importance. In adsorption problems the solid substrate is usually deemed to be rigid on relevant time scales and its boundary smooth and planar on relevant length scales so that it merely acts as a spectator phase exerting an external potential $V(\mathbf{r})$ on atoms in the fluid. The equilibrium properties of the adsorbed fluid can then be calculated using density functional methods. A first-order transition between two distinct adsorbed phases α and β occurs when their grand potentials are equal for given (μ,T). Equivalently, the surface excess grand potential per unit area $\omega^{ex}(\mu,T;V(\mathbf{r}))$, obtained by subtracting the bulk contribution to Ω, must be the same in each phase: $\omega_\alpha{}^{ex} = \omega_\beta{}^{ex}$. The amount adsorbed (the coverage) in each phase is given by the Gibbs adsorption equation $\Gamma = -(\partial\omega^{ex}/\partial\mu)_T$. Almost all density functional studies have employed substrate–fluid potentials for which $V(\mathbf{r}) \equiv V(z)$. Then $\Gamma = \int dz(\rho(z) - \rho_b)$, where $\rho_b(\mu,T)$ is the density of the bulk fluid (usually gas) far from the substrate.

Two types of transition have been investigated: (1) first-order wetting transitions and the accompanying prewetting and (2) layering transitions. We restrict ourselves to examples of some recent results obtained from calculations based on a WDA.

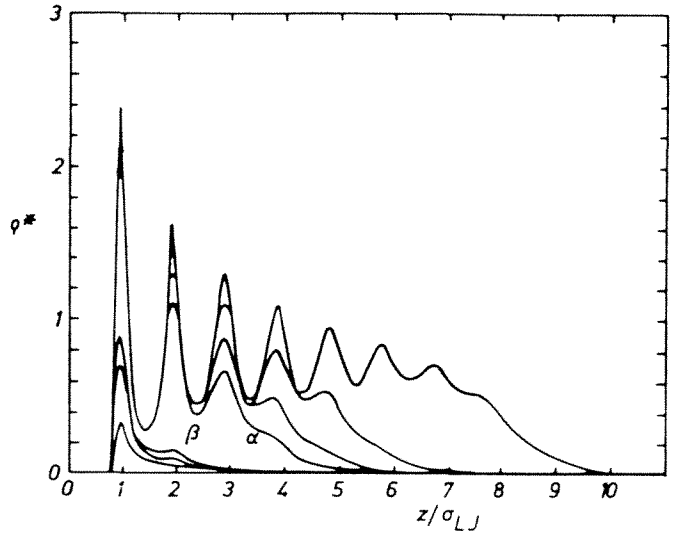
1. Prewetting at a Lennard–Jones 9–3 Substrate

In Section III.B we mentioned that Cahn [57] and Ebner and Saam [60] were the first to discover a first-order wetting transition and the associated prewetting that occurs out of bulk coexistence. The model system considered by Ebner and Saam was a Lennard–Jones 12–6 fluid at a 9–3 substrate, with parameters chosen to mimic argon at the rather weakly adsorbing solid CO_2 substrate. Prewetting is characterized by a discontinuous jump in the adsorption from a value Γ_β, corresponding to a thin adsorbed film, to a value Γ_α typical of a thick liquid film, at some value of chemical potential $\mu_{pw}(T) < \mu_{sat}(T)$. It occurs for T between the wetting transition temperature T_w and the prewetting critical temperature T_{sc}, where the distinction between thin and thick vanishes. Thus the prewetting line $\mu_{pw}(T)$ meets the bulk coexistence curve (tangentially) at T

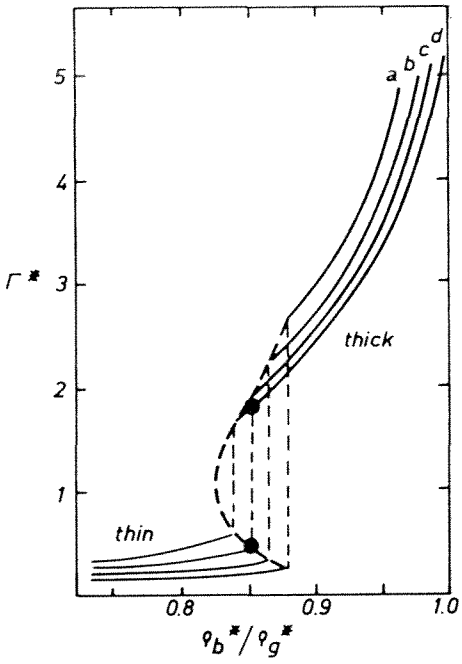
$= T_w$ and terminates at T_{sc}. Figure 8 shows the very recent density functional results of Velasco and Tarazona [97] for the prewetting transition in the model described above. Their version of WDA (Tarazona Mark II) was modified to include a temperature-dependent hard-sphere diameter and a parameter λ that controls the form of the effective attractive interaction. Velasco and Tarazona take ϕ_{att} in (53) to be

$$\phi_{att}(r) = \begin{cases} 0 & r \le \lambda^{1/6}\sigma_{LJ} \\ 4\epsilon\left[\lambda\left(\dfrac{\sigma_{LJ}}{r}\right)^{12} - \left(\dfrac{\sigma_{LJ}}{r}\right)^{6}\right] & \lambda^{1/6}\sigma_{LJ} < r \le r_c \\ 0 & r_c < r \end{cases} \quad (164)$$

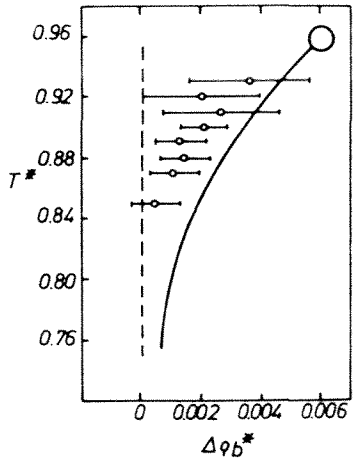
When $\lambda = 1$ ϕ_{att} reduces to the attractive part of the Lennard–Jones potential truncated at $r_c = 2.5\sigma_{LJ}$. By treating λ and the hard-sphere diameter as adjustable parameters, they are able to ensure that the bulk coexistence curve obtained from the theory is identical to that obtained from simulation of the truncated Lennard–Jones fluid. The density profiles in Fig. 8a show the jump in thickness of the adsorbed film. Note that the thick films are highly structured. In Fig. 8b adsorption isotherms are plotted for four different temperatures in the prewetting regime. The location of the prewetting line relative to the bulk coexistence curve is shown in Fig. 8c, where the density functional results are compared with those from the isobaric–isothermal Monte Carlo simulations of Finn and Monson [63] on the same system. The theoretical result for the critical temperature is $T^*_{sc} \equiv T_{sc}/\epsilon = 0.96 \pm 0.01$, while the estimate from simulation is $T^*_{sc} \equiv 0.94 \pm 0.02$. Theory also gives an accurate description of the jump in adsorption at $T^* = 0.88$ (see Fig. 8b). However, theory does appear to overestimate the undersaturation at which prewetting occurs; the transition occurs at lower pressures, relative to bulk saturation, than in simulation and this leads to a lower value for the wetting transition temperature. Velasco and Tarazona predict $T^*_w < 0.74$, which could be below the bulk triple-point temperature, whereas the Monte Carlo results [63] indicate that $T^*_w = 0.84 \pm 0.01$. Whether the differences between theory and simulation can be attributed to possible difficulties in determining an accurate bulk coexistence curve in Monte Carlo simulations is still a matter for debate [97,165]. Nevertheless, it is clear that the density functional approximation does provide a reasonably accurate description of what is a rather subtle and sensitive surface phase transition. Indeed, this example shows how density functional calculations can complement simulation studies. The original predictions for the prewetting transition were made on the basis of fairly simple theories. These predictions stimulated further calculations based on more sophisticated theories and de-


3
2
1
0
ϱ^*
β
α
0 1 2 3 4 5 6 7 8 9 10
z/σ_{LJ}

(a)


5
4
3
2
1
Γ^*
a
b
c
d
thick
thin
0.8 0.9 1.0
ϱ_b^*/ϱ_g^*

(b)


0.96
0.92
0.88
0.84
0.80
0.76
T^*
0 0.002 0.004 0.006
$\Delta\varrho_b^*$

(c)

tailed simulations. As remarked by Finn and Monson, "studying these (wetting) phenomena by Monte Carlo simulation is not yet a routine matter"; observe the error bars in Fig. 8c. By contrast, establishing the existence and location of first-order phase transitions by means of a density functional approach is relatively straightforward. Unlike in simulation, $\omega^{ex}(\mu,T;V)$ is calculated directly.

Before completing this discussion we should note that the nature of prewetting criticality is of special interest. Since the order parameter for the transition is the difference in film thickness (or $\Gamma_\alpha - \Gamma_\beta$) and only the transverse correlation length $\xi_\parallel$ can diverge (the film thickness remains finite at T_{sc}), any criticality should lie in the $d = 2$ Ising universality class. This conjecture has been confirmed by extensive Monte Carlo simulations, in conjunction with a finite-size scaling analysis, for a lattice-gas model that exhibits prewetting [166]. The resulting critical exponents are, to within statistical error, equal to the Onsager values. (It is certainly not feasible to perform an equivalent analysis for prewetting criticality in a continuum fluid.) Consequently, the density functional coexistence curve (dashed line) in Fig. 8b, which corresponds to $\Gamma_\alpha - \Gamma_\beta \sim (T_{sc} - T)^\beta$ with $\beta = \frac{1}{2}$, the mean-field result, is not as flat as the exact curve, which must have $\beta = \frac{1}{8}$.

2. *Layering Transitions in Continuum Fluids*

Layering transition is the name given to the discontinuous increase in adsorption, associated with the growth of a new adsorbed layer of a dense phase, that occurs when the (bulk) gas pressure, or chemical potential, is increased at fixed temperature. The existence of a series of such transitions at temperatures below the bulk triple point has been established

---

Fig. 8 Prewetting as described by WDA. (a) Density profiles for $T^* = 0.88$ and several values of the (reduced) bulk density: $\rho_b^* = 0.01290$, 0.01720, 0.01860, 0.01860, 0.01989, 0.02043, and 0.02150. The two profiles labeled β and α correspond to coexisting thin and thick films at the transition. Bulk liquid–gas coexistence occurs at $\rho_g^* = 0.02344$ for this temperature. (b) Adsorption isotherms for $T^* = 0.80$(a), 0.84(b), 0.88(c), and 0.92(d). The vertical lines indicate the phase transition for each temperature. The two points on the isotherm at $T^* = 0.88$ denote the values of Γ_β and Γ_α obtained for coexisting films in Monte Carlo simulations [63] of the transition at this temperature. (c) Prewetting temperature as a function of the undersaturation of the bulk: $\Delta\rho_b^*(T) \equiv \rho_g^*(T) - \rho_b^*(T)$. The solid line denotes the density functional results, with the large circle marking the prewetting critical point. The data with horizontal error bars are the Monte Carlo results of Finn and Monson [63]. (Redrawn from Ref. 97.)

in many adsorption experiments for rare gases and small molecules adsorbed on graphite. Some recent experiments, along with references to earlier work, are described in [167,168]. If solid-like layers develop, it is not too surprising that the growth takes place in discontinuous jumps. Lattice-gas models of adsorption [169,170] are known to predict an infinite sequence of first-order transitions provided that the substrate potential $V(z)$ is strongly attractive, so that $T_w = 0$ and the gas–substrate interface is wet completely by liquid. Each transition has its own critical temperature $T_{c,n}$, above which the new layers grows continuously with increasing μ. As $n \to \infty$, $T_{c,n}$ approaches the roughening temperature T_R of the lattice model. For $T > T_R$ there are no layering transitions. (If $T_w > T_R$, as occurs for less attractive substrates, layering is replaced by a single prewetting transition.) One might suspect that the discreteness of the transitions is imposed by the discreteness of the imposed lattice. On the other hand, it is clear that very pronounced layering can occur for continuum fluids near substrates (see Figs. 1 and 8a). Thus it is natural to inquire whether discrete *transitions* between layered *liquid* films can develop under favorable circumstances (i.e., for temperatures in the neighborhood of the bulk triple point T_{tr}). Although it should be impossible to have infinitely many transitions for $T > T_{tr}$, since the liquid–gas is always rough, can several transitions occur? This question was addressed by Ball and Evans [99], who used the Tarazona (Mark II) version of WDA to investigate adsorption for various model fluids and substrates. Figure 9 shows some of their results for a Yukawa fluid [see (157)] at a Yukawa wall [see (158)] with decay length $\lambda^{-1} = \sigma$, the hard-sphere diameter. The strength of the wall–fluid potential $\epsilon_w = 4k_B T_c$, where T_c is the critical temperature of the bulk fluid. Similar results are obtained with more realistic potentials [99]. At a low-temperature $T = 0.5T_c$, close to T_{tr}, the liquid wets the wall completely but the film grows layer by layer rather than continuously with μ. Each jump in adsorption (Fig. 9a) results from the addition of (roughly) one dense layer to the adsorbed film (see the density profiles in Fig. 9b). The transitions cluster together more closely as saturation is approached, so that it is increasingly difficult to distinguish individual transitions. In Fig. 9c the adsorption isotherm for $T = 0.6T_c$ is plotted. Only a few "bumps" remain as evidence of the transitions that occurred at lower temperatures. Each layering transition ends at a critical temperature, but these are calculated to be $\leq 0.6T_c$. Since the density functional theory is mean-field like, effects of roughening are omitted. We might expect the incorporation of fluctuation effects in the edge of the growing film to wash out some of the higher n transitions. Although the results in Fig. 9 certainly suggest that discrete transitions between layered liquid films should occur, they raise many queries [99]. Are the first one

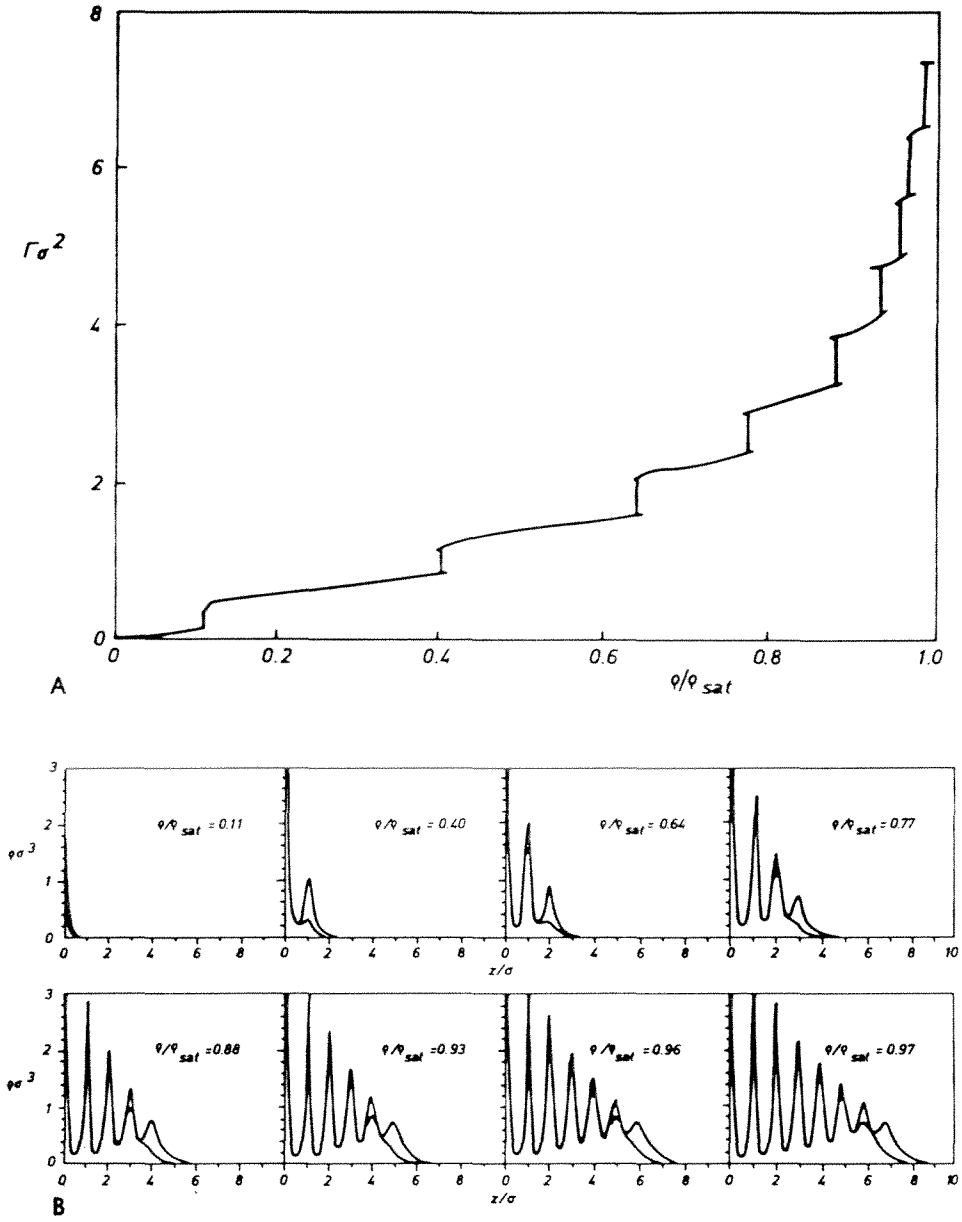


Fig. 9 Layering transitions for a Yukawa fluid at a Yukawa wall. (A) Adsorption isotherm for $T = 0.5T_c$. The first nine transitions are shown. Note that the jumps in Γ become successively larger and that metastable portions are always small. ρ/ρ_{sat} is the ratio of the bulk gas density to that at saturation. (B) Coexisting density profiles at the first eight layering transitions in (A). A new peak, corresponding to a new adsorbed layer, develops at each transition.

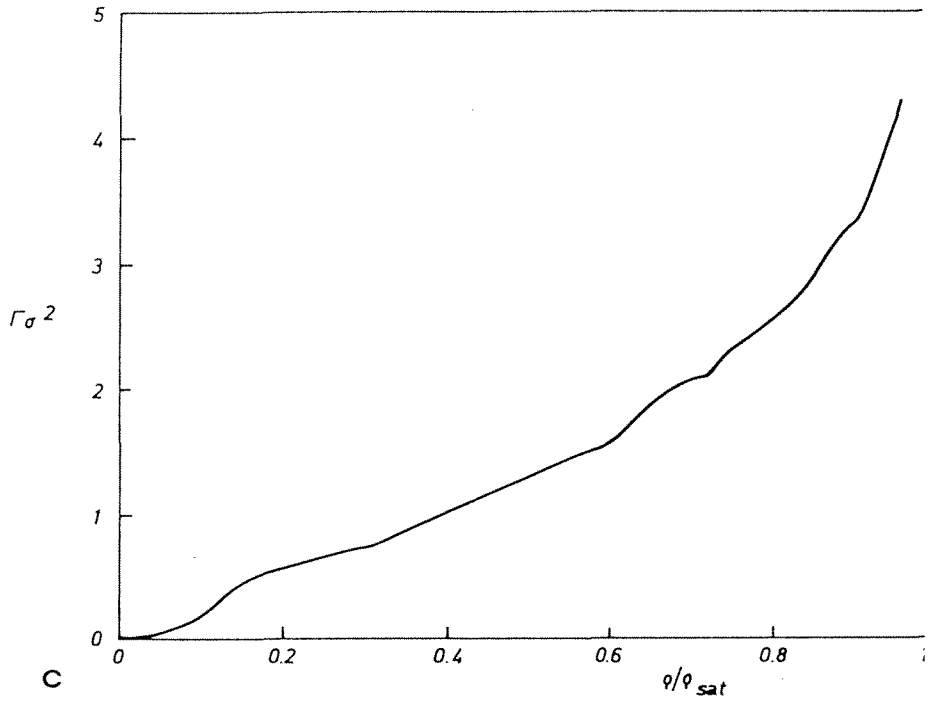


Fig. 9 (Continued) (C) Adsorption isotherm for $T = 0.6T_c$. All first-order transitions have disappeared, implying that the critical temperatures $T_{c,n}$ of (at least) the first nine transitions lie within the range $0.5T_c < T_{c,n} < 0.6T_c$. (Redrawn from Ref. 99.)

or two layers liquidlike or does spontaneous crystalline ordering (parallel to the wall) occur for such large local densities and deep substrate potentials? If so, what is the nature of the crystalline phases? In principle such a query could be answered within the context of the density functional approach [93], but this requires some delicate parametrization of $\rho(\mathbf{r})$. What is the locus of the line of layering critical points $T_{c,n}$? Is this line strongly influenced by T_R and T_{tr}? Does it depend sensitively on the form of the substrate–fluid potential? Simulation studies can provide answers and some important progress has been made [100].

Such issues attract much attention from experimentalists working on multilayer adsorption. There is a growing amount of evidence, for gases adsorbed on graphite, that transitions between layered liquid films do take place. Ethylene [171,172], ethane [173], and oxygen [174] are cases where up to seven or eight discrete steps in the adsorption have been resolved

in ellipsometric or heat-capacity studies for $T > T_{tr}$. Very recent ellipsometric data for argon [168] point to a sequence of first-order transitions for isotherms in the range $72 \leq T \leq 77$ K. These transitions, which correspond to the addition of an "ordered layer displacing the disordered surface outward by one layer," seem to be similar to those described in Fig. 9. However, the critical points $T_{c,4}$ to $T_{c,7}$ are at about 77 K, which lies below $T_{tr} = 83.8$ K. The surface phase diagram in this system is very rich [168], and a detailed interpretation of the various phases and their transitions remains a challenging task.

Finally, we should note that the criticality of layering transitions, like that of prewetting, should lie in the $d = 2$ Ising universality class. Whether fluctuation effects lead to significant departures of $T_{c,n}$ from their mean-field (density functional) values remains to be seen.

D. Phase Transitions in Confined Fluids

There have been many applications of the density functional approach to determining the microscopic structure and the phase equilibria of fluids confined in model pores or capillaries. Open cylinders and slits, with infinite interior interfacial area A, are the most popular models. Usually, one envisages a reservoir of fluid in contact with the pore so that fluid molecules can be adsorbed on the interior substrate of the cylinder and on the two interior substrates of the slit. Those substrates then exert a confining external potential $V(\mathbf{r})$ on the fluid. Minimization of the (approximate) grand potential functional $\Omega_V[\rho]$ yields the equilibrium density profile $\rho(\mathbf{r})$ and grand potential of the confined fluid, which is then a function of μ, T and the pore width or radius.

From the viewpoint of phase transitions, the new feature associated with adsorption in a pore, rather than at a single substrate (wall), is the presence of an extra thermodynamic field. For slit geometry an increment in grand potential is given by

$$d\Omega = -p\, dV - S dT - N\, d\mu + 2\gamma\, dA - Af\, dL \tag{165}$$

The standard (bulk) terms are augmented (see also Chapter 2) by the surface work term: 2γ is the total wall–fluid interfacial tension,* and the new term, which represents the work done when the wall separation L, is increased by an amount dL. The conjugate density to L is the solvation force f, which can be expressed as a pressure difference

$$f = -\frac{1}{A}\left(\frac{\partial \Omega}{\partial L}\right)_{\mu,T,A} - p \tag{166}$$

* Strictly, excess grand potential.

where $p(\mu,T)$ refers to the pressure in the reservoir. Only in the limit $L \to \infty$ does f vanish. In cylindrical geometry the interior radius R_c plays the same role as L but there is no simple analog for f.

The phase transition that first comes to mind for confined fluids is capillary condensation—the phenomenon whereby a gas at chemical potential $\mu < \mu_{sat}$ condenses to a dense, liquidlike phase that fills the pore. This corresponds to a shift of the bulk liquid–gas transition resulting from finite-size effects and the presence of attractive wall–fluid forces. Capillary condensation occurs when the walls favor liquid, so that the contact angle $\theta < \pi/2$. If the walls favor gas, $\theta > \pi/2$, capillary evaporation of liquid occurs for $\mu > \mu_{sat}$. Density functional calculations have provided much insight into the physics of capillary condensation, the nature of its criticality, and how it competes with surface phase transitions such as prewetting and layering, which can still occur for confined fluids. They have also formed a valuable tool for understanding various features of the solvation force for simple fluids. We do not review this subject here but refer the reader to a survey by the present author [71]. In Chapter 8 Lozado-Cassou reviews other theoretical work on confined fluids.

VI. CONCLUDING REMARKS

In this chapter we have reviewed many approximate density functional treatments of nonuniform classical fluids. We should therefore be in a position to answer the following query. Suppose that an enthusiastic graduate student comes to your office with an exciting idea about a new interfacial phenomenon. He or she wishes to develop some microscopic theory for this (equilibrium) problem and has heard that the density functional approach is *the* thing to use but is confused by the plethora of possible approximations. Which recipe or recipes do you recommend? This would be my advice. The simple van der Waals approximation (53) is always a good starting point. It will often indicate what *gross features* of the structure and phase equilibria are to be expected for the particular problem under consideration. Once these have been understood it is usually necessary to adopt a nonlocal treatment of repulsive forces that will incorporate realistic short-range correlations, packing effects, and so on. Naturally, this is crucial for problems that involve layering or require a very detailed knowledge of the microscopic structure of the fluid. Although the WDAs appear to be the most versatile, there are many of these. On the basis of the results and the discussion in Sections III.E and III.F, it is fair to conclude that two of the earlier recipes remain the most successful. (A glance through the list of references will indicate why the author should confess to being a little biased.) Tarazona (Mark II) has

been implemented for the widest variety of problems. Curtin–Ashcroft (CA) seems to lead to similar results but has been less widely applied and requires more computational effort. Although the more recent recipes attempt to improve on these earlier versions, it is not clear that they do—except in a few very specialized cases. (The situation is somewhat different for freezing, where approximations based on position-independent weight functions might be expected to be equally useful. The author offers no advice in this case.)

Compared with theories of bulk liquids, where modern integral equation approaches are embarrassingly accurate, the density functional approximations described here are quite primitive. It is likely that integral equation theories based on the inhomogeneous Ornstein–Zernike equation and some sophisticated closure approximation made at the pair level (see Chapter 4) will prove more accurate for the calculation of the one- and two-body distribution functions of nonuniform fluids. However, these theories have the disadvantage of (1) requiring major computational effort and (2) not yielding the free energy of the fluid directly; this makes their application to the calculation of surface tension and of phase equilibria problematical. Methods based on the variational principle for $\Omega_V[\rho]$ should continue to have the edge as regards the thermodynamic properties. We have emphasized that density functional calculations complement computer simulations, especially in the context of phase transitions. The determination of global phase diagrams via simulation can be a daunting task. The density functional results can point to what types of transition might occur and where they might be located.

What developments can be expected? There is no doubt that the search for more accurate free-energy functionals will continue. For hard-sphere fluids, or fluids with purely repulsive interatomic forces, the WDA does seem to capture the essential physics. Finding an improvement on the simple mean-field treatment of $\mathcal{F}_{\text{att}}[\rho]$ is necessary if a quantitative description of pairwise correlations is to emerge for more realistic fluids. Of course, this is also necessary if a theory is to incorporate the effects of capillary-wave fluctuations, or indeed, of any type of critical fluctuation in the nonuniform fluid. Further work on the structure of uniform liquids, in the spirit of Section IV, might shed some light on what does constitute an accurate approximation for the full functional $\mathcal{F}[\rho]$. However, it seems unlikely that theories based on a free-energy functional will be serious competitors to the modern integral equation theories of uniform liquids. What *is* likely is that existing density functional approximations will find much wider application. For example, work on the crystal–liquid interface, on nucleation (Chapter 10), and on the adsorption of simple *liquids* and their mixtures at substrates is still in its infancy. Techniques devel-

oped for simple fluids are already being extended successfully to more complex fluids. This process will certainly continue, guaranteeing that the subject remains ebullient. (*Note:* This chapter was submitted in November 1990. No attempt has been made to include new material.)

ACKNOWLEDGMENTS

I am grateful to those (many) colleagues and correspondents who have sent preprints and reprints of their papers. They should not be offended if I omitted to mention their work or if I failed to understand it. Correspondence with N. W. Ashcroft, M. Baus, and J. R. Henderson was particularly valuable. My views on the subject were influenced strongly by enjoyable and rewarding collaborations with P. C. Ball, G. P. Brenan, B. Q. Lu, U. Marini Bettolo Marconi, D. W. Oxtoby, A. O. Parry, T. J. Sluckin, P. Tarazona, M. M. Telo da Gama, and J. A. White. Pedro Tarazona must share some of the responsibility for the length of this chapter. He first persuaded me that there was physics beyond a local density approximation. This research was supported by S.E.R.C.

REFERENCES

1. Bogoliubov, N. N. (1946). *J. Phys. URSS 10:* 256, 265.
2. Morita, T., and Hiroike, K. (1961). *Prog. Theor. Phys. Osaka 25:* 537.
3. de Dominicis, C. (1962). *J. Math. Phys. 3:* 983.
4. Stillinger, F. H., and Buff, F. P. (1962). *J. Chem. Phys. 37:* 1.
5. Lebowitz, J. L., and Percus, J. K. (1963). *J. Math. Phys. 4:* 116.
6. Percus, J. K. (1964). In *The Equilibrium Theory of Classical Fluids* (H. L. Frisch and J. L. Lebowitz, eds.), W. A. Benjamin, New York, p. II 33.
7. Stell, G. (1964). In *The Equilibrium Theory of Classical Fluids* (H. L. Frisch and J. L. Lebowitz, eds.), W. A. Benjamin, New York, p. II 171.
8. Hohenberg, P., and Kohn, W. (1964). *Phys. Rev. 136:* B864.
9. Kohn, W., and Sham, L. J. (1965). *Phys. Rev. 140:* A1133.
10. Jones, R., and Gunnarsson, O. (1989). *Rev. Mod. Phys. 61:* 689.
11. Parr, R. G., and Yang, W. (1989). *Density Functional Theory of Atoms and Molecules*, Oxford University Press, Oxford.
12. Mermin, N. D. (1965). *Phys. Rev. 137:* A1441.
13. Ebner, C., Saam, W. F., and Stroud, D. (1976). *Phys. Rev. A 14:* 2264.
14. Saam, W. F., and Ebner, C. (1977). *Phys. Rev. A 15:* 2566.
15. Yang, A. J. M., Fleming, P. D., and Gibbs, J. H. (1976). *J. Chem. Phys. 64:* 3732.
16. Evans, R. (1979). *Adv. Phys. 28:*143.
17. Rowlinson, J. S., and Widom, B. (1982). *Molecular Theory of Capillarity*, Oxford University Press, Oxford.
18. Lovett, R. A., Mou, C. Y., and Buff, F. P. (1976). *J. Chem. Phys. 65:* 570.

19. Wertheim, M. S. (1976). *J. Chem. Phys. 65:* 2377.
20. Weeks, J. D. (1977). *J. Chem. Phys. 67:* 3106.
21. Kalos, M. H., Percus, J. K., and Rao, M. (1977). *J. Stat. Phys. 17:* 111.
22. Evans, R., and Sluckin, T. J. (1980). *Mol. Phys. 40:* 413. Telo da Gama, M. M., Evans, R., and Sluckin, T. J. (1980). *Mol. Phys. 41:* 1355. Sluckin, T. J. (1981). *J. Chem. Soc. Faraday Trans. 2 77:* 1029.
23. Alastuey, A. (1984). *Mol. Phys. 52:* 637.
24. Mier-y-Teran, L., Suh, S. H., White, H. S., and Davis, H. T. (1990). *J. Chem. Phys. 92:* 5087. Tang, Z., Mier-y-Teran, L., Davis, H. T., Scriven, L. E., and White, H. S. (1990). *Mol. Phys. 71:* 369.
25. Sluckin, T. J. (1981). *Mol. Phys. 43:* 817.
26. Woodward, C. E., and Nordholm, S. (1986). *Mol. Phys. 59:* 1177, 1201.
27. Evans, R., and Hasegawa, M. (1981). *J. Phys. C 14:* 5225.
28. Evans, R. (1990). In *Liquids at Interfaces*, Les Houches Session XLVIII (J. Charvolin, J. F. Joanny, and J. Zinn-Justin, eds.), Elsevier, Amsterdam, p. 1.
29. Chayes, J. T., Chayes, L., and Lieb, E. H. (1984). *Commun. Math. Phys. 93:* 57. Chayes, J. T., and Chayes, L. (1984). *J. Stat. Phys. 36:* 471.
30. Hansen, J.-P., and McDonald, I. R. (1986). *Theory of Simple Liquids*, Academic Press, New York.
31. Amit, D. J. (1978). *Field Theory: The Renormalisation Group and Critical Phenomena*, McGraw-Hill, New York.
32. Sullivan, D. E. (1982). *Phys. Rev. A 25:* 1669.
33. Percus, J. K. (1976). *J. Stat. Phys. 15:* 505.
34. Robledo, A. (1980). *J. Chem. Phys. 72:* 1701.
35. Robledo, A., and Varea, C. (1981). *J. Stat. Phys. 26:* 513.
36. Percus, J. K. (1982). *J. Stat. Phys. 28:* 67.
37. Robledo, A., and Rowlinson, J. S. (1986). *Mol. Phys. 58:* 711.
38. Davis, H. T. (1990). *J. Chem. Phys. 93:* 4339.
39. Vanderlick, T. K., Davis, H. T., and Percus, J. K. (1989). *J. Chem. Phys. 91:* 7136.
40. Monson, P. A. (1990). *Mol. Phys. 70:* 401.
41. Percus, J. K. (1981). *J. Chem. Phys. 75:* 1316.
42. Percus, J. K. (1986). *J. Stat. Phys. 42:* 921.
43. Weeks, J. D., Bedeaux, D., and Zielinska, B. J. A. (1984). *J. Chem. Phys. 80:* 3790.
44. Bedeaux, D., and Weeks, J. D. (1985). *J. Chem. Phys. 82:* 972.
45. Bedeaux, D. (1985). In *Fundamental Problems in Statistical Mechanics* (E. G. D. Cohen, ed.), North-Holland, Amsterdam, Vol. 6, p. 125.
46. Evans, R. (1981). *Mol. Phys. 42:* 1169.
47. Fisk, S., and Widom, B. (1969). *J. Chem. Phys. 50:* 3219.
48. Enderby, J. E., Gaskell, T., and March, N. H. (1965). *Proc. Phys. Soc. 85:* 217.
49. Evans, R., and Sluckin, T. J. (1981). *J. Phys. C 14:* 2569.
50. Barker, J. A., and Henderson, J. R. (1982). *J. Chem. Phys. 76:* 6303 and references therein.

51. Lu, B. Q., Evans, R., and Telo da Gama, M. M. (1985). *Mol. Phys. 55:* 1319.
52. de Gennes, P. G. (1981). *J. Phys. Lett. Paris 42:* L377.
53. Telo da Gama, M. M., and Evans, R. (1979). *Mol. Phys. 38:* 367.
54. Davis, H. T., and Scriven, L. E. (1982). *Adv. Chem. Phys. 49:* 357.
55. Telo da Gama, M. M., and Evans, R. (1980). *Mol. Phys. 41:* 1091. (1981). *J. Chem. Soc. Faraday Symp. 16:* 45.
56. Telo da Gama, M. M., and Evans, R. (1983). *Mol. Phys. 48:* 229, 251.
57. Cahn, J. W. (1977). *J. Chem. Phys. 66:* 3667.
58. Sullivan, D. E., and Telo da Gama, M. M. (1986). In *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, New York, p. 45.
59. Dietrich, S. (1988). In *Phase Transitions and Critical Phenomena* (C. Domb and J. L. Lebowitz, eds.), Academic Press, New York, Vol. 12, p. 1.
60. Ebner, C., and Saam, W. F. (1977). *Phys. Rev. Lett. 38:* 1486.
61. Saam, W. F., and Ebner, C. (1978). *Phys. Rev. A 17:* 1768.
62. Evans, R. (1990). In *Liquids at Interfaces*, Les Houches Session XLVIII (J. Charvolin, J. F. Joanny and J. Zinn-Justin, eds.), Elsevier, Amsterdam, p. 1.
63. Finn, J. E., and Monson, P. A. (1989). *Phys. Rev. A 39:* 6402.
64. Grimson, M. J., and Rickayzen, G. (1981). *Mol. Phys. 42:* 47.
65. Evans, R., Tarazona, P., and Marini Bettolo Marconi, U. (1983). *Mol. Phys. 50:* 993.
66. Evans, R., and Marini Bettolo Marconi, U. (1986). *Phys. Rev. A 34:* 3504.
67. Henderson, J. R. (1984). *Mol. Phys. 52:* 1467.
68. Rickayzen, G., and Augousti, A. (1984). *Mol. Phys. 52:* 1355.
69. Powles, J. G., Rickayzen, G., and Williams, M. L. (1988). *Mol. Phys. 64:* 33.
70. Moradi, M., and Rickayzen, G. (1989). *Mol. Phys. 66:* 143.
71. Evans, R. (1990). *J. Phys. Condens. Matter 2:* 8989.
72. Tarazona, P., Marini Bettolo Marconi, U., and Evans, R. (1987). *Mol. Phys. 60:* 573.
73. Zhou, Y., and Stell, G. (1990). *J. Chem. Phys. 92:* 5533, 5544, and references therein.
74. Kjellander, R., and Sarman, S. (1990). *Mol. Phys. 70:* 215.
75. Smithline, S. J., and Rosenfeld, Y. (1990). *Phys. Rev. A 42:* 2434.
76. Curtin, W. A. (1988). *J. Chem. Phys. 88:* 7050.
77. Weeks, J. D., Chandler, D., and Anderson, H. C. (1971). *J. Chem. Phys. 54:* 5237, 5422.
78. Toxvaerd, S. (1971). *J. Chem. Phys. 55:* 3116.
79. Carnahan, N. F., and Starling, K. E. (1969). *J. Chem. Phys. 51:* 635.
80. Fischer, J., Heinbuch, U., and Wendland, M. (1987). *Mol. Phys. 61:* 953.
81. Stell, G. (1977). In *Statistical Mechanics*, Part A, Vol. 5 of *Modern Theoretical Chemistry* (B. J. Berne, ed.), Plenum Press, New York, p. 47.
82. Tarazona, P., and Evans, R. (1983). *Mol. Phys. 48:* 799.
83. Tarazona, P., Telo da Gama, M. M., and Evans, R. (1983). *Mol. Phys. 49:* 283, 301.

84. Tarazona, P., and Evans, R. (1982). *Mol. Phys. 47:* 1033.
85. Evans, R., and Parry, A. O. (1989). *J. Phys. Condens. Matter 1:* 7207.
86. Nordholm, S., Johnson, M., and Freasier, B. C. (1980). *Aust. J. Chem. 33:* 2139. Johnson, M., and Nordholm, S. (1981). *J. Chem. Phys. 75:* 1953, and references therein.
87. Freasier, B. C., and Nordholm, S. (1983). *J. Chem. Phys. 79:* 4431.
88. Freasier, B. C., and Nordholm, S. (1986). *Mol. Phys. 54:* 33.
89. van Kampen, N. G. (1964). *Phys. Rev. 135:* 362.
90. Tarazona, P. (1984). *Mol. Phys. 52:* 81.
91. Tarazona, P., and Evans, R. (1984). *Mol. Phys. 52:* 847.
92. van Swol, F., and Henderson, J. R. (1989). *Phys. Rev. A 40:* 2567.
93. Mederos, L., Tarazona, P., and Navascués, G. (1987). *Phys. Rev. B 35:* 3376, 3384.
94. Tarazona, P. (1985). *Phys. Rev. A 31:* 2672. Erratum, *Phys. Rev. A 32:* 3148.
95. Bruno, E., Caccamo, C., and Tarazona, P. (1987). *Phys. Rev. A 35:* 1210.
96. Velasco, E., and Tarazona, P. (1989). *J. Chem. Phys. 91:* 7916.
97. Velasco, E., and Tarazona, P. (1990). *Phys. Rev. A 42:* 2454.
98. Henderson, J. R., and van Swol, F. (1990). *J. Phys. Condens. Matter 2:* 4537.
99. Ball, P. C., and Evans, R. (1988). *J. Chem. Phys. 89:* 4412.
100. Peterson, B. K., Heffelfinger, G. S., Gubbins, K. E., and van Swol, F. (1990). *J. Chem. Phys. 93:* 679.
101. Ball, P. C., and Evans, R. (1988). *Mol. Phys. 63:* 159.
102. Peterson, B. K., Gubbins, K. E., Heffelfinger, G. S., Marini Bettolo Marconi, U., and van Swol, F. (1988). *J. Chem. Phys. 88:* 6487.
103. Vanderlick, T. K., Scriven, L. E., and Davis, H. T. (1989). *J. Chem. Phys. 90:* 2422.
104. Velasco, E., and Tarazona, P. (1987). *Phys. Rev. A 36:* 979.
105. Heffelfinger, G. S., Tan, Z., Gubbins K. E., Marini Bettolo Marconi, U., and van Swol, F. (1989). *Mol. Simul. 2:* 393.
106. Tan Z., Marini Bettolo Marconi, U., van Swol, F., and Gubbins, K. E. (1989). *J. Chem. Phys. 90:* 3704.
107. Curtin, W. A., and Ashcroft, N.W. (1985). *Phys. Rev. A 32:* 2909.
108. Curtin, W. A. (1987). *Phys. Rev. Lett. 59:* 1228.
109. Curtin, W. A. (1989). *Phys. Rev. B 39:* 6775.
110. Kroll, D. M., and Laird, B. (1990). *Phys. Rev. A 42:* 4806.
111. Curtin, W. A., and Ashcroft, N. W. (1987). *Phys. Rev. Lett. 59:* 2385.
112. Curtin, W. A., and Ashcroft, N. W. (1986). *Phys. Rev. Lett. 56:* 2775.
113. Denton, A. R., and Ashcroft, N. W. (1990). *Phys. Rev. A 42:* 7312.
114. Meister, T. F., and Kroll, D. M. (1985). *Phys. Rev. A 31:* 4055.
115. Groot, R. D. (1987). *Mol. Phys. 60:* 45.
116. Sokolowski, S., and Fischer, J. (1990). *Mol. Phys. 70:* 1097.
117. Sokolowski, S., and Fischer, J. (1990). *Mol. Phys. 71:* 393.
118. Groot, R. D., and van de Eerden, J. P. (1987). *Phys. Rev. A 36:* 4356.
119. Denton, A. R., and Ashcroft, N. W. (1989). *Phys. Rev. A 39:* 4701.

120. Sokolowski, S., and Fischer, J. (1989). *Mol. Phys. 68:* 647.
121. Barrat, J. L., Hansen, J.-P., and Pastore, G. (1988). *Mol. Phys. 63:* 747.
122. Percus, J. K. (1988). *J. Stat. Phys. 52:* 1157 (short review of free-energy density models).
123. Fischer, J., and Heinbuch, U. (1988). *J. Chem. Phys. 88:* 1909.
124. Fischer, J., and Methfessel, M. (1980). *Phys. Rev. A 22:* 2836.
125. Davis, H. T. (1986). *J. Chem. Phys. 85:* 6808.
126. Rosenfeld, Y. (1989). *Phys. Rev. Lett. 63:* 980.
127. Rosenfeld, Y. (1988). *J. Chem. Phys. 89:* 4272.
128. Reiss, H., Frisch, H., and Lebowitz, J. L. (1959). *J. Chem. Phys. 31:* 369.
129. Rosenfeld, Y., Levesque, D., and Weis, J.-J. (1990). *J. Chem. Phys. 92:* 6818.
130. Kierlik, E., and Rosinberg, M. L. (1990). *Phys. Rev. A 42:* 3382.
131. White, J. A., and Evans, R. (1990). *J. Phys. Condens. Matter 2:* 2435.
132. Kim, S.-C., and Jones, G. L. (1990). *Phys. Rev. A 41:* 2222.
133. Denton, A. R., and Ashcroft, N. W. (1990). *Phys. Rev. A 41:* 2224.
134. Zeng, X. C., and Oxtoby, D. W. (1990). *J. Chem. Phys. 93:* 4357.
135. Ohnesorge, R., Löwen, H., and Wagner, H. (1991). *Phys. Rev. A 43:* 2870.
136. Laird, B., and Kroll, D. M. (1990). *Phys. Rev. A 42:* 4810.
137. Baus, M. (1987). *J. Stat. Phys. 48:* 1129 (describes some of the earlier effective liquid approximations).
138. Lutsko, J. F., and Baus, M. (1990). *Phys. Rev. Lett. 64:* 761. (1990). *Phys. Rev. A 41:*6647.
139. Baus, M. (1989). *J. Phys. Condens. Matter 1:* 3131.
140. Zeng, X. C., and Oxtoby, D. W. (1990). *Phys. Rev. A 41:* 7094.
141. Denton, A. R., and Ashcroft, N. W. (1989). *Phys. Rev. A 39:* 426.
142. Kim, S.-C. (1991). *Phys. Rev. A 43:* 4328.
143. Jones, G. L., and Kim, S.-C. (1989). *J. Stat. Phys. 56:* 709.
144. Denton, A. R., and Ashcroft, N. W. (1991). *Phys. Rev. A 44:* 1219.
145. Brenan, G. P., and Evans R. (1991). *Mol. Phys 73:* 789. These authors also calculate B(r).
146. Freasier, B. G., Woodward, C. E., and Nordholm, S. (1989). *J. Chem. Phys. 90:* 5657.
147. Takamiya, M., and Nakanishi, K. (1990). *Mol. Phys. 70:* 767.
148. Yang, A. J. M., Fleming, P. D., and Gibbs, J. H. (1977). *J. Chem. Phys. 67:* 74.
149. Triezenberg, D. G., and Zwanzig, R. (1972). *Phys. Rev. Lett. 28:* 1183.
150. Buff, F. P., Lovett, R. A., and Stillinger, F. H. (1965). *Phys. Rev. Lett. 15:* 621.
151. Sikkenk, J. H., Hillhorst, H. J., and Bakker, A. F. (1985). *Physica A 131:* 587.
152. Parry, A. O., and Evans, R. (1988). *Mol. Phys. 65:* 455.
153. Weeks, J. D., van Saarloos, W., Bedeaux, D., and Blokhuis, E. (1989). *J. Chem. Phys. 91:* 6494.
154. Parry, A. O. (1989). Ph.D. thesis, University of Bristol (unpublished).
155. Weeks, J. D. (1984). *Phys. Rev. Lett. 52:* 2160.

156. Robert, M. (1985). *Phys. Rev. Lett. 54:* 444. (1985). *Phys. Rev. A 32:* 500.
157. Huse, D. A., van Saarloos, W., and Weeks, J. D. (1985). *Phys. Rev. B 32:* 233.
158. Sengers, J. V., and van Leeuwen, J. M. J. (1989). *Phys. Rev. A 39:* 6346.
159. van Leeuwen, J. M. J., and Sengers, J. V. (1989). *Physica A 157:* 839 (one of the latest of a series on this topic by the same authors).
160. Lipowsky, R. (1984). *Phys. Rev. Lett. 52:* 1429. (1985). *Phys. Rev. B 32:* 1731.
161. Evans, R., and Tarazona, P. (1984). *Phys. Rev. Lett. 53:* 400.
162. Sullivan, D. E. (1979). *Phys. Rev. B 20:* 3991. (1981). *J. Chem. Phys. 74:* 2604.
163. Nijmeijer, M. J. P., Bruin, C., Bakker, A. F., and van Leeuwen, J. M. J. (1991). *Phys. Rev. B 44:* 834.
164. Nijmeijer, M. J. P. (1990). Ph.D. thesis, University of Leiden (unpublished).
165. Finn, J. E., and Monson, P. A. (1990). *Phys. Rev. A 42:* 2458.
166. Nicolaides, D. B., and Evans, R. (1989). *Phys. Rev. Lett. 63:* 778.
167. Ser, S., Larher, Y., and Gilquin, B. (1989). *Mol. Phys. 67:* 1077.
168. Youn, H. S., and Hess, G. B. (1990). *Phys. Rev. Lett. 64:* 918.
169. Oliveira, M. J., and Griffiths, R. B. (1978). *Surf. Sci. 71:* 689.
170. Ebner, C. (1980). *Phys. Rev. A 22:* 2776. Ebner, C. (1981). *Phys. Rev. A 23:* 1925. Kim, I. M., and Landau, D. P. (1981). *Surf. Sci. 110:* 415.
171. Drir, M., Nham, H. S., and Hess, G. B. (1986). *Phys. Rev. B 33:* 5145.
172. Kim, H. K., Feng, Y. P., Zhang, Q. M., and Chan, M. H. W. (1988). *Phys. Rev. B 37:* 3511.
173. Nham, H. S., and Hess, G. B. (1988). *Phys. Rev. B 38:* 5166.
174. Drir, M., and Hess, G. B. (1986). *Phys. Rev. B 33:* 4758.

4

Integral Equation Theories for Inhomogeneous Fluids

Douglas Henderson

Utah Supercomputing Institute/IBM Partnership and University of Utah
Salt Lake City, Utah

I. INTRODUCTION: BULK FLUIDS

Bulk fluids are considered first. Integral equations have been employed to calculate the structure and thermodynamic properties of bulk liquids and dense fluids for the past six decades or so. Most of the progress has occurred in the past three decades, during which perturbation theory and integral equations based on the Ornstein–Zernike equation have been developed.

The integral equation approach, although more difficult numerically than perturbation theory, has the advantage of wider applicability. Perturbation theory has been very successful for simple dense fluids composed of spherical molecules interacting with van der Waals forces, but has not been notably successful for molecules fluids composed of nonspherical molecules or for fluids in which there are density inhomogeneities. The integral equation approach is applicable to all of these systems. For a comprehensive review of the theory of liquids, see Barker and Henderson (1976) or Hansen and McDonald (1986).

The integral equation theories give equations for correlation functions. The h-particle correlation function as defined by

$$g^{(h)}(\mathbf{r}_1 \cdots \mathbf{r}_h) = \frac{\rho^{(h)}(\mathbf{r}_1 \cdots \mathbf{r}_h)}{\rho_b{}^h} = \frac{N!}{(N-h)!} \frac{\int e^{-\beta\Phi}\, d\mathbf{r}_{h+1} \cdots d\mathbf{r}_N}{\int e^{-\beta\Phi}\, d\mathbf{r}_1 \cdots \mathbf{r}_N} \quad (1)$$

where the bulk density $\rho_b = N/V$, N and V being, respectively, the number of molecules in the fluid and the volume of the fluid, $\beta = 1/kT$ (k is Boltzmann's constant and T is the temperature), and $\Phi(\mathbf{r}_1 \cdots \mathbf{r}_N)$ is the total potential energy of the system. The generalization of (1) to mixtures is quite straightforward. The case of the pair correlation functions, $g(r_1, r_2)$, is of particular interest. The superscript 2 is dropped for notational convenience.

For simplicity, it has been assumed that the correlation functions are functions of position. For molecular fluids, the correlation functions are functions of position and orientation. Equation (1) still applies to molecular fluids. However, $\mathbf{r}_i$ should be interpreted as a generalized position of the ith molecule including spatial and orientation variables. In actual fact there have been few calculations for inhomogeneous molecular fluids, so that, in general, $\mathbf{r}_i$ will be interpreted as the spatial position of molecule i.

The potential energy can be written as the sum of the *pair* energies, the *triplet* energies, and so on. In this chapter, triplet and higher-order terms will not be considered. Thus

$$\Phi(\mathbf{r}_1 \cdots \mathbf{r}_N) = \sum_{1<j=1}^{N} u(\mathbf{r}_i, \mathbf{r}_j) \tag{2}$$

If (2) is satisfied, the system can be said to be *pairwise additive*. In most applications, the pair energy depends only on the separation of the center of mass of the two molecules, so that the pair potential can be written as $u(r_{ij})$, where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Similarly, the pair correlation function for such a fluid would be $g(r_{12})$.

For the systems considered in this chapter, model potentials will be used. Two widely used model potentials are the hard-sphere potential

$$u(r) = \begin{cases} \infty & r < d \\ 0 & r < d \end{cases} \tag{3}$$

and the Lennard–Jones 6–12 potential

$$u(r) = 4\epsilon_{LJ}\left[\left(\frac{\sigma_{LJ}}{r}\right)^{12} - \left(\frac{\sigma_{LJ}}{r}\right)^{6}\right] \tag{4}$$

For bulk fluids, thermodynamic properties can be calculated from the partition function, the denominator in (1), or from $g(r)$. Thus the equation of state calculated from

$$kT\left(\frac{\partial \rho_b}{\partial p}\right)_T = 1 + \rho_b \int h(r)\, d\mathbf{r} \tag{5}$$

where $h(r) = g(r) - 1$, is called the *total correlation function*,

$$\frac{p}{\rho_b kT} = 1 - \frac{1}{6}\beta\rho \int r \frac{\partial u(r)}{\partial r} g(r)\, d\mathbf{r} \tag{6}$$

or

$$\beta U = \tfrac{3}{2} + \tfrac{1}{2}\beta\rho \int u(r)g(h)\, d\mathbf{r} \tag{7}$$

where p and U are the pressure and internal energy, respectively. Equations (5), (6), and (7) are called the *compressibility*, *pressure* (or *virial*), and *energy equations*, respectively. Equations (6) and (7) are valid only for pairwise additively, whereas (5) is general.

The emphasis in this chapter is on the calculation of correlation functions by means of integral equations. One such integral equation can be obtained by differentiating $g^{(h)}(\mathbf{r}_2 \cdots \mathbf{r}_h)$ with respect to $\mathbf{r}_1$. Using (1) and assuming that only pair potentials contribute to Φ yields

$$-kT\nabla_1 g^{(h)}(1\cdots h) \sum_{j=2}^{h} g^{(h)}(1\cdots h)\nabla_1 u(1j) + \rho_h \int g(1\cdots h+1)\nabla_1 u(1,h+1)\, d\mathbf{r}_{h+1} \tag{8}$$

where $g^{(h)}(1\cdots h) = g^{(h)}(\mathbf{r}_1 \cdots \mathbf{r}_h)$, and so on. For $h = 2$, (8) can be rewritten as

$$-kT \frac{\partial \ln g(12)}{\partial r_{12}} = \frac{\partial u(12)}{\partial r_{12}} + \rho_b \int \frac{g^{(3)}(123)}{g(12)} \frac{r_{12}^2 + r_{13}^2 - r_{23}^2}{2r_{12}r_{13}} \frac{\partial u(13)}{\partial r_{13}} d\mathbf{r}_3 \tag{9}$$

Equation (9) is usually referred to as the *first Born–Green (BG) equation* (Born and Green, 1946).

To solve (9), $g^{(3)}(123)$ must be approximated. The best known and simplest approximation is the *superposition approximation* (Kirkwood, 1935),

$$g^{(3)}(123) = g(12)g(13)g(23) \tag{10}$$

This approximation is exact only in the limit of vanishing density. A similar-looking approximation is

$$g^{(3)}(ijk) = g(ij)g(jk) \tag{11}$$

where r_{ij} and r_{jk} are both smaller than r_{ik}. Equation (11) is not exact at any density. However, it seems to be useful for hard spheres when the centers of the three spheres form a straight line.

As mentioned, the superposition approximation (SA) is valid at low

densities. It may be corrected systematically by means of an expression in powers of ρ_b. Thus

$$g(123) = g(12)g(13)g(23)\left[1 + \rho_b \int f(14)f(24)f(34)\, d\mathbf{r}_4 + \cdots\right] \quad (12)$$

where

$$f(ij) = \exp[-\beta\mu(ij)] - 1 \quad (13)$$

Now

$$h(ij) = f(ij) + \rho_b[1 + f(ij)] \int f(ik)f(jk)\, d\mathbf{r}_k + \cdots \quad (14)$$

Using this, (14) can be rewritten with a *renormalized* integand, giving

$$g(123) = g(12)g(13)g(23)\left[1 + \rho_b \int h(14)h(24)h(34)\, d\mathbf{r}_4 + \cdots\right] \quad (15)$$

or

$$g(123) = g(12)g(13)g(23)\exp\left[\rho_b \int h(14)h(24)h(34)\, d\mathbf{r}_4 + \cdots\right] \quad (16)$$

In addition to approaches based on the BG equation, a second class of approaches can be formulated using the *Ornstein–Zernike* (OZ) *equation* (Ornstein and Zernike, 1914):

$$h(12) = c(12) + \rho_b \int h(13)c(23)\, d\mathbf{r}_3 \quad (17)$$

or

$$h = c + \rho_b h*c \quad (18)$$

where $c(r)$ is the *direct correlation function*. The integral in (17) is called a *convolution*. The asterisk in (18) denotes a convolution integral. The OZ equation is nothing but a definition of $c(r)$. If $h(r)$ is known, $c(r)$ can be calculated. Although the relation between $c(r)$ and $h(r)$ is complex in r space, it is simple in Fourier space, where (17) becomes

$$\tilde{h}(k) = \tilde{c}(k) + \rho_b \tilde{h}(k)\tilde{c}(k) \quad (19)$$

or

$$\tilde{c}(k) = \frac{\tilde{h}(k)}{1 + \rho_b \tilde{h}(k)} \quad (20)$$

where

$$\tilde{h}(k) = \frac{4\pi}{k}\int_0^\infty \sin kr\, h(r)\, dr \tag{21}$$

is the Fourier transform of $h(r)$, and so on. Equation (20) leads to

$$1 - \rho_b\tilde{c}(k) = [1 + \rho_b\tilde{h}(k)]^{-1} \tag{22}$$

Equation (22) means that (5), the compressibility equation, can be written as

$$\beta\left(\frac{\partial p}{\partial \rho_b}\right)_T = 1 - \rho_b \int c(r)\, d\mathbf{r} \tag{23}$$

since the integrals in (5) and (23) are just $\tilde{h}(0)$ and $\tilde{c}(0)$, respectively. The idea behind the OZ equation is that $h(r)$ is divided into two terms: the direct correlation function $c(r)$ and an indirect term, the convolution integral, which gives the correlation between molecules 1 and 2 resulting from the presence of a third molecule. This division is somewhat arbitrary. Presumably, it was chosen for its simplicity in Fourier space. Also, Ornstein and Zernike were interested in light scattering, so Fourier transforms were of natural interest to them. Even so, the division seems appropriate. The direct correlation function, in general, has a shorter range than $h(r)$ and is simple in structure and so is easier to approximate.

Equation (17) is an identity. To obtain an equation, some ansatz relating $h(r)$ and $c(r)$ must be introduced. There are three ansätze in common use. They are the *Percus–Yevick approximation* (PYA) (Percus and Yevick, 1958):

$$h(r) - c(r) = y(r) - 1 \tag{24}$$

and the *hypernetted chain approximation* (HNCA):

$$h(r) - c(r) = \ln y(r) \tag{25}$$

The HNCA was formulated by several authors (Morita, 1958, 1960; van Leeuwen et al., 1959; Morita and Hiroike, 1960, 1961; Meeron, 1960a,b,c; Rushbrooke, 1960; Verlet, 1960). The third approximation is the *mean spherical approximation* (MSA) of Lebowitz and Percus (1966):

$$\begin{aligned} h(r) &= -1 \qquad r < 0 \\ c(r) &= -\beta u(r) \qquad r > 0 \end{aligned} \tag{26}$$

As formulated above, the MSA is used for potentials with a hard core such that $u(r)$ is infinite for $r < 0$. Thus the first part of (26) is an exact statement of the nonoverlapping of the molecules. The second part of (26)

is the approximation and can be regarded as a linearized version of the HNCA, where

$$\ln y(r) = \beta u(r) + \ln g(r) \simeq \beta u(r) + h(r) \tag{27}$$

The PYA is a good approximation for the hard-sphere fluid but is less successful for other systems. The HNCA can be regarded as a complement of the PYA. It is not particularly successful for hard spheres but often describes the corrections to the hard-sphere fluid, due to the longer-range attractive forces, rather well (Barker and Henderson, 1976). One possibility is to combine the two approximations and use the PYA (or even simulation results) for the hard core and the HNCA to compute the corrections due to the attractive forces. Such an approach is called the RHNCA (reference HNCA).

The MSA is identical to the PYA for the hard-sphere fluid since for both approximations

$$\begin{aligned} h(r) &= -1 \qquad r < d \\ c(r) &= 0 \qquad r > d \end{aligned} \tag{28}$$

for the hard-sphere fluid. Thus the MSA is quite accurate for the hard cores and for many systems, at high densities at least, gives the contributions of the attractive forces more accurately than the PYA (but less accurately than the HNCA). If the strength of the attractive forces is large, the MSA can yield negative values for $g(r)$. Even in this case, the thermodynamic functions are often quite reasonable.

Finally, there is the generalization of the HNCA involving *bridge diagrams*. Quite formally,

$$h(r) - c(r) = \ln y(r) - B(r) \tag{29}$$

where $B(r)$ is the sum of all bridge diagrams. This is an infinite sum of diagrams that have not even been enumerated, let alone calculated. However, the lowest-order terms are known. For example,

$$B(r) = \tfrac{1}{2}\rho_b^2 \int f(13)f(14)f(23)f(24)f(34)\, d\mathbf{r}_3\, d\mathbf{r}_4 + \cdots \tag{30}$$

Just as with (15), this equation can be renormalized to give

$$B(r) = \tfrac{1}{2}\rho_b^2 \int h(13)h(14)h(23)h(24)h(34)\, d\mathbf{r}_3\, d\mathbf{r}_4 + \cdots \tag{31}$$

There is evidence that $B(r)$ does not depend strongly on the intermolecular potential. Therefore, if $B(r)$ can be parametrized accurately for some fluid, such as the hard-sphere fluid, the resulting $B(r)$ can be used in many applications (Rosenfeld and Ashroft, 1979).

II. SINGLET THEORY OF INHOMOGENEOUS FLUIDS

At an interface, the density of the fluid is no longer uniform but is a function of position. Thus one can write $\rho(\mathbf{r})$, or if the wall is flat and there is cylindrical symmetry, $\rho(x)$. The OZ equation can be generalized to a mixture

$$h_{ij} = c_{ij} + \sum_k \rho_k h_{ik} * c_{kj} \tag{32}$$

The case of a fluid near a wall can be realized by regarding the wall as a giant molecule dissolved in the fluid. Then the inhomogeneous fluid of n components, say, can be considered as a homogeneous mixture of $n + 1$ components containing a giant molecule (which will not be given a subscript) whose diameter D and density ρ has the property (Perram and Smith, 1976; Henderson et al., 1976a)

$$D \to \infty \qquad \rho \to 0 \qquad \rho D^3 \to 0 \tag{33}$$

Thus the OZ equations divide themselves into three classes:

$$h_{ij} = c_{ij} + \sum_{k=1}^{n} \rho_k h_{ik} * c_{kj} \tag{34}$$

$$h_i = c_i + \sum_{k=1}^{n} \rho_k h_k * c_{ki} \tag{35}$$

and

$$h = c + \sum_{k=1}^{n} \rho_k h_k * c_k \tag{36}$$

The first equation is just the bulk OZ equation. The second equation has the form

$$h_i(x) = c_i(x) + 2\pi\rho_b \int_{-\infty}^{\infty} h_k(t) \mathscr{C}_{ki}(x,t)\, dt \tag{37}$$

where

$$\mathscr{C}_{ki}(x,t) = \int_{|x-t|}^{\infty} c_{ki}(s) s\, ds \tag{38}$$

The function $c_{ki}(s)$ is the bulk direct-correlation function. Its appearance in (35) and (37) is not an approximation. Equations (35) and (37) are merely definitions of $c_i(x)$. Once $h_i(x)$ is known, the density profile can be obtained from

$$\rho_i(x) = \rho_i[1 + h_i(x)] \tag{39}$$

Lozada-Cassou (1981, 1982) has obtained (35) and (37) by an interesting alternative argument.

Equation (36) is also of interest because the force between two giant molecules can be calculated from

$$F = -kT\frac{\partial}{\partial x}\ln g(x) \tag{40}$$

Note that, in contrast to (34) and (35), no iterations are required for a numerical solution of (36).

Equations (34) to (36) can be solved using the different approximations mentioned above. One is under no obligation to use the same approximation throughout. For example, one could use the MSA to solve (34) and the HNC to solve (35) obtaining what is called the HNC1/MSA equation for (35). The number 1 following HNC is meant to denote that the theory is a singlet HNC theory, in contrast to pair theories, which will be considered presently. Similarly, (35), or (37), will be referred to as the OZ1 equation.

Lebowitz (1964) has solved the PY equation for a mixture of hard spheres. His results can be used to obtain analytic results for the Laplace transform of the correlation functions for hard spheres near a hard wall (HS/HW) or for the force between two large hard spheres in a fluid of small hard spheres. These transforms can be inverted (Smith and Henderson, 1970; Henderson and Smith, 1978; Henderson, 1988).

In Fig. 1, the singlet PY (PY1) results (Henderson et al., 1976a) for the HS/HW system are plotted. The agreement with the computer simulations (Snook and Henderson, 1978; Henderson and van Swol, 1984) is quite good except in the immediate neighborhood of the wall. At contact the PY1 (or MSA1) results for the HS/HW system (Carnie et al., 1981) is

$$\rho(0) = \rho_b\left(\beta\frac{\partial p}{\partial \rho_b}\right)^{1/2} \tag{41}$$

where $\beta\partial p/\partial\rho_b$ is the PY compressibility result. Equation (41) is valid, in the MSA1, for any fluid near a hard wall. It is to be compared with the exact result (Henderson et al., 1979), valid for any fluid near a hard wall,

$$\rho(0) = \beta p \tag{42}$$

obtained from force balance considerations. For comparison, the corresponding HNC1 result is

$$\rho(0) = \tfrac{1}{2}\rho_b\left(1 + \beta\frac{\partial p}{\partial \rho_b}\right) \tag{43}$$

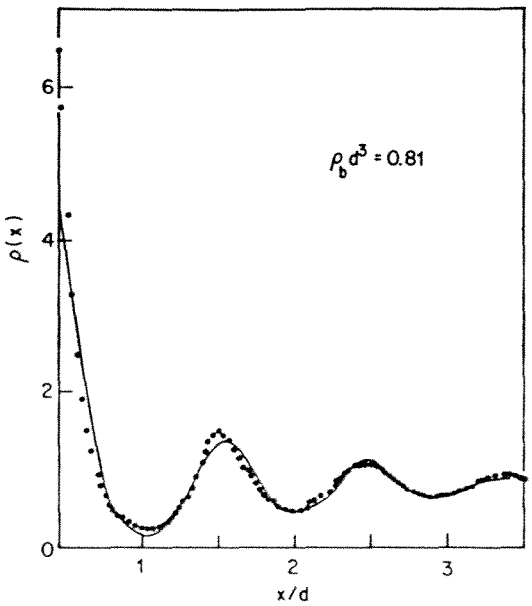


Fig. 1 Density profile of the hard-sphere fluid near a hard wall. The points give the computer simulation values of Snook and Henderson (1978) and the curve gives the PY1 results (Henderson et al., 1976a).

At low densities, (41) and (43) agree with (42). However, as the density is increased, $\rho(0)$, as given by (41), is too small and $\rho(0)$, as given by (43), is too large. Nonetheless, since $\rho(0)$ and $\beta\partial p/\partial\rho_b$ are both large for hard spheres, the PY1 and HNC1 results are at least qualitatively correct for the HS/HW system.

The situation is less favorable for some other systems. For example, (41) to (43) are valid for any fluid at a hard wall. At low densities, (41) and (43) are correct. They are also qualitatively correct at high densities and at high temperatures where the repulsive forces in the fluid dominate. However, for a liquid in equilibrium with its vapor (or at least nearly so), $p \simeq 0$, whereas $\beta\partial p/\partial\rho_b$ is large. Thus (41) and (43) will yield substantial errors. This can be seen in Fig. 2, where the density profile of a Lennard–Jones liquid in equilibrium with its vapor (or at least nearly so) is displayed. The PY1 results (Sullivan and Stell, 1976; Sullivan et al., 1980; Plischke and Henderson, 1986a) fails to show the layer of vapor near the wall (wetting by the vapor) seen in the simulations (Abraham, 1978).

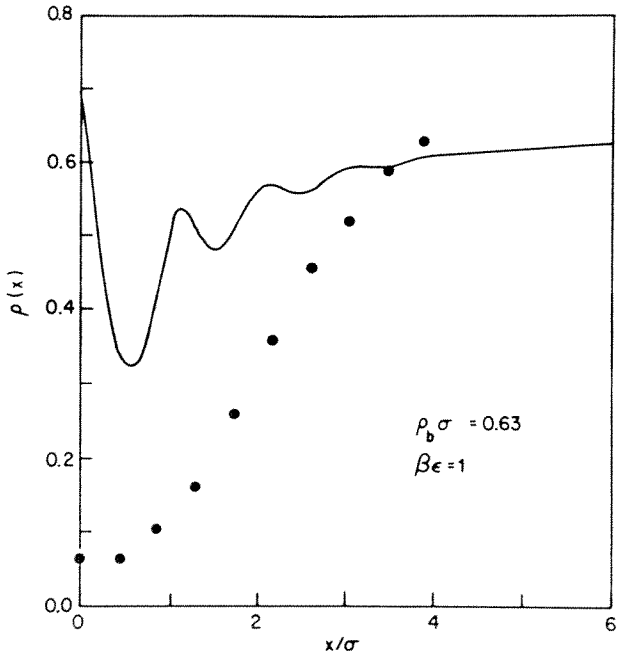


Fig. 2 Density profile for a Lennard–Jones fluid near a hard wall. The points give the computer simulation values of Abraham (1978) and the curve gives the PY1 results (Plischke and Henderson, 1986a).

Molecules interacting by the Yukawa pair potential

$$u(r) = \begin{cases} \infty & r < \sigma \\ -\epsilon\sigma \dfrac{e^{-\lambda(r/\sigma - 1)}}{r} & r > \sigma \end{cases} \tag{44}$$

and near a wall for which the surface/molecule is an exponential

$$V(x) = \begin{cases} \infty & x < 0 \\ -\epsilon_w e^{-\lambda_w x} & x > 0 \end{cases} \tag{45}$$

have been examined by Henderson et al. (1976b) using the MSA1. Although the profiles themselves have the same difficulties as those shown in Fig. 2, the absorption isotherms, defined by

$$\begin{aligned} \Gamma &= \int_0^\infty [\rho(x) - \rho_b]\, dx \\ &= \rho_b \int_0^\infty h(x)\, dx \end{aligned} \tag{46}$$

are more physical, at least in certain regions. The result which they obtained is

$$\Gamma = \rho_b \frac{3\eta/2(1 - \eta)^2 + (C_w/\lambda_w)(1 - \eta)^2}{a} \tag{47}$$

where $a = (\beta\partial p/\partial\rho_b)^{1/2}$ and $\eta = \pi\rho\sigma^3/6$. The expression for the parameters C_w is complex (Henderson et al., 1980) but is simple if $V(x)$ is short range, $\lambda_w \to \infty$ (ϵ_w/λ_w finite). For this case

$$C_w = \frac{\beta\epsilon_w}{(1 - \eta)^2} \tag{48}$$

so that

$$\Gamma = \rho_b \frac{3\eta/2(1 - \eta)^2 + \beta\epsilon_w/\lambda_w}{a} \tag{49}$$

Hence Γ diverges when $a \to 0$. This is not a satisfactory theory of wetting since Γ should diverge near the coexistence curve not at $a = 0$, which is inside this curve. Further, (49) cannot be used to describe wetting by a vapor since Γ should be negative for this case. Even so, (49) might be useful for supercritical isotherms or as a semiempirical isotherm for absorption of a vapor.

There has been some work on inhomogenous molecular fluids. For example, Badiali (1985) has studied dipolar hard spheres near an uncharged flat wall using a cluster expansion and finds that they tend to lie flat near the wall. Torrie et al. (1988) have applied the RHNCA1 to waterlike molecules at a charged hard but not flat wall and find that when the wall particle is large (the diameter of the wall particle is 30 times that of the dipole's hard sphere) there is an "icelike" structure near the "wall" which is not present at smaller diameters of the wall particle. This structure tends to persist even when the "wall" is charged. Unfortunately, Torrie et al. were not able to obtain solutions for infinite wall particle diameter. More recently, Torrie et al. (1989) have used the RHNCA1 for dipolar hard spheres at a charged hard, but not flat, wall and finds that, in contrast to the waterlike fluid, the dipolar fluid is not highly structured near the wall. These are important and fascinating results. However, it should be kept in mind that the relevance of these results to flat walls or electrolytes with significant electrostatic screening has yet to be established.

There has been some work on inhomogeneous molecular fluids using the RISM (reference interaction sphere model). Briefly, the RISM approach considers the molecule fluid to be a collection of spheres with a δ-function interaction between spheres of the same molecule and a regular

pair potential for the intermolecular interaction between spheres in different molecules. One of the integral equation approximations is applied to calculate the sphere–sphere correlation functions. The inhomogeneous case can be realized by considering a mixture consisting of one giant sphere dissolved in the molecular RISM fluid. Sullivan et al. (1981) have used this approach to consider a fluid of hard dumbbells. They find that as the density or the distance between the spheres in the dumbbell is increased, there is an increasing tendency of the spheres near the wall to orient themselves parallel to the wall. A more recent study is that of Borstnik and Janezic (1989), who use RISM to study SO_2 at a soft planar wall.

To complete this section, some recent applications of (36) will be considered. Henderson and Lozada-Cassou (1986) and Henderson (1988) have used the PY results (Lebowitz, 1964) to calculate the force between two large hard spheres in a fluid of small hard spheres. The resulting force is shown in Fig. 3. The oscillations are core exclusion effects and are in agreement with the oscillatory forces measured by Israelachvili and colleagues (Israelachvili, 1985, and references cited therein). Prior to Israelachvili's studies, it was thought that the force between macrospheres would be monotonic.

Strictly speaking, Henderson and Lozada-Cassou used the PYA for (34) and (35) but used the HNCA for (36). If they had used the PYA throughout, the force would be proportional to the $\exp(D)$ rather than D, as should be the case. Since h and c do not appear in the convolutions

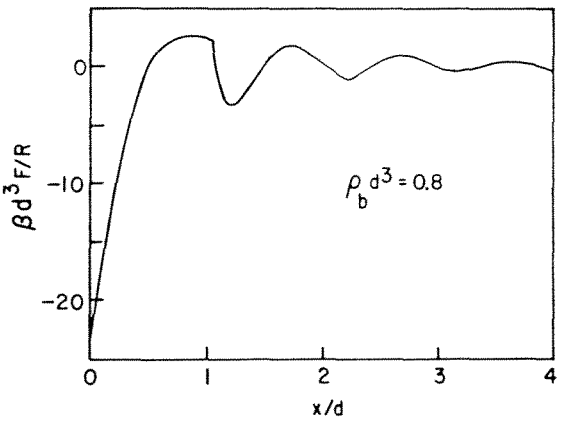


Fig. 3 Force between two giant spheres immersed in a hard-sphere fluid whose density is $\rho d^3 = 0.8$. The quantities d and R are the diameter of the small sphere and the radius of the large spheres, respectively.

in (36), the Lebowitz values for these PYA values for these convolutions can still be employed. In any case, with this hybrid scheme, they obtain correctly $F \propto D$.

In addition to the core exclusion effects discussed above, there are other forces, of course. These might be van der Waals attractive forces, analogous to the r^{-6} term in (4) or electrostatic forces (usually repulsive). The effects of these terms is considered by Henderson and Lozada-Cassou (1986).

Lozada-Cassou (1984) has developed an interesting alternative approach to Eq. (36). Instead, he uses (35) to describe a fluid between two macroparticles, but considers the giant particle in (35) to be a dumbbell of macrospheres. In this way he can calculate the force between two macrospheres, or the pressure between flat walls, from (35). Lozada-Cassou discusses this approach in detail in Chapter 8.

III. PAIR THEORY OF INHOMOGENEOUS FLUIDS

In the preceding section, a theory of inhomogeneous fluids that involves only singlet functions has been outlined. This approach is very useful for many applications (e.g., hard spheres at a hard wall and the electrified interfaces discussed in Chapter 6). However, it fails for some other applications. This failure is not due to the formalism. It is a failure of the ansatz used. Systematically including bridge diagrams would remove these problems.

Another approach is to generalize (17), the OZ equation, to an inhomogeneous fluid by moving the density, which now depends on the position, inside the integral. Thus

$$h(12) = c(12) + \int \rho(3)h(13)c(23)\, d\mathbf{r}_3 \tag{50}$$

To distinguish this equation from the OZ1 equation, (50) will be referred to as the OZ2 equation. This equation can be coupled with one of the PY, HNC, or MSA ansätze to yield the PYA2, HNCA2, or MSA2 equation.

However, in contrast to the previous theories, the pair theories require an additional equation relating $\rho(x)$ to the pair functions. Two relations have been used. The first is the first *Born–Green (BG) equation* (Born and Green, 1946),

$$-kT\frac{\partial \ln \rho(x)}{\partial x} = \frac{\partial V(x)}{\partial x} + 2\pi \int_{-\infty}^{\infty} (x - x')\rho(x')\, dx' \int_{|x-x'|}^{\infty} \frac{\partial u(r)}{\partial r}\, g(x,x',R)\, dr \tag{51}$$

and the second is the *Lovett–Mou–Buff–Wertheim (LMBW) equation* (Lovett et al., 1976; Wertheim, 1976),

$$-kT\frac{\partial \ln \rho(x)}{\partial x} = \frac{\partial V(x)}{\partial x} - 2\pi kT\int_{-\infty}^{\infty}\frac{\partial \rho(x')}{\partial x'}\,dx'\int_0^{\infty} Rc(x,x',R)\,dR \quad (52)$$

In (51) and (52) $V(x)$ is the external potential between the wall and a fluid molecule and $u(r)$ is the usual pair potential among the fluid molecules. The surface has been assumed to be flat so that there is cylindrical symmetry. Thus $\rho(\mathbf{r}) = \rho(x)$ and $g(\mathbf{r}_1,\mathbf{r}_2) = g(x_1,x_2,R_{12})$, where x_1 and x_2 give the perpendicular heights of the two molecules above the surface,

$$r_{12} = |\,\mathbf{r}_1 - \mathbf{r}_2\,| \quad (53)$$

as usual, and R_{12} is the projection of the preparation of the two molecules onto the plane of the surface so that

$$r_{12}^2 = R_{12}^2 + (x_1 - x_2)^2 \quad (54)$$

With the use of the OZ2 equation, the LMBW equation can be written in the equivalent form

$$-kT\frac{\partial \ln \rho(x)}{\partial x} = \frac{\partial V(x)}{\partial x} - 2\pi\int_0^{\infty} R\,dR\int_{-\infty}^{\infty}\rho(x')h(x,x',R)\frac{\partial V(x')}{\partial x'}\,dx' \quad (55)$$

Equations (51) and (52) [or (55)] can be obtained (Henderson and Plischke, 1987) from (9) by generalization to a mixture and allowing one species to be very large but present in infinite dilution. If molecule 1 is the large molecule, the LMBW equation results; whereas if molecule 1 is a fluid molecule, the BG equation results.

For the HS/HW system, (51) becomes

$$\frac{\partial \ln \rho(x)}{\partial x} = 2\pi\int_{[0,x-d]}^{x+d}(x - x')\rho(x')G(x,x')\,dx' \quad (56)$$

where $[0, x - d]$ indicates that the maximum of 0 and $x - d$ is to be used and $G(x,x')$ is the value of $g(x,x',R)$ when the two spheres at a height of x and x' above the surface are in contact. For any simple fluid near a hard wall, (52) becomes

$$\frac{\partial \ln \rho(x)}{\partial x} = \rho(0)K(x,0) + \int_0^{\infty}\frac{\partial \rho(x')}{\partial x'}K(x,x')\,dx' \quad (57)$$

where

$$K(x,x') = 2\pi\int_0^{\infty} Rc(x,x',R)\,dR \quad (58)$$

and (55) becomes

$$\frac{\partial \ln \rho(x)}{\partial x} = 2\pi\rho(0) \int_0^\infty Rh(x,0,R)\, dR \tag{59}$$

Equations (57) and (59) are equivalent as long as h and c are related by the OZ2 equation.

Modified versions of the LMBW equation can be obtained by substituting the contact value theorem, Eq. (42), into the right-hand sides of (57) and (59). This yields

$$\frac{\partial \ln \rho(x)}{\partial x} = \beta p K(x,0) + \int_0^\infty \frac{\partial \rho(x')}{\partial x'} K(x,x')\, dx' \tag{60}$$

and

$$\frac{\partial \ln \rho(x)}{\partial x} = 2\pi\beta p \int_0^\infty Rh(x,0,R)\, dR \tag{61}$$

As long as no approximations have been made, this substitution is only formal. No change has been made. However, if some approximation is used, the LMBW and modified LMBW equations will not be equivalent. Equations (59) to (61) are easily generalized to the case of a hard wall with a longer-range tail, that is,

$$V(x) = \begin{cases} \infty & x < 0 \\ V(x) & x > 0 \end{cases} \tag{62}$$

The contribution of the longer-range part of $V(x)$ in (52) and (55) must be included and the contact value theorem must be generalized (Henderson et al., 1979) to yield

$$kT\rho(0) = p + \int_0^\infty \frac{\partial V(x)}{\partial x} \rho(x)\, dx \tag{63}$$

There is no modified version of the BG equation because $\rho(0)$ does not occur in this equation. Also, the BG equation has the interesting property that this equation yields a $\rho(x)$ that satisfies the exact contact value theorem with the pressure calculated from (6) using whatever the approximation gives for the bulk $g(r)$ (Carnie and Chan, 1981; Henderson et al., 1981). Since (6) often gives poor results, the BG equation may give poor results for $\rho(0)$ even though the exact contact value theorem is satisfied in a formal sense. The LMBW equation does not satisfy the exact contact value theorem. There is no known contact value theorem for the LMBW equation [i.e., there is no explicit formula for the LMBW $\rho(0)$]. Even so, the LMBW equation usually gives good values for $\rho(0)$. They are generally

better than those obtained from the corresponding singlet equation and are often as good, or better, as those obtained from the BG equation, *especially if the modified versions are used.* The LMBW equation involving $c(x_1,x_2,R)$ is often the easiest equation to use numerically because (49) and (53) involve the longer-ranged function $g(x_1,x_2,R)$. Additionally, the BG equation focuses its attention on the region where the derivative of $\exp[-\beta u(r)]$ is large. Also, if the molecules have a hard core, extrapolation to the hard-core diameter may be involved if the BG equation is used.

Sokolowski (1980) has solved the PY2 and LMBW equations for the Sokolowski HS/HW system for low values of ρd^3. Plischke and Henderson (1986a) have solved these same equations for moderate to high hard-sphere fluid densities. As shown in Fig. 4, the resulting density profiles are much better than those of the PY1 theory. Only at contact is there any error in the PY2 results. Whether this is due to numerical errors or to deficiences in the PY approximation is not known. In any case, Plischke and Henderson (1990) have repeated the calculation using the modified

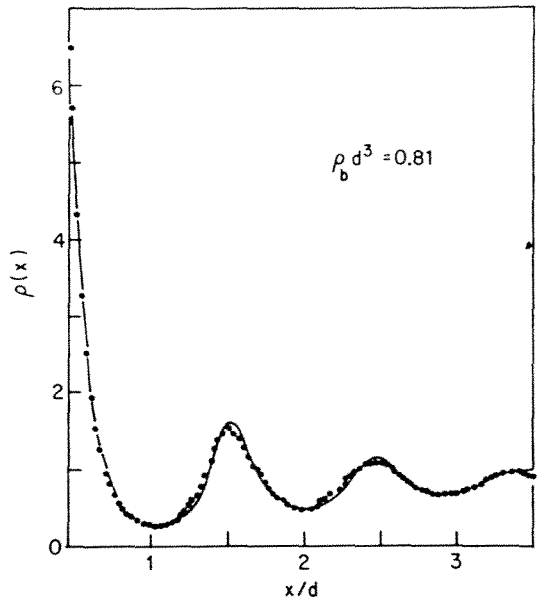


Fig. 4 Density profile of the hard-sphere fluid near a hard wall. The points give the computer simulation values of Snook and Henderson (1978) and the curve gives the PY2 values (Plischke and Henderson, 1986a).

LMBW equation and obtained improved results. The contact values for $\rho(x)$ are nearly exact. Kjellander and Sarman (1988) have also applied the PY2 theory to the hard-sphere/hard-wall (HS/HW) system with good success.

The pair correlation functions for the HS/HW system can also be compared with computer simulations. Such a comparison is made in Figs. 5 and 6 for the contact values, $G(x_1,x_2)$. For the configuration $x_1 = x_2 = x$ shown in Fig. 5, $G(x_1,x)$ is quite featureless. On the other hand, for the configurations shown in Fig. 6, where one sphere is at the wall and the other sphere slides over its surface until the line of center is perpendicular and then the pair moves away from the surface, $G(x_1,x_2)$ oscillates strongly.

The PY2 values at contact are too small (a common feature of the PYA). However, the curves parallel the computer results, so the PYA gives reasonable results. The contact values in Figs. 5 and 6 have been described quite well by Henderson and Plischke (1985). For $x_1 = x_2 = x$ (Fig. 5)

$$g(x,x,R) \simeq 1 + c(x,x)h(r) \tag{64}$$

where $h(r)$ is the bulk total correlation function and $c(x_1,x_2)$ is a function depending on x_1 and x_2 and the density. An examination (Plischke and

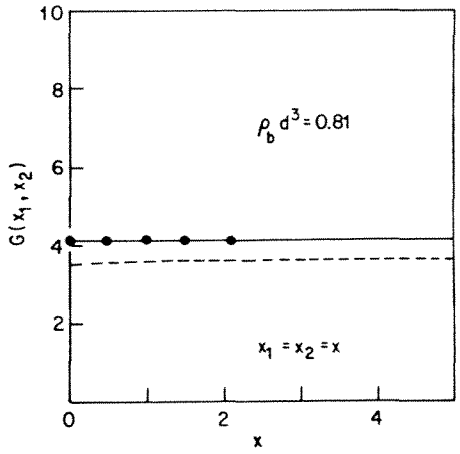


Fig. 5 Values of $G(x_1,x_2)$ for hard spheres near a hard wall for $x_1 = x_2 = x$. The dashed curve gives the results of the Henderson–Plischke fit [Eq. (68)] and the solid curve gives the PY2 results. The points give some of the simulation results of Snook and Henderson (1978).

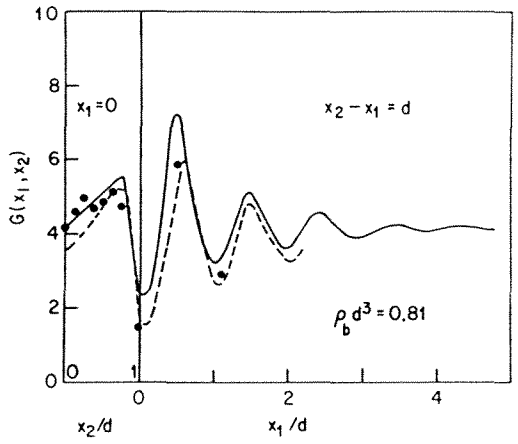


Fig. 6 Values of $G(x_1,x_2)$ for hard spheres near a hard wall. The curves and point have the same meaning as in Fig. 5.

Henderson, 1984) of the density expansion of $g(x_1,x_2,R)$ in the limit of zero density shows that $c(0,0) = \frac{1}{2}$ at low densities. As the density increases, $c(0,0)$ increases. Further, as x_1 and x_2 increase, $c(x_1,x_2) \to 1$. Henderson and Plischke (1985) found that reasonable results could be obtained by using

$$c(x_1,x_2) = \begin{cases} c + \frac{1}{2}(1-c)\left[\dfrac{3(x_1+x_2)}{d} - \left(\dfrac{x_1+x_2}{d}\right)^3\right] & x_1 + x_2 < d \\ 1 & x_1 + x_2 > d \end{cases} \tag{65}$$

They found that c is given quite well by

$$c = 0.5 + 1.03\rho d^3 - 0.5\rho^2 d^6 \tag{66}$$

For $R = 0$ (or $r = |x_1 - x_2|$), the configuration depicted in the singlet panel of Fig. 6, the shielding approximation of Percus (1980),

$$g(0,x_1,x_2) = \frac{\rho}{\rho(x_{\max})}\, g(r) \tag{67}$$

where $x_{\max}$ is the greatest of x_1 and x_2, is quite accurate. The interpolation formula, based on (64) and (67),

$$G(x_1,x_2) = \frac{(x_1 - x_2)\rho g(d)}{d\rho(x_{\max})} + \frac{d - |x_1 - x_2|}{d}\,[1 + c(x_1,x_2)h(d)] \tag{68}$$

gives a good description of $G(x_1,x_2)$ in the left panel of Fig. 6. Further details can be found in Henderson and Plischke (1985).

It is interesting to note that if the surface is regarded as a third (giant) molecule, the shielding approximation is just (11), and for $c = 1$, (64) is just the superposition approximation.

The PY2 theory has also been applied (Plischke and Henderson, 1986b) to a Lennard–Jones fluid near a hard wall. As can be seen in Fig. 7, the PY2 theory correctly predicts the wetting of the wall by the vapor when a liquid is near its vapor pressure and close to a wall.

A related problem is the wetting of the wall by the liquid when a vapor is near a wall. This problem has been studied by Nieminen and Ashcroft (1981), Forbes and Ashcroft (1982), and Hildebrand and Nieminen (1984) using the HNC2 equation with an approximation to the contribution of the bridge diagrams without finding evidence of wetting. However, Bruno et al. (1986, 1987), using a different approximation for the bridge diagrams, find evidence of wetting.

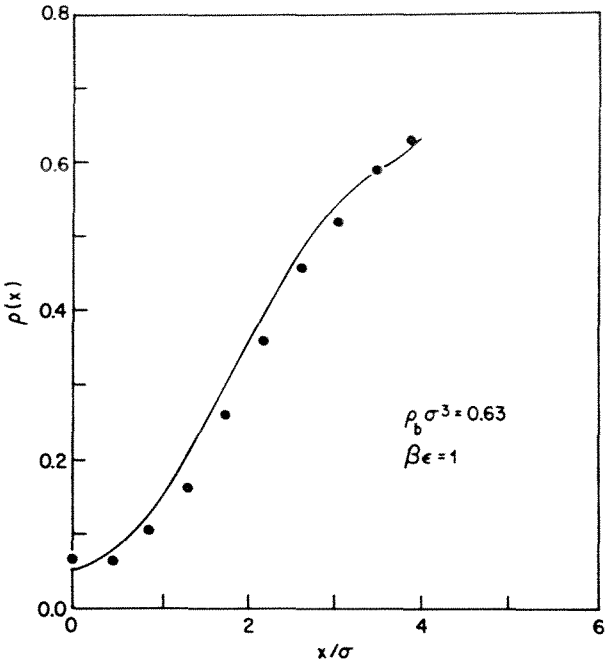


Fig. 7 Density profile for a Lennard–Jones fluid near a hard wall. The points give the computer simulation values of Abraham (1978) and the curve gives the PY2 results (Plischke and Henderson, 1968a).

IV. SIMPLE APPROXIMATIONS FOR THE PAIR FUNCTIONS

If $g(x_1,x_2,R_{12})$ or $c(x_1,x_2,R_{12})$ could be approximated directly and then used with either the BG or LMBW equation, a considerable savings in computer time (and memory!) could be achieved since the OZ2 equation need not be used. The simplest approximation would be the use of the bulk functions $g(r_{12})$ or $c(r_{12})$. The approximation

$$g(x_1,x_2,R_{12}) = g(r_{12}) \tag{69}$$

is just the superposition approximation [Eq. (10)] if the surface is regarded as a third (giant) molecule. If this approximation is substituted into the BG equation, the integration over x can be performed analytically. For the specific case of the HS/HW system, the result is

$$\ln \rho(x) = \ln \rho_b - \pi g(d) \int_{[0,x-d]}^{x+d} [1 - (x - x')^2]\rho(x)\, dx + \frac{4\pi}{3} \rho g(d) \tag{70}$$

where $[0,x - d]$ means that the maximum of 0 and $x - d$ is to be used. Henderson and Plischke (1986) have solved this equation. The results are not very good. The contact value is correct, but the oscillations are out of phase and too large in amplitude.

Equation (69) can also be substituted into the LMBW equation. The result, for the HS/HW system, is

$$\ln \rho(x) = \ln \rho_b - 2\pi\rho(0) \int_x^\infty y(y - x)h(y)\, dy \tag{71}$$

This equation does not even have a solution, except at low densities. However, the modified version of (71),

$$\ln \rho(x) = \ln \rho_b - 2\pi\beta p \int_x^\infty y(y - x)h(y)\, dy \tag{72}$$

does at least have a solution. It is, however, not very good.

The shielding approximation

$$g(x_1,x_2,R_{12}) = \frac{\rho}{\rho(x_{\max})} g(r_{12}) \tag{73}$$

when substituted into the BG equation, also yields an equation for which the x integration can be performed analytically. For the HS/HW system, the result is

$$\rho(x) = \rho_b - \pi\rho_b g(d) \int_{[0,x-d]}^{x} [1 - (x - x')^2]\rho(x')\, dx' \tag{74}$$

Henderson and Plischke (1986) have also solved (74). The results are not very good. The contact value is correct, but the oscillations are damped.

A more promising approach is to correct (69) systematically in the spirit of (15) or (16). In lowest order

$$g(x_1,x_2,R_{12}) = g(r_{12})\left[1 + \int h(13)h(23)\rho(3)\, d\mathbf{r}_3 + \cdots\right] \tag{75}$$

or

$$g(x_1,x_2,R_{12}) = g(r_{12}) \exp\left[\int h(13)h(23)\rho(3)\, d\mathbf{r}_3 + \cdots\right] \tag{76}$$

No results have yet been obtained using this approximation, but it, or its extensions, should be of value.

The HNC approximation to (37) can be obtained from a density functional expression using the approximation

$$c(x_1,x_2,R_{12}) = c(r_{12}) \tag{77}$$

As a result, it is sometimes suggested that improved results could be obtained if the pair direct-correlation function were used in (37). The resulting equation would be an integrated version of the LMBW equation. However, it is not correct! The bridge diagram enters with a factor of unity, not $\frac{1}{2}$, if this *incorrect* procedure is used.

V. BULK FLUID AS A SPECIAL CASE OF AN INHOMOGENEOUS FLUID

Just as a giant molecule can be thought of as a source of an inhomogeneity, one of the molecules in a bulk fluid can be thought of as inducing an inhomogeneity. In this way, the theory of inhomogeneous fluids developed in this chapter can be used to produce improved theories of bulk fluids. For example, we can use the inhomogeneous Ornstein–Zernike equation to calculate the correlation function of two molecules in the presence of a third molecule; call this $g(12/3)$. This inhomogeneous OZ equation can then be coupled with either the BG or LMBW equation to yield the correlation function of one molecule in the presence of a second molecule; call this $g(1/2)$. The pair and triplet function can be calculated from $g(1/2)$ and $g(12/3)$. If the PYA, HNCA, or MSA is employed in this approach, the results are improved over the original formulations of these approximations. For example, in these improved PYA and HNC, some bridge diagrams are included.

Conversely, one could use the singlet OZ equation to yield the correlation function of one molecule in the presence of two molecules $g(1/23)$ and from this, the pair correlation function. The possibilities are end-

less. For example, one could formulate a theory in terms of $g(12/34)$ and use the inhomogeneous OZ equation to obtain a quadruplet correlation function.

Attard (1989) has used the $g(12/3)$ approach to calculate bulk hard-sphere correlation function with excellent results. Lozada-Cassou uses the $g(1/23)$ approach in Chapter 8 to calculate the force between colloidal particles. Stell (1975) has given an alternative formulation of Ornstein–Zernike relations for higher-order correlation functions.

ACKNOWLEDGMENTS

The author is grateful to the Salt Lake City Branch of the IBM Corporation for their help. Particular thanks are due to Wayne Stoker for his support and to Irma Brander, who did an outstanding job of typing the manuscript.

REFERENCES

Abraham, F. F. (1978). *J. Chem. Phys. 84*: 2846.
Attard, P. (1989). *J. Chem. Phys. 91*: 3072.
Badiali, J. P. (1985). *Mol. Phys. 55*: 939.
Barker, J. A., and Henderson, D. (1976). *Rev. Mod. Phys. 48*: 587.
Born, M., and Green, H. S. (1946). *Proc. R. Soc. London A 188*: 10.
Borstnik, B., and Janezic, D. (1989). *Chem. Phys. 130*: 195.
Bruno, E., Caccamo, C., and Tarazona, P. (1986). *Phys. Rev. A 34*: 2513.
Bruno, E., Caccamo, C., and Tarazona, P. (1987). *Phys. Rev. A 35*: 1210.
Carnie, S. L., and Chan, D. Y. C. (1981). *J. Chem. Phys. 74*: 1293.
Carnie, S. L., Chan, D. Y. C., Mitchell, D. J., and Ninham, B. W. (1981). *J. Chem. Phys. 74*: 1472.
Forbes, S. M., and Ashcroft, N. W. (1982). *Phys. Rev. B 25*: 1366.
Hansen, J. P., and McDonald, I. R. (1986). *Theory of Simple Liquids*, 2nd ed., Academic Press, London.
Henderson, D. (1988). *J. Colloid Interface Sci. 121*: 486.
Henderson, D., and Lozada-Cassou, M. (1986). *J. Colloid Interface Sci. 114*: 180.
Henderson, D., and Plischke, M. (1985). *Proc. R. Soc. London A 400*: 163.
Henderson, D., and Plischke, M. (1986). *Proc. Indian Acad. Sci. (Chem. Sci.) 97*: 297.
Henderson, D., and Plischke, M. (1987). *Proc. R. Soc. London A 410*: 409.
Henderson, D., and Smith, W. R. (1978). *J. Stat. Phys. 19*: 191.
Henderson, J. R., and van Swol, F. (1984). *Mol. Phys. 51*: 1991.
Henderson, D., Abraham, F. F., and Barker, J. A. (1976a). *Mol. Phys. 31*: 1291.
Henderson, D., Waisman, E., and Lebowitz, J. L. (1976b). In *Colloidal and Interface Sciences* (M. Kerker, ed.), Academic Press, New York, Vol. 3, p. 37.
Henderson, D., Blum, L., and Lebowitz, J. L. (1979). *J. Electroanal. Chem. 102*: 315.

Henderson, D., Lebowitz, J. L., Blum, L., and Waisman, E. (1980). *Mol. Phys. 39*: 47.
Henderson, D., Blum, O., and Bhuuyan, L. B. (1981). *Mol. Phys. 43*: 1185.
Hildebrand, K., and Nieminen, R. M. (1984). *Surf. Sci. 147*: 599.
Israelachvili, J. N. (1985). *Chem. Ser. 25*: 7.
Kirkwood, J. G. (1935). *J. Chem. Phys. 3*: 300.
Kjellander, R., and Sarman, S. (1988). *Chem. Phys. Lett. 149*: 102.
Lebowitz, J. L. (1964). *Phys. Rev. 133*: A895.
Lebowitz, J. L., and Percus, J. K. (1966). *Phys. Rev. 144*: 251.
Lovett, R., Mou, C. Y., and Buff, F. P. (1976). *J. Chem. Phys. 65*: 570.
Lozada-Cassou, M. (1981). *J. Chem. Phys. 74*: 1412.
Lozada-Cassou, M. (1982). *J. Chem. Phys. 77*: 5258.
Lozada-Cassou, M. (1984). *J. Chem. Phys. 80*: 3344.
Meeron, E. (1960a). *J. Math. Phys. 1*: 192.
Meeron, E. (1960b). *Physica 26*: 445.
Meeron, E. (1960c). *Prog. Theory Phys. 24*: 588.
Morita, T. (1958). *Prog. Theory Phys. 20*: 920.
Morita, T. (1960). *Prog. Theory Phys. 23*: 829.
Morita, T., and Hiroike, K. (1960). *Prog. Theory Phys. 23*: 1003.
Morita, T., and Hiroike, K. (1961). *Prog. Theory Phys. 25*: 537.
Nieminen, R. M., and Ashcroft, N. W. (1981). *Phys. Rev. A 24*: 560.
Ornstein, L. S., and Zernike, F. (1914). *Proc. Acad. Sci. Amsterdam 17*: 793.
Percus, J. K. (1980). *J. Stat. Phys. 23*: 657.
Percus, J. K., and Yevick, G. J. (1958). *Phys. Rev. 110*: 1.
Perram, J. W., and Smith, E. R. (1976). *Chem. Phys. Lett. 39*: 328.
Plischke, M., and Henderson, D. (1984). *J. Phys. Chem. 88*: 6544.
Plischke, M., and Henderson, D. (1986a). *Proc. R. Soc. London A 404*: 323.
Plischke, M., and Henderson, D. (1986b). *J. Chem. Phys. 84*: 2846.
Plischke, M., and Henderson, D. (1990). *J. Chem. Phys. 93*: 4489.
Rosenfeld, Y., and Ashcroft, N. W. (1979). *Phys. Rev. A 20*: 1208.
Rushbrooke, G. S. (1960). *Physica 26*: 259.
Smith, W. R., and Henderson, D. (1970). *Mol. Phys. 19*: 411.
Snook, L. K., and Henderson, D. (1978). *J. Chem. Phys. 68*: 2134.
Sokolowski, S. (1980). *J. Chem. Phys. 73*: 3507.
Stell, G. (1975). In *Phase Transitions and Critical Phenomena* (C. Domb and M. S. Green, eds.), Academic Press, New York, Vol. 5B, Chapter 3.
Sullivan, D. E., and Stell, G. (1976). *J. Chem. Phys. 69*: 5450.
Sullivan, D. E., Levesque, D., and Weiss, J. J. (1980). *J. Chem. Phys. 72*: 1170.
Sullivan, D. E., Barker, R., Gray, C. G., Street, W. B., and Gubbins, K. E. (1981). *Mol. Phys. 44*: 597.
Torrie, G. M., Kusalik, P. G., and Patey, G. N. (1988). *J. Chem. Phys. 85*: 7826.
Torrie, G. M., Perera, A., and Patey, G. N. (1989). *Mol. Phys. 67*: 1337.
van Leeuwen, J. M. J., Groeneveld, J., and de Boer, J. (1959). *Physica 25*: 792.
Verlet, L. (1960). *Nouvo Cimento 18*: 77.
Wertheim, M. (1976). *J. Chem. Phys. 65*: 2377.

5

Inhomogeneous Two-Dimensional Plasmas

B. Jancovici
Université de Paris-Sud
Orsay, France

I. INTRODUCTION

The aim of this chapter is to illustrate some general properties of Coulomb fluids (especially electrolytes) through two-dimensional models having the very peculiar property of being exactly solvable. Thus this chapter may be considered as an introduction to Chapter 6.

The structure of three-dimensional Coulomb fluids is strongly dependent on the special $1/r$ form of the potential, which generates the fundamental laws of electrostatics and specific features of the screening effect. For mimicking three-dimensional Coulomb systems in two dimensions, one must use the *two-dimensional* Coulomb potential, which is logarithmic: The interaction energy between two particles of charges e and e' is chosen as $-ee' \ln(r/L)$, where r is the distance between these particles and L some arbitrary length scale that fixes the zero of energy. This logarithmic potential is the one which, in two dimensions, obeys Poisson's equation

$$\Delta \ln \frac{r}{L} = 2\pi\delta(\mathbf{r})$$

and the other equations of electrostatics (the numerical factors, however, are different in three and two dimensions). This two-dimensional Coulomb

potential has a strength that grows to infinity at large distances; however, this causes little trouble, because screening effects generate *effective* potentials that decrease at large distances.

We deal here with equilibrium properties of classical (i.e., nonquantum) systems. We shall discuss mainly two kinds of models. The one-component plasma (OCP), or jellium, is made of identical particles of charge e embedded in a continuous charged background of the opposite sign. The two-component plasma (TCP), or Coulomb gas, is made of two species of particles, positive and negative, with opposite charges, $\pm e$.

The "charge" e defined here is such that e^2 has the dimensions of an energy. For a system of point particles at temperature T, the only dimensionless coupling constant is $\Gamma = e^2/k_B T$ (where k_B is Boltzmann's constant). This coupling constant is independent of the density, a special feature of *two-dimensional* Coulomb systems. A related property is the occurrence of a very simple equation of state.

For defining a sensible classical model of a three-dimensional two-component plasma, one must introduce some kind of short-range cutoff in the Coulomb interactions; otherwise, the configuration integral diverges when two particles of opposite charges come close to one another. On the contrary, for the two-dimensional case, in the high-temperature region defined by $\Gamma < 2$, the two-component plasma with pure Coulomb interactions is a well-defined system; indeed, a pair of particles of opposite charges $\pm e$ contributes to the Boltzmann factor by a term $\exp[-\Gamma \ln(r/L)] = (L/r)^{\Gamma}$, and this term is integrable, in two-dimensional space, near $r = 0$, provided that $\Gamma < 2$. However, for $\Gamma \geq 2$, this term is no longer integrable, and this signals that a collapse into neutral "molecules" does occur, unless some short-range cutoff of the interaction is introduced. With such a cutoff, the two-component plasma remains well defined for $\Gamma \geq 2$, and near $\Gamma = 4$, it undergoes the celebrated Kosterlitz–Thouless phase transition to a low-temperature insulating phase. This phase transition is outside the scope of the present chapter; there is an enormous literature about it (for a review, see, e.g., Ref. 1). The one-component plasma with pure Coulomb interactions is a well-defined system for any value of Γ. It becomes a two-dimensional Wigner "crystal" for $\Gamma > 142$ [2].

Our main concern is to build models for the electrical double layer (i.e., that arrangement of charges which forms at the interface between two conducting media, for instance at an electrode–electrolyte interface). A real electrical double layer is a complicated three-dimensional system with finite-size ions, a solvent that has a molecular structure, and so on; although an electrolyte is well described by classical mechanics, quantum effects are important for the electrons in an electrode. The two-dimensional solvable models considered here certainly are oversimplified: They are purely classical, the particles are pointlike or almost pointlike, the

solvent is replaced by a continuous medium. Nevertheless, many features of a Coulomb system are closely related to the harmonicity of the Coulomb potential. Therefore, just by using the proper potential, one may expect the models to provide an insight into salient properties of electrical double layers. Also, the exactly solvable models could be used as a test bench for approximate methods.

Before we turn to inhomogeneous fluids, we shall review the solvable models in the homogeneous case. At any temperature, for two-dimensional systems of point particles with pure Coulomb interactions, the equation of state has a very simple form, which is a rather trivial consequence of the mere scale invariance of the logarithmic potential, as explained in Section II. More interestingly, exact results for the other thermodynamical quantities and the structural properties have been obtained, but up to now only for the special temperature such that $\Gamma = e^2/k_BT = 2$; at this temperature, the densities and correlations can be computed not only for homogeneous systems but also for several physically interesting inhomogeneous systems. The solutions at $\Gamma = 2$ are described in Section III for the one-component plasma, in Section IV for the two-component plasma. The results are exploited further in Section V.

II. EQUATION OF STATE FOR SYSTEMS WITH PURE COULOMB INTERACTIONS

The equation of state has the simple form

$$p = \rho(k_BT - \tfrac{1}{4}e^2) \tag{1}$$

where p is the pressure and ρ is the number density of the particles. This equation of state has the peculiarity that it does *not* become the ideal gas law in the low-density limit. The derivation is based on an elementary scaling property [3] of the canonical partition function.

A. Two-Component Plasma

For N positive and N negative particles in a domain D of area Λ, the excess canonical partition function (which exists for $e^2/k_BT = \Gamma < 2$) is

$$Q_{ex} = \frac{1}{\Lambda^{2N}} \int_D \cdots \int_D \exp\left(\frac{1}{k_BT} \sum_{i<j} e_i e_j \ln \frac{r_{ij}}{L}\right) \prod_{i=1}^{2N} d^2\mathbf{r}_i$$

where $e_i = \pm e$. Introduction of rescaled position variables $\mathbf{s}_i = \mathbf{r}_i \Lambda^{-1/2}$ yields

$$Q_{ex} = \exp\left(-\frac{N\Gamma}{2} \ln \frac{\Lambda}{L^2}\right) \int_I \cdots \int_I \exp\left(\frac{1}{k_BT} \sum_{i<j} e_i e_j \ln s_{ij}\right) \prod_{i=1}^{2N} d^2\mathbf{s}_i$$

where now the integral is on a domain I of unit area. Therefore, the dependence of Q_{ex} upon Λ is explicitly displayed in the first factor; correspondingly, the excess free energy is of the form

$$F_{ex} = \tfrac{1}{4}2Ne^2 \ln \frac{\Lambda}{L^2} + F^*(N,T) \tag{2}$$

Derivation with respect to Λ gives (1), where $\rho = 2N/\Lambda$. Equation (1), however, holds only above the critical temperature $T = e^2/2k_B$ ($\Gamma = 2$). Below that temperature, it is reasonable to believe that the system has collapsed into an ideal gas of neutral molecules of density $\rho/2$; indeed, (1) becomes $p = (\rho/2)k_BT$ at $\Gamma = 2$.

B. One-Component Plasma

The same reasoning as above applies, with minor modifications taking into account the presence of a uniform charged background. For a system with N mobile particles, the excess free energy is of the form

$$F_{ex} = \tfrac{1}{4}Ne^2 \ln \frac{\Lambda}{L^2} + F^*(N,T) \tag{3}$$

and the equation of state is also (1), where $\rho = N/\Lambda$ is the number density of the mobile particles.

For the one-component plasma, no collapse occurs, and (1) remains valid at arbitrary low temperatures; thus the pressure becomes negative for $T < e^2/4k_B$. The occurrence of negative pressures is not a special feature of the two-dimensional case; the same phenomenon is present for the three-dimensional one-component plasma. Actually, the negativeness of the pressure causes no special difficulty, since the background must be considered as rigidly maintained at a given density by some external constraint. It may also be noted that the pressure which is considered here is the "thermal" one, defined by deriving the free energy with respect to the area Λ of the whole system; for systems with a background, however, there are other possible definitions of the pressure, some of which are nonequivalent [4].

III. ONE-COMPONENT PLASMA AT $\Gamma = 2$

For the one-component plasma, $\Gamma = 2$ corresponds to a special temperature at which many more exact results are available. It is convenient to work in the canonical formalism.

A. Homogeneous Plasma

We first review the homogeneous case: particles of charge e embedded in a *uniform* neutralizing background. The thermodynamical and correlation functions can be obtained as follows [5,6].

We can start with a system of N particles of charge e confined in a disk of radius R; let $\rho = N/\pi R^2$ be the number density. The disk also contains a uniform background of charge density $-e\rho$. From elementary two-dimensional electrostatics, after some rearrangement one finds for the total potential energy (including particle–particle, background–particle, and background–background interactions)

$$\Phi = \frac{N^2e^2}{2}(\ln R - \tfrac{3}{4}) - \frac{Ne^2}{2}\ln L + \frac{e^2}{2}\pi\rho\sum_j r_j^2 - e^2\sum_{i<j}\ln r_{ij} \quad (4)$$

where $\mathbf{r}_j$ is the position of particle j (the origin is chosen at the center of the disk). Perhaps it is worth to note that the uniform background charge density $-e\rho$ creates the nonuniform background particle interaction cst $+ (e^2/2)\pi\rho r_j^2$, in agreement with Poisson's equation.

When $\Gamma = e^2/k_BT = 2$, the Boltzmann factor is

$$\exp\left(-\frac{\Phi}{k_BT}\right) = A\prod_j e^{-\pi\rho r_j^2}\prod_{i<j} r_{ij}^2$$

where A is a constant.

As alternative labels for the positions $\mathbf{r}_j$ of the particles, it is convenient to introduce the complex numbers $z_j = r_j\exp(i\theta_j)$ (where r_j and θ_j are the polar coordinates of $\mathbf{r}_j$). The Boltzmann factor can be rewritten in terms of a Vandermonde determinant, since

$$\prod_{i<j}|\mathbf{r}_i - \mathbf{r}_j|^2 = \left|\prod_{i<j}(z_i - z_j)\right|^2 = \left|\det\begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_N \\ z_1^2 & z_2^2 & \cdots & z_N^2 \\ \cdots & \cdots & \cdots & \cdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_N^{N-1} \end{bmatrix}\right|^2$$

Taking into account the factors $\exp(-\pi\rho r_j^2)$, we obtain

$$\exp\left(\frac{-\Phi}{k_BT}\right) = A\ |\det[\psi_n(\mathbf{r}_j)]_{n,j=1,\ldots,N}|^2 \quad (5)$$

where

$$\psi_{n+1}(\mathbf{r}) = \exp(-\tfrac{1}{2}\pi\rho r^2)z^n = \exp(-\tfrac{1}{2}\pi\rho r^2)r^n e^{in\theta}$$

The functions $\psi_n(\mathbf{r})$ form an orthogonal set, since their scalar products contain the integrals

$$\int_0^{2\pi} e^{i(n-m)\theta}\, d\theta = 2\pi\delta_{nm}$$

We have now made the crucial step. The determinant in (5) can be regarded as a Slater determinant (i.e., as the wave function for a system of independent fermions occupying orthogonal orbitals). The *classical* probability density $\exp(-\Phi/k_BT)$ for our system of *interacting* particles is proportional to the *quantum mechanical* probability density for a system of *independent* fermions, and we can use well-known simple techniques of quantum mechanics.

1. *Free Energy*

From (5), the configuration integral is

$$Q = \int_{r_1<R} d^2\mathbf{r}_1 \cdots \int_{r_N<R} d^2\mathbf{r}_N \exp\left(\frac{-\Phi}{k_BT}\right) = AN! \prod_{n=1}^{N} \int_{r<R} d^2\mathbf{r}\, |\, \psi_n(\mathbf{r})\, |^2$$

Making the change of variable $t = \pi\rho r^2$, one finds for the excess partition function

$$Q_{\text{ex}} = \frac{Q}{(\pi R^2)^N} = \exp\left(\frac{3}{4}N^2 + N\ln\frac{L}{R}\right) \frac{N!}{N^{N(N+1)/2}} \prod_{n=1}^{N} \gamma(n,N)$$

where

$$\gamma(n,N) = \int_0^N e^{-t} t^{n-1}\, dt$$

is the incomplete gamma function. The corresponding free energy is indeed of the form (3) and generates the pressure (1).

It can rigorously be shown that when taking the thermodynamic limit, it is legitimate to replace at once $\gamma(n,N)$ by $\gamma(n,\infty) = (n-1)!$. Then, by repeated use of Stirling's formula, one finds [5] the excess free energy per particle at $\Gamma = 2$,

$$\lim \frac{F_{\text{ex}}}{N} = \lim\left(-\frac{k_BT}{N}\ln Q_{\text{ex}}\right) = -\frac{e^2}{4}\ln(\pi\rho L^2) + e^2\left(\tfrac{1}{2} - \tfrac{1}{4}\ln 2\pi\right)$$

$$= -\frac{e^2}{4}\ln \pi\rho L^2 + 0.0405e^2$$

2. *Correlations*

For obtaining the one-body and many-body densities associated with the Slater determinant appearing in (5), it is convenient to use normalized

wave functions $\Psi_n(\mathbf{r})$ rather than the unnormalized functions $\psi_n(\mathbf{r})$. Replacing ψ_n by Ψ_n in (5) only changes the multiplicative constant A. The p-body densities associated with the Slater determinant are given by standard expressions in terms of

$$P(\mathbf{r}_1,\mathbf{r}_2) = \sum_n \Psi_n(\mathbf{r}_1)\overline{\Psi}_n(\mathbf{r}_2) \tag{6}$$

One should note, for further reference, that P is the projector on the vector space spanned by the functions $\psi_n(\mathbf{r})$.

For simplicity, we work at once in the thermodynamic limit (the more rigorous approach which starts with a finite system [7,8] gives the same results). The Ψ_n normalized in the whole plane are

$$\Psi_n(\mathbf{r}) = \left(\frac{\rho}{n!}\right)^{1/2} \exp(-\tfrac{1}{2}\pi\rho r^2)(\pi^{1/2}\rho^{1/2}z)^{n-1}$$

An elementary calculation of (6) (where now n runs from 1 to ∞) gives

$$P(\mathbf{r}_1,\mathbf{r}_2) = \rho \exp[\pi\rho(z_1\bar{z}_2 - \tfrac{1}{2}r_1^2 - \tfrac{1}{2}r_2^2)]$$

The one-body density is

$$\rho(\mathbf{r}) = \sum_n |\Psi_n(\mathbf{r})|^2 = P(\mathbf{r},\mathbf{r}) \tag{7}$$

Thus the density is a constant:

$$\rho(\mathbf{r}) = \rho$$

It is remarkable that this sound result comes from a delicate balance between the nonuniform background–particle interaction and the particle–particle interactions.

The two-body density is

$$\begin{aligned}\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) &= \sum_n |\Psi_n(\mathbf{r}_1)|^2 \sum_m |\Psi_m(\mathbf{r}_2)|^2 - \Big|\sum_n \Psi_n(\mathbf{r}_1)\overline{\Psi}_n(\mathbf{r}_2)\Big|^2 \\ &= P(\mathbf{r}_1,\mathbf{r}_1)P(\mathbf{r}_2,\mathbf{r}_2) - |P(\mathbf{r}_1,\mathbf{r}_2)|^2 \end{aligned} \tag{8}$$

where one sees the familiar exchange term. One finds that

$$\rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \rho^2[1 - \exp(-\pi\rho r_{12}^2)]$$

and the truncated two-body density [6], defined as

$$\rho_T^{(2)}(\mathbf{r}_1,\mathbf{r}_2) = \rho^{(2)}(\mathbf{r}_1,\mathbf{r}_2) - \rho^2$$

has the very simple Gaussian form $\rho_T^{(2)} = -\rho^2 \exp(-\pi\rho r_{12}^2)$.

The higher-order distribution functions are given by the general formula

$$\rho^{(p)}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_p) = \det[P(\mathbf{r}_i,\mathbf{r}_j)]_{i,j=1,\ldots,p} \tag{9}$$

while the corresponding truncated distribution functions (Ursell functions) are

$$\rho_T^{(p)}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_p) = (-1)^{p+1} \sum_{(i_1 i_2 \cdots i_p)} P(\mathbf{r}_{i_1},\mathbf{r}_{i_2})P(\mathbf{r}_{i_2},\mathbf{r}_{i_3})\cdots P(\mathbf{r}_{i_p},\mathbf{r}_{i_1}) \quad (10)$$

where the sum runs on all possible cycles $(i_1 i_2 \cdots i_p)$ built with $\{1, 2, \ldots, p\}$.

A somewhat surprising result is that the correlations have a Gaussian decay rather than an exponential decay of the Debye type. There are indications that this Gaussian behavior occurs only at $\Gamma = 2$.

3. *Other Thermodynamic Quantities*

Since all the distribution functions are known in closed form at $\Gamma = 2$, it is possible to compute the derivatives of the free energy with respect to the temperature, at $\Gamma = 2$; these derivatives can be expressed in terms of integrals involving the interaction potential $-\ln(r/L)$ and the distribution functions. By this method, in the thermodynamic limit, the internal energy per particle and the specific heat (at constant volume) per particle have been computed. Their excess values are [6], respectively,

$$\lim \frac{U_{ex}}{N} = -\tfrac{1}{4}e^2[\ln(\pi\rho L^2) + \gamma]$$

where $\gamma = 0.5772\ldots$ is Euler's constant, and

$$\lim \frac{C_{ex}}{N} = k_B\left(\ln 2 - \frac{\pi^2}{24}\right)$$

4. *Other Geometries*

It is possible to start with geometries other than the disk. The particles can be confined on the surface of a sphere [9]. From the start, the partition functions involves factorials rather than incomplete gamma functions, certainly a simplification. In the thermodynamic limit, one finds the same results as above.

Another approach is to start with a rectangular domain and to go first to the limit of a "semiperiodic strip" [10]: a strip of infinite length and finite width, with periodic boundary conditions on its sides. The distribution functions have the remarkable property of being periodic along the strip; in that respect, the strip behaves like a strictly one-dimensional one-component plasma. When the width of the strip is made infinite, the periodicity disappears and, again, one finds the same thermodynamic limits as above.

while the corresponding truncated distribution functions (Ursell functions) are

$$\rho_T^{(p)}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_p) = (-1)^{p+1} \sum_{(i_1 i_2 \cdots i_p)} P(\mathbf{r}_{i_1},\mathbf{r}_{i_2})P(\mathbf{r}_{i_2},\mathbf{r}_{i_3})\cdots P(\mathbf{r}_{i_p},\mathbf{r}_{i_1}) \quad (10)$$

where the sum runs on all possible cycles $(i_1 i_2 \cdots i_p)$ built with $\{1, 2, \ldots, p\}$.

A somewhat surprising result is that the correlations have a Gaussian decay rather than an exponential decay of the Debye type. There are indications that this Gaussian behavior occurs only at $\Gamma = 2$.

3. *Other Thermodynamic Quantities*

Since all the distribution functions are known in closed form at $\Gamma = 2$, it is possible to compute the derivatives of the free energy with respect to the temperature, at $\Gamma = 2$; these derivatives can be expressed in terms of integrals involving the interaction potential $-\ln(r/L)$ and the distribution functions. By this method, in the thermodynamic limit, the internal energy per particle and the specific heat (at constant volume) per particle have been computed. Their excess values are [6], respectively,

$$\lim \frac{U_{ex}}{N} = -\tfrac{1}{4}e^2[\ln(\pi\rho L^2) + \gamma]$$

where $\gamma = 0.5772\ldots$ is Euler's constant, and

$$\lim \frac{C_{ex}}{N} = k_B\left(\ln 2 - \frac{\pi^2}{24}\right)$$

4. *Other Geometries*

It is possible to start with geometries other than the disk. The particles can be confined on the surface of a sphere [9]. From the start, the partition functions involves factorials rather than incomplete gamma functions, certainly a simplification. In the thermodynamic limit, one finds the same results as above.

Another approach is to start with a rectangular domain and to go first to the limit of a "semiperiodic strip" [10]: a strip of infinite length and finite width, with periodic boundary conditions on its sides. The distribution functions have the remarkable property of being periodic along the strip; in that respect, the strip behaves like a strictly one-dimensional one-component plasma. When the width of the strip is made infinite, the periodicity disappears and, again, one finds the same thermodynamic limits as above.

B. Inhomogeneous Plasma: Electrical Double Layer

1. *Problem and Method of Solution*

Our aim is to obtain models describing charged interfaces. For this purpose it is necessary to deal with inhomogeneous one-component plasmas. In the potential energy function, the background–particle interaction of (4), $(e^2/2)\pi\rho r_j^2$, must be replaced by a more general potential $e^2V(\mathbf{r}_j)$, created by a nonuniform background (and perhaps having nonelectrostatic contributions as well).

If the circular symmetry is preserved [i.e., if $V(\mathbf{r})$ depends only on the distance r to the origin], the Boltzmann factor is still a determinant of the form (5), now made of functions

$$\psi_n(\mathbf{r}) = \exp[-V(r)]z^{n-1}$$

which again are orthogonal. Thus the calculation of the one-body and many-body densities proceeds in a way very similar to what has been done for the homogeneous plasma. Cases for which the density varies only in one space direction, say the x-direction, can be solved by a suitable limiting procedure that transforms circles into straight lines: One starts with a circular geometry, and one sends the center of symmetry to infinity in the x-direction. This method has been used for studying a large variety of plane interfaces [11].

Here we describe a more powerful method [12], which applies to a larger class of inhomogeneous cases. For an arbitrary one-body potential $e^2V(\mathbf{r})$, the Boltzmann factor is still proportional to the squared modulus of a determinant:

$$\exp\left(\frac{-\Phi}{k_BT}\right) = A|\det[e^{-V(\mathbf{r})}z_j^{n-1}]|^2$$

but the orbitals $\exp[-V(\mathbf{r})]z^{n-1}$ are not orthogonal to one another in the general case for which $V(\mathbf{r})$ has not the circular symmetry; in terms of nonorthogonal orbitals, there is no standard expression for the densities. Now, the basic remark is that the determinant is changed only by a multiplicative constant if the N orbitals are replaced by N orthonormal linear combinations $\Psi_n(\mathbf{r})$. Therefore, the densities are given by the general expressions (7) to (10) where P is defined by (6) in terms of these orthonormal functions Ψ_n. This P is the projector on the vector space $\mathscr{E}$ spanned by the functions $\exp[-V(\mathbf{r})]z^{n-1}$, and its practical calculation amounts to finding an orthogonal basis for this vector space.

For the cases that are considered here, an essential simplification is

obtained by starting at once with an infinite system, because it is then possible to take advantage of specific translational symmetries and also, as for quantum-mechanical particles in a magnetic field, of a kind of gauge invariance, which here can be described as follows. The one-body potential energy $e^2V(\mathbf{r})$ has an electrostatic part $e^2V_B(\mathbf{r})$. For a given background charge density $-e\rho_B(\mathbf{r})$, the electrostatic potential $eV_B(\mathbf{r})$ is determined by both Poisson's equation $\Delta V_B(\mathbf{r}) = 2\pi\rho_B(\mathbf{r})$ and by boundary conditions. However, because of the screening effect, in an infinite plasma the one-body and many-body densities are expected to be independent of the boundary conditions at infinity; that is, these densities should be invariant under the addition to $V_B(\mathbf{r})$ of an arbitrary function of zero Laplacian (although, in general, the projector P is *not* invariant). This invariance can indeed be explicitly checked. Therefore, it is possible to choose $V_B(\mathbf{r})$ as the simplest possible one for the case under consideration.

2. *Plane Interfaces*

Models for plane interfaces are obtained from a background charge density $-e\rho_B(x)$, which varies only in one space direction, say x. It is then convenient to choose an electrostatic potential of the form $eV_B(x)$, obeying Poisson's equation,

$$\frac{d^2V_B(x)}{dx^2} = 2\pi\rho_B(x) \tag{11}$$

The total one-body potential energy $e^2V(x)$, obtained by adding possible nonelectrostatic contributions to $e^2V_B(x)$, will also depend only on x.

An orthogonal basis for the space $\mathscr{E}$ spanned by the functions $\exp[-V(x)]z^{n-1}$ is provided by the functions

$$\psi_k(\mathbf{r}) = \exp[-V(x) + kz] \qquad k \in \mathbb{R}$$

(at least in the sense of distributions, which is enough for computing a projector). Indeed,

$$z^{n-1} = \frac{d^{n-1}}{dk^{n-1}} e^{kz} \bigg|_{k=0} = \int_{-\infty}^{\infty} (-1)^{n-1}\delta^{(n-1)}(k)e^{kz}\,dk$$

and the functions $\psi_k(\mathbf{r})$ are mutually orthogonal because of the factor $\exp(iky)$ in $\exp(kz)$. Although $\psi_k(\mathbf{r})$ is not normalizable in the y-direction, we can adapt (6), using the standard replacement of a discrete sum $\sum_n$ by an integral $(2\pi)^{-1}\int dk$, and introducing the normalization factor in the x-direction

$$N(k) = \int_{-\infty}^{\infty} |\psi_k(\mathbf{r})|^2\,dx = \int_{-\infty}^{\infty} \exp[-2V(x) + 2kx]\,dx \tag{12}$$

In this way, we obtain the projector on $\mathcal{E}$:

$$P(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{dk}{N(k)}\exp[-V(x_1) + kx_1 - V(x_2) + kx_2 + ik(y_1 - y_2)] \tag{13}$$

The one-body density is $P(\mathbf{r},\mathbf{r})$ and the many-body densities are given by (9) or (10).

The rather general expression (13) can be adapted to more specific cases. For instance, for a homogeneous plasma, the basis can be chosen as $\psi_k = \exp(-\pi\rho x^2 + kz)$, and it can be explicitly checked that this choice gives the same densities as the previous choice $\psi_n = \exp(-\frac{1}{2}\pi\rho r^2)z^{n-1}$. Let us now consider interfaces.

a. Ideally Polarizable Interface. In this model, two different conductors are separated by a plane impermeable to the charges. An externally applied potential difference controls the polarization of the interface: surface charges of opposite magnitudes on each side of the plane. This model grossly mimics systems such as the Hg–FNa(aq) interface, a classical example.

Here we consider a two-dimensional one-component plasma separated into two regions, the half-plane $x > 0$ and the half-plane $x < 0$, by the y-axis assumed to be impermeable to the particles. In general, the background charge density will have different (constant) values $-e\rho_a$ and $-e\rho_b$ in the regions $x > 0$ and $x < 0$, respectively. This system has been studied by a different method in Refs. 13 and 14.

Rather than fixing the numbers of particles in each region, it is equivalent but more convenient to do so only in the average, using a kind of grand-canonical formalism. The impermeability condition can be replaced by the introduction of a potential jump at $x = 0$; this will be seen to be equivalent to a jump in the electrochemical potential. Therefore, to a solution $V_B(x)$ of Poisson's equation (11), we add different constants in each region; without loss of generality, this constant can be chosen as zero in one region. Thus the total one-body potential energy $e^2V(x)$ to be used in (12) and (13) is

$$V(x) = \begin{cases} \pi\rho_a x^2 & x > 0 \\ \pi\rho_b x^2 + V_0 & x < 0 \end{cases}$$

Then the normalization factor (12) has the more explicit form

$$N(k) = \frac{e^{k^2/2\pi\rho_a}}{\sqrt{8\rho_a}}\left[1 + \operatorname{erf}\left(\frac{k}{\sqrt{2\pi\rho_a}}\right)\right] + z_0\frac{e^{k^2/2\pi\rho_b}}{\sqrt{8\rho_b}}\left[1 - \operatorname{erf}\left(\frac{k}{\sqrt{2\pi\rho_b}}\right)\right]$$

where $z_0 = \exp(-2V_0)$ and

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-u^2}\, du$$

is the error function. This gives through (13) and (9) or (10) a one-parameter integral representation for the p-body densities. For instance, the one-body density is

$$\rho(x) = P(\mathbf{r},\mathbf{r}) = \begin{cases} \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} \frac{dk}{N(k)}\, e^{-2\pi\rho_a x^2 + 2kx} & x > 0 \\ \dfrac{1}{2\pi} \displaystyle\int_{-\infty}^{\infty} \frac{dk}{N(k)}\, z_0 e^{-2\pi\rho_b x^2 + 2kx} & x < 0 \end{cases} \tag{14}$$

A typical shape of $\rho(x)$, numerically computed, is shown in Fig. 1.

The departure of $\rho(x)$ from the bulk values ρ_a and ρ_b is localized near $x = 0$. The impermeability causes $\rho(x)$ to have a discontinuity at $x = 0$: $\rho(0^-)/\rho(0^+) = z_0$. It can be checked by explicit calculation that the surface charge densities have opposite values $e\sigma$ and $-e\sigma$ on the $x > 0$ and $x < 0$ sides of the interface, respectively:

$$\sigma = \int_0^{\infty} [\rho(x) - \rho_a]\, dx = -\int_{-\infty}^{0} [\rho(x) - \rho_b]\, dx$$

The electrical potential difference across the interface, $\Delta\phi = \phi(\infty) -$

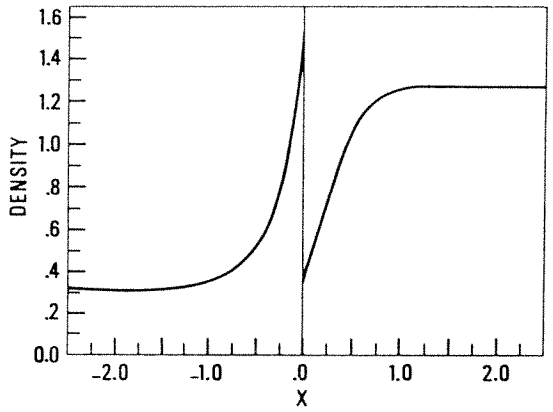


Fig. 1 One-component plasma. Density profile of the ideally polarizable interface, for $\rho_a/\rho_b = 4$, $z_0 = 4.212$; this choice of z_0 gives a surface charge density $e\sigma = -e/\pi$. The unit of length is $(\pi\rho_b)^{-1/2}$. (Adapted from Ref. 14.)

$\phi(-\infty)$, is expected to be related to z_0 by the thermodynamical relation

$$-2V_0 = \ln z_0 = \frac{\mu_b - \mu_a - e\Delta\phi}{k_B T} \tag{15}$$

where μ_a and μ_b are the bulk chemical potentials in the regions $x > 0$ and $x < 0$, respectively; this relation expresses that, in equilibrium, the reversible works needed for extracting a particle from $-\infty(-\mu_b)$, bringing it across the interface $(-e^2V_0 + e\Delta\phi)$, and injecting it into $+\infty(\mu_a)$ add up to zero. Here $\Delta\phi$ can also be computed microscopically: from elementary electrostatics, $\Delta\phi$ is given by the first moment of the charge density profile as

$$\Delta\phi = 2\pi e \left\{ \int_{-\infty}^{0} [\rho(x) - \rho_b] x \, dx + \int_{0}^{\infty} [\rho(x) - \rho_a] x \, dx \right\}$$

and an explicit calculation using (14) gives

$$z_0 = \sqrt{\frac{\rho_b}{\rho_a}} \exp\left(\frac{-2\Delta\phi}{e}\right)$$

This is indeed equivalent to (15), since $\sqrt{\rho_b/\rho_a} = \exp[(\mu_b - \mu_a)/k_B T]$, a consequence of the equation of state (1). Thus we have explicitly checked that both thermodynamics and the exact statistical mechanics of the model give the same relation (15).

By now it should be clear that $\Delta\phi$ (or equivalently, z_0) is an independent parameter, which controls the properties of the interface such as the density profile $\rho(x)$, the total surface charge density on each side $\pm e\sigma$, and so on. The bulk properties, far away from $x = 0$ in each region, are independent of $\Delta\phi$.

b. Nonpolarizable Interface. We now consider two different conductors separated by a plane permeable to the charges. The potential difference is no longer an adjustable parameter; it is determined by the equilibrium condition. This model mimics systems such as the calomel electrode.

Here, we very simply obtain a nonpolarizable interface as a special case of the previous one, by setting $V_0 = 0$ (i.e., $z_0 = 1$) in the previous equations; suppressing the potential jump V_0 amounts to letting the particles move freely through the line $x = 0$. This system has been studied by a different method in Ref. 15.

Now, $\rho(x)$ is continuous at $x = 0$. The relation (15) becomes

$$\mu_b - \mu_a - e\Delta\phi = 0$$

expressing the equality of the electrochemical potentials $\mu_b + e\phi(-\infty)$

and $\mu_a + e\phi(\infty)$. For given bulk densities ρ_a and ρ_b, $\Delta\phi$ is fully determined.

c. Primitive Electrode. This simplest model of an electrolyte–electrode interface pictures the electrode as a structureless impenetrable wall, carrying a given surface charge density. Thus we consider a two-dimensional one-component plasma, with a constant background charge density $-e\rho$, occupying the region $x > 0$. The region $x < 0$ represents an impenetrable electrode: It contains neither background nor particles. The line $x = 0$ carries a given surface charge density $-e\sigma$. This system has been studied by a different method in Refs. 16 and 17.

A somewhat subtle point is that it is illegitimate to apply the general formalism of Section III.B.1 directly to the present geometry, because this formalism has been built for infinite systems with a screening conducting medium at infinity in *all* directions. A correct approach is to describe the primitive electrode as a limiting case, starting with the geometry of the ideally polarizable interface plus a surface charge density $-e\sigma$ at $x = 0$. Thus, introducing the proper discontinuity of dV/dx across the charged surface at $x = 0$, we can choose

$$V(x) = \begin{cases} \pi\rho x^2 - \alpha x & x > 0 \\ \pi\rho_b x^2 - 2\pi(\sigma + \alpha)x + V_0 & x < 0 \end{cases}$$

As discussed in Section III.B.1, the densities should be independent of the constant α, and it can explicitly be checked that this is so; thus we can choose $\alpha = 0$. We then repeat the calculation of Section III.B.1.a, take the limit $\rho_b \to 0$, and at the end only take the limit $V_0 \to \infty$ (i.e., $z_0 \to 0$). In the limit $\rho_b \to 0$, one finds that $[N(k)]^{-1}$ vanishes if $k < -2\pi\sigma$. The result for the projector P, which is nonzero only for x_1 and x_2 nonnegative, then is

$$P(\mathbf{r}_1,\mathbf{r}_2) = \frac{\sqrt{2\rho}}{\pi} \int_{-2\pi\sigma}^{\infty} dk \, \frac{\exp[-\pi\rho(x_1 - k/2\pi\rho)^2 - \pi\rho(x_2 - k/2\pi\rho)^2 + ik(y_1 - y_2)]}{1 + \operatorname{erf}(k/\sqrt{2\pi\rho})} \times x_1,x_2 \geq 0 \quad (16)$$

Typical shapes of the density profile $\rho(x) = P(\mathbf{r},\mathbf{r})$ are shown in Fig. 2. The departure of $\rho(x)$ from its bulk value ρ is localized near the electrode at $x = 0$. It can be checked by an explicit calculation that the surface polarization charge in the plasma is the opposite of the charge $-e\sigma$ on the electrode:

$$e \int_0^{\infty} [\rho(x) - \rho] \, dx = e\sigma$$

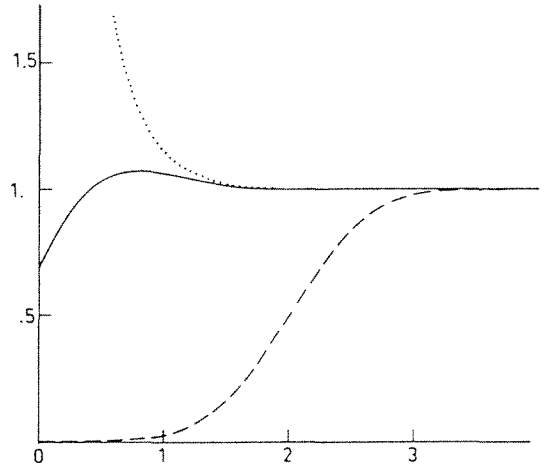


Fig. 2 One-component plasma. Density profiles $\rho(x)/\rho$ in the primitive electrode model. The solid line is for an uncharged electrode, the dashed line for a very positive surface charge density on the electrode, the dotted line for a very negative surface charge density on the electrode. The unit of length is $(\pi\rho)^{-1/2}$.

d. Miscellaneous Other Cases. Many special cases or variants of the examples just discussed have been studied explicitly, starting from a circular geometry, as skeched at the beginning of Section III.B.1. The interested reader is referred to the literature [8,11,15,18–20]. The results can be retrieved in an easier and more systematic way as special cases or limits of (13).

Some cases with image forces have been solved. The primitive electrode model can be modified by considering a plasma along a wall made of a material of either zero or infinite dielectric constant. The zero dielectric constant case is not completely academic, because what matters is the ratio of the dielectric constant of the container wall to the dielectric constant of the electrolyte solvent, and this ratio is actually small for real systems; the zero dielectric constant case has been solved by introducing the proper image forces into the Hamiltonian of the plasma [7,18]. The infinite dielectric constant model pictures the electrode as an ideal conductor; this case also has been solved by introducing the proper image forces into the Hamiltonian of the plasma [21]; it is, however, much easier to take the limit of the polarizable interface when on one side the plasma density becomes infinite [22], because this plasma then behaves as an ideal conductor (its correlation length, which is of the order of $\rho^{-1/2}$, goes to zero).

3. *Interfaces with Adsorption Sites*

An electrode in contact with an electrolyte may adsorb ions; there is an equilibrium between the adsorbed ions and the electrolyte ions. The adsorption sites are likely to form a surface lattice, which reflects the periodic structure of the electrode.

It is possible to mimic these adsorption effects, yet still have a solvable model, by adding equidistant adsorption sites on the interfaces studied in Section III.B.2. The potential $v_{ad}(\mathbf{r})$ created at $\mathbf{r}$ by an adsorption site at the origin is assumed to be the sticky potential of Baxter [i.e., such that $\exp(-v_{ad}/k_BT) = 1 + \lambda\delta(\mathbf{r})$]. Special cases have been solved by a *tour de force* of expansion resummations [23]. It is, however, simpler and more efficient to use [24] the general method described in Section III.B.1.

The Boltzmann factor associated to the total one-body potential now is $\exp[-2V(x)][1 + \lambda\delta(x) \sum_{m\in\mathbb{Z}} \delta(y - ma)]$, where a is the distance between adjacent adsorption sites; as in Section III.B.2, $e^2V(x)$ is the one-body potential energy associated to the background and possibly includes other contributions from walls, potential steps, and so on. The densities are given by (9) or (10) in terms of the projector P on the vector space $\mathcal{E}$ spanned by the functions

$$\varphi_k(\mathbf{r}) = [N(k)]^{-1/2} e^{-V(x)+kz}[1 + \lambda\delta(x) \sum_{m\in\mathbb{Z}} \delta(y - ma)]^{1/2}$$

For convenience, we have already included

$$N(k) = \int_{-\infty}^{\infty} e^{-2V(x)+2kx}\, dx$$

which is a normalization factor when $\lambda = 0$.

For $\lambda \neq 0$, the basis formed by the $\varphi_k(\mathbf{r})$ is not an orthogonal one. However, the matrix of the scalar products is block diagonal, as a consequence of the periodicity in y of the total one-body potential; indeed, writing $k = 2\pi(\zeta + n)/a$, with $\zeta \in [0,1]$ and $n \in \mathbb{Z}$, we find scalar products of the form

$$\int d\mathbf{r}\, \overline{\varphi}_k(\mathbf{r})\varphi_{k'}(\mathbf{r}) = a\delta(\zeta - \zeta')A_\zeta(n,n')$$

If we are able to find the orthogonal eigenvectors $(a_n^{(i)})$ and the associated eigenvalues α_i of the matrix A_ζ, such that

$$\sum_{n'} A_\zeta(n,n')a_{n'}^{(i)} = \alpha_i a_n^{(i)}$$

we obtain an orthogonal basis of $\mathscr{E}$: the functions

$$\psi_{\zeta,i}(\mathbf{r}) = \sum_n a_n^{(i)} \varphi_{2\pi(\zeta+n)/a}(\mathbf{r})$$

Then the projector P is

$$P(\mathbf{r}_1,\mathbf{r}_2) = \frac{1}{a}\int_0^1 d\zeta \sum_i \frac{\psi_{\zeta,i}(\mathbf{r}_1)\bar{\psi}_{\zeta,i}(\mathbf{r}_2)}{\alpha_i \sum_n [a_n^{(i)}]^2}$$

The peculiar form of the Baxter potential gives

$$A_\zeta(n,n') = \delta_{nn'} + f_\zeta(n) f_\zeta(n')$$

where

$$f_\zeta(n) = \left[\frac{\lambda}{aN[(2\pi/a)(\zeta+n)]}\right]^{1/2} e^{-V(0)}$$

and the problem for a matrix of this form is a standard one. A_ζ has two eigenvalues: $\alpha_1 = 1 + \sum_n [f_\zeta(n)]^2$, associated with the eigenvector $[a_n^{(1)} = f_\zeta(n)]$, and $\alpha_2 = 1$, associated with all the other eigenvectors that span the subspace orthogonal to $a^{(1)}$. After some rearrangement, one finds

$$P(\mathbf{r}_1,\mathbf{r}_2) = [1 + \lambda\delta(x_1)\sum_m \delta(y_1 - ma)]^{1/2}[1 + \lambda\delta(x_2)\sum_m \delta(y_2 - ma)]^{1/2} P^*(\mathbf{r}_1,\mathbf{r}_2)$$

with

$$P^*(\mathbf{r}_1,\mathbf{r}_2) = P_0(\mathbf{r}_1,\mathbf{r}_2) - e^{-V(x_1)-V(x_2)} \int_0^1 \frac{d\zeta}{a} \frac{\mu G(\zeta,z_1)G(\zeta,\bar{z}_2)}{1+\mu G(\zeta,0)}$$

where P_0 is the projector in the absence of the adsorption sites, $\mu = \lambda \exp[-2V(0)]/a$, and

$$G(\zeta,z) = \sum_{n\in\mathbb{Z}} \left[N\left(\frac{2\pi}{a}(\zeta+n)\right)\right]^{-1} \exp\left[\frac{2\pi}{a}(\zeta+n)z\right]$$

Obviously, when $\mathbf{r}_1$ and $\mathbf{r}_2$ are not on adsorption sites, $P = P^*$, and therefore P^* describes the nonadsorbed particles; their density is $\rho^*(\mathbf{r}) = P^*(\mathbf{r},\mathbf{r})$. As to the adsorbed particles, they are described by the singular part of P; the mean occupation number of an adsorption site can be written as

$$n_{ad} = \lambda\rho^*(0) = \int_0^1 d\zeta \frac{\mu G(\zeta,0)}{1+\mu G(\zeta,0)}$$

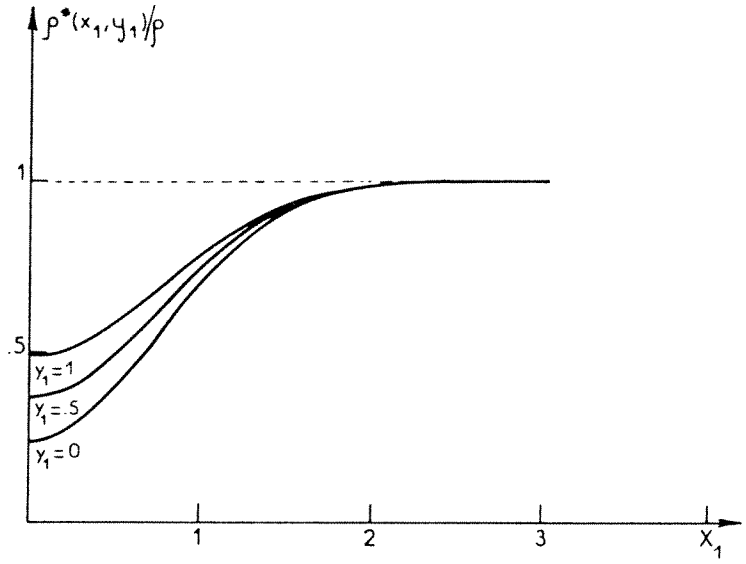


Fig. 3 One-component plasma. Density profiles near a line of adsorption sites; $\lambda = 10$ and $a = 2$. The unit of length is $(\pi\rho)^{-1/2}$. (From Ref. 23.)

More detail can be found in Refs. 23 and 24. Here, let us only mention that, in the simplest case of a line of adsorption sites in a homogeneous background of charge density $-e\rho$, we can take $V(x) = \pi\rho x^2$, and $G(\zeta,z)$ can be expressed in terms of the Jacobi theta function*

$$\theta_3(u,t) = \left(\frac{\pi}{t}\right)^{1/2} \sum_{n\in\mathbb{Z}} \exp\left[\frac{-\pi^2(u+n)^2}{t}\right]$$

as

$$G(\zeta,z) = a\rho \exp\left(\frac{\pi\rho z^2}{2}\right)\theta_3\left(\zeta - \frac{\rho a z}{2}, \frac{\pi\rho a^2}{2}\right)$$

The density $\rho^*(\mathbf{r})$ of nonadsorbed particles is shown in Fig. 3.

IV. TWO-COMPONENT PLASMA AT $\Gamma = 2$

Also for the two-component plasma, $\Gamma = 2$ corresponds to a special temperature at which many exact results are available. However, since this

* The present definition of the theta function is the same as in Ref. 23.

temperature is precisely the one at which a collapse occurs for the pure Coulomb interaction system, which then has a divergent configuration integral, it is necessary to introduce some kind of short-distance cutoff, for instance by representing the particles as charged hard disks of diameter d; exact results can be obtained near the limit $d \to 0$. This small hard-disk model will be obtained by considering the continuum limit of a lattice model [25] (for which there are no divergences). It is now convenient to work in the grand-canonical formalism.

A. Lattice Model

We represent the position $\mathbf{r}$ of a particle by the complex number $z = x + iy$, where (x,y) are the Cartesian components of $\mathbf{r}$. For a system of N positive and N negative particles, the complex coordinates of which are u_i and v_i, respectively, the Boltzmann factor is, at $\Gamma = 2$,

$$\exp\left\{2\sum_{i<j}\left[\ln\left|\frac{u_i-u_j}{L}\right| + \ln\left|\frac{v_i-v_j}{L}\right|\right] - 2\sum_{ij}\ln\left|\frac{u_i-v_j}{L}\right|\right\}$$

$$= L^{2N}\left|\frac{\Pi_{i<j}(u_i-u_j)(v_i-v_j)}{\Pi_{ij}(u_i-v_j)}\right|^2$$

$$= L^{2N}\left|\left[\det\frac{1}{u_i-v_j}\right]_{i,j=1,\ldots,N}\right|^2$$

where the last equality stems from the Cauchy double alternant determinant formula. Two interwoven lattices U and V are introduced. The positive (negative) particles sit on the sublattice $U(V)$; each lattice site is occupied by no or one particle. A possible external potential is described by position-dependent fugacities $\lambda(u_i)$ and $\lambda(v_i)$. Then the grand partition function (here defined as a sum including only neutral system) is

$$Z = 1 + L^2\sum_{\substack{u\in U\\ v\in V}}\lambda(u)\lambda(v)\left|\frac{1}{u-v}\right|^2$$

$$+ L^4\sum_{\substack{u_1,u_2\in U\\ v_1,v_2\in V}}\lambda(u_1)\lambda(u_2)\lambda(v_1)\lambda(v_2)\left|\det\left[\frac{1}{u_i-v_j}\right]_{i,j=1,2}\right|^2 + \cdots$$

(where the sums are defined with the prescriptions that configurations which differ only by a permutation of identical particles are counted only once). It can be easily seen that this grand partition function is the ex-

pansion of a determinant built on *all* lattice sites:

$$Z = \det \begin{bmatrix} 1 & 0 & \cdots & \dfrac{L\lambda(u_1)}{u_1 - v_1} & \dfrac{L\lambda(u_1)}{u_1 - v_2} & \cdots \\ 0 & 1 & \cdots & \dfrac{L\lambda(u_2)}{u_2 - v_1} & \dfrac{L\lambda(u_2)}{u_2 - v_2} & \cdots \\ & & \cdots & & & \cdots \\ \dfrac{L\lambda(v_1)}{\bar{v}_1 - \bar{u}_1} & \dfrac{L\lambda(v_1)}{\bar{v}_1 - \bar{u}_2} & \cdots & 1 & 0 & \cdots \\ \dfrac{L\lambda(v_2)}{\bar{v}_2 - \bar{u}_1} & \dfrac{L\lambda(v_2)}{\bar{v}_2 - \bar{u}_2} & \cdots & 0 & 1 & \cdots \\ & & \cdots & & & \end{bmatrix}$$

A more compact notation can be used: Each lattice site is characterized by its complex coordinate z (or, alternatively, the corresponding vector $\mathbf{r}$) and an isospin index s; $s = +1$ (-1) if the site at $\mathbf{r}$ belongs to the positive sublattice U (negative sublattice V). The position-dependent fugacities will be called $\lambda_+(\mathbf{r})$ for positive sites and $\lambda_-(\mathbf{r})$ for negative sites. A matrix built on all lattice sites can be regarded as a direct product of a matrix in coordinate space and a 2×2 matrix in isospin space; a matrix element taken only in coordinate space is still a 2×2 matrix in isospin space. In this notation, introducing the usual Pauli matrices 1, σ_x, σ_y, σ_z operating in isospin space, we write Z as

$$Z = \det\left\{1 + \left[\lambda_+(\mathbf{r})\frac{1 + \sigma_z}{2} + \lambda_-(\mathbf{r})\frac{1 - \sigma_z}{2}\right]\langle \mathbf{r} \mid K \mid \mathbf{r}' \rangle\right\}$$

where

$$\langle \mathbf{r} \mid K \mid \mathbf{r}' \rangle = \frac{\sigma_x + i\sigma_y}{2}\frac{L}{z - z'} + \frac{\sigma_x - i\sigma_y}{2}\frac{L}{\bar{z} - \bar{z}'}$$

Along these lines, in the case of an homogeneous system (no external potential), the equation of state and the correlation functions have been explicitly computed [25] in terms of integrals involving elliptic functions. Here, rather than describing these results about the lattice system, we shall consider the continuum limit [26], which is substantially simpler.

B. Continuum Limit

The great simplification that occurs in the continuum limit (i.e., as the lattice spacing goes to zero) is that the *inverse* of matrix K becomes a simple differential operator. Indeed, $\langle \mathbf{r} \mid K \mid \mathbf{r}' \rangle$ can be written as

$$\langle \mathbf{r} \mid K \mid \mathbf{r}' \rangle = L(\sigma_x \partial_x + \sigma_y \partial_y) \ln \mid \mathbf{r} - \mathbf{r}' \mid$$

and since $\nabla^2 \ln r = 2\pi\delta(\mathbf{r})$, it is obvious that the inverse operator is

$$K^{-1} = \frac{S}{2\pi L}(\sigma_x\partial_x + \sigma_y\partial_y)$$

(the area per site S appears when discrete sums are replaced by integrals). Thus an alternative form of Z is

$$Z = \det\left\{\left[\sigma_x\partial_x + \sigma_y\partial_y + m_+(\mathbf{r})\frac{1+\sigma_z}{2} + m_-(\mathbf{r})\frac{1-\sigma_z}{2}\right][\sigma_x\partial_x + \sigma_y\partial_y]^{-1}\right\}$$

where $m_\pm(\mathbf{r}) = (2\pi L/S)\lambda_\pm(\mathbf{r})$ is a rescaled fugacity that has the dimensions of an inverse length) and

$$\ln Z = \mathrm{Tr}\left\{\ln\left[\sigma_x\partial_x + \sigma_y\partial_y + m_+(\mathbf{r})\frac{1+\sigma_z}{2} + m_-(\mathbf{r})\frac{1-\sigma_z}{2}\right] - \ln[\sigma_x\partial_x + \sigma_y\partial_y]\right\} \tag{17}$$

These forms of Z and $\ln Z$ involve the two-dimensional Dirac operator and express an equivalence, well known to field theorists, between the Coulomb gas at $\Gamma = 2$ and a *free* Fermi field. It is remarkable that introduction of an external potential through a position-dependent fugacity is equivalent to a position-dependent mass for the Fermions.

Although the last expression of $\ln Z$ is a formal one that is not properly defined without some regularization, we use it for obtaining the one-body densities and the many-body truncated densities, in the usual way, by taking functional derivatives with respect to the fugacities $m_\pm(\mathbf{r})$. Marking the sign of the particle at $\mathbf{r}_i$ by the index $s_i = \pm 1$, and defining the matrix

$$G_{s_1s_2}(\mathbf{r}_1,\mathbf{r}_2) = \left\langle \mathbf{r}_1 s_1 \left| \left[\sigma_x\partial_x + \sigma_y\partial_y + m_+(\mathbf{r})\frac{1+\sigma_z}{2} + m_-(\mathbf{r})\frac{1-\sigma_z}{2}\right]^{-1} \right| \mathbf{r}_2 s_2 \right\rangle \tag{18}$$

we obtain the one-particle densities

$$\rho_{s_1}(\mathbf{r}_1) = m_{s_1}(\mathbf{r}_1)G_{s_1s_1}(\mathbf{r}_1,\mathbf{r}_1) \tag{19}$$

the truncated two-body densities

$$\rho^{(2)T}_{s_1s_2}(\mathbf{r}_1,\mathbf{r}_2) = -m_{s_1}(\mathbf{r}_1)m_{s_2}(\mathbf{r}_2)G_{s_1s_2}(\mathbf{r}_1,\mathbf{r}_2)G_{s_2s_1}(\mathbf{r}_2,\mathbf{r}_1) \tag{20}$$

and more generally the truncated p-body densities

$$\rho^{(p)T}_{s_1 s_2 \ldots s_p}(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_p) = (-1)^{p+1} m_{s_1}(\mathbf{r}_1) m_{s_2}(\mathbf{r}_2)\cdots m_{s_p}(\mathbf{r}_p) \times \sum_{(i_1 i_2 \cdots i_p)} G_{s_{i_1} s_{i_2}}(\mathbf{r}_{i_1},\mathbf{r}_{i_2})\cdots G_{s_{i_p} s_{i_1}}(\mathbf{r}_{i_p},\mathbf{r}_{i_1}) \quad (21)$$

where the summation runs over all cycles $(i_1 i_2 \cdots i_p)$ built with $\{1, 2, \ldots, p\}$.

Therefore, the calculation of the one-body and n-body densities reduces to obtaining G, which is a Green's function, the solution of a system of four coupled partial differential equations. In a 2×2 matrix notation,

$$\left[\sigma_x \partial_{x_1} + \sigma_y \partial_{y_1} + m_+(\mathbf{r}_1)\frac{1+\sigma_z}{2} + m_-(\mathbf{r}_1)\frac{1-\sigma_z}{2}\right] G(\mathbf{r}_1,\mathbf{r}_2) = \mathbb{1}\delta(\mathbf{r}_1 - \mathbf{r}_2) \quad (22)$$

Since the correlations (21) have to decay at large distances, the boundary conditions for (22) are that $[m_{s_1}(\mathbf{r}_1) m_{s_2}(\mathbf{r}_2)]^{1/2} G_{s_1 s_2}(\mathbf{r}_1,\mathbf{r}_2) \to 0$ as $|\mathbf{r}_1 - \mathbf{r}_2| \to \infty$.

By using the formal expansion of G in powers of $\sigma_x \partial_x + \sigma_y \partial_y$ it is easy to derive the useful symmetry relations

$$G_{ss}(\mathbf{r}_1,\mathbf{r}_2) = \overline{G_{ss}(\mathbf{r}_2,\mathbf{r}_1)}$$
$$G_{s-s}(\mathbf{r}_1,\mathbf{r}_2) = -\overline{G_{-ss}(\mathbf{r}_2,\mathbf{r}_1)}$$

C. Homogeneous Plasma

The simplest case of a homogeneous two-component plasma is described by taking constant fugacities $m_+ = m_- = m$. In that case, $G_{++} = G_{--}$, and by combining the elements of the matrix equation (22) one obtains

$$(m^2 - \nabla_1^2) G_{++}(\mathbf{r}_1,\mathbf{r}_2) = m\delta(\mathbf{r}_1 - \mathbf{r}_2)$$

and

$$m G_{-+}(\mathbf{r}_1,\mathbf{r}_2) = -(\partial_{x_1} + i\partial_{y_1}) G_{++}(\mathbf{r}_1,\mathbf{r}_2)$$

The solution is

$$G_{++}(\mathbf{r}_1,\mathbf{r}_2) = \frac{m}{2\pi} K_0(m|\mathbf{r}_1 - \mathbf{r}_2|) \quad (23a)$$

and

$$G_{-+}(\mathbf{r}_1,\mathbf{r}_2) = \frac{m}{2\pi} \frac{x_1 - x_2 + i(y_1 - y_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} K_1(m|\mathbf{r}_1 - \mathbf{r}_2|) \quad (23b)$$

where K_0 and K_1 are modified Bessel functions. The many-body truncated

densities are then given by (20) and (21); the two-body truncated densities, for instance, have the simple forms

$$\rho_{++}^{(2)T}(r) = -\left(\frac{m^2}{2\pi}\right)^2 [K_0(mr)]^2$$

$$\rho_{+-}^{(2)T}(r) = \left(\frac{m^2}{2\pi}\right)^2 [K_1(mr)]^2$$

Since, for large r, the Bessel functions K_0 and K_1 essentially behave like $\exp(-mr)$, the rescaled fugacity m is the inverse correlation length. It is very remarkable that for a given value of the rescaled fugacity m (or of the correlation length m^{-1}), the correlation functions [i.e., the n-body ($n \geq 2$) *truncated* densities] are well-defined quantities for the presently considered continuum-point particle system with pure Coulomb interactions.

The catastrophic collapse of the nonregularized model, however, shows up in the one-body density (19) which is infinite since $K_0(mr)$ diverges logarithmically as $r \to 0$. For suppressing this divergence, we replace the point particles by charged hard disks of diameter d. Near the limit $md \to 0$, we can keep the point-particle expression for the correlation functions, for separations larger than d, and compute the one-body density by using the perfect-screening rule

$$\rho_+ = \rho_- = \int_{r>d} d^2\mathbf{r}[\rho_{+-}^{(2)T}(r) - \rho_{++}^{(2)T}(r)]$$

(which expresses that the charge in the polarization cloud around a given particle is opposite to the charge of that particle). In this way we obtain the behavior of the density ρ_s of particles of one sign, as $md \to 0$:

$$\rho_+ = \rho_- \sim \frac{m^2}{2\pi}\left(\ln\frac{2}{md} - \gamma\right) \sim \frac{m^2}{2\pi} K_0(md) \tag{24}$$

where $\gamma = 0.5772\ldots$ is Euler's constant. Thus for small hard disks of diameter d, one obtains ρ_s simply by replacing the infinite $K_0(0)$ by the finite $K_0(md)$ before using (23) in (19).

Either by integrating $2m^{-1}\rho_s(m)$ or by using a regularized form of (17), one obtains for the pressure p

$$\frac{p}{k_B T} = \frac{m^2}{2\pi}\left(\ln\frac{2}{md} - \gamma + \frac{1}{2}\right)$$

As $d \to 0$, one finds the finite limit $p/2k_BT\rho_s \to \frac{1}{2}$. This is indeed the expected result for an ideal gas of collapsed neutral pairs.

D. Inhomogeneous Plasma: Electrical Double Layer

1. Method

For computing the densities in the general inhomogeneous case, one has to solve (22). Here we consider only plane interfaces. The interface is assumed to be along the y-axis, and the system is translationally invariant in the y direction [i.e., the position-dependent fugacities $m_\pm(\mathbf{r})$ actually depend only on x]. In such a geometry, the standard technique is to Fourier-transform $G(\mathbf{r}_1,\mathbf{r}_2)$ with respect to $y_2 - y_1$ (G depends on y_1 and y_2 only through their difference):

$$G(\mathbf{r}_1,\mathbf{r}_2) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \hat{G}(x_1,x_2,k) e^{ik(y_1 - y_2)} \tag{25}$$

In terms of $\hat{G}$, one obtains ordinary differential equations, in one variable x_1, such as

$$m_+(x_1)\hat{G}_{++}(x_1,x_2,k) + \left(\frac{d}{dx_1} + k\right)\hat{G}_{-+}(x_1,x_2,k) = \delta(x_1 - x_2) \tag{26a}$$

$$\left(\frac{d}{dx_1} - k\right)\hat{G}_{++}(x_1,x_2,k) + m_-(x_1)\hat{G}_{-+}(x_1,x_2,k) = 0 \tag{26b}$$

For the cases under consideration, these equations can be solved easily and explicitly. In general, the fugacities $m_\pm(x_1)$ will be discontinuous on the interface at $x_1 = 0$; the solutions of (26) in the regions $x_1 > 0$ and $x_2 < 0$ must be connected by the conditions that $\hat{G}_{++}$ and $\hat{G}_{-+}$ be continuous at $x_1 = 0$.

2. Plane Interfaces

We now consider a few simple more specific models.

a. Primitive Electrode. In this two-component plasma version of the primitive electrode model, the half-space $x < 0$ represents an impenetrable electrode; the fugacities $m_\pm(x)$ vanish in that region. The line $x = 0$ carries a given surface charge density $-e\sigma$. This electrode, and another one of opposite charge assumed to be at $x = +\infty$, generate an electrostatic potential which is 0 for $x < 0$ and $2\pi e\sigma x$ for $x > 0$; correspondingly, the fugacities are of the form $m_s(x) = m \exp(-4\pi s\sigma x)$ in the half-space $x > 0$ where the Coulomb gas is.

We want to solve (26) for a source point at $x_2 > 0$. In the wall region $x_1 < 0$, as $m_\pm \to 0$, the solution that vanishes at $x_1 = -\infty$ becomes such that $\hat{G}_{++} = 0$ if $k < 0$, $\hat{G}_{-+} = 0$ if $k > 0$. Thus, for the solution in the region $x_1 \geq 0$, continuity at $x_1 = 0$ gives the boundary conditions at x_1

$= 0$: $\hat{G}_{++}(x_1 = 0, x_2, k) = 0$ if $k < 0$, $\hat{G}_{-+}(x_1 = 0, x_2, k) = 0$ if $k > 0$. The effect of these boundary conditions is to add a "reflected" exponential to the free-space exponential solution of (26). The result is

$$[m_s(x_1)m_s(x_2)]^{1/2}\hat{G}_{ss}(x_1, x_2, k) = \frac{m^2}{2\kappa(k)}\{\exp[-\kappa(k) \mid x_1 - x_2 \mid + A_s(k)\exp[-\kappa(k)(x_1 + x_2)]\} \tag{27}$$

where

$$\kappa(k) = [m^2 + (k - 2\pi\sigma)^2]^{1/2}$$

and

$$A_+(k) = -1 \qquad \text{if } k < 0$$

$$A_+(k) = \frac{\kappa(k) - k + 2\pi\sigma}{\kappa(k) + k - 2\pi\sigma} \qquad \text{if } k > 0$$

$$A_-(k) = \frac{\kappa(k) + k - 2\pi\sigma}{\kappa(k) - k + 2\pi\sigma} \qquad \text{if } k < 0$$

$$A_-(k) = -1 \qquad \text{if } k > 0$$

$\hat{G}_{-+}$ is given by (26b). This gives through (25) and (21) a one-particle integral representation for the n-body densities. However, in the one-body density $\rho_s(x)$, there is a divergent bulk contribution ρ_s [from the first term on the right-hand side of (27)] for which one must use the regularized form (24).

The density profiles are shown in Fig. 4 for an uncharged and a charged wall. At $x = 0$, a further divergence has to be regularized: $\rho_s(0) - \rho_s$ must be replaced by $\rho_s(d/2) - \rho_s$, which means that the center of a hard disk has its radius $d/2$ as its minimum abscissa. Away from the interface, the densities go to their common bulk value (24), governed by m, while the electrode charge density $-e\sigma$ affects only the structure near the electrode. By an explicit calculation, one can check the expected screening rule

$$e\int_0^\infty [\rho_+(x) - \rho_-(x)]\,dx = e\sigma$$

which expresses that the surface polarization charge in the Coulomb gas is the opposite of the surface charge $-e\sigma$ on the electrode.

b. Ideally Polarizable Interface. The model is a two-component plasma separated into two regions (1) (the half-plane $x > 0$) and (2) (the half-plane $x < 0$) by a membrane (the y-axis) impermeable to the particles. As in the one-component model, the impermeability condition can be

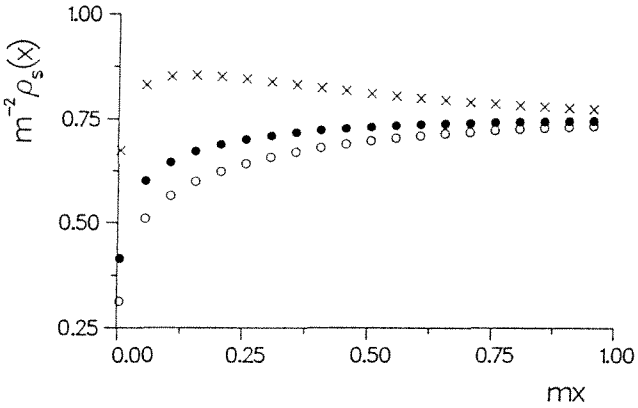


Fig. 4 Two-component plasma. Density profiles in the primitive electrode model. For an uncharged wall ($\sigma = 0$), $\rho_+(x) = \rho_-(x)$ (black circles). For a charged wall ($2\pi\sigma = m$), $\rho_+(x)$ (crosses) and $\rho_-(x)$ (white circles). The cutoff is $md = 0.01$.

described by using different constant fugacities on each side. Altogether, there are four constant fugacities; without loss of generality, they can be expressed in terms of the four constants m_a, m_b, ϕ_a, ϕ_b, defined through

$$m_s(x) = \begin{cases} m_a \exp\left(\dfrac{2s\phi_a}{e}\right) & \text{if } x > 0 \\ m_b \exp\left(\dfrac{2s\phi_b}{e}\right) & \text{if } x < 0 \end{cases} \tag{28}$$

With this choice, (26) reduces in each region to a system of linear differential equations with constant coefficients, the solutions of which have to be connected at $x_1 = 0$ by the continuity requirement. The result is

$$[m_+(x_1)m_+(x_2)]^{1/2}\hat{G}_{++}(x_1,x_2,k) = \frac{m_a^2}{2\kappa_a}\left\{\exp[-\kappa_a \mid x_1 - x_2 \mid] + \frac{(\kappa_a - k)m_b - (\kappa_b - k)e^{-2\Delta\phi/e}m_a}{(\kappa_a + k)m_b + (\kappa_b - k)e^{-2\Delta\phi/e}m_a}\exp[-\kappa_a(x_1 + x_2)]\right\}$$

if $x_1, x_2 > 0$

$$[m_+(x_1)m_+(x_2)]^{1/2}\hat{G}_{++}(x_1,x_2,k) = \frac{m_b^2}{2\kappa_b}\left\{\exp[-\kappa_b \mid x_1 - x_2 \mid]\right.$$

$$+ \frac{(\kappa_b + k)m_a - (\kappa_a + k)e^{2\Delta\phi/e}m_b}{(\kappa_b - k)m_a + (\kappa_a + k)e^{2\Delta\phi/e}m_b} \exp[-\kappa_b \mid x_1 + x_2 \mid]\Bigg\}$$

if $x_1, x_2 < 0$

$$[m_+(x_1)m_+(x_2)]^{1/2}\hat{G}_{++}(x_1,x_2,k) = (m_a m_b)^{3/2}[(\kappa_a + k)m_b e^{\Delta\phi/e} + (\kappa_b - k)m_a e^{-\Delta\phi/e}]^{-1} \exp[-\kappa_a \mid x_1 \mid - \kappa_b \mid x_2 \mid]$$

if $x_1 > 0, \quad x_2 < 0$

where $\kappa_a = (m_a^2 + k^2)^{1/2}$, $\kappa_b = (m_b^2 + k^2)^{1/2}$, $\Delta\phi = \phi_a - \phi_b$. One obtains $[m_-(x_1)m_-(x_2)]^{1/2}\hat{G}_{--}(x_1,x_2,k)$ by changing the signs of $\Delta\phi$ and k in the expressions above. One obtains $\hat{G}_{-+}$ from (26b). Again, this gives through (25) and (21) an integral representation for the densities; a finite one-body density is obtained by an appropriate regularization.

The density profiles are shown in Fig. 5. By an explicit calculation, one can check that the total charges on each side of the interface have opposite values:

$$\sigma = \int_0^\infty dx[\rho_+(x) - \rho_-(x)] = -\int_{-\infty}^0 dx[\rho_+(x) - \rho_-(x)]$$

Far away from the interface, all the densities are the bulk ones, determined by the fugacities m_a and m_b in regions (a) and (b), respectively. Near the interface, the densities depart from their bulk values; this departure depends on the additional parameters ϕ_a and ϕ_b, although only the difference

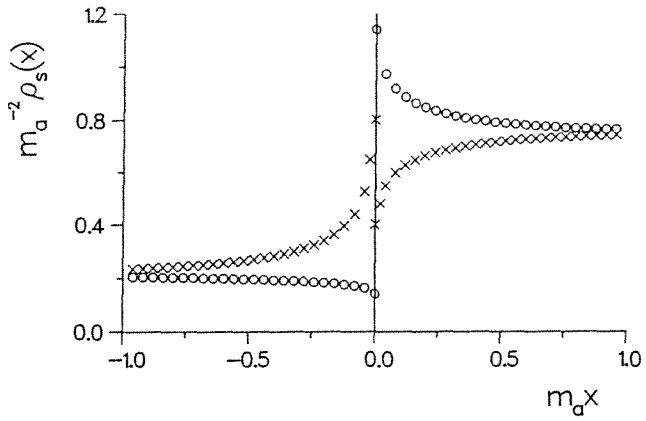


Fig. 5 Two-component plasma. Density profiles of the ideally polarizable interface, for $m_b = 0.5m_a$ and $\exp(-2\Delta\phi/e) = 4$: $\rho_+(x)$ (crosses) and $\rho_-(x)$ (white circles). The cutoff is $md = 0.01$.

$\Delta\phi = \phi_a - \phi_b$ appears in the formulas. These results are not unexpected. If we define bulk chemical potentials μ_a and μ_b by $m_a = m_0 \exp(\mu_a/k_BT)$ and $m_b = m_0 \exp(\mu_b/k_BT)$ where m_0 is some inverse length, the fugacities of the particles of sign s in regions (a) and (b) can be written as $m_0 \exp[(\mu_a + se\phi_a)/k_BT]$ and $m_0 \exp[(\mu_b + se\phi_b)/k_BT]$, respectively. This means that $\mu_a + se\phi_a$ and $\mu_b + se\phi_b$ can be interpreted as electrochemical potentials, including an electrical part $se\phi_a$ or $se\phi_b$, and ϕ_a and ϕ_b can be interpreted as constant electrical potentials in the bulks of regions (a) and (b). This interpretation is corroborated by an explicit calculation of the first moment of the charge density profile across the interface which indeed reproduces the correct potential difference:

$$2\pi e \int_{-\infty}^{\infty} dx\, x[\rho_+(x) - \rho_-(x)] = \Delta\phi$$

c. Miscellaneous Other Cases. If, in the results for the impermeable membrane, we take the limit $m_b \to 0$, we retrieve the hard-wall expressions. Another interesting limit is $m_b \to \infty$. Then, the correlation length m_b^{-1} in region (b) vanishes, and region (b) becomes an ideal conductor.

As a special case of the impermeable membrane, we can obtain a solvable model of a semipermeable membrane (i.e., a membrane permeable to one species, say the positive particles, and impermeable to the other species). Now, there are only two control parameters, which can be chosen as the bulk fugacities on each side, m_a and m_b. We expect the electrochemical potential of the positive particles to be the same on both sides, since these positive particles can freely cross the membrane; this condition, $m_a \exp(e\phi_a/k_BT) = m_b \exp(e\phi_b/k_BT)$, then determines $\Delta\phi = \phi_a - \phi_b$, which is no longer a free parameter. With this condition, it is easy to check that the formalism of Section III.D.1.b gives indeed a density of positive particles that is continuous at $x = 0$, while in general the density of negative particles has a jump.

V. FURTHER EXPLOITATION OF THE RESULTS

The results that we have obtained for solvable models will now be used as illustrations of quantities and properties of general interest in the physics of Coulomb fluids.

A. Macroscopic Physics of Interfaces

1. Relevant Control Parameters

A fundamental law of the electrostatics of conductors says that they "want" to remain neutral inside. In the bulk, the charge density vanishes;

if there is any excess charge, it concentrates on the surface. At an interface between two conductors, there is in general a local nonzero charge density (the electrical double layer), but its net charge *is* zero. For our solvable models, these basic facts are rooted in statistical mechanics.

Indeed, for the homogeneous one-component plasma, the formalism gives [see (7)] a particle charge density opposite to the background charge density. The presence of external electrodes at infinity does not change anything; these electrodes create inside the plasma an additional potential of zero Laplacian, which is irrelevant, as discussed in Section III.B.1. For one-component models of interfaces, the formalism gives bulk properties far away from the interface and opposite charges on each side of the interface. In addition to the background densities, for polarizable interfaces there is one additional control parameter that governs the surface properties. This parameter has been chosen as the surface charge density for the primitive electrode model and as the potential difference across the interface [see (15)] for the ideally polarizable interface model; in any case, the potential difference and the surface charge density on one side of the interface are related to one another, and only one of them can be chosen as an independent variable.

The two-component plasma models have been studied in the grand-canonical formalism, where the "desire to remain neutral" appears as a restriction on the number of relevant chemical potentials or fugacities. For the homogeneous two-component plasma, one can solve (22) with independent constant fugacities m_+ and m_- for the positive and negative particles, respectively; nevertheless, the system remains locally neutral, and the one-body and many-body densities are found to depend only on $m = (m_+ m_-)^{1/2}$, in agreement with general rigorous results about S-component Coulomb systems [27]. For the two-component model of the ideally polarizable interface, we have introduced four electrochemical potentials ($\mu_a \pm e\phi_a$ on one side, $\mu_b \pm e\phi_b$ on the other side), but we have found that only three combinations are relevant: μ_a and μ_b control the bulk densities, while for controlling the electrical double layer at the interface $\phi_a - \phi_b$ is also needed.

2. *Differential Capacity*

On a polarizable interface separating two regions (a) and (b), the surface charge density $e\sigma$ on side (a) and the potential difference $\Delta\phi = \phi_a - \phi_b$ depend on one another. The derivative $C = \partial(e\sigma)/\partial(\Delta\phi)$, taken at constant bulk properties on each side, is called the differential capacity. The function $C(\Delta\phi)$ is usually accessible to measurements, and it plays an important role in electrochemistry. The usual experimental shape for the curve representing $C(\Delta\phi)$ is parabola-like, with a minimum.

For the models considered here, the charge density profile $q(x)$ is known, and one can obtain the differential capacity from

$$e\sigma = \int_0^\infty q(x)\, dx$$

and

$$\Delta\phi = 2\pi \int_{-\infty}^{\infty} q(x)x\, dx$$

For the one-component plasma models [14], the computed $C(\Delta\phi)$ does not reproduce the experimental behavior. In the primitive electrode model of Section III.B.2.c, one finds that

$$\frac{\Delta\phi}{e} = -\frac{1}{4} - \frac{\pi\sigma^2}{\rho} - \frac{1}{2}\ln\frac{1 - \operatorname{erf}(\sigma\sqrt{2\pi/\rho})}{2}$$

The corresponding $C(\Delta\phi)$ is monotonic, increasing from 0 to $+\infty$ as $\Delta\phi$ goes from $-\infty$ to $+\infty$. In the ideally polarizable interface of Section III.B.2.a, one finds a more complicated expression [14]; the corresponding $C(\Delta\phi)$ is found to have a maximum. These unphysical behaviors are obtained because there is in the model only one species of mobile particles: When a strong $\Delta\phi$ pushes these particles away from the interface, the thickness of the equivalent plane condenser becomes large, and its capacity becomes small.

For the two-component plasma models, the differential capacity is in qualitative agreement with the usual experimental results, with a minimum in the curve representing $C(\Delta\phi)$. For the primitive electrode model of Section IV.D.2.a, one finds the simple result

$$C(\Delta\phi) = \frac{m}{\pi}\cosh\frac{2\Delta\phi}{e}$$

For the ideally polarizable interface model of Section IV.D.2.b, the general results are more complicated; however, in the special case of equal bulk densities on each side (i.e., equal bulk fugacities m), one again finds a simple result:

$$C(\Delta\phi) = \frac{m}{4}\cosh\frac{\Delta\phi}{e}$$

Thus, allowing for a microscopic structure of the electrode does not change much the general behavior of the differential capacity curve.

3. *Surface Tension and Lippmann Equation*

In a system where there is an interface of area A, the surface tension γ is defined in terms of the free energy F or the grand potential Ω (depending on which set of independent variables has been chosen) by $\gamma = \partial F/\partial A$ or $\gamma = \partial\Omega/\partial A$. The surface tension plays an important role in electrochemistry, especially when the interface is between two conducting liquids.

For an electrified interface, by a thermodynamical argument, one can derive the Lippmann equation, which relates γ to the electrical properties:

$$\frac{\partial\gamma}{\partial\Delta\phi} = -e\sigma \tag{29}$$

In statistical mechanics, there are several possible routes for computing γ. One can start with a finite system, compute the free energy F including the surface corrections, and use $\gamma = \partial F/\partial A$; this route has actually been used for one-component plasma models [7,14,28] and it has the advantage that one can make an explicit check of relations such as the Lippmann equation (29). Another possible route would be to use expressions of γ in terms of integrals involving the correlation functions [28]; this has not yet been done for the solvable models considered here. Finally, it is possible to use known derivatives of γ, such as (29); in this way, γ can be computed from the density profiles, and this is the route that will be briefly described here.

For the one-component plasma model of a primitive electrode, the surface tension can be regarded as a function $\gamma(\rho,\sigma,T)$ (here the temperature is fixed by the condition $\Gamma = 2$). When $\sigma = 0$, γ obeys the sum rule [28]

$$\frac{\partial\gamma}{\partial\rho} = (\nu - 1)\pi e^2 \int_0^\infty dx\, x^2[\rho(x) - \rho] \qquad (\sigma = 0) \tag{30}$$

where $\nu(\nu = 2,3)$ is the dimensionality. For the present model, one can compute the right-hand side of (30) as a function of ρ, integrate (30) from ($\rho = 0$, $\gamma = 0$) for obtaining

$$\frac{2}{e^2}\sqrt{\frac{\pi}{\rho}}\,\gamma(\rho,\sigma = 0) = -\frac{1}{\sqrt{2}}\int_0^\infty \ln\frac{1 + \operatorname{erf}(t)}{2}\,dt \approx 0.239$$

and integrate (29) [using the known function $\Delta\phi(\sigma)$] for obtaining

$$\gamma(\rho,\sigma) - \gamma(\rho,\sigma = 0) = e^2\left[\frac{2\sigma^3}{3\rho} + \frac{1}{2}\sigma\ln\frac{1 - \operatorname{erf}(\sigma\sqrt{2\pi/\rho})}{2}\right]$$

By a similar method, one obtains a more complicated expression for the one-component plasma model of an ideally polarizable interface [14].

For the two-component plasma model of a primitive electrode, it is more convenient to use as an independent variable the bulk fugacity rather than the bulk density ρ_s. When $\sigma = 0$, instead of (30) one has the more familiar Gibbs adsorption equation:

$$-\frac{1}{k_B T} m \frac{\partial \gamma}{\partial m} = \int_0^\infty dx[\rho_+(x) + \rho_-(x) - \rho_+ - \rho_-] \qquad (\sigma = 0) \quad (31)$$

[this is just the surface analog of the bulk relation $(k_B T)^{-1} m(\partial p/\partial m) = \rho_+ + \rho_-$]. For the present model, the computation of the right-hand side of (31) gives $[(1/2\pi) - \frac{1}{4}]m$; the integration of (31) gives $\gamma(m,\sigma = 0)$, and finally, the integration of (29) [using the known function $\sigma(\Delta\phi)$] gives the simple result

$$\gamma = \frac{e^2 m}{2}\left(\frac{1}{4} - \frac{1}{2\pi}\cosh\frac{2\Delta\phi}{e}\right)$$

By a similar method, one would obtain a more complicated expression for the two-component plasma model of an ideally polarizable interface; in the special case of equal bulk densities on each side (i.e., equal bulk fugacities m) one again finds a simple result,

$$\gamma = -\frac{e^2 m}{4}\left(\cosh\frac{\Delta\phi}{e} - 1\right)$$

Since in the models considered here the interface is a rigid one, the computed surface tension is not necessarily positive, but the general shape of the electrocapillarity curve $\gamma(\Delta\phi)$ is correct for all these models: The curve looks like an inverted parabola, with a maximum. However, it must be admitted that this is not a very severe test, since (29) gives $\partial^2\gamma/(\partial\Delta\phi)^2 < 0$ as a simple consequence of a positive differential capacity $C = \partial(e\sigma)/\partial\Delta\phi > 0$.

B. Contact Theorems

These general theorems express a balance between the bulk pressure(s) and the forces exerted by a wall or interface (and the background, if any).

Among the solvable models considered here, the simplest case is the two-component plasma near a charged hard wall (primitive electrode model). Then, in dimension $\nu(\nu = 2,3)$, the contact theorem is [29]

$$k_B T[\rho_+(0) + \rho_-(0)] = p + (\nu - 1)\pi e^2\sigma^2$$

This is a balance between the kinetic pressure associated to the particle–wall collisions (the left-hand side), the bulk pressure p, and the electrostatic pressure $(\nu - 1)\pi e^2\sigma^2$. At $\Gamma = 2$, the pressure p has about one-half of the ideal gas value $k_BT(\rho_+ + \rho_-)$, and that explains why, at $\sigma = 0$, the contact densities are about half the bulk ones (Fig. 4).

For a one-component plasma near a charged hard wall, there is an additional term that takes into account the force exerted by the background; the contact theorem becomes [4,30]

$$k_BT\rho(0) = p + (\nu - 1)\pi e^2\sigma^2 + e\rho\Delta\phi \qquad (\nu = 2,3)$$

Similar theorems hold at the interface between two plasmas [15,26]. Adsorption sites can be taken into account, providing an additional term [24,31].

For the solvable models under consideration, these contact theorems can be explicitly checked using the computed pressures and density profiles.

C. Correlations and Sum Rules

The correlation functions of Coulomb systems obey a variety of sum rules that are consequences of the screening properties. A review of these sum rules has recently been published [32]. By explicit calculations, which will not be detailed here, it can be shown that each sum rule is actually obeyed by the correlation functions of the two-dimensional solvable models. We shall only quote a few examples. Although the sum rules are of very general validity and hold for an S-component plasma, here we shall write them for a one-component plasma, for the sake of a simpler notation.

1. Screening of a Particle of the System

A particle of the system is surrounded by a polarization cloud of opposite charge:

$$\rho(\mathbf{r}) = -\int d^2\mathbf{r}'\, \rho_T^{(2)}(\mathbf{r},\mathbf{r}') \tag{32}$$

2. Charged Hard Wall (Primitive Electrode)

Near a hard wall made of insulating material carrying a surface charge density $-e\sigma$, the response of the density to a variation of σ is related to the dipole moment of the correlation function:

$$\frac{\partial\rho(x;\sigma)}{\partial\sigma} = -\frac{2\pi(\nu - 1)e^2}{k_BT}\int_{x'>0} d^2\mathbf{r}'(x' - x)\rho_T^{(2)}(\mathbf{r},\mathbf{r}') \qquad (\nu = 2,3) \tag{33}$$

The total charge near the electrode is $+e\sigma$. Thus

$$\frac{2\pi(\nu - 1)e^2}{k_B T} \int_0^\infty dx \int_{x'>0} d^2\mathbf{r}'(x' - x)\rho_T^{(2)}(\mathbf{r},\mathbf{r}') = -1 \qquad (34)$$

the integrations must be performed in the indicated order.

Near the wall, the dipole moment of the polarization cloud, which appears in (33), does not vanish: The wall makes the screening less good. A related effect is that the correlations have only a slow (algebraic) decay along the wall instead of the fast decay found in the bulk (Gaussian in the present two-dimensional one-component model, exponential in the two-component model). With positions noted as $\mathbf{r} = (x,\mathbf{y})$, where $\mathbf{y}$ stands for the components(s) of $\mathbf{r}$ parallel to the wall, $\rho_T^{(2)}$ is a function $\rho_T^{(2)}(x,x', |\mathbf{y} - \mathbf{y}'|)$ which has the asymptotic behavior

$$e^2\rho_T^{(2)}(x,x', |\mathbf{y} - \mathbf{y}'|)_{|\mathbf{y}-\mathbf{y}'|\to\infty} \sim \frac{f(x,x')}{|\mathbf{y} - \mathbf{y}'|^\nu} \qquad (\nu = 2,3)$$

where $f(x,x')$ is a function that is important only for small x and x' and which obeys the sum rule

$$\int_0^\infty dx \int_0^\infty dx' \, f(x,x') = -\frac{k_B T}{2[(\nu - 1)\pi]^2} \qquad (\nu = 2,3)$$

3. *Ideal Conductor Wall*

Suppose now that the region $x < 0$ is occupied by an ideal conductor, while the plasma is in region $x > 0$ (this case can be obtained as the limit of an interface between two plasmas when the one in the region $x < 0$ becomes of infinite density). There is a given potential difference $\Delta\phi = \phi(+\infty) - \phi(0)$. Then the analog of (33) is

$$\frac{\partial\rho(x;\Delta\phi)}{\partial(\Delta\phi)} = \frac{e}{k_B T}\left[\rho(x) + \int_{x'>0} d^2\mathbf{r}' \, \rho_T^{(2)}(\mathbf{r},\mathbf{r}')\right]$$

[The right-hand side does not vanish, notwithstanding (32), because now there are also polarization charges on the surface of the ideal conductor, and they complete the screening.] The analog of (34) is

$$\frac{2\pi(\nu - 1)e^2}{k_B T} \int_0^\infty dx \, x\left[\rho(x) + \int_{x'>0} d^2\mathbf{r}' \, \rho_T^{(2)}(\mathbf{r},\mathbf{r}')\right] = 1$$

Here also, the integrations must be performed in the order indicated. Now the whole space is filled with conducting material, there is a good screening, and the correlations have a fast decay, even along the wall.

VI. FINAL REMARKS AND CONCLUSION

The two-dimensional classical Coulomb systems which have been studied here are solvable models at $\Gamma = 2$ because they are equivalent to quantum systems of *free* (i.e., noninteracting) fermions, as discussed in Sections III.A and III.B for the one-component plasma and in Section IV.B for the two-component plasma.

In the case of the homogeneous one-component plasma, the individual "wave functions" $\psi_n = \exp(-\frac{1}{2}\pi\rho r^2)z^{n-1}$ are the wave functions of the (degenerate) ground state of a charged quantum particle in a magnetic field $\mathbf{B}$ normal to the plane, when the vector potential has been chosen as $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$. The alternative choice $\psi_k = \exp(-\pi\rho x^2 + kx + iky)$ corresponds to another gauge, with a vector potential $\mathbf{A} = (0,Bx,0)$. Thus there is a close analogy between the classical one-component plasma and a quantum magnetic problem, and the freedom about the choice of the background potential and the associated basis, discussed in Section III.B.1, is indeed related to the gauge invariance of quantum electrodynamics. The analogy can be extended to nonhomogeneous cases; the quantum problem then is about a particle in position-dependent magnetic and electric fields.

A more general model can be solved: a two-component plasma with a background (i.e., a system made of positive particles, negative particles, and a charged continuous background). It suffices to use in (17) and (18) position-dependent fugacities of the form $m_s(\mathbf{r}) = m \exp[-2sV_B(\mathbf{r})]$, where $eV_B(\mathbf{r})$ is the background electrostatic potential. Explicit results have been obtained for a uniform* background [12] and near a hard wall [33]. In the limit $m \to 0$, the one-component plasma is recovered. This is an alternative route for obtaining the p-body densities of the one-component plasma.

Although the techniques of resolution that have been used rely on the equivalence with free fermions and therefore are valid only at $\Gamma = 2$, it is likely that $\Gamma = 2$ is not a singular point (provided that one defines the two-component plasma with some short-distance cutoff).

It has been shown in Section V.A.2 that the two-component models give a better account of the differential capacity than the one-component models.

The simple two-dimensional models that have been studied in this chapter are not expected to provide detailed pictures of real systems. However, they contain one essential ingredient, Coulomb's law, and it is remarkable that this is enough for illustrating many generic qualitative features of charged fluids.

* There are sign mistakes in Ref. 12.

ACKNOWLEDGMENT

This chapter is based in part on a Ph.D. thesis (in French) [33].

REFERENCES

1. Minnhagen, P. (1987). The two-dimensional Coulomb gas, vortex unbinding, and superfluid-superconducting films, *Rev. Mod. Phys. 59*: 1001.
2. Choquard, Ph., and Clérouin, J. (1983). Cooperative phenomena below melting of the one-component two-dimensional plasma, *Phys. Rev. Lett. 50*: 2086, and references quoted there.
3. Hauge, E. H., and Hemmer, P. C. (1981). The two-dimensional Coulomb gas, *Phys. Norvegica 5*: 209.
4. Choquard, Ph., Favre, P., and Gruber, Ch. (1980). On the equation of state of classical one-component systems with long-range forces, *J. Stat. Phys. 23*: 405.
5. Alastuey, A., and Jancovici, B. (1981). On the classical two-dimensional one-component Coulomb plasma, *J. Phys. (France) 42*: 1.
6. Jancovici, B. (1981). Exact results for the two-dimensional one-component plasma, *Phys. Rev. Lett. 46*: 386.
7. Smith, E. R. (1982). Effects of surface charge on the two-dimensional one-component plasma. I. Single double layer structure, *J. Phys. A 15*: 1271.
8. Forrester, P. J., and Smith, E. R. (1982). Effects of surface charge on the two-dimensional one-component plasma. II. Interacting double layers, *J. Phys. A 15*: 3861.
9. Caillol, J. M. (1981). Exact results for a two-dimensional one-component plasma on a sphere, *J. Phys. Lett. (France) 42*: L-245.
10. Choquard, Ph., Forrester, P. J., and Smith, E. R. (1983). The two-dimensional one-component plasma at $\Gamma = 2$: the semiperiodic strip, *J. Stat. Phys. 33*: 13.
11. Alastuey, A., and Lebowitz, J. L. (1984). The two-dimensional one-component plasma in an inhomogeneous background: exact results. *J. Phys. (France) 45*: 1859.
12. Cornu, F., and Jancovici, B. (1988). Two-dimensional Coulomb systems: a larger class of solvable models, *Europhys. Lett. 5*: 125.
13. Rosinberg, M. L., and Blum, L. (1984). An exactly solvable model for ideally polarizable interfaces, *Chem. Phys. Lett. 106*: 48.
14. Rosinberg, M. L., and Blum, L. (1984). The ideally polarizable interface: a solvable model and general sum rules, *J. Chem. Phys. 81*: 3700.
15. Jancovici, B. (1984). Surface properties of a classical two-dimensional one-component plasma: exact results, *J. Stat. Phys. 34*: 803.
16. Jancovici, B. (1981). Charge distribution and kinetic pressure in a plasma: a soluble model, *J. Phys. Lett. (France) 42*: L-223.
17. Jancovici, B. (1982). Classical Coulomb systems near a plane wall, I, *J. Stat. Phys. 28*: 43.

18. Forrester, P. J., Jancovici, B., and Smith, E. R. (1983). The two-dimensional one-component plasma at $\Gamma = 2$: behavior of correlation functions in strip geometry, *J. Stat. Phys. 31*: 129.
19. Blum, L., and Jancovici, B. (1984). Exactly solvable model for the interaction of two parallel charged plates in an ionic medium, *J. Phys. Chem. 88*: 2294.
20. Blum, L. (1984). A model for the interaction of two electric double layers in two dimensions: the metal electrolyte interface and the Donnan membrane, *J. Chem. Phys. 80*: 2953.
21. Forrester, P. J. (1985). The two-dimensional one-component plasma at $\Gamma = 2$: metallic boundary, *J. Phys. A 18*: 1419.
22. Alastuey, A., Jancovici, B., Blum, L., Forrester, P. J., and Rosinberg, M. L. (1985). The ideally polarizable interface: the metallic boundary limit, *J. Chem. Phys. 83*: 2366.
23. Rosinberg, M. L., Lebowitz, J. L., and Blum, L. (1986). A solvable model for localized adsorption in a Coulomb system, *J. Stat. Phys. 44*: 153.
24. Cornu, F. (1989). Two-dimensional models for an electrode with adsorption sites, *J. Stat. Phys. 54*: 681.
25. Gaudin, M. (1985). L'isotherme critique d'un plasma sur réseau ($\beta = 2$, $d = 2$, $n = 2$), *J. Phys. (France) 46*: 1027.
26. Cornu, F., and Jancovici, B. (1989). The electrical double layer: a solvable model, *J. Chem. Phys. 90*: 2444.
27. Lieb, E. H., and Lebowitz, J. L. (1972). The constitution of matter: existence of thermodynamics for systems composed of electrons and nuclei, *Adv. Math. 9*: 316.
28. Russier, V., Badiali, J. P., and Rosinberg, M. L. (1985). The one-component plasma near a hard wall: surface tension and sum rules, *J. Phys. C 18*: 707.
29. Henderson, D., Blum, L., and Lebowitz, J. L. (1979). An exact formula for the contact value of the density profile of a system of charged hard spheres near a charged wall, *J. Electroanal. Chem. 102*: 315.
30. Totsuji, H. (1981). Distribution of charged particles near a charged hard wall in a uniform background, *J. Chem. Phys. 75*: 871.
31. Blum, L., Rosinberg, M. L., and Badiali, J. P. (1989). Contact theorems for models of the sticky electrode, *J. Chem. Phys. 90*: 1285.
32. Martin, Ph. A. (1988). Sum rules in charged fluids, *Rev. Mod. Phys. 60*: 1075.
33. Cornu, F. (1989). Mécanique statistique classique de systèmes coulombiens bidimensionnels: résultats exacts, Thèse, Orsay, France.

6

Statistical Mechanics of Electrolytes at Interfaces

L. Blum
University of Puerto Rico
Rio Piedras, Puerto Rico

Douglas Henderson
Utah Supercomputing Institute/IBM Partnership and University of Utah
Salt Lake City, Utah

I. INTRODUCTION

The study of the interface between two phases that are charged and/or conducting is of relevance to a number of systems that occur in nature: colloids, micelles, membranes, solid–solution interfaces in general, and metal–solution interfaces in particular. These form a bewildering array of systems of enormous complexity. The investigation of the structure of these systems poses considerable difficulties, both experimentally as well as theoretically. The experimental problem is that the interface has 10^{-8} particle relative to the bulk, solid or liquid, phases. For this reason one needs a surface-specific method, which is able to discriminate between the signals from the surface and the bulk. Electrons do not penetrate into solids and for that reason have been used extensively for the ex-situ determination of the surface structure of solids. They must be used in vacuum and that precludes their use in the in-situ study of the liquid–solid interface. The study of electrode surfaces removed from the liquid cell under various conditions has provided an enormous wealth of useful data that we will not try to review here. The only way to understand the relation between the ex-situ and in-situ structures is to measure both, something that is only now becoming possible. The in-situ structural determination methods are x-ray scattering or diffraction techniques and the scanning

tunnel microscope (STM). There are three x-ray-based techniques that have been used to determine the structure in well-characterized single-crystal metal electrolyte interfaces:

1. *Extended x-ray absorption fine structure (EXAFS)* [1,2]. This method permits the determination of the distances of the nearest neighbors of a given target atom, not necessarily in a regular structural arrangement, and also yields information about the electronic state of that atom when it is adsorbed at the surface.
2. *Grazing incidence x-ray diffraction (GIXS)* [3,4]. This method permits determination of the in-plane structure of an adsorbed monolayer on the surface. It is an accurate technique and requires a regular structure of the adsorbed monolayer which should be different of that of the substrate.
3. *Standing wave methods*. An x-ray standing wave is set up at the interface of the solid and the fluid. There are several modes in which these standing waves can be formed; they allow determination of the distance from the surface of the solid into the fluid phase [5,6].

The STM [7,8] is a very promising technique since in principle it permits the direct determination of the structure of any surface that is sufficiently conducting. It has been shown recently by various groups that the STM is capable of resolving structural details of metal surfaces in contact with electrolytic solutions. However, when the electrochemical potential is scanned, the tunnel voltage of the STM also changes. This does not affect the study of surface geometry, since the images are relatively independent of the tunnel voltage. The resolution of the STM pictures of the metal–electrolyte interface is on the order of 0.5 Å. A first attempt at the theory of the STM in an electrolyte has recently been developed by Schmickler and Henderson [9].

Other in-situ techniques give information that is thermodynamic in nature since it comprises the average over a number of atoms. One technique that has been established recently is the quartz microbalance [10–12]; this instrument can measure small changes in the mass of a metallic electrode that is attached to a quartz oscillator. The electrosorption valency, for example, can be calculated directly, by measuring the mass deposited at the electrode and the amount of charge from voltamogram. The interpretation of the results of this instrument requires electrode surfaces that have large molecularly smooth regions. Spectroscopic methods using [13–15] ultraviolet, visible, or [16] Raman spectroscopy are very useful in situ probes because a large number of organic molecules can be studied. Interesting information about changes in bonding and symmetry can be

extracted. The optical spectroscopic methods do not require installations such as the synchrotron, and are most useful for complex molecular species. The techniques are surface-enhanced Raman, surface infrared spectroscopy, and second harmonic generation, which permits us to discriminate between different geometries of the adsorbates on single-crystal surfaces.

Among the optical techniques there are also the more traditional methods, such as the ellipsometry, electroreflectance, and particularly, surface plasmons, where experimental and theoretical advances have made it possible to offer a picture of the surface electronic states of the metal in some selected cases, such as the silver (111) phase. We should mention here the measurement of image-potential-induced surface states by electroreflectance spectroscopy. In this case, besides the normal surface states that arise from termination of the crystal lattice, there are discrete states due to the existence of an image potential for charges near the conducting interface.

A method that has yielded very interesting information about the structure and interactions in the diffuse part of the double layer is the direct measurement of forces between colloid particles [17]. The forces between two mica plates are measured directly in the presence of different solutions: These forces show pronounced oscillations of a period similar to the dimensions of the molecules enclosed between the plates. Last, but certainly not least, there is a very extensive and important literature on the differential capacitance of solutions near either solid (polycrystalline or single-crystal) or liquid (mercury) electrodes which we will not try to cover. We should mention recent work on the influence of the crystallographic orientation of silver on the potential of zero charge of the electrodes, in which a detailed mapping of the influence of the crystal face on the differential capacitance of the inner layer is made [18–20].

The complexity of the system described by the experimental methods defies any simple theoretical interpretation. Yet these are needed for an understanding of what is actually going on at the charged interface. It is clear that the simplest theory should discuss two kinds of forces: the long-range Coulomb forces and the short-range forces that are at the origin of the chemical bonds and are also responsible of the repulsion between atomic cores. There are important quantum effects at the interface due to the quantum nature of the electrons in a metal [21–29]. These effects are very difficult to compute in a proper way, and in most theoretical discussions only very sketchy models of the liquid side of the interface are discussed when attempting to describe the metal side of the interface.

For this reason we have organized the theoretical discussion starting with very simple model systems about which much is known, and going

to systems that are much more realistic but difficult to handle. The emphasis of the theoretical treatment will be on the structure functions, or distribution functions $\rho(1)$, $\rho(2)$, . . . which give the probability of finding an ion(s) or solvent molecule(s) at specified position(s) near the interface. The properties of the interface can then be calculated from these distribution functions. One of the very interesting theoretical developments of recent years has been the exactly solvable model developed by Jancovici, Cornu, and co-workers. This is a two-dimensional model at a particular value of the reduced temperature and is particularly useful to elucidate the subtle properties of the long-range Coulomb forces. This model is discussed in detail in Chapter 5. For the nonprimitive model with solvent molecules there is a one-dimensional exactly solvable model. Exactly solvable models serve as benchmarks for approximate theories and to test exact and general sum rules.

We start with the simplest model of the interface, which consists of a smooth charged hard wall near a ionic solution that is represented by a collection of charged hard spheres, all embedded in a continuum of dielectric constant ϵ. This system is fairly well understood when the density and coupling parameters are low. Then we replace the continuum solvent by a molecular model of the solvent. The simplest of these is the hard sphere with a point dipole [30,31], which can be treated analytically in some simple cases. More elaborate models of the solvent introduce complications in the numerical discussions. A recently proposed model of ionic solutions uses a solvent model with tetrahedrally coordinated sticky sites. This model is still analytically solvable. More realistic models of the solvent, typically water, can be studied by integral equations and computer simulations, which, however, is very difficult for charged interfaces. The full quantum mechanical treatment of the metal surface does not seem feasible at present. The jellium model is a simple alternative for the discussion of the thermodynamic and also kinetic properties of the smooth interface [21–29].

II. THEORIES FOR THE SINGLET AND PAIR DISTRIBUTION FUNCTIONS

We have a mixture of ions of density ρ_i, charge e_i, and diameter σ_i. $\rho_i(z)$ is the number density profile of species i at a distance z from the electrode, which is always assumed to be flat and perfectly smooth. Typical values of $\rho_i(z)$ obtained from both theory and simulation [32–37] are shown in Fig. 1. The singlet distribution function is

$$g_i(z) = \frac{\rho_i(z)}{\rho_i} = h_i(z) + 1 \tag{1}$$

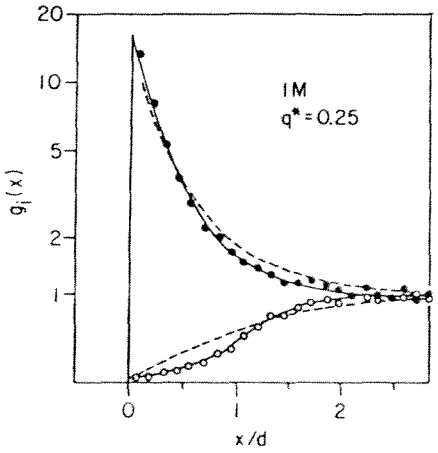


Fig. 1 Density profile for charged hard spheres near a hard-charged electrode for a $1M$, 1:1 electrolyte. The reduced charge density $q^* = q_s\sigma^2/e$. The model parameters are $\sigma = 4.21$ Å, $T = 298$ K, and $\epsilon = 78.5$. The points give the simulation values and the solid and dashed curves give the HNC1/MSA and Gouy–Chapman values. The value of q^* is 0.25.

The charge density $q(z)$ is given by

$$q(z) = \sum_{i=1}^{m} e_i \rho_i(z) \tag{2}$$

where m is the number of ionic species. The electrostatic potential ϕ is obtained by integration of Poisson's equation,

$$\nabla^2\phi(z) = \frac{\partial_2\phi(z)}{\partial z^2} = \frac{-4\pi q(z)}{\epsilon} \tag{3}$$

This equation can be integrated to obtain the alternative relation between the charge and potential profiles

$$\phi(t) = -\frac{2\pi}{\epsilon}\int_0^\infty |z - t|\, q(z)\, dz \tag{4}$$

The total potential drop $\Delta\phi$ is obtained from (4) by either letting $\phi(z) = 0$ at $z = 0$ or $z \to \infty$, depending on the reference potential of the model. In general, the latter choice is adopted. An important quantity is the differential capacitance C_d, defined by

$$C_d = \frac{\partial q_s}{\partial(\Delta\phi)} \tag{5}$$

where q_s is the surface charge on the electrode. This quantity is difficult to measure directly and is inferred either from surface tension measurements or frequency-dependent ac measurements of the capacitance. The surface charge satisfies the electroneutrality condition

$$q_s = -\int_0^\infty q(z)\,dz = \frac{E_0\epsilon}{4\pi} \tag{6}$$

where E_0 is the external or applied field. That is $E(z = 0)$.

Consider the Poisson equation (3). If we approximate the density of the ions by Boltzmann's distribution formula [38,39]

$$\rho_i(z) = \rho_i e^{-\beta e_i \phi(z)} \tag{7}$$

and substitute into (3), we obtain the Poisson–Boltzmann equation,

$$\nabla^2\phi(z) = \frac{-4\pi}{\epsilon}\sum_{i=1}^{m} e_i\rho_i e^{-\beta e_i\phi(z)} \tag{8}$$

A first integral of this differential equation can be obtained multiplying both sides by $\nabla\phi(z)$ and integrating. For the planar electrode this yields

$$E^2(z) = [\nabla\phi(z)]^2 = -\frac{8\pi kT}{\epsilon}\sum_{i=1}^{m}\rho_i\,[e^{-\beta e_i\phi(z)} - 1] \tag{9}$$

Using the definition of C_d and the electroneutrality relation (6) we get the formula for differential capacitance:

$$C_d = \sqrt{\frac{2\pi}{\epsilon kT}}\,\frac{\sum_{i=1}^{m}\rho_i[e^{-\beta e_i\phi(0)}]}{\sum\rho_i[e^{-\beta e_i\phi(0)} - 1]} \tag{10}$$

where $\phi(0) = \phi(z)|_{z=0} = \Delta\phi$ is the potential at the origin and is equivalent to the total polarization potential of the electrode. At this point it is convenient to make a change in the variable

$$\chi(z) = e^{-\beta e\phi(z)} \tag{11}$$

where $e_i = z_i e$. e is the elementary charge and z_i is the electrovalence of species i. We integrate (9) to get

$$\int_{\chi(0)}^{\chi(z)} d\chi\,\frac{1}{\chi\sqrt{\sum_i \rho_i\chi^{z_i} - \rho_0 A}} = \sqrt{\frac{8\pi e_2}{\epsilon kT}}\,(z - z_0) \tag{12}$$

where χ and A are integration constants. The electrovalence z_i is always a small number and the integration of the left-hand side is always possible

ı terms of elliptic functions [40]. When $z_1 = -z_2 = 1$ the radicand of ıe left-hand side of (12) is a perfect square and the integral can be per-ɔrmed explicitly. For the potential drop $\Delta\phi$ we obtain the implicit relation

$$\frac{\beta e E_0}{\kappa\epsilon} = 2 \sinh \frac{\beta e \Delta\phi}{2} \tag{13}$$

he density profile is given by

$$\rho_i(z) = \rho_i \left(\frac{1 + z_i \alpha \epsilon^{-\kappa z}}{1 - z_i \alpha \epsilon^{-\kappa z}} \right)^2 \tag{14}$$

/here

$$\kappa^2 = \frac{4\pi}{\epsilon kT} \sum_{i=1}^{m} \rho_i e_i^2 \tag{15}$$

efines the Debye screening parameter and α is given by

$$\alpha = \tanh \frac{\beta e \Delta\phi}{4} \tag{16}$$

Some remarks about the Gouy–Chapman theory are useful: Despite ıe apparent oversimplification, the Poisson–Boltzmann equation satis-es approximately an overall dynamic equilibrium condition that fixes the ɔntact density at the electrode surface. The Poisson–Boltzmann ap-roximate contact theorem is

$$kT \sum_{i=1}^{m} \rho_i(0) = \frac{\epsilon}{8\pi} E_0^2 + kT \sum_{i=1}^{m} \rho_i \tag{17}$$

'he contact theorem, as well as other sum rules that are valid for the harged interface, will be discussed in Section III. The density profiles ɔbtained from the Gouy–Chapman theory are monotonous; that is, they how no oscillations. Since, in this theory, the electroneutrality condition ; satisfied and the contact theorem is almost satisfied, $\rho_i(z)$ is pinned at ıe origin and has a fixed integral so that the density profile cannot deviate ɔo much from the correct result. When the contact theorem is not sat-ɔfied, such as in the case of mixtures of unequal-size ions at low electrode harge or for high density, the profiles are oscillatory, and we expect eviations from the GC theory. This is also true for the nonprimitive ıodel, in which the solvent is a fluid of finite-size molecules.

II. EXACT RESULTS AND THEOREMS

'he sum rules for the charged interfaces can be classified in two cate-ɔories:

1. The screening sum rules, which are specific to Coulomb forces. Because of the very long range of the electrostatic forces, the stability of the system requires that all charges surround themselves with a neutralizing cloud. The surface charge satisfies the electroneutrality condition

$$-q_s = \int_0^\infty dz \sum_{i=1}^{m} e_i \rho_i(z) = \frac{\epsilon E_0}{4\pi} \tag{18}$$

where E_0 is the external or applied field. In the one-sided models this is the independent variable. In homogeneous systems of molecules interacting with Coulomb forces the screening of charges and multipoles by the conducting media is intuitive because of the isotropy of the system. In the homogeneous solution every charge is surrounded by an ionic cloud of exactly the opposite charge. Also every dipole is surrounded by a charged cloud that has a dipole moment exactly opposite that of the original dipole, and in general, one can show [41–43] for any charge distribution in an homogeneous system that any arbitrary multipole is perfectly screened by the ion distribution. The fact is that it is also true in the inhomogeneous case, which is not intuitively obvious, and has been confirmed by the beautiful work of Jancovici [44–49], for an exactly solvable two-dimensional model. This model is discussed in detail in Chapter 6. The demonstration of these theorems is based on the Born–Green–Yvon hierarchy (BGY) and an assumption on the clustering of the correlation functions.

2. The dynamic sum rules that are derived from balance-of-force considerations. Systems interacting with conservative forces must satisfy momentum conservation and force balance. This apparently trivial requirement is not satisfied by some of the approximate theories used in the description of the electrode interface. For interfaces consisting of hard walls plus some soft interactions (which could be either attractive or repulsive) and for planar systems, the force balance equation reads [50, 52]

$$p = kT \sum_{i=1}^{M} \rho_i(0) - \sum_{i=1}^{m} \rho_i \int_0^\infty dz \frac{\partial w_i(z)}{\partial z} g_i(z) \tag{19}$$

where $w_i(z)$ is the unscreened total interaction potential between the ions and the electrode wall.

We consider a system that is limited by an arbitrarily rough planar but charged surface. The precise mathematical requirement is that there is a prism with an arbitrarily large cross-sectional area S and height L (the volume $V = SL$) such that the force through the walls parallel to z is of $O(S^{1-\delta})$, where $\delta \to 0$ as $S \to \infty$. We integrate the BGY equation in the volume of a prism of the same section S but smaller height $L_1 < L$. Sum-

ming over all species i, we get

$$-kT \sum_{i=1}^{m} \int d1\ \nabla_1 \rho_i(1) = \sum_{i=1}^{m} \int d1\ [\rho_i(1)\ \nabla_1\ u_i^0(1) + E_i \rho_i(1)\ \nabla_1\ \phi(1)] + \sum_{i=1}^{m} \int d1 \sum_{j=1}^{m} \int d2\ \rho_{ij}(1,2)\ \nabla_1\ u_{ij}^0(1,2) + \sum_{i=1}^{m} e_i \int d1\ \rho_i(1) \sum_{j=1}^{m} e_i \int d2\ \rho_j(2) h_{ij}(1.2)\ \nabla_1\ w_{ij}(1.2) \quad (20)$$

where $w_{ij}(1,2)$ is the soft part of the pair potential. Since for the planar geometry $\nabla_1 = \partial/\partial z_1$ the integral of the left-hand term can be easily performed. Thus

$$\sum_{i=1}^{m} \int d1\ \nabla_1\ \rho_i(1) = S\ [\rho_i(L_1) - \tilde{\rho}_i(0)] \quad (21)$$

where we have used the definition

$$\tilde{\rho}_i(0) = \frac{1}{S} \int dx_1\ dy_1\ \rho_i(x_1, y_1, z_s) \quad (22)$$

where z_s is the position of the surface at (x_1, y_1). To integrate the second term in the right-hand side we use Poisson's equation,

$$\nabla^2\ \phi(1) = \frac{-4\pi}{\epsilon} \sum_{i=1}^{m} e_i \rho_i(1) = -\nabla_1 \cdot \mathbf{E}(1) \quad (23)$$

where we have used the electric field

$$\mathbf{E}(1) = -\nabla_1 \phi(1) \quad (24)$$

Substitution into the second term of the right-hand side leads to

$$\int d1 \sum_{i=1}^{m} e_i \rho_i(1)[E_z(1)] = -\epsilon/4\pi \int d1\ \nabla^2\ \phi(1)[E_z(1)] = \frac{\epsilon}{4\pi} \int d1 [\nabla_1 \cdot \mathbf{E}(1)] E_z(1) \quad (25)$$

where $\mathbf{E}(1)$ is the electric field at position $\mathbf{r}_1$, and $E_z(1)$ is the z-component of this field. Using Poisson's equation and integrating by parts, we get

$$\int d1 \sum_{i=1}^{m} e_i \rho_i(1)[E_z(1)] = \frac{\epsilon}{8\pi} \int_S dx_1\ dy_1 \int_{z_s(x_1,y_1)}^{L_1} dz_1 \frac{\partial}{\partial z} E_z^2(1) + \frac{\epsilon}{4\pi} \int_V dz_1\ dy_1\ dz_1 \left[E_z(1) \frac{\partial E_x(1)}{\partial x_1} + E_z(1) \frac{\partial E_y(1)}{\partial y_1} \right]$$

The second term on the right-hand side is zero: For a periodic interface in the x and y directions, if we take S to be the surface of a unit cell, then the terms $\partial E_x(1)/\partial x_1$ will be of equal magnitude but with opposite sign for neighboring cells. For the general random interface we conjecture that this term is finite; then is the limit $S \to \infty$ the contribution vanishes. We have

$$\frac{1}{S}\int d1 \sum_{i=1}^{m} e_i \rho_i(1) E_z(1) = \frac{-\epsilon}{8\pi} \langle [E_z(1)]^2 \rangle_S \tag{27}$$

where the average square field in the z direction is

$$\langle [E_z(1)]^2 \rangle_S = \frac{1}{S}\int_S dx_1\, dy_1\, E_z^2(x_1, y_1, z_s) \tag{28}$$

The other single-particle term containing the short-range interactions between the molecules and ions and the wall yields

$$\left\langle \rho_i(1) \frac{\partial u_{i0}(1)}{\partial z_1} \right\rangle_S = \frac{1}{S}\int_S dx_1\, dy_1\, dz_1\, \rho_i(1)\, \nabla_1\, u_i^0(1) \tag{29}$$

where

$$\left\langle \rho_i(1) \frac{\partial u_{i0}(1)}{\partial z_1} \right\rangle_S = \frac{1}{S}\int_S dx_1\, dy_1 \int_{z_s(x_1,y_1)}^{L_1} dz_1\, \rho_i(1) \frac{\partial u_i^0}{\partial z_1} \tag{30}$$

The last term on the right-hand side of (20) is the average of the pair forces. We write

$$I_p = \int_S dx_1\, dy_1 \int_S dx_2\, dy_2 \int_{z_s(x_1,y_1)}^{L_1} dz_1 \int_{z_s(x_2,y_2)}^{L} dz_2 \sum_{ij} F_{ij}(1,2)$$

$$I_p = \int_S dx_1\, dy_1 \int_S dx_2\, dy_2 \int_{z_s(x_1,y_1)}^{L_1} dz_1 \left[\int_{z_s(x_2,y_2)}^{L_1} dz_2 + \int_{L_1}^{L} \right] \sum_{ij} F_{ij}(1,2) \tag{31}$$

where

$$F_{ij}(1,2) = \rho_{ij}(1,2) \nabla_1 u_{ij}^0(1,2) + \rho_i(1)\rho_j(2) h_{ij}(1,2)\, \nabla_1 \frac{e_i e_j}{\epsilon r_{12}} \tag{32}$$

Since the forces of a pair are equal in magnitude and opposite in sign,

$$F_{ij}(1,2) = -F_{ji}(2,1) \tag{33}$$

from where the first double integral in (31) vanishes. Since the interactions are short range, the second double integral yields the virial contribution

to the bulk pressure p:

$$p_{virial} = \frac{1}{S}\int_S dx_1\, dy_1 \int_S dx_2\, dy_2 \int_{z_s(x_1,y_1)}^{L_1} dz_1 \int_{L_1}^{L} dz_2 \sum_{ij} F_{ij}(1,2)$$

$$= \frac{2\pi}{3}\sum_{ij}\rho_i\rho_j \int_0^\infty dr\, r^3 \left[\frac{[e_ie_j]h_{ij}(r)}{\epsilon r} + g_{ij}(r)\frac{\partial u_{ij}^0(r)}{\partial r}\right] \tag{34}$$

Putting it all together yields the general contact theorem for a planar on the average, but not necessarily smooth, surface:

$$kT\sum_{i=1}^{m}\bar{\rho}_i(0) = p + \frac{\epsilon\langle[E_z(1)]^2\rangle_S}{8\pi} - \sum_{i=1}^{m}\left\langle\rho_i(1)\frac{\partial u_i^0(1)}{\partial z_1}\right\rangle_S \tag{35}$$

This theorem [52] is a generalization of the previously derived contact theorems to the realistic case of nonsmooth electrode surfaces. It contains the previous results as particular cases. If the interface is a smooth hard wall, the surface averages become the surface values of the parameters and we get

$$kT\sum_{i=1}^{m}\rho_i(0) = p + \frac{\epsilon[E_z(1)]^2}{8\pi} - \sum_{i=1}^{m}\left\langle\rho_i(1)\frac{\partial u_i^0(1)}{\partial z_1}\right\rangle_S \tag{36}$$

where the last term is now

$$\left\langle\rho_i(1)\frac{\partial u_i^0(1)}{\partial z_1}\right\rangle_S = \int_0^\infty dz_1\,\rho_i(1)\frac{\partial u_i^0(1)}{\partial z_1} \tag{37}$$

When $u_i^0(1)$ is zero, we obtain the contact theorem:

$$kT\sum_{i=1}^{m}\rho_i(0) = p + \frac{\epsilon\overline{E}_0^2}{8\pi} \tag{38}$$

for the primitive model with a continuum solvent of dielectric constant ϵ, and

$$kT\sum_{i=1}^{m}\rho_i(0) + kT\sum_{i=1}^{s}\rho_s(0) = \frac{\overline{E}_0^2}{8\pi} + p \tag{39}$$

for the case of an electrolyte in a molecular solvent. Comparison of (38) and (17) shows that the Poisson–Boltzmann theory satisfies the contact theorem with the perfect gas approximation

$$p = kT\sum_{i=1}^{m}\rho_i \tag{40}$$

For a surface with an array of sticky adsorption sites, such as in the

case of the sticky site model (SSM) model discussed elsewhere [51], the adsorption potential has the form

$$\exp[-\beta u_a(\mathbf{r})] = 1 - \lambda_a(\mathbf{R})\delta(z) \tag{41}$$

with

$$\lambda_a(\mathbf{R}) = \sum_{n_1 n_2} \lambda_a \delta(\mathbf{R} - n_1\mathbf{a}_1 - n_2\mathbf{a}_2) \tag{42}$$

Here $\mathbf{R} = x,y$ is the position at the electrode surface, and z the distance to the contact plane, which is at a distance $\sigma/2$ from the electrode. In (42), n_1 and n_2 are natural numbers, and $\mathbf{a}_1$ and $\mathbf{a}_2$ are the lattice vectors of the adsorption sites on the surface. The parameter λ_a represents the fugacity of an adsorbed atom of species a. Define now the regular part of the density function

$$y_i(1) = \rho_i(1) \exp[\beta u_i^0(1)] \tag{43}$$

Substituting into the general contact theorem, Eq. (35), gives [52]

$$kT \sum_{i=1}^{m} \tilde{\rho}_i(0) = p + \frac{\epsilon \langle [E_z(1)]^2 \rangle_s}{8\pi} - \lambda_a \sum_{i=1} \left\langle \frac{\partial y_i(1)}{\partial z_1} \right\rangle_s \tag{44}$$

This theorem has been verified recently [53] for the exactly solved model of a one-component plasma in two dimensions.

IV. SCREENING SUM RULES

In electrically neutral systems any fixed arrangement of charges is screened by the mobile charges of the system. In homogeneous bulk phases this is an intuitively natural fact because if the long-range Coulomb forces would not be screened, the partition function would not exist (it would diverge) and matter would not be stable [54]. This is expressed by the fact that the charge distribution around a given charge e is of equal value but opposite sign. Thus

$$-e_i = \int d2 \sum_j e_j \rho_j h_{ij}(1,2) \tag{45}$$

Rotational invariance in bulk fluids requires that not only charges but also multipoles of arbitrary order should be screened by the mobile charges of the media [41,42]. This fact is much less intuitive in the neighborhood of charged objects, in particular in the neighborhood of a charged electrode. However, the theorems hold, and in classical mechanics, at least, perfect screening of all multipoles occurs in both homogeneous and inhomogeneous systems. These conclusions are supported by the results

of the exactly solved Jancovici model [48,49] (see Chapter 5). However, perfect screening of all multipoles does not occur in quantum systems or in systems out of equilibrium [55]. As a consequence of the screening, the second moment of the pair distribution function must be normalized. This is the Stillinger–Lovett [56] moment relation. As was shown by Outhwaite [57], it can be written in the form of a normalization condition for the electrostatic potential

$$\phi_i(r) = \left[\frac{e_i}{r} - \sum_j e_j\rho_i \int d\mathbf{r}_i \frac{h_{ij}(r_1)}{|\mathbf{r} - \mathbf{r}_1|}\right] \Big/ \epsilon \tag{46}$$

which satisfies the sum rule

$$1 = \beta \sum e_j\rho_j \int d\mathbf{r}_1 \, \phi_j(1) \tag{47}$$

Carnie and Chan [58] have shown that this normalization condition is also valid for the inhomogeneous systems of charged particles. Consider the inhomogeneous Ornstein–Zernike (OZ) equation (see Chapter 4)

$$h_{ij}(1,2) - c_{ij}(1,2) = \sum_{k=1}^{m} \int d3 \, h_{ik}(1,3)\rho_k(3)c_{kj}(3,2) \tag{48}$$

where $c_{ij}(1,2)$ is the direct correlation function and the singlet density function $\rho_i(1)$ satisfies the electroneutrality condition

$$q_s = -\frac{1}{S}\int_0^\infty d\mathbf{r} \sum_{i=1}^{m} e_i\rho_i(\mathbf{r}) \tag{49}$$

where q_s is the surface charge density and S is the area of the interface. The inhomogeneous pair distribution function satisfies the sum rule

$$-e_i = \int d2 \sum_j e_j\rho_j(2)h_{ij}(1,2) \tag{50}$$

From the diagram expansion we write the direct correlation function as the sum of a short-range part and the pair potential, which in our case is long range:

$$c_{ij}(1,2) = c_{ij}^0(1,2) - w_{ij}(1,2) \tag{51}$$

where we recall that the electrostatic interaction of ions is

$$w_{ij}(1,2) = \frac{e_ie_j}{\epsilon r_{12}} \tag{52}$$

media, using the first screening relation (50), we have

$$e_j\phi_i(1) = -w_{ij}(1,2) + \sum_{k=1}^{m} \int d3\; w_{ik}(1,3)\rho_k(3)h_{kj}(3,2) \tag{53}$$

and from the OZ equation (48) we get

$$h_{ij}(1,2) = -\beta e_i\phi_j(1) + c_{ij}^0(1,2) + \sum_{k=1}^{m} \int d3\; c_{ik}^0(1,3)\rho_k(3)h_{kj}(3,2) \tag{54}$$

Multiplying this equation by $e_j\rho_j(2)$, integrating over r_2, and summing over r yields

$$1 = \beta \sum_j e_j \int d\mathbf{r}_2\; \rho_j(2)\phi_j(2,1) \tag{55}$$

which is the generalization of the Outhwaite formula of the Stillinger–Lovett sum rule for inhomogeneous charged systems.

V. OTHER SUM RULES

For flat hard electrode surfaces there are number of other sum rules. A complete review of these rules was recently made by Martin [59]. A relevant sum rule for the calculation of density profiles in the electric double layer is the dipole rule [60],

$$kT \frac{\partial \ln \rho_i(1)}{\partial E_0} = \int d2 \sum_j e_j\rho_j(2)h_{ij}(1,2)(z_1 - z_2) \tag{56}$$

where E_0 is the bare field at the electrode surface. The differential capacity, which is defined by

$$C_d = \frac{\partial q_s}{\partial(\Delta\phi)} \tag{57}$$

where q_s is the surface charge, $q_s = E_0\epsilon/4\pi$, and $\Delta\phi$ is the potential drop, satisfies the sum rule

$$\frac{1}{C_d} = \frac{8\pi^2}{\epsilon^2 S} \int d1\; d2 \sum_{ij} e_ie_j\rho_i(1)\rho_j(2)h_{ij}(1,2)(z_1 - z_2)^2 \tag{58}$$

The surface tension γ obeys relations that can be given in terms of the direct correlation function [61,62] (see Chapter 2).

$$\gamma = \frac{\pi kT}{2} \int dz_1\; dz_2 \sum_{ij} \frac{\partial\rho_i(1)}{\partial z_1} \frac{\partial\rho_j(2)}{\partial z_2} \int_0^\infty dr_{12}[r_{12}]^3 c_{ij}(1,2) \tag{59}$$

Using the Ornstein–Zernike equation we get the form that contains the pair correlation function

$$\gamma = \frac{\pi}{2kT} \int dz_1\, dz_2 \sum_{ij} \rho_i(1) \frac{\partial u_i(1)}{\partial z_1} \rho_j(2) \frac{\partial u_j(2)}{\partial z_2} \int_0^\infty dr_{12} [r_{12}]^3 h_{ij}(1,2) \quad (60)$$

These sum rules provide ways of asserting the accuracy of the different approximations used to compute the charge and ion density near charged walls.

VI. INTEGRAL EQUATIONS: THE PRIMITIVE MODEL

There have been a large number of papers dealing with ways to improve the Gouy–Chapman equation. In the regime of low density and high temperature (or large dielectric constant) the GC theory is quite good despite its simplifications because it satisfies the contact theorem (38) asymptotically for $E_0 \to \infty$, and it satisfies the electroneutrality condition (49). However, in real systems with molecular solvents the density and coupling constant are large, significant deviations from the behavior predicted by the GC theory occur. For this reason it is interesting to determine the accuracy of the integral equations for the primitive model for high coupling constants beyond the parameters that correspond to experimental situations, because it will indicate which theory can be used for the nonprimitive model of the electric double layer. These theories can be formulated as integral equations for the density profile $\rho_i(1)$, or as a differential or integrodifferential equation for the potential $\phi(1)$. The central quantity of our discussion is [63] the one-particle direct correlation function, from which the integral equations will be deduced:

$$c_i(1) = \ln \frac{\rho_i(1)}{z_i} + \beta u_i(1) \quad (61)$$

where $c_i(1)$ is the one-particle direct correlation function, z_i the fugacity of species i, and $u_i(1)$ the external potential. The function $c_i(1)$ is a member of the family of direct correlation functions $c_{ij} \cdots 1,2, \ldots)$, which is the sum of all irreducible graphs with density factors $\rho_i(1)$ for every field point. (For a detailed discussion of correlation functions, see, for example, Hansen and McDonald [64].) Functional series differentiation [65,66] produces approximations, such as the hypernetted chain (HNC), and its modifications, and the mean spherical (MSA), and its modifications, that are used in conjunction with the Ornstein–Zernike equation. A different set of approximations is obtained by spatial differentiation of $c_i(1)$, which gives the Born–Green–Yvon (BGY) and Wertheim–Lovett–Mou–Buff

(WLMB) equations. Finally, the Kirkwood equation is obtained by differentiation with respect to the chemical potential. Some of these methods have been discussed in Chapter 4.

A. Ornstein–Zernike-Based Approximations

At the interface between an electrode and a fluid the density of the fluid is a function of the distance of the point to the surface $\rho_i(z)$. The Ornstein–Zernike equation for this system can be obtained as a limit of a system that is a homogeneous mixture in which there are some large ions, of radius $R_w \to \infty$, such that $\rho_w R_w^3 \to 0$. In this limit the planar [67,68] HAB (Henderson–Abraham–Barker) OZ equation is

$$h_i(1) - c_i^w(1) = \sum_{j=1}^{m} \int d2\, h_j(2)\rho_j c_{ij}^B(1,2) \tag{62}$$

where $h_i(1)$ is the density profile correlation function of ion i; $c_i^w(1)$ is not the single-particle direct correlation function $c_i(1)$, but a different magnitude defined below (64), and $c_{ij}^B(1,2)$ is the bulk direct correlation function.

$$h_i(1) = g_i(1) - 1 = \frac{\rho_i(1) - \rho_i}{\rho_i} \tag{63}$$

The function $c_{ijk}^B\cdots(1,2,3,\ldots)$ is a very complicated function and, in general, does not admit a simple diagram expansion. To get insight about the meaning of this function, we use functional series expansion. Consider the functional power series expansion of $\ln \rho_i(1)$ around the uniform density [65,66] ρ_i:

$$\beta u_i(1) + \ln \rho_i(1) = \ln \rho_i + \sum_{j=1}^{m} \int d2\, h_j(2)\rho_j c_{ij}^B(1,2) + \sum_n \frac{1}{n!} \sum_{j,k,\cdots=1}^{m} \rho_j\rho_k\cdots \int d2\, d3\cdots h_j(2)h_k(3)c_{ijk}^B\cdots(1,2,3,\ldots) \tag{64}$$

The direct correlation functions are defined by the functional derivative

$$c_{ijk\cdots}^B(1,2,3,\ldots) = \frac{\delta^n c_i{}^B(1)}{\delta\rho_j(2)\,\delta\rho_k(3)\cdots} \tag{65}$$

The superscript B stands for the bulk functions. We now introduce the

function $c_i^w(1)$, defined by

$$c_i^w(1) = -\beta u_i(1) - \ln g_i(1) + h_i(1) + \sum_{j,k=1}^{m} \left[\rho_j \rho_k \cdots \int d2\, d3\, h_j(2) h_k(3) c_{ijk}^B(1,2,3) \right] + \cdots \tag{66}$$

The inhomogeneous potential is of the form

$$u_i(1) = u_i^0(1) + w_i(1) \tag{67}$$

with $u_i^0(1)$ short range, and for a hard, smooth charged electrode, the electrostatic part is

$$w_i(1) = -\frac{e_i E_0 z_1}{2} \tag{68}$$

Combining this definition with the functional expansion (64) we get the HNC1 equation for the flat wall electrode:

$$\beta w_i(1) + \ln \rho_i(1) = \sum_{j=1}^{m} \rho_j \int d2\, h_j(2) c_{ij}^B(1,2) \tag{69}$$

Equation (69) has a deceivingly simple aspect, but because of the long-range character of $w_i(1)$ is not convergent and therefore not amenable to numerical solution. Using (51),

$$c_{ij}(|r_{12}|) = c_{ij}^0(|r_{12}|) - \beta w_{ij}(|r_{12}|) \tag{70}$$

with (52),

$$w_{ij}(|r_{12}|) = \frac{e_i e_j}{\epsilon |r_{12}|} \tag{71}$$

and replacing into (69) yields

$$\beta e_i \phi(1) + \ln \rho_i(1) = \sum_{j=1}^{m} \rho_j \int d2\, h_j(2) c_{ij}^0(1,2) \tag{72}$$

where $\phi(1)$ is defined by

$$\phi(1) = E_0 z_1 + \int d2 \sum_{j=1}^{m} \frac{e_j \rho_j(2)}{\epsilon r_{12}} \tag{73}$$

Equation (73) when combined with (6) yields (4). This equation is the plane electrode version of the hypernetted chain equation, called the HNC1 [69]. Note that the Gouy–Chapman theory results if the right-hand side of (72) is neglected. It is completely defined in terms of short-range

quantities, which is not the case for the first form of (69). The HNC1 is the theory that has the closure with the largest number of graphs. It satisfies the electroneutrality relations and the Stillinger–Lovett sum rules (see below). One important observation about the HNC1 is that it does not satisfy the contact theorem (38), but rather,

$$kT \sum_{i=1}^{m} \rho_i(0) = \frac{\epsilon}{8\pi} E_0^2 + \frac{\rho_0 kT[1 + \beta \partial p/\partial \rho_0]}{2} \tag{74}$$

where

$$\rho_0 = \sum_{i=1}^{m} \rho_i \tag{75}$$

For high fields and low concentrations the fact that we get the compressibility rather than the pressure is not very important and the HNC1 is still a reasonably good theory, as will be shown below. However, for dense systems at low field, this is a rather severe shortcoming. Specifically, when we are dealing with a molecular (dipolar) solvent the density is very large and the dielectric constant ϵ is on the order of 1 (instead of 80 in water), which makes the electrostatic term, which satisfies the contact theorem (38), small in comparison to the contact density term. The consequence is that the HNC1 will put more counterions near the electrode than the exclusion of the hard cores will permit. Eventually, thermodynamic stability conditions will be violated, and we get a negative capacitance, reflected by a decreasing potential drop $\Delta\phi$ with increasing applied external field E_0.

The HNC is the most accurate theory for bulk electrolytes. One would expect that this fact would remain true in the plane electrode limit. However, because of the inaccuracy of the HNC for uncharged hard-sphere fluids, the HNC1 does not do well in representing the exclusion volume of the ions and is not, on the whole, such a good approximation for the electric double layer. The bulk direct correlation function

$$c_{ij}^{B}(|\, r_{12} \,|) \tag{76}$$

which should be used in solving the HNC1 equation (72), is that obtained of the bulk HNC equation for the same system. However, this HNC/HNC approximation yields poor results when compared to computer simulations [69]. Generally, better results are obtained if instead of the HNC bulk direct-correlation function the corresponding MSA functions are used [58,70–74]. Some results of the GC and HNC1/MSA theory are compared with simulations in Fig. 1. The next term to be considered is the third term of (64), which is a three-particle contribution. The three-

particle direct-correlation function is in general a very complex function and must be approximated. The simplest of these approximations is to include the first diagram of the density expansion of the three-point direct-correlation function, the bridge diagram [75]. Ballone et al. [76] performed this calculation with good success. The density profile for the 1 M, 1–1 electrolyte at a surface charge $q_s^* = q_s\sigma^2/e = 0.7$ is shown in Fig. 2. This is the highest surface density simulated and shows charge oscillations due to the hard core of the electrolyte, which is not seen in the HNC1 approximation. In this calculation the bridge diagrams were computed directly from the product of the three bulk pair correlation functions, which is the first term in the density expansion of the bulk triplet direct-correlation function

$$c^B_{ijk}(1,2,3) = h^B_{ij}(1,2)h^B_{ik}(1,3)h^B_{kj}(3,2) \tag{77}$$

The results of this calculation are shown in Fig. 2. Since there are no adjustable parameters, the agreement is very good. An alternative, less

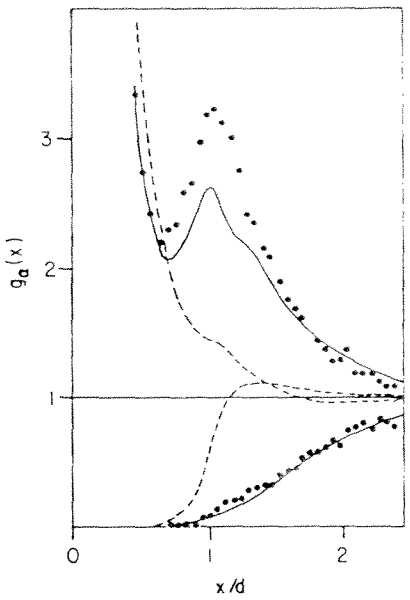


Fig. 2 Density profile for charged hard spheres near a charged hard wall for a 1M, 1:1 electrolyte. The definition of q^* and the model parameters are the same as those of Fig. 1. The solid and dashed curves give the HNC1 results with and without the contribution of the bridge diagrams. The value of q^* is 0.7.

laborious procedure was suggested by Rosenfeld and Blum [77], but actual calculations were not performed.

Another way of improving the HNC1 approximation was introduced by Forstmann and co-workers [78–82]. In their method the HNC1 equation is used as described above, but instead of taking the bulk direct-correlation function, as prescribed by (72), a local density-dependent $c_{ij}^{B}(r,\tilde{\rho})$ is taken. The local density is defined by

$$\tilde{\rho}_i(z) = \frac{1}{2\sigma\delta}\int_{x-\delta}^{x+\delta} dx \int_{x-\sigma/2}^{x+\sigma/2} dy\ \rho_i(y) \tag{78}$$

where σ is the diameter of the ion and δ is an adjustable parameter. The bulk correlation function is then

$$c_{ij}^{B}(|\ r_{12}\ |)|_{\rho_i=\tilde{\rho}_i(z)} \tag{79}$$

For the test case with surface charge, $q_s^* = 0.7$; the results of this method are similar to those shown in Fig. 2. A recent calculation by Mier et al. [83], based on a density functional expansion, also gives results comparable to those of Fig. 2 but does not require the use of adjustable parameters.

B. BGY-Based Approximations

The BGY equations [84,85] can be derived from the one-particle direct-correlation function $c_i(1)$. Consider again (64). Letting the gradient ∇ act on the Mayer f function of the graphical expansion of $c_i(1)$, we get the BGY equation; the first member of this hierarchy is

$$-kT\,\nabla_1\,\rho_i(1) = \rho_i(1)\,\nabla_1\,u_i(1) + \sum_{j=1}^{m}\int d2\ \rho_{ij}(1,2)\,\nabla_1\,u_{ij}(1,2) \tag{80}$$

Using (67) and (70) to eliminate the long-range terms, we obtain (80) in a different form:

$$\begin{aligned} -kT\,\nabla_1\,\rho_i(1) &= \rho_i(1)\,\nabla_1\,u_i^0(1) + e_i\rho_i(1)\,\nabla_1\,\phi(1) \\ &\quad + \sum_{j=1}^{m}\int d2\ \rho_{ij}(1,2)\,\nabla_1\,[u_{ij}^0(1,2)] \\ &\quad + \rho_i(1)e_i\sum_{j=1}^{m} e_j\int d2\ \rho_j(2)h_{ij}(1,2)\,\nabla_1\left[\frac{1}{\epsilon r_{ij}}\right] \end{aligned} \tag{81}$$

This equation can be integrated from ∞ to z, to yield

$$\ln[g_i(z)] = -e_i[\phi(z) + \psi_j(z)] + J_{ij}(z) \tag{82}$$

which together with the Poisson equation (3) forms a closed system of

equations that is very convenient for numerical solutions. This equation is of the same type as the one derived from the HNC1 equation (72). The right-hand term consists of three contributions: the potential $\phi(1)$, which is determined by the single-particle distribution function $\rho_i(z)$; and the terms $\psi_j(z)$ and J_{ij}, which are functions of the pair distribution function $h_{ij}(1,2)$. From (81) we get

$$\psi_j(z) = \int_z^\infty dz_1\, \rho_i(1) e_i \sum_{j=1}^{m} e_j \int d2\, \rho_j(2) h_{ij}(1,2)\, \nabla_1 \frac{1}{\epsilon r_{ij}}$$

$$J_{ij}(z) = \int_z^\infty dz_1 \sum_{j=1}^{m} \int d2\, \rho_{ij}(1,2)\, \nabla_1\, u_{ij}^0(1,2) \tag{83}$$

We remark that in (81) [and also in (82)], if the fluctuation terms $J_{ij}(z)$ and $\psi_j(z)$ are neglected, we recover the Gouy–Chapman equation (8), which has a known analytical solution. In the BGY-based theories the pair correlation function $h_{ij}(1,2)$ must be given by some approximation. The interesting feature of the BGY equation is that no matter what the closure, it satisfies the contact theorem (36). The simplest approximation is equivalent to Kirkwood's superposition approximation and consists in writing

$$h_{ij}(1,2) = h_{ij}^B(|\, r_{12}\,|) \tag{84}$$

where $h_{ij}^B(|\, r_{12}\,|)$ is the bulk pair correlation function. It fails to satisfy the electroneutrality condition (45)

$$-e_i = \int d2 \sum_j e_j \rho_j(2) h_{ij}(1,2) \tag{85}$$

and gives very poor results when compared to the computer simulations. The approximation [86]

$$h_{ij}(1,2) = \begin{cases} f_i(1) f_j(2) h_{ij}^B(|\, r_{12}\,|) & r_{12} > \sigma_{ij} \\ -1 & r_{12} < \sigma_{ij} \end{cases} \tag{86}$$

where

$$\sigma_{ij} = \frac{\sigma_j + \sigma_i}{2} \tag{87}$$

is constructed so that the functions $f_i(1)$ are required to satisfy the electroneutrality condition (85) for the inhomogeneous pair distribution function. This inhomogeneous pair correlation function will for high densities give negative values of $g_{ij}(1,2)$. The problem is specially severe for salt concentration greater than 2 M or also for high surface density charge. A simple way to circumvent this problem was suggested by Caccamo et

al. [87]. The approximation

$$h_{ij}(1,2) = \begin{cases} f_i(1)f_j(2)h_{ij}^B(|\, r_{12}\,|) + A_{ij}(1,2) & r_{12} > \sigma_{ij} \\ -1 & r_{12} < \sigma_{ij} \end{cases} \tag{88}$$

where the function $A_{ij}(1,2)$ takes into account the hard-core exclusion effect of the ions near the electrode surface. A sensible guess is

$$A_{ij}(1,2) = Ah_{ij}^B(|\, r_{12}\,|) \qquad z_1, z_2 < 2\sigma_{ij} \tag{89}$$

The parameter A is adjusted to eliminate negative values of $g_{ij}(1,2)$. The results are, however, not very sensitive to the exact value of A. Comparison with the Monte Carlo simulations is again good; for the test case with $q_s^* = q_s\sigma^2/e = 0.7$, this method yields results that are similar to those shown in Fig. 2.

Another method, which is probably the most systematic and has the great advantage of formally satisfying all known sum rules, is to use the inhomogeneous pair correlation function $h_{ij}(1,2)$ obtained from the inhomogeneous OZ equation

$$h_{ij}(1,2) - c_{ij}(1,2) = \sum_{k=1}^{m} d3\; h_{ik}(1,3)\rho_k(3)c_{kj}(3,2) \tag{90}$$

with a suitable closure for the direct-correlation function. Thus we have the MSA2 approximation

$$c_{ij}(1,2) = -\beta w_{ij}(1,2) \qquad r_{12} > \sigma_{ij} \tag{91}$$

and the HNC2 approximation

$$c_{ij}(1,2) = -\beta w_{ij}(1,2) + h_{ij}(1,2) - \ln g_{ij}(1,2) \qquad r_{12} > \sigma_{ij} \tag{92}$$

C. WLMB-Based Equations

Yet another integral equation is derived from the one-particle direct-correlation function $c_i(1)$, Eq. (65), by introducing relative coordinates in the diagram representation and taking the derivatives with respect to those coordinates. This yields an exact hierarchy of equations which is related to the BGY hierarchy. The first member is the Wertheim–Lovett–Mou-Buff (WLMB) equation,

$$\nabla_1\rho_i(1) + \beta\rho_i(1)\,\nabla_1 u_i(1) = \rho_i(1)\sum_{j=1}^{m}\int d2\; c_{ij}(1,2)\,\nabla_2\rho_j(2) \tag{93}$$

This equation contains long-range, divergent terms. Introducing the local

potential $\phi(1)$ [Eq. (4)] we have

$$\nabla_1 \ln \rho_i(1) + \beta \nabla_1 \phi_i(1) = \sum_{j=1}^{m} \int d2\, c_{ij}^{SI}(1,2) \nabla_2 \rho_j(2) \tag{94}$$

The MSA2 approximation cannot be integrated explicitly, as is the case for the homogeneous MSA. However, when the ions are approximated by charged points, then, for some specific form of the density profiles, $\rho_i(1)$, the OZ equation can be integrated [72,88]. One interesting feature of the results above is that the pair correlation function g_{12} decays as $1/r^3$ if the charges are held at a fixed distance from the electrode. This behavior is exact and was first pointed out by Jancovici [48] (see Chapter 5).

The numerical solution of the HNC2 and MSA2 approximations has been extensively studied by Plischke and Henderson [89,90]. For $q_s^* = q_s\sigma^2/e = 0.7$ they obtain results very similar to those shown in Fig. 2. Approximate solutions of this type satisfy the local electroneutrality condition (45) and dipole sum rule (56).

Plischke and Henderson solve the HNC2 and MSA2 equations by Fourier transform techniques. To facilitate the calculation they subtract from $c(1,2)$ functions containing the long-range behavior and which are analytically transformable. Colmenares and Olivares [91,92] have used the WLMB equation together with the approximation

$$c(1,2) = c_B(1,2) \tag{95}$$

and obtained reasonably good results. However, it has been shown that this procedure is equivalent to solving the HNC1 equation.

D. Kirkwood's Equation

An interesting approach has been suggested by Kjellander and Marcelja [93–96], based on the observation that for the HNC approximation the chemical potential can be obtained explicitly as a function of the pair potential $h_{ij}(|\, r_{12}\,|)$ for a homogeneous fluid. Then, within the HNC, the function $c_i(1)$ can be explicitly evaluated. The central idea is to divide the three-dimensional space into two-dimensional layers that are homogeneous. The three-dimensional OZ equation can be mapped into coupled set of N two-dimensional OZ equations for a mixture of N components; each component is an ion in a different layer. The particles interact with a species-dependent interaction pair potential. In the limit of an infinite number of layers this procedure yields the correct inhomogeneous OZ

equation. The chemical potential $\mu_i(\alpha)$ of the ith ion the αth layer is given by Kirkwood's equation:

$$\mu_i(\alpha) = kT \ln \rho_i(\alpha) + kT \ln \frac{\Lambda_0}{\Delta z} + V_i(\alpha) + \sum_{j=1}^{m} \rho_j(\beta) \int_0^1 d\lambda \int dR\, g_{ij}(R,\alpha\beta;\lambda) \frac{\partial[u_{ij}(R,\alpha\beta;\lambda)]}{\partial\lambda} \tag{96}$$

where λ is the coupling parameter, Δz the thickness of the layer, Λ_0 the ideal gas fugacity, $V_i(\alpha)$ the interaction between a particle in layer α and the wall, and R the two-dimensional distance. In the HNC closure

$$c_{ij}(R,\alpha\beta) = -\beta w_{ij}(R,\alpha\beta) + h_{ij}(R,\alpha\beta) - \ln g_{ij}(R,\alpha\beta) \tag{97}$$

Kirkwood's equation can be integrated to yield

$$\rho_i(\alpha) = \frac{\Delta z}{\Lambda_0} \exp\left[\beta\mu_i(\alpha) + \sum_{\beta,j=1} \rho_j(\beta) \int dR \left[\frac{1}{2} h_{ij}^2(R,\alpha\beta) - c_{ij}(R,\alpha\beta) - \beta w_{ij}(R,\alpha\beta)\right] - \left[\frac{1}{2} \ln\left[g_{ij}(R,\alpha\beta)\right] - \frac{\beta w_{ij}(R,\alpha\beta)}{2}\right]_{R=0} - \Phi_i(\alpha)\right]$$

where $\Phi_i(\alpha)$ is the average potential for layer α.

$$\Phi_i(\alpha) = \frac{2\pi e^2}{\epsilon} \sum_{\beta} \rho_i(\beta) | z_\alpha - z_\beta | \tag{98}$$

The results of their calculation for $q_s^* = q_s\sigma^2/e = 0.7$ are again similar to those shown in Fig. 2.

The modified Poisson–Boltzmann (MPB) equation [98,99] is an approximate solution of Kirkwood's equation [100]. As has been shown by Levine, Outhwaite, and Bhuiyan, it can be cast in the form of an inhomogeneous linearized Debye–Hückel for the pair potential

$$\nabla^2\phi_j(1,2) = \kappa^2(z_2)\phi_j(1,2) \tag{99}$$

where the inhomogeneous Debye parameter is

$$\kappa^2(z_2) = \frac{4\pi}{\epsilon kT} \sum_{i=1}^{m} \rho_i e_i^2 g_i(z_2) \tag{100}$$

They obtain an approximate solution, valuable at low concentration and small electrode charge. It does not yield the layering seen in Fig. 2, but is useful for its regime of validity because of its simplicity.

VII. INTEGRAL EQUATIONS: NONPRIMITIVE MODELS

In contrast to the large number of studies for the primitive model, using various approximations, there have been few investigations of inhomogeneous electrolytes in which a molecular model of the solvent is employed. The earliest study is that of Carnie and Chan [73] and of Blum and Henderson [97], who used the hard sphere with a point dipole model for the solvent and the MSA1 approximation. This model is useful and instructive because it is analytic and because, as has been shown recently, the MSA1 is a "controlled approximation" in the sense that the internal energy is always a bound of the exact value of the internal energy [101].

The solution of the MSA1 for the ion–dipole mixture is analytic but implicit. Explicit results can be obtained in limiting cases, such as the dilute solution limit. We obtain in this limit a series of powers of Debye's parameter κ. The potential difference is

$$\Delta\phi = \frac{E_0}{\epsilon\kappa} + \frac{E_0}{2\epsilon}\left(\alpha_i + \frac{\epsilon - 1}{\lambda}\sigma_s\right) + \cdots \tag{101}$$

where σ_i is the ionic diameter, σ_s the solvent diameter, ϵ the dielectric constant of the solvent, and the parameter λ is obtained from ϵ using the MSA equation [102]. The result is

$$\lambda(1 + \lambda)^2 = 4\sqrt{\epsilon} \tag{102}$$

Since this is a cubic equation for λ, an explicit expression of this parameter can be obtained for any solvent. It is related to an effective dipole moment, which can be computed from Wertheim's work [102] or can be obtained from the experimental dielectric constant. The parameter λ is a slowly varying function of ϵ: For $\epsilon = 1$, $\lambda = 1$ and for $\epsilon = 80$, $\lambda = 2.7$. The effective dipole moment for this last case is 2.2 D, compared to the value of 1.85 for the water molecule.

The first term is just the linearized Gouy–Chapman result, while the second term is independent of concentration. If the series in (101) is terminated after the second term, the differential capacitance will have the form

$$C^{-1} = C_{GC}^{-1} + C_H^{-1} \tag{103}$$

where C_{GC} is the Gouy–Chapman capacitance and C_H, often called the inner layer or Helmholtz capacitance, is independent of concentration. As shown in Fig. 3, (103) is verified experimentally [103], although there are small deviations that are usually ascribed to experimental errors. We will return to this point shortly.

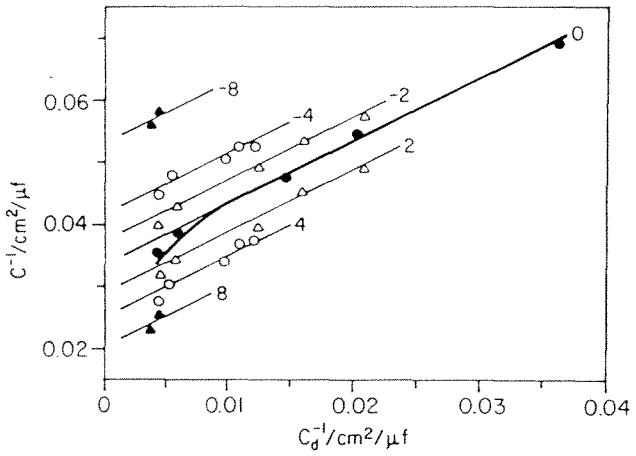


Fig. 3 Inverse capacitance of an aqueous solution of NaH_2PO_4 at 25°C near a hanging mercury electrode as a function of the inverse Gouy–Chapman capacitance, C_d^{-1}. The light straight lines of unit slope give the results obtained from (103). The heavy lines give the full MSA results obtained using $\sigma = \sigma_s = 2.76$ Å. The metallic contribution of -0.026 cm$^2/\mu$F has been included.

Experimentally, C_H does not depend on the properties of the ions. Because of this, in the past C_H has been regarded as the capacitance of an inner layer of water molecules, tightly bound to the electrode surface. This is highly unlikely from the emerging theoretical and experimental evidence, which shows that the structure of the first adsorbed layer (which is not the classic inner layer responsible for the capacitance behavior) is rather sensitive to changes in the potential and not at all simple [104]. Thus the idea that the total capacitance is the result of coupling the inner layer in series with the diffuse layer is unlikely.

In the old picture C_H is extracted from the experimental data and is just a parameter. One needs to introduce a dielectric constant for the inner layer, usually very low ($\epsilon_i \simeq 3$ to 5), with no explanation of why the solvent would separate sharply in two phases for all values of the potential bias. The dielectric constant is a bulk property and has no meaning for a monolayer, and although phase transitions may occur at the inner Helmholtz layer, they will be very sensitive to the applied potential bias [52] and therefore not at all constant.

The ion–dipole charged wall model described above is much more

satisfactory. It gives rise to (103) at low to moderate concentrations. Because

$$\frac{\epsilon - 1}{\lambda} \gg 1 \tag{104}$$

the theory correctly predicts that C_H will be determined by the solvent properties to a good approximation. The model also gives a formula for C_H with no adjustable parameters that yields values in general agreement with experiment. However, the underlying interpretation is different from the old picture, because now the additional constant term is due to the fact that the solvent has a molecular structure, and the additional shift is due to the finite size of the solvent molecule. It is easy to see from (101) that this effect will disappear in the limit $\sigma_s \to 0$. The solvent polarization by the field of the diffuse layer is an effect that occurs over the entire double layer and is not restricted to the monolayer adjacent to the electrode. This is seen in Fig. 4, where the orientational part of the profile of the solvent molecules, $\Delta h_s(z)$, is plotted. The dashed curve is a plot of $e^{-\kappa x}$. It is clear that as required by theory, for large z, $\Delta h_s(z)$ behaves asymptotically as $e^{-\kappa x}$.

We have mentioned that (101) and (103) are just the leading terms in the expansion of the MSA result. It is fairly simple to solve these equations

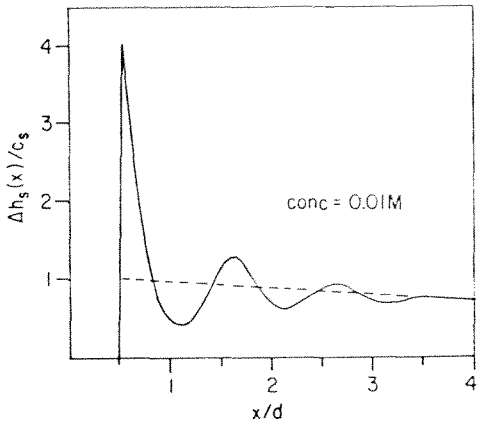


Fig. 4 $\Delta h_s(z)$ as a function of z in the MSA for a 0.01 M, 1:1 model electrolyte solution. The parameters of Fig. 3 are used. The parameter c_s is a normalization parameter whose value is not significant for the present calculation.

numerically and obtain the MSA capacitance near the point of zero charge for finite concentrations. In Fig. 3 the MSA capacitances are compared with experimental values, which now show similar deviations from linearity.

To get good agreement with experiment, (101) must be coupled with a more realistic model of a metal. A metal is not a charged hard wall; it contains electrons. The simplest model of a metal is the jellium model, in which the metal is regarded as a free fermion gas of electrons in a positive background potential arising from the metallic ions. In principle, the neutralizing background should be the lattice of the metal ions. But this is the problem of an electron in a three-dimensional lattice, which is periodic in only two dimensions, and therefore has difficult boundary conditions. A much simpler model is obtained by smearing the charge of the ions. Then the problem is much easier, since we have electrons in a continuous background that terminates at the electrode surface. Clearly, the electrons can spill out into the electrolyte. The problem of the electron at such an interface has been studied extensively in the literature [105]. The addition of the electrolyte complicates the problem considerably. For this reason in the early treatments the electron density is parameterized and the parameters are found minimizing the surface free energy (or equivalently, the surface energy, since the electron gas is completely degenerate). Several refinements of this model have been made: The metallic lattice structure is taken into account by means of a pseudopotential. In any case, the inhomogeneous region of this electron gas gives rise to a potential difference that has to be added to the one of the electrolyte. This results in generally good agreement of the work functions and potentials of zero charge for most non-*d*-metals, which unfortunately excludes the noble metals. A more refined model of jellium in which the Kohn–Lang equations [106] were solved were performed by Price and Halley [28,29]. In their calculation the electron density profile shows the required Friedel oscillations.

The modified Poisson–Boltzmann (MPB) approximation has been used to study the ion–dipole hard-wall system [98,99]. Unfortunately, solutions of these equations can be obtained only at low solvent densities and small dielectric constant. Within its region of validity, the results are interesting and similar to that of the MSA.

Torrie and colleagues [107] have used the HNC1 approximation together with a model of the solvent that includes dipoles and quadrupoles to study solvent effects in the double layer. For reasons that are not yet understood, they are unable to obtain convergence for planar walls. However, they are able to examine spherical electrodes which are 50 times the size of the ions and solvent molecules. These calculations are very

difficult, but they are worthwhile since the HNC1 approximation is the most reliable simple approximation for electrolytes and direct computer simulations of these systems in the presence of ions are not yet feasible. Torrie et al. find considerable structure near the electrode when the solvent is pure, but this ordering is inhibited when the ions are present; presumably, this is due to the screening of the interactions between the solvent molecules by the ions.

VIII. INTERACTION BETWEEN MACROIONS IN AN ELECTROLYTE

Although experimental studies of capacitances provide some support for the modern theories of the double layer discussed herein, the support is not overwhelming. Semiempirical descriptions can always be brought into agreement with experiment. The problem is that the capacitance and the potential are integral quantities, which are averages over large regions of the interface and therefore insensitive to the details of the structure of the interface. Furthermore, only very recently have there been direct structural studies of electrode interfaces, which have revealed that the metal–electrode interface undergoes structural changes as the potential is changed.

Israelachvili and colleagues [108] have measured the force between crossed cylindrical mica surfaces in an electrolyte solution. The prediction of Poisson–Boltzmann theory is that the force between the plates should be repulsive and monotonically increasing as the separation between the cylinders is decreased. When the short-range van der Waals forces are also included, as is done in the Derjaguin, Landau, Verwey, and Overbeek theory (DLVO), the force is repulsive at large separations but attractive at short range, as shown, as shown in Fig. 5. Israelachvili's experimental results agree with the DLVO theory at large separation but show strong oscillations at short separation.

Since one of the differences between the modern theories of the double layer and the Gouy–Chapman theory is the appearance of oscillations in the correlation functions predicted by these modern theories might account for the oscillations seen by Israelachvili. Ideally, one would like to obtain the solution of the MSA for large macroions in an ion–dipole solution. This has been done in principle [109], but the numerical results are not yet available.

Numerical results have been obtained by Henderson and Lozada-Cassou [110], who used a similar but simplified model. They considered charged hard-sphere ions and a hard-sphere solvent in a dielectric continuum whose dielectric constant equals that of the solvent. This result,

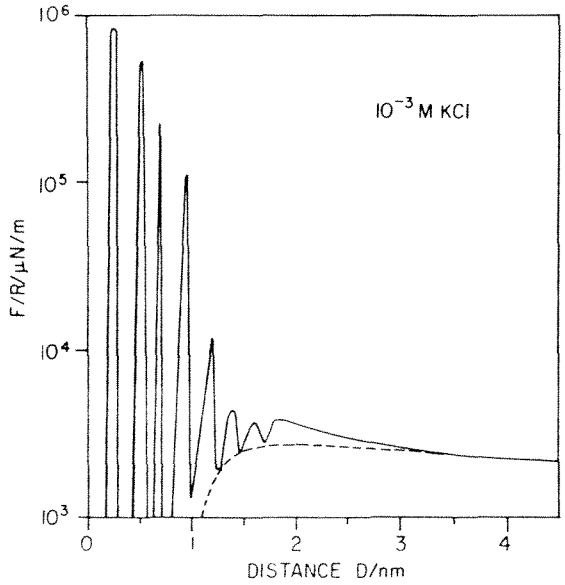


Fig. 5 Force between macroions in an electrolyte. The solid curve gives the experimental results of Israelachvili et al. The dashed curve gives the DLVO result.

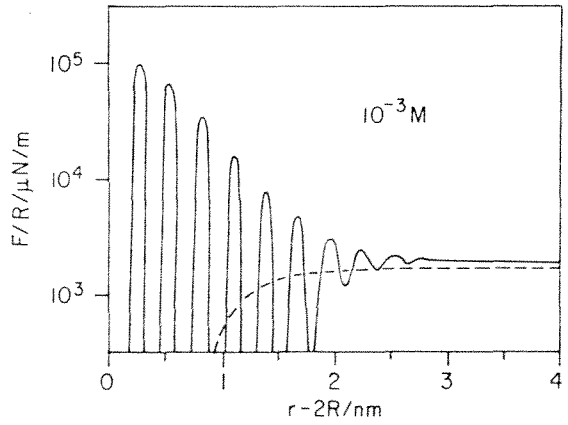


Fig. 6 Force between macroions in an electrolyte. The solid curve gives the theoretical results of Henderson and Lozada–Cassou. The dashed curve gives the DLVO result.

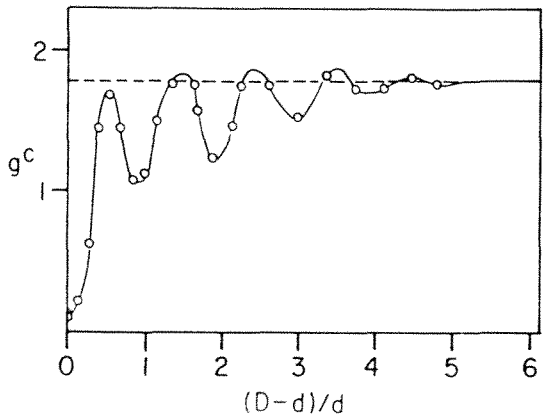


Fig. 7 Pressure between two parallel plates as a function of the plate separation. The fluid between the plates is the tetrahedral sticky water model. The results are obtained by grand canonical ensemble Monte Carlo simulation.

shown in Fig. 6, is strikingly similar to the experimental results shown in Fig. 5.

Yet another model of the solvent has recently been proposed by Blum, Bratko, Cummings, and Luzar [111–113]. This model, which is solved analytically, consists of a hard sphere with a point dipole and a tetrahedral sticky potential that mimics hydrogen bonding. A computer simulation [114] of this solvent between parallel plates clearly yields the pressure oscillations seen in the experiments of Israelachvili, as shown in Fig. 7.

IX. SUMMARY

In most of this century our understanding of inhomogeneous electrolytes has been based on the Poisson–Boltzmann approximation (the Gouy–Chapman or the DLVO theories). These theories predict rather featureless correlation functions. Until recently, experimental results were not sufficiently sensitive to expose the deficiencies of these theories. This is rapidly changing. Modern theories predict much more interesting correlation functions. Further, the integral quantities, such as the differential capacitance, agree with the experiment without semiempirical adjustment. Moreover, such venerable concepts as the inner layer have been found to have a limited validity in the explanation of capacitance experiments.

REFERENCES

1. Blum, L., Abruna, H. D., White, J., Gordon, J. G., II, Borges, G. L., Samant, M., and Melroy, O. R. (1986). Study of underpotential deposited copper on gold by fluorescent EXAFS, *J. Chem. Phys. 85*: 6732.
2. Samant, M. G., Borges, G. L., Gordon, J. G., II, Melroy, O. R., and Blum, L. (1987). In situ surface *EXAFS* of a lead monolayer at a silver (111) electrode/electrolyte interface, *J. Am. Chem. Soc. 109*: 5970.
3. Samant, M. G., Toney, M. F., Borges, G. L., Blum, L., and Melroy, O. R. (1988). In situ grazing incidence x-ray diffraction study of electro-chemically deposited Pb monolayers on Ag (111), *Surf. Sci. Lett. 193*: L29.
4. Melroy, O. R., Toney, M. F., Borges, G. L., Samant, M. G., Kortright, J. B., Ross, P. N., and Blum, L. (1989). An in-situ grazing incidence x-ray scattering study of the initial stages of electrochemical growth of lead on silver (111). *J. Electroanal. Chem., 158*: 403.
5. Materlik, G., Zegerhagen, J., and Uelhof, W. (1985). *Phys. Rev. B 32*: 5502. Materlik, G., Schmah, M., Zegerhagen, J., and Uelhof, W. (1987). *Ber. Bunsenges. Phys. Chem. 91*: 292.
6. Bedzyk, M. J. (1990). Measuring the diffuse double layer at an electrochemical interface with long period x-ray standing waves. *Synchrotron Radiat. News 3*: 25.
7. Schardt, B. C. (1986). *Appl. Phys. Lett. 49*: 1172. (1990). *Interdisciplinary Conference on Electrified Interfaces*, Asilomar, Calif.
8. Wiechers, J., Twomey, T., Kolb, D. M., and Behm, R. J. (1988). *J. Electroanal. Chem. 248*: 451. (1990). *Interdisciplinary Conference on Electrified Interfaces*, Asilomar, Calif.
9. Schmickler, W., and Henderson, D. (1990). A model for the scanning tunnel microscope operating in an electrolyte solution, *J. Electroanal. Chem. 290*: 283.
10. Bruckenstein, S., and Shay, M. (1985). *Electrochim. Acta 30*: 1295.
11. Kanazawa, K. K., and Gordon, J. G., II (1985). *Anal. Chem. 57*: 1770.
12. Deakin, M. R., and Melroy, O. R. (1989). Monitoring the growth of an oxide film on aluminum in situ with the quartz crystal microbalance. *J. Electrochem. Soc. 136*: 349.
13. Fleischmann, M., Graves, P., Hill, I., Oliver, A., and Robinson, J. (1983). Raman spectroscopic and x-ray diffraction studies of electrode solution interfaces, *J. Electroanal. Chem. 150*: 33.
14. Jeanmaire, D. I., and Van Duyne, R. P. (1977). *J. Electroanal. Chem. 84*: 1.
15. Corn, R. M., and Philpott, M. (1984). The interpretation of very low frequency Raman scattering from roughened silver electrodes: clusters, cavities or complexes? *J. Chem. Phys. 81*: 4138.
16. Richmond, G. L., Robinson, J. M., and Shannon, V. L. (1988). Second harmonic generation studies of interfacial structure and dynamics, *Prog. Surf. Sci. 28*: 1.
17. Israelachvili, J. N. (1985). *Chem. Scr. 25*: 7.

18. Vallette, G., Hamelin, A., and Parsons, R. (1978). *Z. Phys. Chem. (Frankfurt) 113*: 71.
19. Vallette, G. (1981). *J. Electroanal. Chem. 122*: 285.
20. Vallette, G. (1982). *J. Electroanal. Chem. 138*: 37.
21. Badiali, J. P., and Goodisman, J. (1978). Thermodynamics and statistical mechanics of the ideally polarisable electrode: Gibbs isotherm and surface tension, *J. Electroanal. Chem. 91*: 151.
22. Badiali, J. P., Rosinberg, M. L., and Goodisman, J. (1981). Effect of solvent on properties of liquid metal surface, *J. Electroanal. Chem. 130*: 31.
23. Badiali, J. P., Rosinberg, M. L., and Goodisman, J. (1983). Contribution of the metal to the differential capacitance of an ideally polarisable electrode, *J. Electroanal. Chem. 143*: 73.
24. Badiali, J. P., Rosinberg, M. L., and Goodisman, J. (1983). The metals in the polarisable interface coupling with the solvent phase, *J. Electroanal. Chem. 150*: 25.
25. Badiali, J. P., Rosinberg, M. L., Vericat, F., and Blum, L. (1983). A microscopic model of the liquid metal-ionic solution interface, *J. Electroanal. Chem. 158*: 253.
26. Kornyshev, A. A., Schmickler, W., and Vorontytsev, M. (1982). Nonlocal electrostatic approach to the problem of a double layer at a metal electrolyte interface, *Phys. Rev. B 26*: 5244.
27. Schmickler, W., and Henderson, D. (1984). The interphase between jellium and a hard sphere electrolyte: a model for the electric layer, *J. Chem. Phys. 80*: 3381. (1984). Approximate solution for the electronic density profile at the surface of jellium, *Phys. Rev. 30*: 3081. (1986). The interphase between jellium and a hard sphere electrolyte: capacity-charge characteristics and dipole potentials, *J. Chem. Phys. 85*: 1650.
28. Price, D. L., and Halley, J. W. (1983). A new model of the differential capacitance of the double layer, *J. Electroanal. Chem. 150*: 347.
29. Price, D. L., and Halley, J. W. (1988). Electronic structure of metal electrolyte interfaces: three-dimensional calculation, *Phys. Rev. B 38*: 9357.
30. Blum, L. (1974). Solution of the mean spherical model for a mixture of charged hard spheres and dipoles, *Chem. Phys. Lett. 26*: 200.
31. Blum, L., Cummings, P. T., and Bratko, D. (1990). A general solution of the molecular Ornstein–Zernike equation for spheres with anisotropic surface adhesion, *J. Chem. Phys. 92*: 3741.
32. Torrie, G. M., and Valleau, J. P. (1979). Electrical double layers, *Chem. Phys. Lett. 65*: 343.
33. Torrie, G. M., and Valleau, J. P. (1980). Electrical double layers. I. Monte Carlo study of a uniformly charged surface, *J. Chem. Phys. 73*: 5807.
34. Torrie, G. M., Valleau, J. P., and Patey, G. N. (1982). Electrical double layers. II. Monte Carlo and HNC studies of image effects, *J. Chem. Phys. 76*: 4615.
35. Torrie, G. M., and Valleau, J. P. (1982). Electrical double layers. III. Monte Carlo and HNC studies of image effects, *J. Phys. Chem. 86*: 3251.

36. Valleau, J. P., and Torrie, G. M. (1984). Electrical double layers, IV, *J. Chem. Phys. 81*: 6291.
37. Torrie, G. M., Valleau, J. P., and Outhwaite, C. W. (1984). Electrical double layers, V, *J. Chem. Phys. 81*: 6296.
38. Gouy, G. (1910). *J. Phys. 9*: 457.
39. Chapman, D. L. (1913). *Philos. Mag. 25*: 475.
40. Levine, S. (1951). *Proc. Phys. Soc. 64*: 781; *66*: 357.
41. Gruber, Ch., Lebowitz, J. L., and Martin, Ph. A. (1981). Sum rules for the inhomogeneous Coulomb system, *J. Chem. Phys. 75*: 944.
42. Blum, L., Gruber, Ch., Lebowitz, J. L., and Martin, Ph. A. (1982). On perfect screening for charged systems, *Phys. Rev. Lett. 48*: 1769.
43. Blum, L., Gruber, Ch., Henderson, D., Lebowitz, J. L., and Martin, Ph. A. (1983). On the properties of inhomogeneous charged systems, *J. Chem. Phys. 78*: 3195.
44. Jancovici, B. (1981). Exact results for the two-dimensional one-component plasma, *Phys. Rev. Lett. 46*: 386.
45. Jancovici, B. (1981). Charge distribution and kinetic pressure in a plasma: a soluble model, *J. Phys. Lett. 42*: L223.
46. Alastuey, A., and Jancovici, B. (1981). On the classical two-dimensional one-component Coulomb plasma, *J. Phys. 42*: 1.
47. Jancovici, B. (1982). Classical Coulomb system near a plane wall. I, *J. Stat. Phys. 28*: 43.
48. Jancovici, B. (1982). Classical Coulomb system near a plane wall. II, *J. Stat. Phys. 29*: 263.
49. Jancovici, B. (1984). Surface properties of a classical two-dimensional one-component plasma: exact results, *J. Stat. Phys. 34*: 803.
50. Henderson, D., Blum, L., and Lebowitz, J. L. (1979). An exact formula for the contact value of the density profile of a system of charged hard spheres near a charged wall. *J. Electroanal. Chem., 102*: 315.
51. Huckaby, D. A., and Blum, L. (1990). Exact results for the adsorption of a dense fluid onto a triangular lattice of sticky sites, *J. Chem. Phys. 92*: 2646.
52. Blum, L. (1990). Structure of the electric double layer, *Adv. Chem. Phys. 78*: 171.
53. Cornu, F. (1989). Two-dimensional models for an electrode with adsorption sites, *J. Stat. Phys. 54*: 681.
54. Lieb, E. H., and Lebowitz, J. L. (1972). The existence of thermodynamics for real matter with Coulomb forces, *Adv. Math. 9*: 316.
55. Alastuey, A., and Martin, Ph. A. (1988). Absence of exponential clustering for the static quantum correlations and the time displaced correlations in charged fluids, *Europhys. Lett. 6*: 385.
56. Stillinger, F. H., and Lovett, R. (1968). Ion pair theory of concentrated electrolytes: basic concepts, *J. Chem. Phys. 48*: 3858. (1968). General restrictions on the distribution of ions in electrolytes, *J. Chem. Phys. 49*: 1991.

57. Outhwaite, C. W. (1973). Comment on the second moment condition of Stillinger and Lovett, *Chem. Phys. Lett. 24*: 73.
58. Carnie, S. L., and Chan, D. Y. C. (1981). The Stillinger and Lovett condition for nonuniform electrolytes, *Chem. Phys. Lett. 77*: 437.
59. Martin, Ph. A. (1988). Sum rules in charged fluids, *Rev. Mod. Phys. 60*: 1075.
60. Blum, L., Gruber, Ch., Henderson, D., Lebowitz, J. L., and Martin, Ph. A. (1981). A sum rule for an inhomogeneous electrolyte, *J. Chem. Phys. 75*: 5974.
61. Triezenberg, D. G., and Zwanzig, R. W. (1972). *Phys. Rev. Lett. 28*: 1183.
62. Henderson, J. R., and van Swol, F. (1984). On the interface between a planar wall: theory and simulations of a hard sphere fluid at a hard wall, *Mol. Phys. 51*: 991.
63. Lebowitz, J. L., and Percus, J. K. (1963). Statistical mechanics of nonuniform fluids, *J. Math. Phys. 4*: 116, 248.
64. Hansen, J. P., and McDonald, I. R. (1976). *Theory of Simple Liquids*, Academic Press, New York.
65. Blum, L., and Stell, G. (1976). Solution of the Ornstein–Zernike equation for the wall particle distribution function, *J. Stat. Phys. 15*: 439.
66. Sullivan, D. E., and Stell, G. (1977). Correlations at a fluid surface, *J. Chem. Phys. 67*: 2567.
67. Henderson, D., Abraham, F. F., and Barker, J. A. (1976). The Ornstein–Zernike equation for a fluid in contact with a surface, *Mol. Phys. 31*: 1291.
68. Percus, J. K. (1976). Model for density variation at a fluid surface, *J. Stat. Phys. 15*: 423.
69. Henderson, D., Blum, L., and Smith, W. R. (1979). Application of the hypernetted chain approximation to the electric double layer at a charged planar interface, *Chem. Phys. Lett. 63*: 381.
70. Lozada-Cassou, M., Saavedra-Barrera, R., and Henderson, D. (1982). The application of the hypernetted chain approximation to the electrical double layer: comparison with Monte Carlo results for symmetric salts, *J. Chem. Phys. 77*: 5150.
71. Lozada-Cassou, M., and Henderson, D. (1983). Application of the hypernetted chain approximation to the electrical double layer: comparison with Monte Carlo results for 2:1 and 1:2 salts, *J. Phys. Chem. 87*: 2821.
72. Carnie, S. L., and Chan, D. Y. C. (1984). Correlations in inhomogeneous Coulomb systems, *Mol. Phys. 51*: 1047.
73. Carnie, S. L., and Chan, D. Y. C. (1980). The structure of electrolytes at charged surfaces: ion–dipole mixtures, *J. Chem. Phys. 73*: 2949.
74. Carnie, S. L., and Torrie, G. M. (1984). The statistical mechanics of the electric double layer, *Adv. Chem. Phys. 56*: 141.
75. Rosenfeld, Y., and Ashcroft, N. W. (1979). Theory of simple classical fluids: universality in the short range structure, *Phys. Rev. A 20*: 1208.
76. Ballone, P., Pastore, G., and Tosi, M. P. (1986). Restricted primitive model for electrical double layers: modified HNC theory of density profiles and Monte Carlo study of differential capacitance, *J. Chem. Phys. 81*: 2943.

77. Rosenfeld, Y., and Blum, L. (1986). Fluids in contact with a hard surface: universality of the bridge functions for the density profile, *J. Chem. Phys. 85*: 2197.
78. Marconi, U. M. B., Wichen, J., and Forstmann, F. (1984). The structure of size symmetric electrolytes at charged surfaces: the restricted primitive model in the HNC/MSA approximation, *Chem. Phys. Lett. 107*: 609.
79. Nielaba, P., and Forstmann, F. (1985). Packing of ions near an electrolyte–electrode interface in HNC/LMSA approximation to the rpm model, *Chem. Phys. Lett. 117*: 46.
80. Alts, T., Nielaba, P., D'Aguanno, B., and Forstmann, F. (1987). A local density functional approximation for the ion distribution near a charged electrode in the restricted primitive model electrolyte, *Chem. Phys. 111*: 223.
81. Nielaba, P., Alts, T., D'Aguanno, B., and Forstmann, F. (1986). The local hypernetted chain approximation for the 2:1 RPM electrolyte near a charged wall, *Phys. Rev. A 34*: 1505.
82. D'Aguanno, B., Nielaba, P., Alts, T., and Forstmann, F. (1986). A local HNC/HNC approximation for the 2:2 restricted primitive model electrolyte near a charged wall, *J. Chem. Phys. 85*: 3476.
83. Mier y Teran, L., Suh, S. H., White, H. S., and Davis, H. T. (1990). A nonlocal free energy density functional approximation for the electrical double layer, *Chem. Phys. 92*: 5087.
84. Born, M., and Green, H. S. (1938). *Proc. R. Soc. London 188*: 10.
85. Yvon, J. (1935). La théorie statistique des fluides et l'equation d'etat, in *Actualités scientifiques et industrielles*, No. 203, Hermann, Paris.
86. Croxton, T. L., and McQuarrie, D. A. (1981). The electrical double layer in the Born–Green–Yvon equation, *Mol. Phys. 42*: 141.
87. Caccamo, C., Pizzimenti, G., and Blum, L. (1986). An improved closure for the Born–Green–Yvon equation for the electric double layer, *J. Chem. Phys. 84*: 3327.
88. Blum, L., Hernando, J., and Lebowitz, J. L. (1983). A numerical method and general discussion of integral equations for the primitive model of the electric double layer, *J. Phys. Chem. 87*: 2825.
89. Plischke, M., and Henderson, D. (1989). The primitive model of the electric double layer: nonsymmetric electrolytes, *J. Chem. Phys. 90*: 5738.
90. Henderson, D., and Plischke, M. (1988). Pair and singlet correlation functions of inhomogeneous fluids calculated using the Ornstein–Zernike equation, *J. Phys. Chem. 92*: 7177.
91. Colmenares, P. J., and Olivares, W. (1986). A noniterative solution of the WLMB integral equation for the electric double layer, *J. Chem. Phys. 90*: 1977.
92. Colmenares, P. J., and Olivares, W. (1988). Numerical solution of two new integral equations for the electric double layer, *J. Chem. Phys. 88*: 3221.
93. Kjellander, R., and Marcelja, S. (1984). Correlation and image effects in electric double layers, *Chem. Phys. Lett. 112*: 49.

94. Kjellander, R., and Marcelja, S. (1985). Inhomogeneous Coulomb fluids with image interactions between planar surfaces, I., *J. Chem. Phys. 82*: 2122.
95. Kjellander, R., and Marcelja, S. (1985). Inhomogeneous Coulomb fluids with image interactions between planar surfaces II on the anisotropic hypernetted chain approximation, *J. Chem. Phys. 82*: 2122.
96. Kjellander, R., and Marcelja, S. (1986). Interaction of charged surfaces in electrolyte solutions, *Chem. Phys. Lett. 127*: 402.
97. Blum, L., and Henderson, D. (1981). Mixtures of hard ions and dipoles against a charged hard wall: Ornstein–Zernike equation, some exact results and the mean spherical approximation, *J. Chem. Phys. 74*: 1902.
98. Levine, S., Outhwaite, C. W., and Bhuiyan, L. B. (1981). *J. Electroanal. Chem. 123*: 105.
99. Outhwaite, C. W., and Bhuiyan, L. B. (1983). An improved modified Poisson–Boltzmann in electric-double-layer theory, *J. Chem. Soc. Faraday Trans. 2 79*: 707.
100. Kirkwood, J. G. (1934). *J. Chem. Phys. 2*: 767.
101. Rosenfeld, Y., and Blum, L. (1986). Statistical mechanics of charged objects: general method and application to simple systems, *J. Chem. Phys. 85*: 1556.
102. Wertheim, M. S. (1971). Exact solution of the mean spherical model for fluids of hard spheres with permanent electric dipole moments, *J. Chem. Phys. 55*: 4291.
103. Parsons, R., and Zobel, F. G. R. (1965). *J. Electroanal. Chem. 9*: 333.
104. Melroy, O. R., Toney, M. F., Borges, G. L., Samant, M. G., Kortright, J. B., Ross, P. N., and Blum, L. (1988). Two-dimensional compressibility and growth of electrochemically adsorbed lead on silver (111), *Phys. Rev. B 38*: 10962.
105. Lundquist, S., and March, N. H. (1983). *Theory of the Inhomogeneous Electron Gas*, Plenum Press, New York.
106. Lang, N. D., and Kohn, W. (1970). Theory of metal surfaces: charged density and surface energy, *Phys. Rev. B 1*: 4555.
107. Torrie, G. M., and Patey, G. N. (1991). Molecular solvent models of electrical double layers, *Electrochimica Acta 36*: 1677.
108. Israelachvili, J. N. (1985). *Chem. Scr. 25*: 7.
109. Holovko, M. F., Trokhynchuk, A. D., Protsykhevich, P., and Henderson, D. (1991). The Laplace transform of the MSA-pair distribution functions between macroions in the ion–dipole fluid, USI Report No. 18, University of Utah.
110. Lozada Cassou, M., and Henderson, D. (1986). A simple theory for the force between spheres immersed in a fluid, *J. Colloid Interface Sci. 114*: 180.
111. Bratko, D., Blum, L., and Luzar, A. (1985). A simple model for the intermolecular potential of water, *J. Chem. Phys. 83*: 6367.
112. Cummings, P. T., and Blum, L. (1986). Analytic solution of the molecular

Ornstein–Zernike equation for nonspherical molecules: spheres with anisotropic surface adhesion, *J. Chem. Phys. 84*: 1833.
113. Blum, L., Cummings, P. T., and Bratko, D. (1990). A general solution of the molecular Ornstein–Zernike equation for spheres with anisotropic surface adhesion, *J. Chem. Phys. 92*: 3741.
114. Luzar, A., Bratko, D., and Blum, L. (1986). Monte Carlo simulation of hydrophobic interaction, *J. Chem. Phys. 86*: 2995.

7

Wetting Experiments

Carl Franck
Cornell University
Ithaca, New York

I. INTRODUCTION

The observation of wetting phenomena is something which for many of us started early in our lives. Everyday experiences include the capillary action occurring when a paper towel wipes up a spill and the behavior of water on clean versus greasy surfaces. In this chapter I review some of the recent work (through 1990) on experimental wetting both in and out of thermal equilibrium [1]. Our discussion will be divided up into the answers to four questions: (1) What are the intermolecular forces that produce wetting layers? (2) What are the different thermodynamic states of wetting, and what is the nature of the transitions between them? (3) What are the effects of thermal fluctuations on wetting layers? (4) How can wetting systems be driven out of thermal equilibrium?

II. HOW DO WETTING LAYERS REVEAL INTERMOLECULAR FORCES?

By a wetting layer, we mean a region of fluid at a substrate/fluid interface which is perturbed by the substrate from its bulk structure. We will label the thickness of the wetting layer by l. In Fig. 1 we picture such a system. We call the bulk fluid adjacent to the wetting layer the spectator phase.

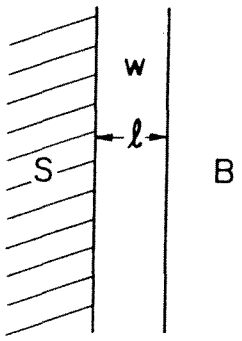


Fig. 1 Schematic diagram of a boundary fluid. S, substrate; W, wetting layer; B, bulk (spectator) phase; l, thickness of wetting layer.

There is great variety in the possible ways a wetting arrangement may be realized. If the bulk fluid is a one-component vapor, the wetting layer could be a liquid or a solid form of the same material. If the bulk system is a homogeneous binary liquid mixture, the wetting layer could be a phase with a different composition from that of the bulk.

We seek to learn the nature of the intermolecular forces that are responsible for establishing the wetting layer. The means by which this will be accomplished is to arrange for competition between these forces and an external control parameter that is set to be unfavorable to formation of the wetting layer. The idea is to vary the strength of the external parameter and observe the thickness of the wetting layer. In this section we consider only cases where the substrate-fluid interaction is always favorable toward increasing the thickness of the wetting layer. As a first example of such a measurement, we examine the noble gas adsorption experiments of Krim et al. [2]. In one of their experiments, the substrate was a gold (111) surface and the bulk fluid was argon. The fluid was kept at a fixed temperature. The surface was found to be loaded with liquid argon, but since the pressure was below that of liquid–vapor coexistence (i.e., the vapor is undersaturated), the amount of adsorbed liquid on the substrate was limited by the deviation of the pressure from the pressure at coexistence. Because the system temperature was kept above the triple temperature, we expect that the wetting and spectator phases are both fluids. The amount of adsorbed material was measured by vibrating the substrate with an attached quartz mechanical oscillator. The shift in frequency turns out to be proportional to the change in the adsorbed mass. The sensitivity of this technique is so great that much less than a monolayer of adsorption can readily be detected. Their measurements of fre-

quency shift versus gas pressure yields the result seen in Fig. 2. The power-law growth in the amount adsorbed with the deviation of the pressure from the saturation pressure was found to have an exponent in agreement with the expectation that the intermolecular force responsible for the formation of the wetting layer was a nonretarded dispersion force of attraction between the argon atoms and the substrate. The intermediate steps needed to make this assertion are worked out in Ref. 3.

As a second example of how the forces responsible for wetting have been established in a one-component fluid, we consider the experiment of Kayser et al. for sulfur hexafluoride at bulk liquid–vapor coexistence [4]. In this experiment, with wetting layers that were considerably thicker (tens of nanometers), they tested the validity of the Dzyalosinskii, Lifshitz, and Pitaevskii (DLP) theory of dispersion forces [5]. The competition was between gravity and the substrate–molecule attraction. To arrange for a range of the gravity control parameter, the substrate surface was vertical. Thus the experimenters could scan the wetting layer as func-

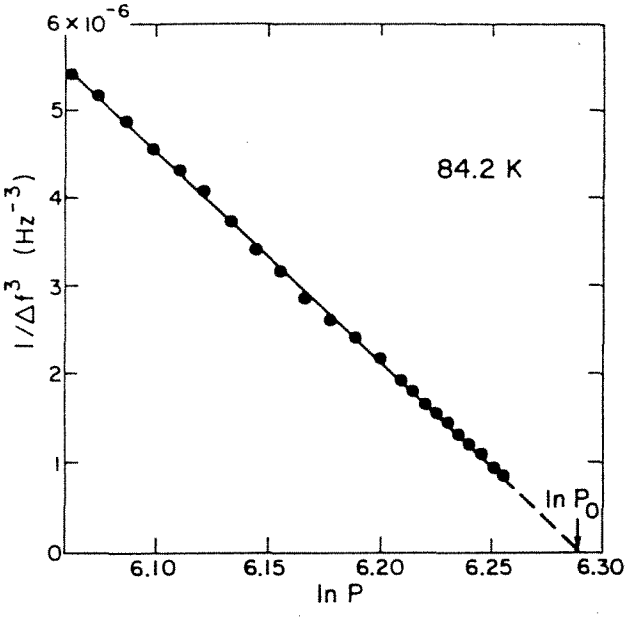


Fig. 2 Adsorption of argon gas by a favorable substrate for wetting. The oscillator frequency shift, Δf, is proportional to the amount of adsorption, and P is the vapor pressure (P_0 corresponds to saturation). The system temperature is fixed above the triple temperature. (From Ref. 2.)

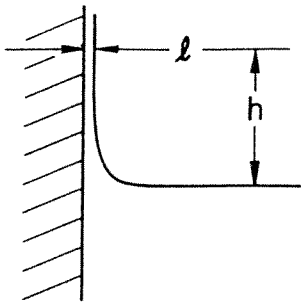


Fig. 3 Gravity-controlled wetting experiment. Vapor is on top, liquid is on the bottom. See the text for symbol definitions.

tion of height (see Fig. 3) to look for variation in the thickness of the wetting layer. This was elegantly described as follows: $F(l) + \Delta\rho gh = 0$, where l is the thickness of the wetting layer, h the height of the point of measurement above the bulk liquid–vapor interface, $\Delta\rho$ the mass density difference between the liquid and the vapor, g the acceleration of gravity, and F a function supplied by the DLP theory. The wetting layer thickness was measured through the polarization properties of the light reflected (i.e., ellipsometry) from the substrate (fused silica)–liquid interface. By minimizing the temperature gradients, thermal equilibrium was attained. The substrate also served as the window for the probe light beam. Its internal strain caused the limiting feature in the measurement: extraneous optical birefringence. The agreement achieved with the DLP theory is demonstrated: extraneous optical birefringence. The agreement achieved with the DLP theory is demonstrated in Fig. 4. Vertical scans were made at two values of reduced temperatures [$t \equiv (T - T_c)/T_c$, where T_c is the temperature of the liquid–vapor critical point]. Between these two temperatures there is no significant variation in the prediction. We see that the complete DLP theory, including retardation effects, is supported. As in the earlier work of Sabisky and Anderson on adsorbed helium [6], we have a successful quantitative comparison between theory and experiment.

Turning to wetting by binary liquid mixtures, for the reasons put forth by Kayser et al. [7], it is very difficult to get adsorption measurements in thermal equilibrium when the bulk is at two-phase liquid–liquid coexistence. Among the problems are great sensitivity to thermal gradients and extremely long relaxation times. However, interesting results have been gotten readily when the bulk is in the single-phase region. As an example, we consider the experiment of Ripple et al. on the liquid mixture

carbon disulfide + nitromethane on a borosilicate glass substrate [8]. As shown in Fig. 5, the experiment was performed upon approaching liquid–liquid coexistence by changing temperature at a fixed chemical composition. As the system approached the phase separation temperature T^* [= T_c (critical mixing temperature) − 1.24 K], the amount of adsorption on nitromethane in the presence of the bulk phase, which was rich in carbon disulfide, increased. This was indicated by an increase in the reflectivity of the substrate–liquid interface, as shown in Fig. 6. The theory used was based on the assertion by Kayser [9] that ionic forces should be considered as part of the substrate–liquid interaction. In fact, they were found to dominate the dispersion forces completely in this system, despite the apparent nonionic character of the materials involved. Besides providing a good account of the observations, the ionic theory was further

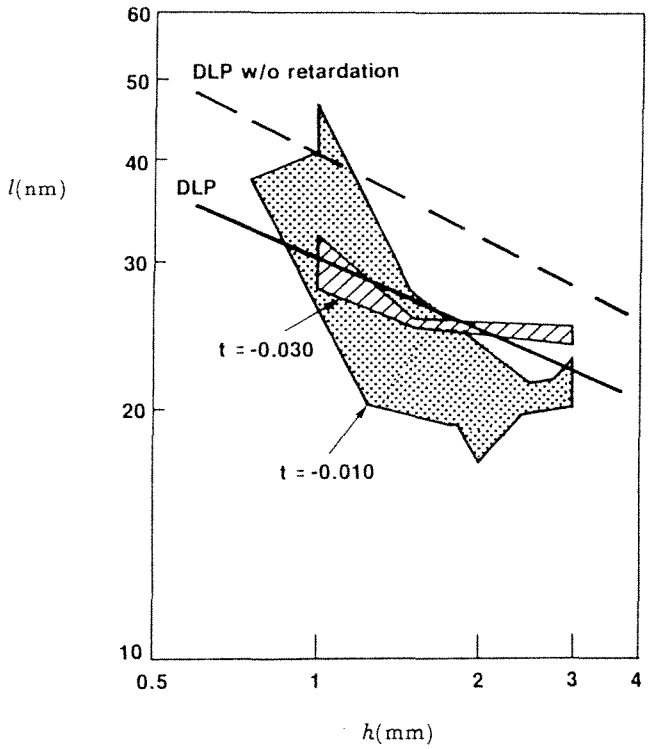


Fig. 4 Wetting layer thickness l versus vertical height h. The sample is sulfur hexafluoride. Two values of reduced temperature (t; see the text) are shown. (From Ref. 4.)

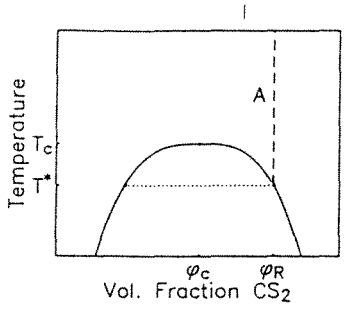


Fig. 5 Thermodynamic trajectory (marked "A") for the study of wetting by a binary liquid mixture. The system is a mixture of carbon disulfide + nitromethane at a fixed composition (ϕ_R) which is rich in the nonwetting material. The phase separation temperature is T^* and the critical mixing point is located at composition ϕ_c and temperature T_c. (From Ref. 8.)

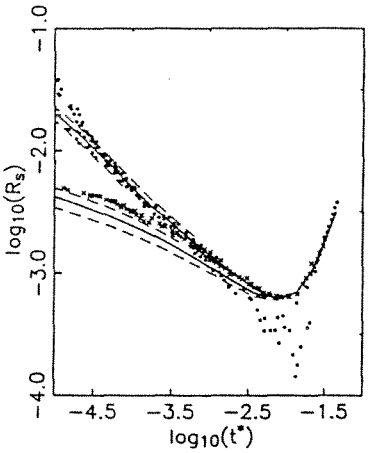


Fig. 6 Reflectivity R_s of a liquid–glass interface versus reduced temperature with respect to the bulk transition temperature [$t^* = (T - T^*)/T^*$]. The liquid is nitromethane + carbon disulfide in the single-phase region. Dots correspond to an untreated system. Crosses correspond to a sample with a soluble salt (tetrabutyl ammonium iodide) capable of screening electrostatic substrate–liquid interactions. The upper solid curve indicates the theory for the untreated sample. The lower solid curve is the theory for the treated sample. Dashed lines show the extent of uncertainty in the theory. The increase at large t^* is due to thermal expansion. (From Ref. 8.)

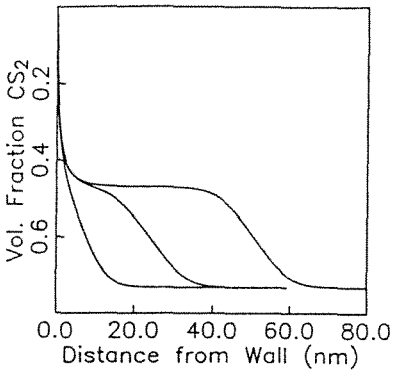


Fig. 7 Model structure of the wetting layer as coexistence is approached in the experiment of Fig. 6. Reading left to right: $t^* = 3.71 \times 10^{-3}$, 1.94×10^{-4}, and 2.56×10^{-5}. (From Ref. 8.)

confirmed by the success at explaining the large decrease in adsorption produced by the addition of a soluble salt to the system. The salt had the effect of "shorting out" the ionic force. The remaining adsorption was found to be due to an effective force arising from bulk critical fluctuations, as discussed below. Finally, the theory used here shows in Fig. 7 how the wetting layer becomes more distinct as coexistence is approached.

To summarize, by arranging competition between substrate–interaction forces favorable to the formation of a wetting and external thermodynamic forces, equilibrium studies of the degree of adsorption are being used to discover the nature of substrate–liquid interactions.

III. DISCOVERY OF INCOMPLETE WETTING AND WETTING TRANSITIONS

For many scientists the recent excitement in the subject of wetting began with the insight of Cahn [10] that in the wetting situation at bulk two-phase coexistence described by Fig. 1, there could be situations in which the thickness of the wetting layer was not limited simply by external forces, but rather by the thermodynamics of the liquid–surface interaction. To use different language, which will be especially useful in dynamical discussions below, one says that the liquid–liquid interface that defines the boundary between the wetting layer and the bulk liquid is "pinned" at the liquid–substrate interface [11]. Furthermore, Cahn argued that if this "incomplete wetting" situation were to occur then closer to the bulk critical point and still at bulk coexistence, the system would

inevitably switch to a state in which the thickness of the wetting phase was limited only by external forces (e.g., gravity). The latter state is called "complete wetting." It was the "approach to complete wetting" that we in fact studied in all the examples in Section II. In this section we examine experimental results that have established the existence of states of incomplete wetting, and the "Cahn" transitions between incomplete and complete wetting, all in thermal equilibrium.

For an example of a wetting transition in a one-component system, let us return to the noble gas adsorption experiments of Krim et al. [2] (see Fig. 8). This time, by going below the triple temperature they considered the wetting of the gold substrate by krypton solid, with a bulk vapor phase.

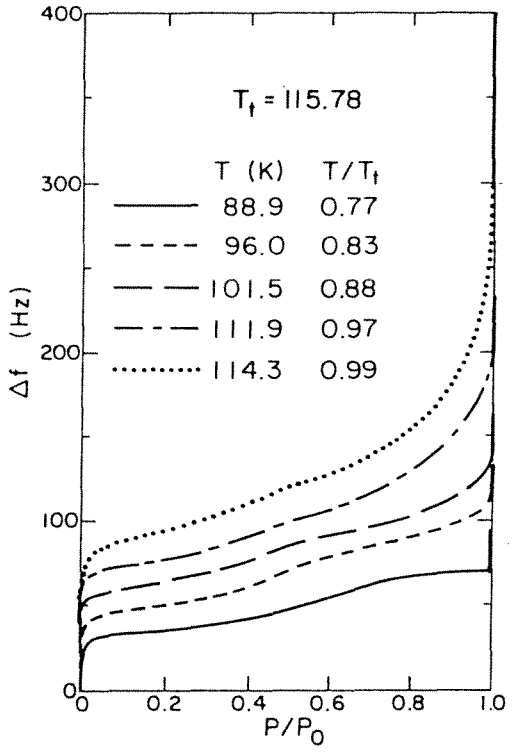


Fig. 8 Adsorption of krypton at constant temperature as measured by frequency shift of oscillating substrate, as a function of the bulk gas pressure. P_0 is the saturation pressure. Below the triple temperature, T_t, the system is incompletely wet. At or above the triple temperature, the system is completely wet. (From Ref. 2.)

For temperatures below the triple temperature, the adsorption was finite as the system approached the saturation pressure. However, this finite value was ever increasing as the temperature reached toward the triple temperature, above which the adsorption at saturation pressure diverged (not shown). Thus we see that there is a Cahn transition between incomplete and complete wetting at the point at which the wetting phase melts.

The study of wetting by *binary* liquid mixtures has yielded wetting transitions that, by contrast, are not associated with bulk phase transitions. The first example was the study of the wetting of the liquid–vapor interface by mixtures of cyclohexane + methanol at liquid–liquid coexistence (as well as liquid–vapor coexistence) by Moldover and Cahn [12]. In this experiment the methanol-rich phase was the wetting phase and the cyclohexane-rich phase was the bulk spectator phase. By measuring the dihedral angle, θ, called the contact angle, formed by the wedge of the wetting phase as it slips between the substrate (vapor phase) and the spectator phase, they assessed the state of wetting. As shown in Ref. 13, the connection is as follows:

$$\cos \theta = \frac{\sigma_S - \sigma_W}{\sigma} \tag{1}$$

where σ_W, σ_S, and σ are the interfacial free energies among the following: the wetting phase in the bulk and the substrate; the spectator phase in the bulk and the substrate; and between the two bulk liquid phases (wetting and spectator phase), respectively. The crucial result contained in this formula is that as long as the vapor is incompletely wet by the methanol-rich phase, $\cos \theta < 1$, but as the preference of the vapor for the wetting phase increases, θ decreases, until it saturates at 0 and the system becomes and remains completely wet. Moldover and Cahn tuned the system through such a wetting transition by reducing the amount of water added to the mixture at a fixed temperature (22°C), giving the result shown in Fig. 9.

It is interesting to compare the Moldover–Cahn observations with the earlier work of Zisman in which he developed a scheme to characterize the energy of surfaces [14]. Zisman measured the advancing (in the sense that the wetting layer wedge was moving toward the direction where it is thinnest*) contact angle of liquids in air and plotted it against the liquid–air surface tension, all for a fixed temperature. In this manner he noticed a straight-line relationship yielding a wetting transition from incomplete to complete at what he called the critical surface tension, his measure of the surface energy of the substrate.

* Note that this is potentially a non-thermal-equilibrium phenomenon.

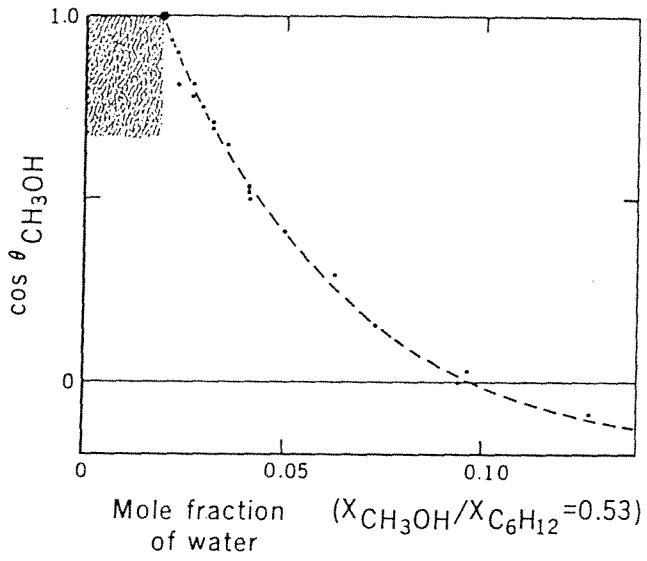


Fig. 9 Contact angle, θ, for methanol + cyclohexane at liquid–liquid coexistence at the vapor/liquid interface versus added water concentration. (From Ref. 12; copyright 1980 by the AAAS.)

Much of the early theoretical work following Cahn focused attention on critical wetting transitions [15]. Not only was a new universality class discovered, but it had the intriguing feature that the critical exponents depended on a nonuniversal parameter. Finally, the transition's upper critical dimensionality was found to be three. It was expected that the corrections to mean field behavior would be unusually strong. By contrast, if a wetting transition is first order, then by comparison with the case of critical wetting, the effects of thermal fluctuations are insignificant and the transition should have mean field behavior [11]. It is thus very important to be certain of the order of wetting transitions. It can in fact be seen that the Moldover–Cahn transition of Fig. 9 is first order by virtue of the discontinuity of the slope [16] in cos θ versus water concentration (we follow Moldover and Cahn and take the shaded region to correspond to cos θ = 1).

Schmidt and Moldover [17] pursued the question of the order of wetting in an experiment that provided the first direct look at the thickness of the wetting layer undergoing a Cahn transition. They performed an ellipsometric measurement of the interface between a mixture of perfluoromethylcyclohexane + isopropanol at liquid–liquid coexistence and its

vapor. The results of measurements of the thickness of the wetting layer as a function of temperature while the bulk liquid was at liquid–liquid coexistence for two different vertical heights of the spectator phase are shown in Fig. 10. The increase in wetting layer thickness with decreasing vertical height of the spectator phase is significant since it is evidence for gravity thinning of the wetting layer. As pointed out by Kayser et al. in Ref. 7, this is no mean feat.

An alternative experimental attack on the Cahn transition in a binary

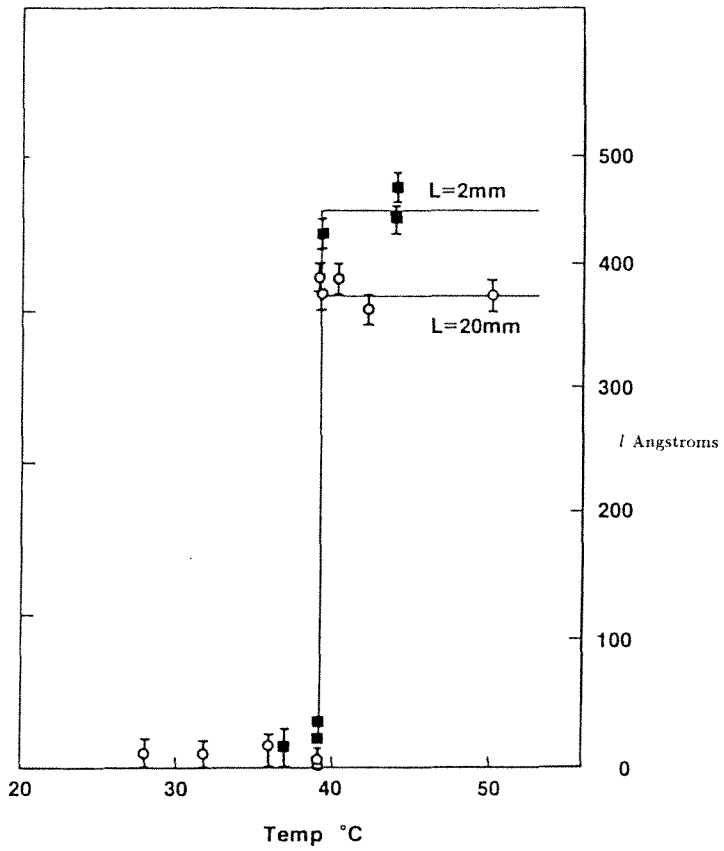


Fig. 10 Thickness of wetting layer of isopropanol-rich phase at the liquid/vapor interface of the isopropanol + perfluoromethylcyclohexane binary liquid at liquid–liquid coexistence versus temperature. The jump at ≈ 39°C is a Cahn transition. L is the vertical height of the spectator phase. The greater L is, the greater the cost in gravitational energy to form the wetting layer. (From Ref. 17.)

liquid mixture at liquid–liquid coexistence was pioneered by Pohl and Goldburg [18] using the method of capillary rise. The idea is to turn the substrate of interest into a hollow cylinder, then immerse it completely in the binary liquid mixture under study, with the vertically oriented tube passing through the free liquid–liquid interface. As soon as you say the words "sap rises up a tree," one can find from Ref. 19 the difference in the height of the liquid–liquid interface inside the tube and far from it to be given by

$$h = \frac{2\sigma \cos \theta}{\Delta \rho g R} \tag{2}$$

where $\Delta\rho$ is the mass density between the lower and the upper liquid phases and R is the inner radius of the tube. This result is valid only for $h \gg R$. Pohl and Goldburg compared the rise in the system 2,6-lutidine + water on borosilicate glass with what the rise would have been had the substrate been completely wet. They accomplished this by measuring the shape of the liquid–liquid interface inside the tube. This provided a direct measurement of $\sigma/(\Delta\rho g)$, which could then, by substitution into Eq. (2), give the rise expected from a completely wet system. Note that a completely wet system has the rise with the greatest possible absolute magnitude for a given temperature. Pohl and Goldburg found that closer than about 15 K to the critical mixing temperate, the system switched to complete wetting, being incompletely wet at further distances in temperature. It is worthwhile to note the apparent extreme sensitivity in the state of wetting to the nature of the substrate. While Pohl and Goldburg found that the wetting layer was composed of the water-rich phase, other researchers [20,21] have found wetting by the lutidine-rich phase on other silica-based systems.

A next important step in capillary rise work was the realization by Sigl and Fenzl [22] of a test of the connection between the Cahn transition and the critical mixing transition. They used the same system that was studied by Pohl and Goldburg with the addition of a salt. The idea was to exploit the following scaling hypothesis [3,23]:

$$\cos \theta = F(h_1 t^{-\Delta_1}) \tag{3}$$

The assumption is that there is a short-range interaction between substrate and liquid of strength h_1 which is responsible for the attraction of the wetting material to the substrate (the letter h is used to recall the equivalent Ising model magnetic problem). We further assert that the scaling function behaves in the manner expected for small values of its argument [23], $F(x) \sim x$ up to the transition point, beyond which $F(x) = 1$, since the system is in the completely wet state. Using this reasoning, Sigl and

Fenzl discovered a value for the exponent Δ_1 that was consistent with renormalization group calculations. They had found that by adding potassium chloride to the 2,6-lutidine + water system, the wetting transition temperature could be brought to the critical temperature. This provided as wide as possible a temperature range to test the scaling prediction.

Another test of the above "short-range" scaling hypothesis was provided in subsequent capillary rise experiments by Durian and Franck [24]. In this work the wetting transition of carbon disulfide + nitromethane on borosilicate glass was studied over a wide range of wetting transitions produced by gradually altering the surface chemistry of the substrate. The scheme was to change the degree of short-range attraction for the surface by replacing with nonattractive groups, surface sites attractive to nitromethane. The chemical procedures followed were originally used to control the adsorption properties of surfaces used in chromatography. The capillary rise results are shown in Fig. 11. The interpretation was as follows: The chemically unaltered surface gives a reference system that is

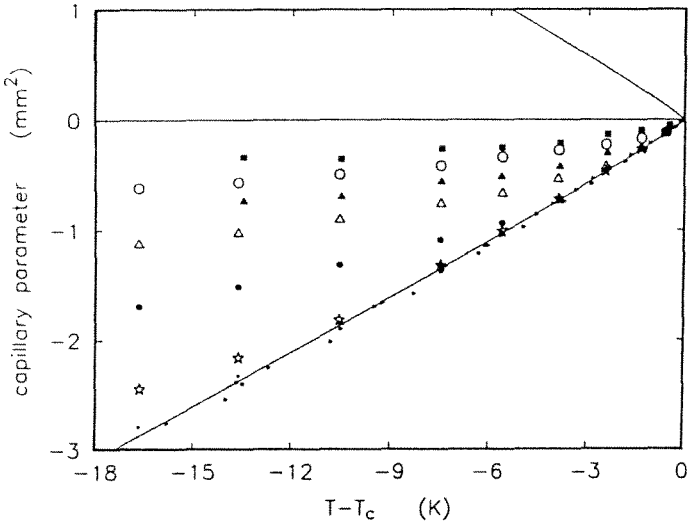


Fig. 11 Product of the capillary rise and the inner capillary radius versus distance in temperature from the critical point for carbon disulfide + nitromethane on borosilicate glass. The small symbols correspond to chemically unaltered surface. This is taken to indicate complete wetting at all temperatures. Each type of large symbol corresponds to a surface dosed with hexamethyldisilazane by a different amount in order to reduce the attraction of the nitromethane. The joining of the rise versus temperature plots into the complete wetting reference curve are the signatures of Cahn transitions. (From Ref. 24.)

completely wet at all the temperatures measured. If we fix the temperature, we see that chemical alteration of the substrate over (probably) less than a monolayer pushes the system into an incompletely wet state (less than "complete" capillary depression). However, as the temperature is raised toward the critical temperature, the rise observations for each particular preparation shown return to the completely wet reference state, in a second Cahn transition. To check the scaling hypothesis, Eq. (3), a replacement had to be found for the unmeasured quantity, h_1. This was accomplished by recognizing that the Cahn transition itself must occur at a universal value of the argument of F. That is, $h_1 t_W^{-\Delta_1}$ is a constant, where $t_W = (T_c - T_W)/T_c$ is the reduced wetting transition temperature, T_W being the wetting transition temperature itself. This relationship allows one to find h_1 by knowing T_W. Figure 12 shows the scaling plot resulting from this philosophy. The collapse of the data (on a straight line) is in agreement with the scaling hypothesis. With the additional assumption that $F(x) \sim x$, the slope of the line gives the scaling exponent Δ_1 in good agreement with the Sigl and Fenzl experiment and renormalization group theory. Note that the abrupt change in slope of F at the transition point indicates that the transitions are clearly first order since they have a discontinuity in the slope of cos θ as a function of a thermodynamic control parameter.

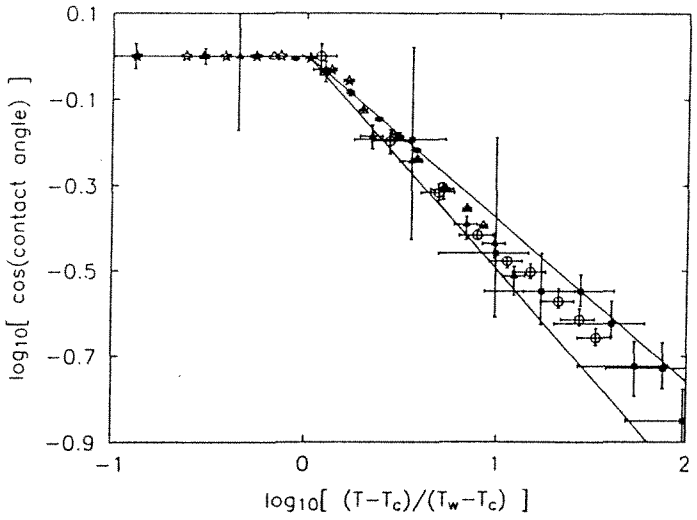


Fig. 12 Scaling plot of the same data as in Fig. 11. T_W is the wetting transition temperature. (From Ref. 24.)

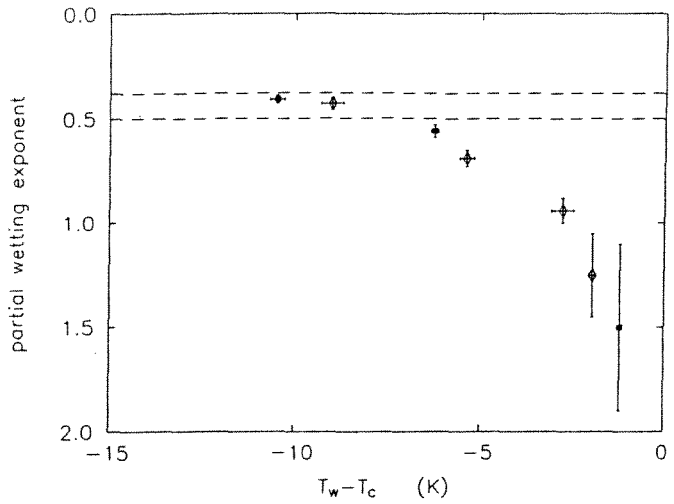


Fig. 13 Exponents extracted for acetic anhydride + cyclohexane on a chemically altered glass as a function of the difference between the wetting transition temperature T_W and the critical temperature. The value of the universal exponent (Δ_1), indicated with its uncertainty by the dashed lines, is measured by the nitromethane + carbon disulfide experiments. Note that this universal behavior occurs in acetic anhydride + cyclohexane when the wetting transition temperature is the furthest from the critical temperature. (From Ref. 24.)

Durian and Franck got a surprise when the examined a second binary liquid mixture, acetic anhydride + cyclohexane, on similar substrates. Not only did highly modified surfaces show a crossover from wetting by one phase to wetting by the other, but in the final approach to the transition, the scaling hypothesis failed in a very suggestive manner. They fit the results for the incompletely wet state to a power law as it was about to undergo a transition. The exponent thus extracted is plotted in Fig. 13 as a function of the wetting transition temperature. They found the intriguing result that the universal exponent is not reached when the wetting transition is close to the critical point, but rather, appears to be obeyed when the wetting transition is *furthest* from the critical point. These observations are as yet unexplained in the literature.

Most of the early theoretical work on wetting was directed to systems with short-range interactions only. But as the experimental evidence for the importance of long-range forces in the approach to complete wetting presented in Section II was accumulating, parallel theoretical effort was made to understand what effect long-range forces had on the Cahn tran-

sition. One very clear message was given by de Gennes [25] and Nightingale and Indekeu [26]: a long-range force unfavorable to wetting was expected to eliminate the completely wet state. Possible evidence for this effect was found by Abeysuriya et al. [27], in measurements of highly altered glass surfaces with capillary rise experiments using nitromethane + carbon disulfide. These authors interpreted the alteration of the surface chemistry by the addition of methyl groups in place of hydroxyl groups, as changing the sign of the surface field h_1. They interpreted the disappearance of Cahn transitions as evidence for the long-range force effect cited above. However, airtight application of the theoretical ideas is made difficult by our lack of knowledge as to the nature of the long-range interactions for these chemically altered surfaces. In related work, Durian and Franck [28] found evidence for the remnant of a Cahn transition suppressed by a long-range force. Consistent with the prediction of Ebner and Saam [29], they found evidence of what could happen to a Cahn transition between incomplete and complete wetting when a long-range force unfavorable to wetting was introduced. It could transform into a transition between two states of incomplete wetting.

Thus far our experimental discussion of wetting transitions has been limited completely to behavior when the bulk is at two-phase coexistence between the wetting and spectator phases. In his original work, Cahn already recognized that a first-order wetting transition could not exist in isolation on the bulk coexistence line, but rather, must be attached to a line of first-order incomplete–incomplete wetting transition extending into the bulk two-phase region. To this date, this line of "prewetting" transitions has remained undetected, presumably because the line of prewetting transitions is so short and/or it hugs the bulk coexistence curve so closely. The strongest experimental effort has been that of Schmidt and Moldover, in which special care was taken to discriminate between possible prewetting transitions and the bulk phase separation transition [30].

In summary, the concept of transitions between different states of wetting has provided an exciting period of activity. Even now, the wetting phase diagram remains a very incomplete picture from the point of view of both observation and theory.

IV. THERMAL FLUCTUATIONS AND WETTING: EFFECT OF BULK CRITICALITY

The free energy of a wetting system includes contributions besides the direct intermolecular forces discussed in Section II. The effect we consider in this section is that due to thermal fluctuations. While the hope

of discovering thermal fluctuation effects in the form of critical wetting transitions and capillary waves (as presented in Ref. 11) on the interface between the spectator and wetting phases has remained unfulfilled, an exciting branch of wetting research was opened up when Fisher and de Gennes [31] considered the effects of *bulk* criticality on wetting layers. We have already met this problem, most recently in the connection established in Section III between wetting transitions and bulk critical transitions. Now we confront it head on by studying the behavior of wetting layers as the bulk phase is brought to its critical point. The Fisher–de Gennes theory was a natural extension of bulk scaling theory through introduction of the surface contact field of strength h_1 that we discussed earlier. Most of the experiments we discuss will be optical measurements along the lines of the earlier measurements of the thickness of wetting layers. We can argue that the boundary liquid will be in for a dramatic effect in its spatial extent. To see this, consider what is occurring in the bulk fluid while it approaches the critical point from the disordered side. Throughout the fluid the average value of the order parameter is zero. Assume that the boundary forces the order parameter to be nonzero at the point of contact with the bulk fluid. Now, just as the healing length for spontaneous (and transient) fluctuations in the bulk order parameter is the correlation length, ξ, we would expect the boundary disturbance to trail off into the bulk phase with the same characteristic length. A vital difference between the two problems is that in the wetting system, the fluctuation is not transient, but is sustained. As the critical point is approached, we would expect this wetting structure, called *critical adsorption*, to grow in spatial extent as ξ diverges. We will describe the structure in terms of the average order parameter, m, as a function of perpendicular distance from the substrate, z. It is worthwhile first to compare the expected structure of critical adsorption with that of complete wetting when the bulk is at two-phase coexistence between the spectator and wetting phases. In the latter case, there are distinctive regions over which the order parameter (e.g., chemical concentration in a binary liquid mixture) is unchanging. This trend is well described in Fig. 7 for the case were the temperature is closest to the coexistence temperature. By contrast, the critical adsorption profile is expected to be continuously varying over all regions of space where it is appreciably different from the bulk. Fisher and de Gennes predicted that there would be two spatial regimes in the critical adsorption profile. For z much greater than ξ, they expected $m \sim \exp(-z/\xi)$, the *exponential regime*, and for z smaller than ξ, a *power-law regime* with $m \sim z^{-\beta/\nu}$, where β and ν are the bulk critical exponents for the order parameter and correlation length as a function of reduced temperature, respectively.

In subsequent experimental activity, the critical adsorption feature was detected by Beaglehole [32], Beysens and Leibler [33], and Franck and Schnatterly [34] in binary liquid mixtures, with vapor, glass, and glass substrates, respectively. The Beaglehole and Franck–Schnatterly experiments used reflected light as the probe, in ellipsometry and reflectivity measurements, respectively. The Beysens–Leibler experiment used an interesting twist: A light wave confined to the liquid–solid interface was used to excite fluorescence radiation from one of the components of the boundary liquid. The fluorescence was then monitored. In each of these experiments, the idea was to sweep the temperature into the bulk critical regime, monitoring the optical signal. It was clear from the interpretation of the Franck–Schnatterly experiment that the feature being detected was the long-range exponential feature of the profile. Recent ellipsometry measurements by Schmidt [35] have given a close look at the adsorption at the liquid–vapor interface of methylcyclohexane + perfluoromethylcyclohexane as the critical mixing point is approached, as shown in Fig. 14. The measured optical quantity, the ellipticity, is the difference of the phase shifts of light polarized in and perpendicular to the plane including the incoming and outgoing wavevectors of light. The angle of incidence was tracking Brewster's angle. Ellipticity is particularly sensitive to interfacial structure. As explained by Beaglehole [36], this is because it directly measures the deviation of the profile from a perfect step. The

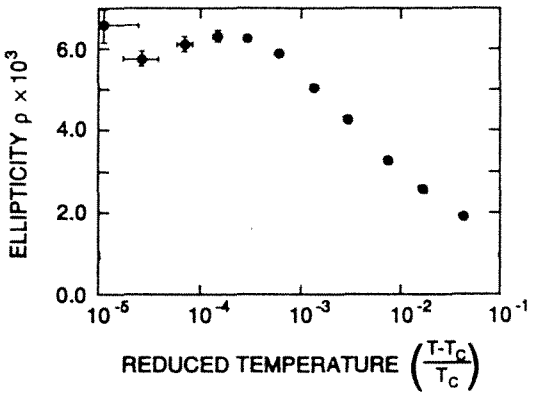


Fig. 14 Ellipsometric measurements of the adsorption of the binary liquid mixture methylcyclohexane + perfluoromethylcyclohexane approaching its bulk critical point at the vapor/liquid interface. The ellipticity (discussed in the text) is measured as a function of the reduced temperature. System is at the critical composition. (From Ref. 35.)

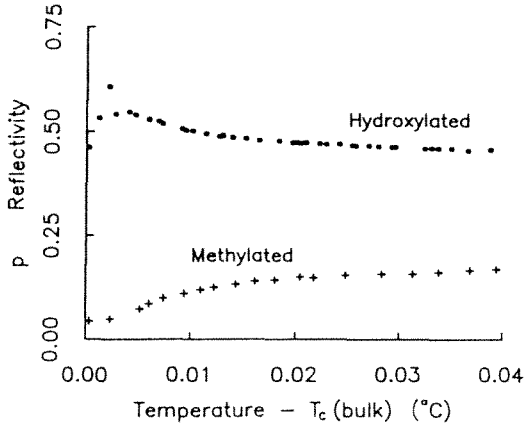


Fig. 15 Critical adsorption under control of surface chemistry. An increase (decrease) in the reflectivity indicates critical adsorption in the direction of the nitromethane (carbon disulfide) richness in the mixture as it approaches the bulk critical point. (From Ref. 39.)

interpretation of these experiments by Schmidt [35] and Liu and Fisher [37] involved a complete numerical solution to Maxwell's equations for an inhomogeneous dielectric (as was the case in the earlier wetting reflected light studies mentioned here). In the analysis of these data by Liu and Fisher [37], it was found that the preferred interpretation included the power-law regime of critical adsorption. These authors also found that their best interpretation of the experiment was consistent with a response of the fluid to the substrate that was saturated in the sense that increasing the strength of h_1 would not increase the magnitude of the critical adsorption.

Ripple et al. demonstrated [8] that away from the critical point, critical adsorption effects can successfully compete with long-range forces due to dispersion. This is demonstrated by the agreement with the expected critical adsorption signal with the measurement of the salted sample in Fig. 6. The modification of critical adsorption effects by long-range forces is in need of investigation, although early theoretical work asserted that it was not too significant [38].

In another avenue of study of critical adsorption, Dixon et al. [39] showed that control of the surface field by surface chemistry exercised in the capillary rise experiments discussed earlier could be used to reverse the sign of the order parameter in critical adsorption. Figure 15 shows the result that making a glass surface switch from hydroxyl to methyl

character was enough to change the long-range distortion of the order parameter from one component of the mixture to the other. Such control of the surface field may prove useful for future critical adsorption studies.

Besides these microscopic observations of the structure of the fluid near a boundary, a macroscopic adsorption experiment near the liquid–vapor critical point of sulfur hexafluoride absorbed on graphite by Blumel and Findenegg [40] has also tested the theory of Fisher and de Gennes. The intriguing result is that although a clear power law adsorption at a fixed density was found, it was not in a agreement with the theory. The experiment was performed ingeniously as follows: In order to keep the experimental system of fluid + absorbent at a fixed density, while the temperature was scanned, a closed reference cell of fixed volume and hence fixed density (the system was always in the bulk single-phase regime of pressure and temperature) was installed in the same thermostat. The experimental cell was adjustable in volume. The volume was set to keep the pressure in the two cells the same. Since it is possible to make a very sensitive null differential measurement of pressure, the trajectory of the system could be made to carefully follow a constant density as temperature was scanned. When the data were converted to adsorption on the substrate as a function of reduced temperature $[t = (T - T_c)/T_c]$, they found the result given in Fig. 16. The exponent for the power law expected by the authors was $\nu - \beta \approx 0.31$; instead, they found an exponent of 0.504 ± 0.005 at a density that was 1.007 times the critical density. A possible reason for the discrepancy is that the theory is correct, but the range of the experiment included the exponential regime of decay at long distances compared to the correlation length. As the authors point out, this would give rise to a larger exponent than that measured, so this does not seem likely. Another issue to be examined is the condition of the substrate. Since the graphite sample is not a single sheet, but rather, a collection of platelets, perhaps new physics takes over when the size of the adsorbed layer reaches a distance scale on the order of the distance between platelets.

To summarize, the most thoroughly investigated experimental aspect of the effect of thermal fluctuations on wetting has been the study of boundary fluids while the bulk is approaching criticality. The scaling theory of Fisher and de Gennes, which has received support from renormalization group calculations [41] and analytic calculations in two dimensions [42], has passed a recent quantitative optical test of the structure of critical adsorption but not an earlier adsorption test. As Liu and Fisher [37] point out, there is much more opportunity for experimental work on the problem.

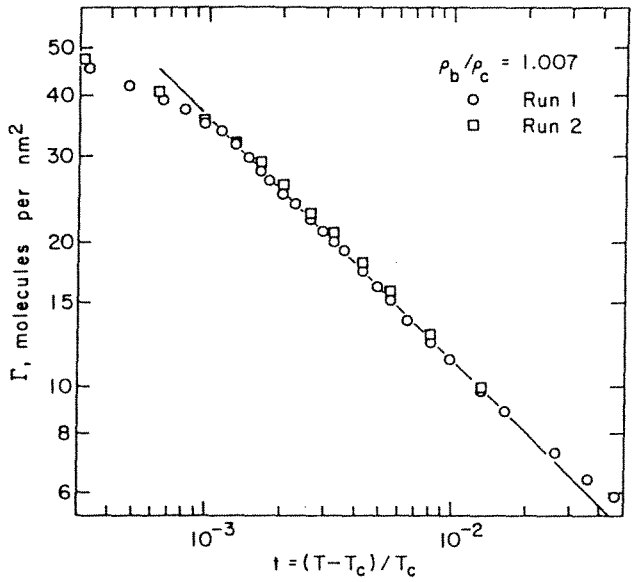


Fig. 16 Adsorption of sulfur hexafluoride on graphite versus reduced temperature. This system is close to the critical density. (From Ref. 40.)

V. WETTING OUT OF EQUILIBRIUM

Finally, we turn to nonequilibrium wetting phenomena. We are concerned first with how wetting layers of the sort we described in Section II, i.e. completely wet, can be taken out of thermal equilibrium by fluid motion and changes in temperature. Second, we examine transport experiments that demonstrate the role of wetting in controlling dissipative fluid motion.

Kayser, Moldover, and Schmidt [7] argued that motion of the bulk fluid could upset the thermal equilibrium of a wetting system at bulk coexistence because of the delicacy of the gravity-produced chemical concentration and pressure gradients that controlled the thickness of the wetting layer. Their idea was that by stirring the liquid, the effective vertical height of the spectator phase was reduced to the size of the region of diffusion-dominated (instead of convection dominated) transport close to the substrate. To test this, and better understand earlier experiments, Wu et al. [43] measured the thicknesses of the gravity-thinned wetting layers in a stirred binary liquid mixture. They found that higher stir rate gave a thicker wetting layer, as Kayser et al. [7] had predicted.

Durian and Franck [44] found another means of perturbing wetting layers from thermal equilibrium: by sudden changes in system temperature. They found that the wetting layer thickness responded with a slow but dramatic change in magnitude in response to slight temperature changes. As shown in Fig. 17, successive temperature jumps could be used to perturb and keep the wetting layer out of thermal equilibrium. The results were interpreted in terms of diffusion-limited growth, following the theory of Lipowsky and Huse [45]. The asymmetric response to temperature changes suggested that the wetting layer was being perturbed sufficiently far from its free-energy minimum so as to detect the anharmonicity in the effective potential in which the wetting layer–spectator liquid interface was trapped.

In hydrodynamic transport measurements, Stokes et al. have demonstrated the importance of wetting structures in controlling dissipation and fluctuations in fluid motion [46]. The elegant observation that a partially wet interface in the form of a contact line dragging across a substrate must at some small length scale violate the no-slip boundary condition of hydrodynamics [47] has been a major factor in attracting interest to this problem.

Recognizing the similarity of the behavior of a moving contact line and the response of a charge density wave (CDW) with respect to depinning

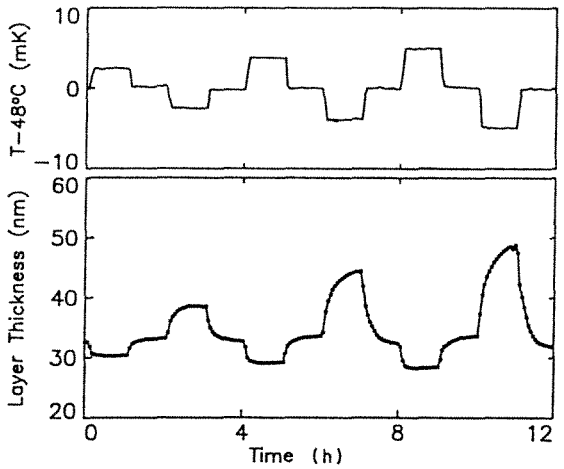


Fig. 17 Wetting layer thickness and temperature versus time for the nitromethane + carbon disulfide system at bulk coexistence wetting a glass surface. (From Ref. 44.)

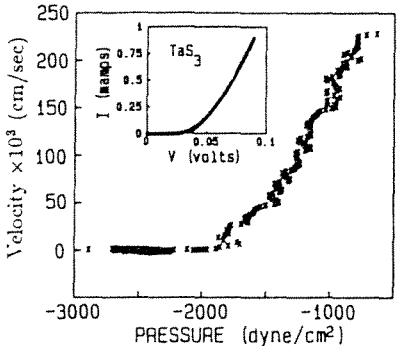


Fig. 18 Velocity of an interface between two pure liquids: decane pushing into water in a glass bead pack. The bulk dissipation has been subtracted away. We see the conducting as well as pinned behavior. Inset is an analogous current–voltage relation for a charge density wave conductor. (From Ref. 46.)

and hysteretic friction, they carried out an innovative series of wetting transport measurements patterned after the electrical measurements of the CDW systems. We will examine two of their experiments: the study of the depinning phenomena in liquid–liquid interfaces moving through random media and the generation of broadband noise by dc biasing.

The first experiment was performed with a substrate consisting of a pack of 0.5-mm glass beads. A nonwetting fluid, decane, was pushed into a wetting fluid, water, in which the beads were initially immersed. Figure 18 shows the remarkable similarity in the velocity versus pressure curve with the current–voltage relation for a charge density wave. In both cases we witness a depinning phenomenon. In the fluid experiment, the dissipation due to ordinary Stokes dissipation of the liquid far from the interface has been discounted.

In the second Stokes, Kushnick, and Robbins experiment, they "listened" to the broadband noise generated in the same fluid system when the system was depinned and set to run at a fixed velocity. Again in perfect analogy with the generation of voltage noise in a dc current-biased charge density wave conductor, the system exhibits pressure fluctuations with time. Figure 19 shows this observation.

In summary, nonequilibrium wetting experiments on binary liquids at liquid–liquid coexistence are exploiting the extreme sensitivity of these systems to driving forces. Also, a new class of fluid experiments has appeared in the form of transport measurements.

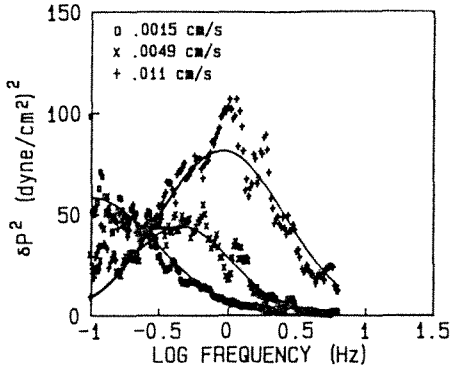


Fig. 19 Noise power spectrum for pressure fluctuation in the same system as in Fig. 18. The system is set to have constant liquid velocity. We see results for different velocities. (From Ref. 46.)

VI. FUTURE EXPERIMENTAL WETTING RESEARCH

In each of the areas of activity we have examined we have seen that major research opportunities exist. The problem of characterizing intermolecular forces with wetting is as rich as the variety of such forces. Answers to these questions are needed to help us complete the phase diagram of wetting with particular preparations. The possibility then exists that we could design our way into discovering both critical wetting transitions and prewetting (see Ref. [48]). The problem of thermal fluctuation effects in wetting is still in need of experiments to study critical adsorption. In addition, new fluctuation phenomena in wetting layer structure and in wetting transitions are still to be observed. The area of nonequilibrium studies beckons, with many systems that are easily driven out of equilibrium. The effects of wetting on transport has led to a new methodology in wetting experiments. Finally, with regard to sample preparation, wetting experimenters dealing with bulk liquids will have to invent a lot of new techniques to match the level of the ultrahigh vacuum surface science. But any single new technique, whether it be the liquid scientist's version of a pump (and filter) to purify the sample, or a new analytic probe of surface composition such as Raman spectroscopy, will give us new opportunities to make interesting wetting experiments.

ACKNOWLEDGMENTS

It is with great pleasure that I acknowledge all that I have learned about this subject from numerous colleagues. I want to thank in particular S.

Schnatterly, M. Schlossman, X.-l. Wu, J. Dixon, K. Abeysuriya, D. Durian, D. Ripple, S. Watson, T. VanVechten, P. de Gennes, B. Widom, M. E. Fisher, P. Nightingale, R. Lipowsky, S. Leibler, W. Goldburg, V. Privman, M. Schick, M. Moldover, R. Kayser, W. Saam, J. Schmidt, D. Beysens, G. Forgacs, G. Findenegg, A. Liu, M. Gelfand, N. Ashcroft, and H. Desai. I have found M. Schick's review [11] especially useful in preparing this chapter. The support of the National Science Foundation through Low Temperature Physics Grant DMR-9001106 and the Materials Science Center of Cornell University is gratefully acknowledged.

REFERENCES

1. Experimental wetting has also been reviewed recently by D. Beysens (1990), in *Liquids at Interfaces* (J. Charvolin, J. F. Joanny, and J. Zinn-Justin, eds.), North-Holland, Amsterdam, p. 499; and by D. Beaglehole (1986), in *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, Chichester, West Sussex, England, p. 523.
2. Krim, J., Dash, J. G., and Suzanne, J. (1984). *Phys. Rev. Lett. 52*: 640.
3. Pandit, R., Schick, M., and Wortis, M. (1982). *Phys. Rev. B 26*: 5112.
4. Kayser, R. F., Schmidt, J. W., and Moldover, M. R. (1985). *Phys. Rev. Lett. 54*: 707.
5. Dzyalosinskii, I. E., Lifshitz, E. M., and Pitaevskii, L. P. (1961). *Adv. Phys. 10*: 165.
6. Sabisky, E. S., and Anderson, C. H. (1973). *Phys. Rev. A 7*: 790.
7. Kayser, R. F., Moldover, M. R., and Schmidt, J. W. (1986). *J. Chem. Soc. Faraday Trans. 2 82*: 1701.
8. Ripple, D., Wu, X.-L., and Franck, C. (1988). *Phys. Rev. B 38*: 9054.
9. Kayser, R. F. (1986). *Phys. Rev. Lett. 56*: 1831. (1986). *Phys. Rev. B 34*: 3254. (1988). *J. Phys. (Paris) 49*: 149.
10. Cahn, J. W. (1977). *J. Chem. Phys. 66*: 3667.
11. Schick, M. (1990). In *Liquids at Interfaces* (J. Charvolin, J. F. Joanny, and J. Zinn-Justin, eds.), North-Holland, Amsterdam, p. 415.
12. Moldover, M. R., and Cahn, J. W. (1980). *Science 207*: no. 4435, 1073–1075.
13. Rowlinson, J. S., and Widom, B. (1982). *Molecular Theory of Capillarity*, Oxford University Press, New York, p. 211.
14. Adamson, A. W. (1982). *Physical Chemistry at Surfaces*, Wiley, New York, p. 351.
15. Brezin, E., Halperin, B. I., and Leibler, S. (1983). *Phys. Rev. Lett. 50*: 1387.
16. Teletzke, G. F., Scriven, L. E., and Davis, H. T. (1983). *J. Chem. Phys. 78*: 1431.
17. Schmidt, J. W., and Moldover, M. R. (1983). *J. Chem. Phys. 79*: 379.
18. Pohl, D. H., and Goldburg, W. I. (1982). *Phys. Rev. Lett. 48*: 1111.
19. Rowlinson, J. S., and Widom, B. (1982). *Molecular Theory of Capillarity*, Oxford University Press, New York, p. 10.
20. Gurfein, V., Beysens, D., and Parrot, F. (1989). *Phys. Rev. A 40*: 2543.

21. Wiltzius, P., Dierker, S. B., and Dennis, B. S. (1989). *Phys. Rev. Lett. 62*: 804.
22. Sigl, L., and Fenzl, W. (1986). *Phys. Rev. Lett. 57*: 2191.
23. Sullivan, D. E., and Telo da Gama, M. M. (1986). In *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, Chichester, West Sussex, England, p. 45.
24. Durian, D. J., and Franck, C. (1987). *Phys. Rev. Lett. 59*: 555. (1987). Erratum. *Phys. Rev. Lett. 59*: 1492.
25. de Gennes, P.-G. (1983). *C.R. Acad. Sci. (Paris) Ser. II 297*: 9.
26. Nightingale, M. P., and Indekeu, J. O. (1985). *Phys. Rev. B 32*: 3364.
27. Abeysuriya, K., Wu, X.-L., and Franck, C. (1987). *Phys. Rev. B 35*: 6771.
28. Durian, D. J., and Franck, C. (1987). *Phys. Rev. B 36*: 7307.
29. Ebner, C., and Saam, W. F. (1987). *Phys. Rev. Lett. 58*: 587. (1987). Phys. Rev. B 35: 1822.
30. Schmidt, J. W., and Moldover, M. R. (1986). *J. Chem. Phys. 84*: 4563.
31. Fisher, M. E., and de Gennes, P.-G. (1978). *C.R. Acad. Sci. (Paris) Ser. B 287*: 207.
32. Beaglehole, D. (1980). *J. Chem. Phys. 73*: 3366. (1981). *J. Chem. Phys. 75*: 1544.
33. Beysens, D., and Leibler, S. (1982). *J. Phys. Lett (Paris) 43*: 13.
34. Franck, C., and Schnatterly, S. E. (1982). *Phys. Rev. Lett. 48*: 763.
35. Schmidt, J. W. (1990). *Phys. Rev. A 41*: 885.
36. Beaglehole, D. (1986). In *Fluid Interfacial Phenomena* (C. A. Croxton, ed.), Wiley, Chichester, West Sussex, England, p. 523.
37. Liu, A. J., and Fisher, M. E. (1989). *Phys. Rev. A 40*: 7202.
38. Peliti, L., and Leibler, S. (1983). *J. Phys. C 16*: 2635.
39. Dixon, J. A., Schlossman, M., Wu, X.-L., and Franck, C. (1985). *Phys. Rev. B 31*: 1509.
40. Blumel, S., and Findenegg, G. H. (1985). *Phys. Rev. Lett. 54*: 447.
41. Rudnick, J., and Jasnow, D. (1982). *Phys. Rev. Lett. 49*: 1595. Brezin, E., and Leibler, S. (1983). *Phys. Rev. B 27*: 594.
42. Fisher, M. E., and Au-Yang, H. (1980). *Physica (Utrecht) 101A*: 255. Au-Yang, H., and Fisher, M. E. (1980). *Phys. Rev. B 21*: 3956.
43. Wu, X.-L., Ripple, D., and Franck, C. (1987). *Phys. Rev. A 36*: 3975.
44. Durian, D. J., and Franck, C. (1989). *Phys. Rev. A 40*: 5220.
45. Lipowsky, R. and Huse, D.A. (1986). Phys. Rev. Lett. 57: 353.
46. Stokes, J. P., Kushnick, A. P., and Robbins, M. O. (1988). *Phys. Rev. Lett. 60*: 1386.
47. *Physics Today*, May 1988, p. 17.
48. Proof note: Very recently, evidence for new wetting transitions, including prewetting, with helium have been reported, see Ketola, K.S., Wang, S. and Hallock, R.B. (1992) Phys. Rev. Lett. 68: 201 and Taborek, P. and Rutledge, J.E. (1992) Phys. Rev. Lett. 68: 2184 and references therein.

8

Fluids Between Walls and in Pores

Marcelo Lozada-Cassou
Universidad Autónoma Metropolitana–Iztapalapa
Mexico City, Mexico

I. INTRODUCTION

In this chapter a general approach to improving existing homogeneous liquid theories and extending them to the case of inhomogeneous fluids is presented briefly. Although the approach is general as to the type of fluid, in most of the chapter emphasis is given to charged liquids. As an application of this general approach, the structure of a hard-sphere fluid and an electrolyte inside a slit and a cylindrical pore is surveyed. In particular, the force between two charged plates in an electrolyte is analyzed. From the structure of a confined electrolyte its transport properties can be calculated. The transport coefficients of an electrolyte in a narrow slit are analyzed in terms of the electrolyte structure inside the slit.

To study homogeneous liquids, two main approaches have traditionally being used: hierarchy equations and the Ornstein–Zernike (OZ) equation [1–4]. In hierarchy equations such as the Born–Green–Yvon (BGY) equation, the n-particle distribution function is expressed in terms of the $(n + 1)$-particle distribution function. A shortcoming of this approach is that a superposition approximation has to be used to solve the equation. In the Ornstein–Zernike equation an approximate closure for the direct correlation function is used. From the analysis of the graphical expansion of the direct correlation function, depending on which types of clusters

are ignored, different liquid theories are obtained [1–4]. Particularly, for charged fluids the hypernetted chain (HNC) approximation compares better than others with computer simulation results [5–9]. In the HNC equation the so-called bridge diagrams are not taken into account. Efforts to include bridge diagrams have been made [10,11]. However, the graphical analysis is quite involved and the correction terms are difficult to calculate.

Another approach to the study of homogeneous liquids is the modified Poisson–Boltzmann (MPB) equation [12–14]. In every case the liquid is defined through a mathematical model (i.e., the particle interaction potential must be given). The particles can be of a single species or a mixture. All of the various approaches (to the study of classical fluids) mentioned above are quite general in relation to the type of the particle interaction potential and the number of species in the fluid. The various microscopic liquid theories are basically mathematical manipulations of the species probability densities, starting from a conservation equation or a probability density definition. These facts have been used in the past by Lozada–Cassou and others to extend, in a straightforward manner, existing homogeneous liquid theories to be applicable to a fluid in an external field of planar [5,6,15,16], cylindrical [17–19], and spherical [20] geometries. The idea behind these extensions is really very simple: Since there are almost no restrictions in the various theories as to the type of particle interaction potential, and/or the number of species, and/or their concentrations for homogeneous liquids, one can consider the external field in an inhomogeneous fluid to be just another particle. Thus an n-species inhomogeneous fluid can be taken as an $(n + 1)$-species homogeneous liquid and a homogeneous liquid theory used. The new "interparticle" potential will in general introduce geometrical factors which have to be taken into account. This procedure of deriving inhomogeneous liquid theories has been called the direct method (DM) [15,16].

The opposite is also true: One can think a particle in a homogeneous liquid as being a source of an external field. This equivalent approach has been used by Percus to study inhomogeneous fluids [21], through a "turning on" of the external field procedure, and very successfully by Henderson et al. [5,6,22–24], through an asymptotic procedure in which one of the species in the fluid becomes infinitely large while its concentration goes to zero. However, these limit procedures or asymptotic methods (AM) are in general somewhat more mathematically involved [15,16]. It is apparent from the above that from the point of view of existing liquid theories, it is to some extend artificial to talk of an inhomogeneous liquid theory. In fact, often, inhomogeneous liquid theories are simple restatements of already existing homogeneous liquid theories. This is the case

of the Stillinger–Kirkwood inhomogeneous liquid theory [25], which is equivalent to the Kirkwood–Poirier theory for homogeneous liquids [15,26], the BGY equation [27–31], the one-particle distribution function OZ equation [15,16,21–23], and as shown by Outhwaite et al., the MPB equation [32,33].

Another way of taking advantage of the freedom one has to define the interparticle interaction potential is to define artificially a species (at infinite dilution) made of clusters of two or more particles taken from the fluid. Thus an n-species homogeneous fluid will become an $(n + 1)$-species homogeneous fluid, where the new species is taken to be at infinite dilution. In this way we gain a better microscopic description of the system: That is, this procedure provides a simple way to calculate the unsymmetrical $(n + 1)$th-particle distribution function for a homogeneous fluid, when a cluster of n particles is considered. The unsymmetrical $(n + 1)$th-particle distribution function is proportional to the probability of finding the $(n + 1)$th particle at a certain position with respect to the position of the other n particles, provided that the relative positions of the first n particles are fixed. These ideas, together with those in the paragraphs above, have been used in the past to obtain integral equation theories for a fluid inside a slit [34–38], by defining in the fluid a "dumbbell particle" made of two infinite plates, and a fluid around two colloidal particles [39], by defining in the fluid a "dumbbell" made of two large spheres.

Another approach to the study of a fluid inside a slit is through the inhomogeneous two-particle OZ (OZ2) equation, where the external field is produced by the two plates [40]. If in a homogeneous fluid the OZ2 equation is applied, by assuming that one of the regular particles in the fluid is the source of the external field, the unsymmetrical three-particle distribution function can be calculated. This is a particular application of the ideas given in the paragraphs above. Attard has successfully used this approach to calculate the three-particle Percus–Yevick (PY) distribution function for a hard-sphere fluid [41].

Recently, the DM has been used to derive the inhomogeneous one-particle OZ equation for a hard-sphere fluid inside spherical [42] and cylindrical [43] pores. Depending on the type of fluid (e.g., hard spheres, Lennard–Jones fluid, electrolyte, liquid crystal), the structure of a confined fluid (e.g., in a slit or a spherical or cylindrical pore) is relevant to the study of colloidal stability [34–40,44–47], electrokinetic phenomena [48–55], and/or optical properties of some materials, as for example in liquid-crystal devices. If the confined fluid is an electrolyte, two electrolyte models have been studied extensively: the point-ion model (PIM) and the restrictive primitive model (RPM). In the point-ion model the ions are

taken as point charges, whereas in the RPM the ions are taken as charged hard spheres of diameter a. In both models the solvent is taken through a uniform dielectric constant, ϵ. Most of the discussion in this chapter is given in terms of the point-ion and restrictive primitive models.

In confined fluids the atomic or ionic size plays a relevant role. However, only recently have studies of confined RPM fluids been made. From the structure of an electrolyte inside a slit the force between the walls can be calculated. On the other hand, the application of pressure and/or electrical potential gradients across charged capillaries, in equilibrium with a reservoir containing an electrolyte solution, gives an origin to electrokinetic phenomena. The Onsager transport coefficients can be calculated from the electrolyte structure inside the slit or pore, provided that the system has sufficiently small Reynolds numbers. In this chapter we study the force between two charged walls immersed in an electrolyte and the electrokinetic properties of an electrolyte inside a charged slit, as well as the electrolyte structure inside a cylindrical pore. As a limit situation a confined hard-sphere fluid is also studied.

II. GENERAL APPROACH TO LIQUIDS STRUCTURE: HOMOGENEOUS AND INHOMOGENEOUS LIQUIDS

As pointed out in Section I, most of the existing homogeneous liquid theories can easily be extended to the study of the structure of a fluid next to an external field of arbitrary shape. As examples of the general procedure, in this section the inhomogeneous one-particle OZ equation for a RPM electrolyte next to electrical fields of differents geometries will be obtained. Emphasis will be given to the hypernetted chain/mean spherical (HNC/MS) approximation. As particular cases, the HNC/MS equations for an electrolyte between two charged plates and inside a charged cylindrical pore will be obtained. From these equations, the corresponding equations for a point-ion electrolyte or a hard-sphere fluid can be trivially obtained by making the ionic size or ionic valence, respectively, equal to zero.

A. Homogeneous Liquid

For simplicity, let us consider a three-species homogeneous liquid. The OZ equations for this fluid are

$$h_{ij}(r_{21}) = c_{ij}(r_{21}) + \sum_{m=1}^{3} \rho_m \int c_{jm}(r_{13}) h_{im}(r_{23})\, dv_3 \qquad i,j = 1,2,3 \quad (1)$$

where $h_{ij}(r)$ [$= g_{ij}(r) - 1$] is the total correlation function [$g_{ij}(r)$ is the radial distribution function], $c_{ij}(r)$ is the direct correlation function, ρ_m is the bulk concentration of species m, and dv_3 is the volume element around particle 3. $\rho_m g_{im}(r)$ gives the local concentration of particles of species m at the distance r of a particle of species i.

To be able to do anything with (1), a closure for the direct correlation function has to be given. Different nonexact closures have been proposed in the past [1–4]. Among them are

$$c(r) = \frac{f(r)g(r)}{e(r)} \tag{2}$$

$$c(r) = -\beta U(r) \tag{3}$$

$$c(r) = h(r) - \ln g(r) - \beta U(r) \tag{4}$$

where $\beta = 1/KT$, K is the Boltzmann constant, T is the temperature, $U(r)$ is the particle interaction potential,

$$e(r) \equiv e^{-\beta U(r)} \tag{5a}$$

and

$$f(r) \equiv e(r) - 1 \tag{5b}$$

Equations (2), (3), and (4) are referred to as the Percus–Yevick (PY), mean spherical (MS), and hypernetted chain (HNC) approximations, respectively. $f(r)$ is called the Mayer f-function.

If in (1) $c_{ij}(r_{21})$ is approximated through (4), and $c_{jm}(r_{13})$ is approximated through (3) or (4), the HNC/MS or HNC/HNC integral equations are obtained, respectively. Other combinations are, of course, possible. The PY/PY approximation has proved to be better for a hard-sphere fluid [1–4], and the HNC/HNC [56,57] and the HNC/MS [7] for an RPM electrolyte.

Density expansions of the direct correlation function, the radial distribution function, and the total correlation function can be made. These expansions can be expressed in terms of diagrams [3,4]. In the diagrammatic language, the various approximated integral equations theories correspond to keep different sets of diagrams out of the total of them, in the exact expansion. The density expansion for the radial distribution function of a homogeneous one-component fluid is given by

$$g(r_{21}) = \left[1 + f(r_{21})\right]\left[1 + \sum_{n=1}^{\infty} \rho^n g_n(r_{21})\right] \tag{6}$$

In Fig. 1, the diagrams (or graphs) corresponding to the first two terms in the density expansion are shown. In the HNC approximation one dia-

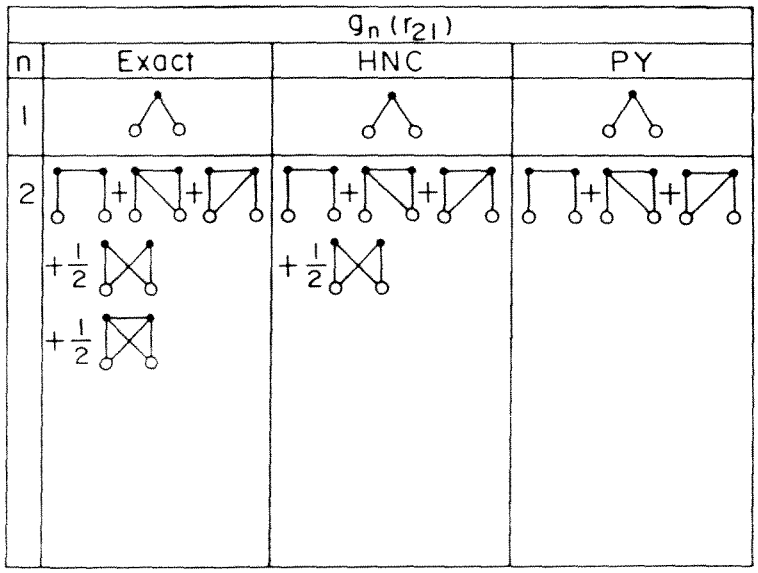


Figure 1 Diagrams involved in the first two terms of the density expansion for the radial distribution function of a homogeneous one-component fluid.

gram is missing, and in the PY approximation two diagrams are missing, with respect to the exact expansion. The diagram missing in the HNC column belongs to a class of diagrams called *bridge diagrams*, whereas in the PY column the bridge diagram and a *product diagram* are missing. At all orders in the density expansion, in the HNC approximation bridge diagrams are missing, whereas in the PY approximation in addition to the bridge diagrams, product diagrams are missing. In standard diagrammatic nomenclature [3], bridge diagrams are those h-allowed diagrams with no cutting points and with product diagrams excluded, and product diagrams are those whose value is the product of two or more simpler diagrams. Although the HNC approximation has more diagrams than the PY approximation, it is not necessarily better than the PY approximation. Particularly, as mentioned earlier, the PY approximation is better than the HNC approximation for hard spheres [1–4]. This is probably due to a cancellation of errors. Nevertheless, the HNC approximation is better if applied to charged liquids. Therefore, depending on the system studied, a liquid theory with more diagrams included is not necessarily better.

For a three-species RPM electrolyte, the HNC equations can be ob-

tained from (1) and (4): That is,

$$g_{ij}(r_{21}) = \exp\left[-\beta U_{ij}(r_{21}) + \sum_{m+1}^{3} \rho_m \int c_{jm}(r_{13}) h_{im}(r_{23})\, dv_3 \right] \qquad i,j = 1,2,3 \quad (7)$$

where

$$U_{ij}(r_{21}) = U^{*}_{ij}(r_{21}) + U^{el}_{ij}(r_{21}) \quad (8)$$

$$U^{*}_{ij}(r_{21}) = \begin{cases} \infty & r_{21} < a \\ 0 & r_{21} > a \end{cases} \quad (9)$$

$$U^{el}_{ij}(r_{21}) = \frac{e^2 z_i z_j}{e r_{21}} \qquad r_{21} > a \quad (10)$$

where e is the proton's charge and z_i and z_j are the ionic valence of species i and j, respectively. $U^{*}_{ij}(r_{21})$ is the hard-sphere part of the potential and $U^{el}_{ij}(r_{21})$ is the electrostatic part.

Since in the RPM the ions are taken as charged hard spheres, the volume element in (7) is given in spherical coordinates:

$$dv_3 = r_3^2 \sin\theta\, d\theta\, d\varphi\, dr_3 \quad (11)$$

If we conveniently choose particle 2 to be in the coordinates' origin,

$$dv_3 = r_{23}^2 \sin\theta\, d\theta\, d\varphi\, dr_{23} \quad (12)$$

and (7) can be written as

$$g_{ij}(r_{21}) = \exp\left[-\frac{\beta e^2 z_i z_j}{\epsilon r_{21}} + \sum_{m=1}^{3} \rho_m \int_a^{\infty} h_{im}(r_{23}) \zeta_{jm}(r_{21}, r_{23})\, dr_{23} \right] \qquad i,j = 1,2,3 \quad (13)$$

such that $r_{21} > a$ and

$$\zeta_{jm}(r_{21}, r_{23}) \equiv \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\varphi\, r_{23}^2 c_{jm}(r_{13}) \quad (14)$$

In bipolar coordinates, using the relation

$$r_{13}^2 = r_{21}^2 + r_{23}^2 - 2 r_{21} r_{23} \cos\theta \quad (15)$$

Equations (12) and (14) become

$$dv_3 = \frac{r_{13}r_{23}}{r_{21}}\, dr_{13}\, dr_{23}\, d\varphi \tag{16}$$

and

$$\zeta_{jm}(r_{21},r_{23}) = \frac{2\pi r_{23}}{r_{21}} \int_{|r_{21}-r_{23}|}^{r_{21}+r_{23}} c_{jm}(r_{13}) r_{13}\, dr_{13} \tag{17}$$

respectively. If the MS approximation is used in (17), an analytical expression is obtained. However, if the HNC approximation is used, (17) has to be evaluated numerically. In both cases numerical methods have to be used to solve (13) [58]. In (13) and (17), the hard-sphere potential, (9), is implicitly taken into account in the integration limits.

In the limit of point ions the HNC/MS equation reduces to the PB differential equation [59]. In fact, for point ions, (13) is the integral version of the PB differential equation.

In the MPB integrodifferential equation, the PB equation is modified to take into account the fluctuation terms due to the ionic size. In these equations, boundary conditions given by the size and charge of the central ion are needed to solve them. In (13) these conditions are implicit in the ionic electrostatic potential, first term in the exponential, and in the hard-sphere potential, lower integration limit of the second term in the exponential.

B. Inhomogeneous Fluids

In the introduction it was pointed out that since, in general, in liquid theories there are no restrictions on the particle interaction potential and the species concentration, an external field can be taken as just another particle (DM). To illustrate this, let us first briefly outline the procedure to obtain from (13) the HNC equations for an electrolyte next to an external field of spherical, cylindrical, and planar geometries.

1. Spherical Electrode

If one assumes species 3 to be large charged spherical particles of radius R at infinite dilution, the ions of species 1 and 2 have the interaction potential given by (9) and (10). However, their interaction with a particle of species 3 is given by

$$U_{3j}^{*}(r_{21}) = \begin{cases} \infty & r_{21} < R + \dfrac{a}{2} \\ 0 & r_{21} > R + \dfrac{a}{2} \end{cases} \tag{18}$$

and

$$U_{3j}^{el}(r_{21}) = \frac{4\pi R^2 \sigma e z_j}{\epsilon r_{21}} \qquad r_{21} > R + \frac{a}{2} \qquad j = 1,2 \tag{19}$$

where σ is the surface charge density on the spherical electrode. From (7), (18), and (19) we find that

$$g_{3j}(r_{21}) = \exp\left[- \frac{\beta 4\pi R^2 \sigma e z_j}{\epsilon r_{21}} + \sum_{m=1}^{2} \rho_m \int_{R+a/2}^{\infty} h_{3m}(r_{23}) \zeta_{jm}(r_{21}, r_{23})\, dv_{23} \right] \qquad j = 1,2 \tag{20}$$

where $\zeta_{vm}(r_{21}, r_{23})$ is given by (14) or (17). Since species 3 has been taken at infinite dilution, the sum in the exponential of (20) is only over the two ionic species, and (20) decouples from those for the ionic solution, which are given by (13) but with $i,j,m = 1,2$. As we shall see later, under certain conditions, this is an approximation. Equation (20) has been solved in the HNC/MS case [20].

2. *Planar Electrode*

For a planar external field, species 3 is made of infinite plates of width d and a surface charge density σ, and it is at infinite dilution. Thus the ion-plate potential is given by

$$U_{3j}^{*}(r_{21}) = \begin{cases} \infty & r_{21} < \dfrac{d + a}{2} \\ 0 & r_{21} > \dfrac{d + a}{2} \end{cases} \tag{21}$$

and by applying Gauss's law,

$$U_{3j}^{el}(r_{21}) = - \frac{4\pi\sigma e z_j}{\epsilon} r_{21} \qquad r_{21} \geq \frac{d + a}{2} \tag{22}$$

where σ is the surface charge density. In cylindrical coordinates $dv_3 = z\, dz\, dr_{23}\, d\varphi$, where dr_{23} and $d\varphi$ are the length and angle elements along and around, respectively, an axis perpendicular to the plate. z is measured along an axis parallel to the plate. In these coordinates, and taking advantage of the symmetry of the system,

$$r_{13}^2 = z^2 + | r_{21} - r_{23} |^2 \tag{23}$$

Hence, in bipolar coordinates,

$$dv_3 = dr_{23} r_{13}\, dr_{13}\, d\varphi \tag{24}$$

Therefore, from (7) and (21) to (24) we find that

$$g_{3j}(r_{21}) = \exp\left[\frac{\beta 4\pi\sigma e z_j}{\epsilon} r_{21} + \sum_{m=1}^{2} \rho_m \int_{-\infty}^{\infty} h_{3m}(r_{21})\zeta_{jm}(r_{21},r_{23})\, dr_{23}\right] \quad j = 1,2 \tag{25}$$

In these coordinates, (17) becomes

$$\zeta_{jm}(r_{21},r_{23}) = 2\pi \int_{|r_{21}-r_{23}|}^{\infty} r_{13} c_{jm}(r_{13})\, dr_{13} \tag{26}$$

Equation (25) for the HNC/MS [5,6] and HNC/HNC [24] approximations has been solved numerically. The solution is independent of the width d of the plate, and thus the solution is equivalent to that obtained with the AM [15,16].

3. *Cylindrical Electrode*

For a cylindrical external field, species 3 is made of infinitely long cylinders of radius R and surface charge density σ, and it is at infinite dilution. The ion-cylinder potential is given by

$$U_{3j}^{*}(r_{21}) = \begin{cases} \infty & r_{21} < R + \dfrac{a}{2} \\ 0 & r_2 < R + \dfrac{a}{2} \end{cases} \tag{27}$$

and

$$U_{3j}^{el}(r_{21}) = -\frac{4\pi R\sigma e z_{ij}}{\epsilon} \ln r_{21} \tag{28}$$

Equation (28) is easily obtained by applying Gauss's law to the charged cylinder. In cylindrical coordinates $dv_3 = r_{23}\, dr_{23}\, dz\, d\varphi$, where dz and $d\varphi$ are the length and angle elements along and around the cylinder, respectively. In these coordinates,

$$r_{13}^2 = z^2 + r_{21}^2 + r_{23}^2 - 2r_{21}r_{23}\cos\varphi \tag{29}$$

From (7), the HNC equation for the cylindrical double layer is

$$g_{3j}(r_{21}) = \exp\left[\left(\frac{\beta 4\pi R\sigma e z_j}{\epsilon}\right)\ln r_{21} + \sum_{m=1}^{2} \rho_m \int_{R+a/2}^{\infty} h_{3m}(r_{21})\zeta_{jm}(r_{21},r_{23})\, dr_2\right] \quad j = 1,2 \tag{30}$$

The kernel is given by

$$\zeta_{jm}(r_{21},r_{23}) = r_{23}\int_0^{2\pi} d\varphi \int_{-\infty}^{\infty} c_{jm}(r_{13})\, dz \tag{31}$$

In all the expressions for the kernels of the nonlinear integral equations so far discussed [Eqs. (14), (17), (26), and (31)], care should be taken of considering in the integration limits the ion–ion exclusion volume due the hard-sphere interaction. Equation (30) has been solved for the HNC/MS approximation [17,18].

4. *Slit Pore*

To obtain the HNC equation inside and around a slit pore, species 3 is made of dumbbell particles formed by two infinite plates of width d each, separated by a distance τ, and it is at infinite dilution. The ion–dumbbell electrostatic interaction potential is given by [34–38]

$$U_{3j}^{el}(r_{21}) = -\frac{2\pi e z_j}{\epsilon}[(r_{21} + t + d + |r_{21} - t - d|)\sigma_{out} + (r_{21} + t + |r_{21} - t|)\sigma_{in}] \tag{32}$$

for $0 \le r_{21} < t$ or $t + d < r_{21}$, where $t = \tau/2$, and σ_{out} and σ_{in} are the charge densities on the outside and inside surfaces of the plates, respectively. The ion–dumbbell hard sphere/hard wall interaction is infinite whenever the volume of an ion of diameter a and a hard wall of width d overlap, and zero elsewhere.

As for the planar electrode, $dv_3 = z\, dz\, dr_{23}\, d\varphi$, where z is measured along an axis parallel to the plates and dr_{23} and $d\varphi$ are the length and angle elements along and around, respectively, an axis perpendicular to the plates. The origin of the coordinate r_{21} is taken at half the distance between the plates (see Fig. 2). In this geometry, (23) and (24) are still valid, and, therefore, from (7) the HNC equations for the slit pore are

$$g_{3j}(r_{21}) = \exp\left[-\beta U_{3j}^{el}(r_{21}) + \sum_{m=1}^{2} \rho_m \int_{-\infty}^{\infty} h_{3m}(r_{21})\zeta_{jm}(r_{21},r_{23})\, dr_{23}\right] \qquad j = 1,2 \tag{33}$$

The kernel $\zeta_{jm}(r_{21},r_{23})$ has the same general expression as for the planar electrode, (26). However, because the boundary conditions for these two systems are different, the integral in (26) will in general give different results [37,60].

Associated with the charge densities σ_{out} and σ_{in} are the surface mean electrostatic potentials φ_0 and ψ_0, respectively. Sometimes the boundary

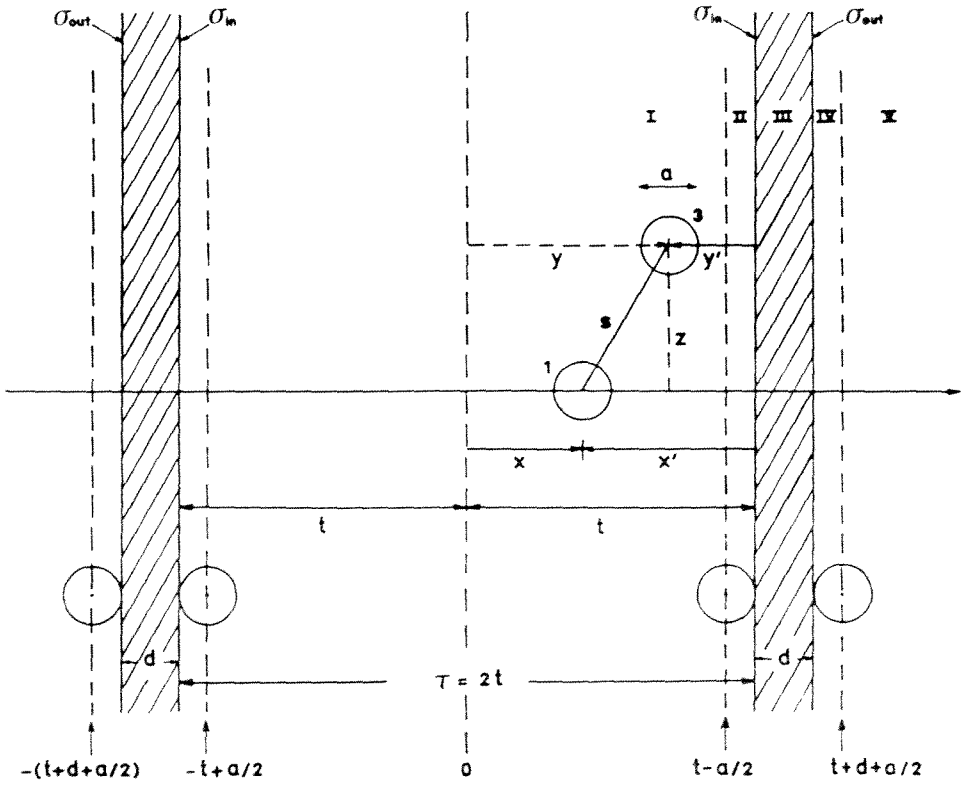


Figure 2 Geometry for a slit pore.

conditions are given in terms of the surface potentials instead of the surface charge densities. To find the relations among these quantities, as well as with those that characterize the five different regions shown in Fig. 2, the Laplace and Poisson equations of the electrostatic theory have to be integrated, that is,

$$\nabla^2\psi(x) = 0 \tag{34}$$

$$\nabla^2\psi(x) = -\frac{4\pi}{\epsilon}\rho_c(x) \tag{35}$$

where $\psi(x)$ is the local mean electrostatic potential and $\rho_c(x)$ is the local charge density, defined by

$$\rho_c(x) = \sum_m ez_m\rho_m g_{3m}(x) \equiv \sum_m ez_m\rho_m(x) \tag{36}$$

For simplicity in the notation, let us define $x \equiv r_{21}$, $y = r_{23}$, and $s = r_{13}$.

From (34) and (35) some of the relations that can be found are [37]

$$\varphi_0 = \frac{4\pi}{\epsilon} \int_{t+d+a/2}^{\infty} (t + d - y)\rho_c(y)\, dy \tag{37}$$

$$\psi_0 = \varphi_0 + \frac{4\pi\sigma_{in}}{\epsilon} d + \frac{4\pi d}{\epsilon} \int_0^{t-a/2} \rho_c(y)\, dy \tag{38a}$$

$$= \varphi_0 - \frac{4\pi\sigma_{out}}{\epsilon} d - \frac{4\pi d}{\epsilon} \int_{t+d+a/2}^{\infty} \rho_c(y)\, dy \tag{38b}$$

$$ez_j\psi(x) = U_{3j}^{el}(x) - \frac{2\pi e z_j}{\epsilon} \int_0^{\infty} (x + y + |x + y|)\rho_c(y)\, dy \tag{39}$$

$$\psi(x) = \psi_0 - \frac{2\pi}{\epsilon} \int_0^{t-a/2} (x + y - \tau + |x - y|)\rho_c(y)\, dy \tag{40}$$

$$\psi(x) = \zeta - \frac{2\pi}{\epsilon} \int_0^{t-a/2} (x + y - \tau + a + |x - y|)\rho_c(y)\, dy \tag{41}$$

where $\zeta \equiv \psi(t - a/2)$, and the last two equations are valid for $0 \le x \le t - a/2$. For this system the electroneutrality condition is

$$\sigma_{in} + \sigma_{out} = -\int_0^{t-a/2} \rho_c(y)\, dy - \int_{t+d+a/2}^{\infty} \rho_c(y)\, dy \tag{42}$$

In studies of these types of systems a boundary condition that is often imposed is either $\varphi_0 = \psi_0$ or $\sigma_{in} = \sigma_{out}$. If $\varphi_0 = \psi_0$, from (38) and (42) it is found that

$$\sigma_{in} = -\int_0^{t-a/2} \rho_c(y)\, dy \tag{43a}$$

and

$$\sigma_{out} = -\int_{t+d+a/2}^{\infty} \rho_c(y)\, dy \tag{43b}$$

That is, the electrolyte inside the slit neutralizes the electric field produced by the inside surfaces of the slit, and the electrolyte outside the slit neutralizes the electric field produced by the outside surfaces of the slit. This is not true for every situation if the condition $\sigma_{in} = \sigma_{out} \neq 0$ is imposed; then the more general electroneutrality condition (42), should be used. Although many chemists would expect (43) to be valid always, it should be remembered that the electroneutrality condition, and for that matter the chemical potential, are in part ruled by the competition between the counterion–counterion repulsion and the counterion–wall attraction. The physics of the model clearly indicates that even if the counterion–coun-

terion repulsion is disregarded, the maximum number of counterions (of diameter a) that can be put inside a slit (of width τ) will not be able to neutralize a fixed and sufficiently high value of σ_{in}. It is possible, of course, that such a value of σ_{in} might be unphysically high. Apparently, most studies of electrolytes confined in slits have not taken this fact into consideration. If the condition $\psi_0 = \psi_0$ is imposed to (33), it can be shown [34,37] that the electrolyte structures inside and outside the slits are independent of each other and of the parameter d, whereas if the condition $\sigma_{in} = \sigma_{out}$ is imposed, they are not independent of each other and of the parameter d.

If we define the short-range functions $c_s(s)$ and $c_D^{SR}(s)$ by writing the ion–ion direct-correlation function as

$$c_{jm}(s) = c_s(s) + z_j z_m c_D^{SR}(s) + \frac{z_j z_m e^2}{\epsilon s} \tag{44}$$

then using (39) and (44), after some algebra, (33) becomes

$$g_{3j}(x) = \exp[-\beta e z_j \psi(x) - \beta J_j(x)] \tag{45}$$

where $J_j(x)$ is a potential that contains the short-range contributions to the total ion-slit potential of mean force.

The function $J_j(x)$ can conveniently be given in terms of sum and difference functions,

$$J_j(x) = J_s(x) + z_j J_D(x) \tag{46}$$

where

$$\beta J_s(x) = 2\pi\rho_T \int_{t-a/2}^{t+d+a/2} K(x,y)\, dy - 2\pi \int_0^{t-a/2} K(x,y)\rho_s(y)\, dy \tag{47}$$

$$\beta J_D(x) = -2\pi \int_0^{t-a/2} L(x,y)\rho_D(y)\, dy \tag{48}$$

$$\rho_s(x) = \sum_m \rho_m h_{3m}(x) \tag{49}$$

$$\rho_D(x) = \sum_m \rho_m z_m h_{3m}(x) \tag{50}$$

$$\rho_T = \sum_m \rho_m \tag{51}$$

The kernels in the integrals are defined as

$$K(x,y) = \int_{|x-y|}^{\infty} c_s(s) s\, ds \tag{52}$$

$$L(x,y) = \int_{|x-y|}^{\infty} c_D^{SR}(s) s\, ds \tag{53}$$

To solve (45), an approximation for $c_s(s)$ and $c_D^{SR}(s)$ has to be given. The analytical expressions for $K(x,y)$ and $L(x,y)$, if the MS approximation is used, are given in Ref. 37.

It is straightforward to see that for point ions, since in this case $c_s(s) = c_D^{SR}(s) = 0$, the kernels $K(x,y)$ and $L(x,y)$ are zero and $J_j(x)$ vanishes. By construction, the mean electrostatic potential, $\psi(x)$, satisfies the exact Poisson equation, (35). Therefore, for point ions, substitution of (45) in (35) gives the Poisson–Boltzmann equation for a point-ion electrolyte inside and around a charged slit. Equation (45), with $J_j(x) = 0$, is the integral version of the PB differential equation. The function $J_j(x)$ can be looked at as a particular way of taking into account the fluctuations terms. Different ways of taking these terms into account give origin to the different versions of the MPB integrodifferential equations [32,33, 61,62].

Equation (45) has been solved numerically for the case in which $\varphi_0 = \psi_0$. Unfortunately, no Monte Carlo (MC) data are available for this system. Apparently, it is more difficult to obtain MC for this case than for the constant charge case ($\sigma_{in} = \sigma_{out}$). However, (45) reduces to the planar double-layer case in the limit of infinite separation between the plates [34], and in this case agreement with MC data is good [5,6]. If $z_1 = z_2 = 0$ and $\psi_0 = 0$, (45) reduces to the HNC equation for a hard-sphere fluid. Recently, molecular dynamics (MD) calculations for a hard-sphere fluid confined between two hard walls have been obtained [63] and compared to those from (45). The results are good, even though, as is well known, the HNC approximation is not the best for a hard-sphere fluid. In Section III, results for an electrolyte confined by a charged slit are shown.

5. *Cylindrical Pore*

As a last application of the DM to the OZ equation to obtain inhomogeneous liquid theories, the HNC equation for an electrolyte inside and around a cylindrical pore will be now derived. As in the previous cases, in (7) a suitable species 3 is defined. For a cylindrical pore, species 3 is taken to be made of hollow, infinitely long cylinders of internal and external radius of t and $t + d + a/2$, respectively, where d is the thickness of the solid part of the hollow cylinder. The cylinders are charged on both surfaces and are at infinite dilution.

By application of Gauss's law, the ion–hollow cylinder electrostatic potential is given by

$$U_{3j}^{el}(x) = -\frac{2\pi e z_j}{\epsilon}\{(t + d)\sigma_{out} \ln[x^2 + (t + d)^2 + |x^2 - (t + d)^2|] + t\sigma_{in} \ln[x^2 + t^2 + |x^2 - g^2|]\} \quad (54)$$

for $0 \leq x < t$ or $t + d < x$, and x is measured perpendicular to the cylinder axis (we have defined $x \equiv r_{21}$). In cylindrical coordinates $dv_3 = y\, dy\, dz\, d\varphi$, where dz and $d\varphi$ are the length and angle elements along and around the cylinder, respectively, and $y \equiv r_{23}$ (see Fig. 3).

Therefore, from (7) and (54), the HNC equations for a cylindrical pore are

$$g_{3j}(x) = \exp\left[-\beta U_{3j}^{cl}(x) + \sum_{m=1}^{2} \rho_m \int_0^{\infty} h_{3m}(y)\zeta_{jm}(x,y)\, dy \right] \qquad j = 1,2 \quad (55)$$

where the general expression for the kernel $\zeta_{jm}(x,y)$ has the same general expression as for the solid cylinder case, (31).

From the integration of (34) and (35) for this system, we find that [64]

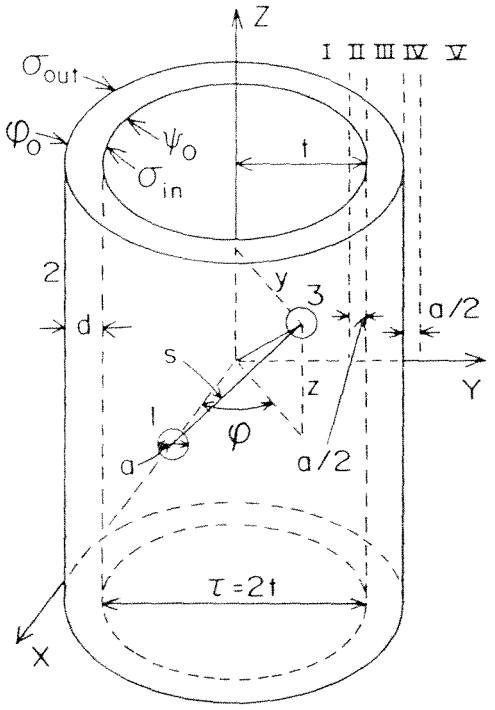


Figure 3 Geometry for a cylindrical pore.

$$\varphi_0 = \frac{4\pi}{\epsilon} \int_{t+d+a/2}^{\infty} \ln \frac{t+d}{y} \rho_c(y) y \, dy \tag{56}$$

$$\psi_0 = \varphi_0 + \frac{4\pi\sigma_{in}}{\epsilon} t \ln \frac{t+d}{t} + \frac{4\pi}{\epsilon} \ln \frac{t+d}{t} \int_0^{t-a/2} \rho_c(y) y \, dy \tag{57a}$$

$$= \varphi_0 - \frac{4\pi\sigma_{out}}{\epsilon} (t+d) \ln \frac{t+d}{t} - \frac{4\pi}{\epsilon} \ln \frac{t+d}{t} \int_{t+d+a/2}^{\infty} \rho_c(y) y \, dy \tag{57b}$$

$$ez_j\psi(x) = U_{3j}^{el}(x) - \frac{2\pi e z_j}{\epsilon} \int_0^{\infty} \ln \frac{x^2 + y^2 + |x^2 - y^2|}{2} \rho_c(yy \, dy \tag{58}$$

$$\psi(x) = \psi_0 - \frac{2\pi}{\epsilon} \int_0^{t-a/2} \ln \frac{x^2 + y^2 + |x^2 - y^2|}{2t^2} \rho_c(y) y \, dy \tag{59}$$

Equation (59) is valid only for $0 \leq x \leq t - a/2$. For this geometry the electroneutrality condition is

$$t\sigma_{in} + (t+d)\sigma_{out} = -\int_0^{\infty} \rho_c(y) y \, dy \tag{60}$$

φ_0 and ψ_0 are the mean electrostatic potentials on the outside and inside surfaces of the hollow cylinder, respectively, and σ_{out} and σ_{in} are their corresponding surface charge densities. As for the slit case, if $\psi_0 = \varphi_0$,

$$\sigma_{in} = -\frac{1}{t} \int_0^{t-a/2} \rho_c(y) y \, dy \tag{61a}$$

and

$$\sigma_{out} = -\frac{1}{t+d} \int_{t+d+a/2}^{\infty} \rho_c(y) y \, dy \tag{61b}$$

If the condition $\sigma_{out} = \sigma_{in}$ is imposed, (60) should be used.

Using (44) and (58), (55) can be written as

$$g_{3j}(x) = \exp[-\beta e z_j \psi(x) - I_j(x)] \tag{62}$$

where $I_j(x)$ is a potential that contains the short-range contributions to the total ion–hollow cylinder potential of mean force. Although (62) looks identical to (45), for the slit pore, it is analytically different and its solution, of course, is in general different to that for the slit pore. Analytical expressions for $I_j(x)$, when the MS approximation is used, as well as a full account of the equations relevant for this system, will be presented later elsewhere [64]. Nevertheless, some results are shown in Section III. The DM has also been applied by Zhou and Stell [43] to the OZ homogeneous

equation to obtain the PY/PY and HNC/PY equations for a hard-sphere fluid in a cylindrical pore.

As for all the other geometries, in the limit of point ions, (62) reduces to the PB equation, and in the zero charge limit it reduces to the HNC equation for a hard-sphere fluid in a cylindrical pore.

C. Three-Point Extension to Integral Equation Theories

The OZ equation for a one-component bulk fluid of species i is given by

$$h_{ii}(r_{21}) = c_{ii}(r_{21}) + \rho_i \int c_{ii}(r_{13})h_{ii}(r_{23})\, dv_3 \tag{63}$$

If artificially we assume this one-component fluid to be made of two species, one of which is the species i and the other, say species α, is made of dumbbells, at infinite dilution, formed by two particles of the species i, fixed at a certain center-to-center distance τ between them, the OZ equation for this system is

$$h_{\alpha i}(\mathbf{r}_{21};\tau) = c_{\alpha i}(\mathbf{r}_{21};\tau) + \rho_i \int c_{ii}(r_{13})h_{\alpha i}(\mathbf{r}_{23};\tau)\, dv_3 \tag{64}$$

where $c_{ii}(r_{13})$ can be obtained from (63). The vector notation in $\mathbf{r}_{21}$ and $\mathbf{r}_{23}$ is introduced to emphasize that the distribution of particles of species i around particles of species α will, in general, not be spherically symmetric. Any of the existing closures for the direct correlation function can be taken for $c_{\alpha i}(\mathbf{r}_{21};\tau)$ and $c_{ii}(r_{13})$. If, for example, we use the HNC approximation, we find that

$$c_{ii}(r_{13}) = h_{ii}(r_{13}) - \ln g_{ii}(r_{13}) - \beta U_{ii}(r_{13}) \tag{65}$$

and

$$c_{\alpha i}(\mathbf{r}_{21}) = h_{\alpha i}(\mathbf{r}_{21};\tau) - \ln g_{\alpha i}(\mathbf{r}_{21};\tau) - \beta U_{\alpha i}(\mathbf{r}_{21};\tau) \tag{66}$$

An important difference between (65) and (66) is in their interaction potential functions, which, in general, will be different. If (66) is substituted in (64), we obtain

$$g_{\alpha i}(\mathbf{r}_{21};\tau) = \exp[-\beta U_{\alpha i}(\mathbf{r}_{21};\tau) + \rho_i \int c_{ii}(r_{13})h_{\alpha i}(\mathbf{r}_{23};\tau)\, dv_3] \tag{67}$$

Equation (67) look very much like the usual homogeneous multicomponent HNC equation. However, it is not. Note that although $g_{\alpha i}(\mathbf{r}_{21};\tau)$ can be taken as a regular molecular two-particle distribution function, because the dumbbell is made of two regular particles of the fluid, it is also the

unsymmetrical three-particle distribution function, defined as proportional to the probability of finding a third particle at a given distance of two other particles, which are at a fixed distance between them [65]. In mathematical notation, the unsymmetrical homogeneous three-particle distribution function $g_{iii}^{[3]}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ of three given particles 1, 2, 3 is defined as

$$g_{iii}^{[3]}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3) = \frac{g_{iii}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)}{g_{ii}^{(2)}(\mathbf{r}_1,\mathbf{r}_2)} \tag{68}$$

where $g_{iii}^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)$ is the homogeneous three-particle distribution function and $g_{ii}^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ is the homogeneous two-particle distribution function. Therefore, from (67), the unsymmetrical three-particle distribution function can be calculated, within some approximation. However, one would like to have the two-particle distribution function for particles of species i.

The BGY equation is an exact theorem which relates the two-particle distribution function with the unsymmetrical three-particle distribution function [1,3,4,65]:

$$\frac{1}{\beta}\frac{d}{dr_{21}}\ln[g_{ii}^{(2)}(r_{21})] = -\frac{d}{dr_{21}}U_{ii}(r_{21}) - \rho_i\int\frac{dU_{ii}}{dr_{13}}\cos\theta\, g_{iii}^{[3]}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)\,dv_3 \tag{69}$$

where θ is the angle formed by the position vectors $\mathbf{r}_{21}$ ($\equiv \mathbf{r}_2 - \mathbf{r}_1$) and $\mathbf{r}_{31}$ ($\equiv \mathbf{r}_3 - \mathbf{r}_1$). In the notation of (67), (69) can be written as

$$\frac{1}{\beta}\frac{d}{d\tau}\ln[g_{ii}(\tau)] = -\frac{d}{d\tau}U_{ii}(\tau) - \rho_i\int\frac{dU_{ii}(r_{13})}{dr_{13}}\cos\theta\, g_{\alpha i}(\mathbf{r}_{23};\tau)\,dv_3 \tag{70}$$

From (67), $g_{\alpha i}(\mathbf{r}_{21};\tau)$ can be calculated. Then, after a simple integration, from (70) the mean force, $F_{ii}(\tau)$, between two particles of species i can be obtained:

$$F_{ii}(\tau) \equiv -\frac{d}{d\tau}W_{ii}(\tau) = \frac{1}{\beta}\frac{d}{d\tau}\ln[g_{ii}(\tau)] \tag{71}$$

After solving (67) for a sufficient number of values of the parameter τ, a numerical integration of (70) can be made, and thus $g_{ii}(\tau)$ can be obtained. This approach has been used to calculate the force between two charged plates [34–38] and between charted spherical particles [39].

The extension of this approach to a multicomponent fluid is straightforward. Suppose now that we have a nine-component fluid with species

1, 2, 3 and α_{ij}, with $i,j = 1,2,3$, such that $\alpha_{ij} = \alpha_{ji}$, and the α_{ij} species, at infinite dilution, are made of dumbbells formed by a particle of species i and a particle of species j (a distance τ_{ij} apart from each other). Then (67) and (70) become

$$g_{\alpha_{ij}k}(\mathbf{r}_{21};\tau_{ij}) = \exp\left[-\beta U_{\alpha_{ij}k}(\mathbf{r}_{21};\tau_{ij}) + \sum_{m=1}^{3} \rho_m \int c_{km}(r_{13}) h_{\alpha_{ij}m}(\mathbf{r}_{23};\tau_{ij})\, dv_3 \right] \qquad i,j,k = 1,2,3 \quad (72)$$

and

$$\frac{1}{\beta}\frac{d}{d\tau_{ij}} \ln[g_{ij}(\tau_{ij})] = -\frac{d}{d\tau_{ij}} U_{ij}(\tau_{ij}) - \sum_{m=1}^{3} \rho_m \times \int \frac{dU_{im}(r_{13})}{dr_{13}} \cos\theta\, g_{\alpha_{ij}m}(\mathbf{r}_{23};\tau_{ij})\, dv_3 \qquad i,j = 1,2,3 \quad (73)$$

Through $c_{km}(r_{13})$ all the equations are coupled. An important simplification to (72) and (73) can be achieved if $\rho_3 \simeq 0$, and $c_{km}(r_{13})$ is calculated through the usual HNC equation for a two-component fluid, (7). Then the only dumbbell that survives is the species $\alpha \equiv \alpha_{33}$, and (72) and (73) become

$$g_{\alpha j}(\mathbf{r}_{21};\tau) = \exp\left[-\beta U_{\alpha j}(\mathbf{r}_{21};\tau) + \sum_{m=1}^{2} \rho_m \int c_{jm}(r_{13}) h_{\alpha m}(\mathbf{r}_{23};\tau)\, dv_3 \right] \qquad j = 1,2 \quad (74)$$

and

$$\frac{1}{\beta}\frac{d}{d\tau} \ln[g_{33}(\tau)] = -\frac{d}{d\tau} U_{33}(\tau) - \sum_{m=1}^{2} \rho_m \int \frac{dU_{3m}(r_{13})}{dr_{13}} \cos g_{\alpha m}(\mathbf{r}_{23};\tau)\, dv_3 \quad (75)$$

For two large charged spheres, species α, immersed in a point-ion electrolyte, (74) reduces to the integral version of the PB equation. From (74) the electrolyte structure around the two spheres is obtained, and from (75) their mean force can be calculated. Thus this is the integral equation version of the Verwey–Overbeek (VO) [44] theory. Exact numerical solution to (74) and (75) for this system has been obtained [39]. The differential nonlinear PB equation for this system had been solved in the past only in an approximated manner or in a very restricted way [66–69].

If in (74) species 1 and 2 are taken as a two-species RPM electrolyte, and α is taken to be made of two charged infinitely large plates of width d (DM), we again get the HNC equation for the electrolyte structure in and around a slit pore, (33). For this model the force between the two plates of the slit can be calculated through (75). In Section III some results for this force will be shown.

Equation (25) for a RPM electrolyte next to a planar electrode has been found to fail for very high electrode's surface charge density σ. While this fact might be irrelevant for real system studies, since the values of σ above which (25) fails are probably unphysically high [60], it is academically relevant. The problem is that in (25) a bulk pair correlation function is used in the expression for $\zeta_{jm}(r_{21},r_{23})$ [Eq. (26)]. Therefore, no angular or concentration dependence of the direct-correlation function as a function of particle distance to the electrode are taken into account. To solve this shortcoming of (25), Nielaba and Fortsmann [70] have used a local density approximation in place of the bulk direct-correlation function. They have obtained very good results. However, a shortcoming of their theory is the use of an adjustable parameter. Plischke and Henderson [71–73] and Kjellander and Marčelja [40,74,75] have numerically solved the two-particle inhomogeneous OZ (OZ2) equation. In this equation the angular and concentration dependence on the ion–ion direct-correlation function as a function of their distance to the electrode are taken into account. Up to now, the OZ2 results are the more accurate. Unfortunately, because of the equation's complexity, their numerical solution demands very large computer resources, and for some values of the parameters, a numerical solution is difficult or impossible to achieve. Mier y Terán et al. [76–78] have solved (25) through a nonlocal free-energy density functional approach, where different nonlocal hard-sphere models for the weighting functions [79–81] of the generic density functional proposed by Percus [82,83] are used. Of them, that of Tarazona [80] seems to be the most successful. An advantage of this approach is its relatively low computer resource demands. A shortcoming is neglect of the angular and concentration dependence of the Coulombic part of the ion–ion direct-correlation function. Also, a better model for the hard-sphere weighting functions is needed. However, results of this theory are comparable in accuracy to those of Plischke and Henderson and Kjellander and Marčelja.

If in (72), the α-species is chosen to be made of dumbbells, at infinite dilution, formed by a charged infinitely large plate, of width d, and an ion, a two-particle distribution function next to a plate (i.e., the inhomogeneous two-particle distribution function) is obtained. This equation

will be comparable in accuracy to that of Plischke and Henderson. From (73), the one-particle distribution function can be calculated. To improve further the accuracy in calculation of the electrolyte structure next to a plate, the α-particle can be taken to be formed by a cluster of two (three, etc.) ions joined to a plate. This approach provides a systematic way of improving a given homogeneous liquid theory [34–39] and its corresponding inhomogeneous liquid theory. Clearly, this approach reduces to the conventional liquid theories if the α-particle is made of just one particle of the fluid.

In the context of the PB or MPB theories, this approach is equivalent to solving the PB or MPB differential equations with a boundary condition that increasingly takes into account more ions as fixed at a certain distance from the electrode. If two ions are taken into account, the two-particle inhomogeneous distribution function will be obtained, and so on.

Apparently, the larger the number of particles in the α-species, the better the accuracy of a given liquid theory (homogeneous or inhomogeneous). This conclusion seems to be supported by recent calculations where α-particles made of one and two particles of the fluid are compared [38] and by the fact that OZ2 comes out to be better than OZ in electrical double-layer studies [71–73].

Let us now introduce the following notation. If the α-particle procedure is applied to a given theory such that the α-particle is made of two particles of the system, we will call this new approximation the three-point extension (TPE) to that theory. If the α-particle is made of three particles of the fluid, we will call the approximation the four-point extension (FPE), and so on. For example, if the TPE is applied to the HNC approximation for homogeneous liquids, this new approximation will be called the TPE HNC. This notation comes after the implications that this approach has in terms of the diagrammatic representation of the distribution function density expansions. As an example of these implications, let us take the TPE HNC approximation. Let us assume that we have a one-component homogeneous fluid of species i. Then, in the TPE, the α-particle is made of a dumbbell formed by two particles of species i. For this system the HNC closure is given by (66), and the TPE HNC equation is given by (67). In (66) and (67), particle 2 is the α-particle and particles 1 and 3 are of species i. In diagrammatic notation, particles 2 and 1 are root (or white) points and all the other particles in the fluid are field (or black) points. Since particle 2 is really made out of two regular particles in the fluid, a two-root-point diagram in the two-component system becomes a three-root-point diagram, where two of these root points represent particle 2 and one root point is for particle 1. This is the origin of the term *three-point extension*. After integration of $g_{\alpha i}(\mathbf{r}_{21};\tau)$ through (73), the root point

of particle 1 becomes a field point, and a new bond is defined as

$$\frac{d}{dr_{13}} \ln[1 + f(r_{13})] \cos\theta = \begin{matrix} 0 & - - - & 0 \\ 1 & & 3 \end{matrix} \tag{76}$$

Note that for a one-component fluid of species i, (73) becomes

$$\frac{1}{\beta}\frac{d}{d\tau} \ln[g_{ii}(\tau)] = -\frac{d}{d\tau} U_{ii}(\tau) - \rho_i \int \frac{dU_{ii}(r_{13})}{dr_{13}} \cos\theta \, g_{\alpha i}(\mathbf{r}_{23};\tau) \, dv_3 \tag{77}$$

After this process, the HNC approximation becomes the TPE HNC approximation. It is interesting that although the HNC approximation does not include bridge diagrams, the TPE HNC does. As an example, let us see how the first-order density expansion diagram, and some of the second-order, for the HNC approximation are modified by (67) and (77). In the HNC approximation we have two diagrams: one shown in Fig. 1, with no direct bond between the two white circles, and corresponding, a second with a direct bond between the two white circles [see (6)]. These diagrams are shown in Fig. 4a. In Fig. 4b, the resulting diagrams after the TPE is introduced, through (67), are shown. In Fig. 4c the resulting diagrams after integration through (77) are shown. In Fig. 4c the lower diagram is a bridge diagram. Unfortunately, there does not seem to be a simple way of expressing these diagrams only in terms of f-bonds. In Fig. 4d, e, and f, the corresponding modifications due to the TPE in some of the second-order density diagrams are shown.

III. CONFINED FLUIDS

In Section II.C it is shown that the accuracy of a given liquid theory for homogeneous fluids can be improved systematically by increasing the number of fluid particles in the clusters of the α-species. Thus, through the direct method (DM) presented in Section II.B, the accuracy of the corresponding liquid theory for inhomogeneous fluids can also be improved systematically. From the discussion of Section II it is perhaps clear that a homogeneous fluid can be looked upon as an inhomogeneous fluid and vice versa, depending on the scope with which one looks at the fluid structure. For example, in a homogeneous dilute lyophobic colloidal dispersion one can look at the electrolyte structure around one colloidal particle as an inhomogeneous fluid around, say, a spherical external field, and if two colloidal particles are sufficiently close to each other, the external field would be that produced by a charged dumbbell. The electrolyte structure around one colloidal particle is called the electrical double layer (EDL), and as two such colloidal particles approach each other, one talks

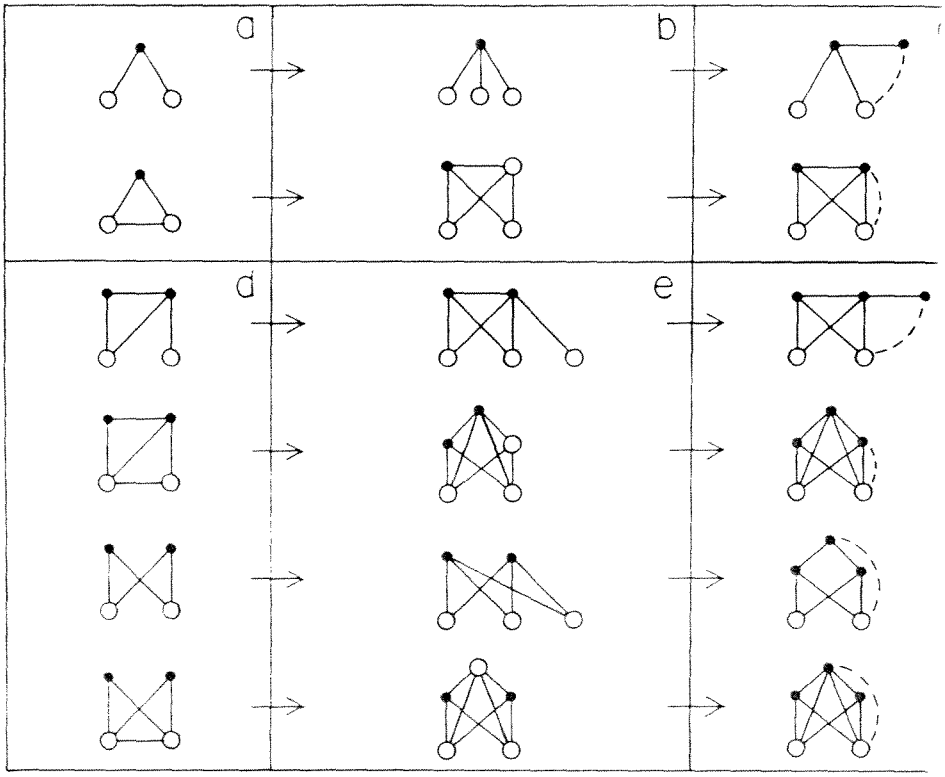


Figure 4 (a) HNC first-order density expansion diagrams for the radial distribution function; (b) resulting diagrams after the three-point extension (TPE) is introduced through (67); (c) resulting diagrams after the integration indicated in (77) is made; (d) some of the HNC second-order density expansion diagrams for the radial distribution function. Their corresponding transformations due to the application of the TPE as in (b) and (c) are shown in (e) and (f), respectively.

of EDL interactions. If a spherical colloidal particle is very large, the spherical EDL becomes a planar EDL, and the interaction of two such particles can be studied through the interaction of planar EDLs. The electrolyte between the two very large spheres can be taken as a confined fluid in chemical equilibrium with the bulk electrolyte (i.e., the electrolyte in a slit). Study of the force on the slit walls for different slit widths is relevant in the study of equilibrium conditions of colloidal dispersions, whereas the electrolyte structure inside a slit or cylindrical pore is relevant in the study of transport properties of electrolytes in porous media, among

other applications. In Section III.A the HNC and TPE HNC predictions for the force between planar EDLs will be compared and analyzed. In Section III.B, the structure of an electrolyte inside a cylindrical pore will be shown and compared to that of an electrolyte in a slit.

A. Slit Pore

For a colloidal dispersion of spherical colloid particles immersed in a two-species RPM electrolyte, the HNC equations are given by (13). If, for simplicity, the colloidal particles are taken to be at infinite dilution, (13) become

$$g_{ij}(r_{21}) = \exp\left[-\frac{\beta e^2 z_i z_j}{\epsilon r_{21}} + \sum_{m=1}^{2} \rho_m \int_a^\infty h_{im}(r_{23})\zeta_{jm}(r_{21},r_{23})\, dr_{23} \right] \qquad i,j = 1,2 \quad (78)$$

$$g_{3j}(r_{21}) = \exp\left[-\frac{\beta e^2 z_3 z_j}{\epsilon r_{21}} + \sum_{m=1}^{2} \rho_m \int_{R-a/2}^\infty h_{3m}(r_{23})\zeta_{jm}(r_{21},r_{23})\, dr_{23} \right] \qquad j = 1,2 \quad (79)$$

$$g_{33}(r_{21}) = \exp\left[-\frac{\beta e^2 z_3^2}{\epsilon r_{21}} + \sum_{m=1}^{2} \rho_m \int_{R-a/2}^\infty h_{3m}(r_{23})\zeta_{3m}(r_{21},r_{23})\, dr_{23} \right] \quad (80)$$

where R is the colloidal particle radius and a is the ionic diameter. Equations (78) and (79) are decoupled from (80) because $\rho_3 = 0$. However, for significantly large values of ρ_3, (13) must be solved. Beresford-Smith and Chan [84–85] have solved (13) for a concentrated colloidal dispersion dispersed in a point-ion electrolyte. Patey [86] has solved (78) to (80) for a RPM electrolyte dispersion medium. Medina-Noyola and McQuarrie [87] have solved the equivalent of (78) to (80) for the MS approximation.

From (80) one obtains the colloid–colloid potential of mean force, $W_{33}(r_{21})$ [see (71)]. From the analysis of the colloid–colloid interaction potential, the stability conditions for a colloidal dispersion can be found. A well-established theory for the study of colloidal stability is the Deryaguin–Landau–Verwey–Overbeek (DLVO) theory [44]. In this theory, the colloidal particles are taken to be at infinite dilution. Thus the interaction of only two of such colloidal particles is analyzed. The ions are taken to be point ions, and their structure around the colloidal particles is calculated through the PB equation. Between the two colloidal particles, two types of forces are considered: electrostatic forces and London–van der Waals (LVW) forces. For this model, the electrostatic forces, calculated through the VO theory, are always repulsive, whereas the LVW

forces are attractive. In Patey's calculations, for two spherical colloidal particles in a RPM electrolyte [86], no LVW forces were taken into account. However, he did found an attractive region for $W_{33}(r_{21})$. This implies that long-range attractive forces of electrostatic origin are present. If this force is real, the traditional approach in colloidal stability studies would have to be reviewed.

In the limit of point ions, the HNC equations for a bulk electrolyte [Eq. (78)] and the spherical EDL [Eq. (79)] reduce to the PB equation [20]. This is also the case in planar and cylindrical geometries [17]. However, Teubner [88] showed that (80) does not reduce to the PB equation in this limit. This was puzzling and raised some doubts about the existence of this long-range attractive force. Nevertheless, experimental evidence of the existence of this force has been found by Ise et al. [89–92]. On the other hand, as has been pointed out before, the TPE HNC for the interaction of two colloidal particles, (74), does reduce to the PB equation, in the limit of point ions, and (75) gives the electrostatic force of the VO theory, in the same limit. The DLVO theory reduces to the VO theory if no attractive LVW forces are considered. The TPE HNC equation has not been solved for two spherical colloidal particles in a RPM electrolyte. Therefore, no direct comparison with Patey's results has been made. However, if the colloidal particles are taken to be flat plates, the TPE HNC equations are much simpler to solve.

If in (74) and (75) and (78) to (80), the MS approximation is taken for the ion–ion direct-correlation function, the TPE HNC/MS and HNC/MS equations are obtained, respectively. The force between two charged plates immersed in a RPM electrolyte has been calculated through the TPE HNC/MS and HNC/MS equations [36–38]. To obtain the force between the plates from (80), the spatial derivative of $W_{33}(r_{21})$ must be taken.

For two plates the TPE HNC/MS equation is given by (74) and (75) such that the electrostatic part of $U_{aj}(\mathbf{r}_{21};\tau)$ is given by (32), and that of $U_{33}(\tau)$ and $U_{3m}(r_{13})$ are given by [34]

$$U_{33}^{el}(\tau) = -\frac{2\pi A}{\epsilon}(\sigma_{in} + \sigma_{out})^2\tau \tag{81}$$

and

$$U_{3m}^{el}(r_{13}) = -\frac{2\pi e z_m}{\epsilon}(\sigma_{in} + \sigma_{out})r_{13} \tag{82}$$

where A is the area of the plates. The hard sphere–hard plate part of the interaction potential is taken as usual [34]. For two plates, the HNC/MS equations can be derived by applying the DM to (79) and (80). If species

3 is taken to be made of charged plates, (79) becomes (25), and (80) becomes

$$-\frac{W_{33}(\tau)}{A}=\frac{8\pi\sigma^2}{\epsilon}\left(\tau+\frac{d}{2}\right)+\frac{1}{\beta}\sum_{m=1}^{2}\rho_m\int_{-\infty}^{\infty}c_{3m}(r_{13})h_{3m}(r_{23})\,dr_{23} \tag{83}$$

where τ is the plate separation, r_{13} and r_{23} the ion distance to each of the two plates, d the width of the plates, σ the plate surface charge density, and $W_{33}(\tau)/A$ the plate–plate potential of mean force per unit area of the plates. To be noted is that, apparently, (83) has not been deduced through the AM described in the introduction, because of the difficulties involved in taking the limit of $R\to\infty$ in (80). In the TPE HNC/MS equation two charge density parameters, σ_{in} and σ_{out}, are necessary, since when one plate approaches the other, charge polarization occurs, as is, of course, to be expected. A shortcoming of the HNC/MS equations is that no such polarization appears on the plates. Thus just one charge density parameter is necessary (i.e., in the HNC/MS equations $\sigma_{in}=\sigma_{out}\equiv\sigma$).

Equation (75) is the exact BGY equation. If a superposition approximation is used [see (68)],

$$g_{\alpha m}(\mathbf{r}_{23};\tau)=g_{3m}(r_{13})g_{3m}(r_{23}) \tag{84}$$

where $g_{3m}(r_{13})$ and $g_{3m}(r_{23})$ are the ionic distribution around plates 1 and 2, respectively. If the solution to (25) is used together with (75) and (84), the plate–plate force obtained from the superposition of HNC/MS EDLs can be calculated. In this superposition approximation, (75) for the two plates interaction becomes

$$\frac{d}{d\tau}\left[\frac{W_{33}(\tau)}{A}\right]=\frac{8\pi\sigma^2}{\epsilon}+\sum_{m=1}^{2}\rho_m\int_{-\infty}^{\infty}\frac{dU_{3m}(r_{13})}{dr_{13}}\cos\theta\,g_{3m}(r_{13})g_{3m}(r_{23})\,dr_{23} \tag{85}$$

Let us call this approximation the BGY/HNC/MS approximation. On the other hand, the point-ion limit of the TPE HNC/MS, HNC/MS, and BGY/HNC/MS approximations will be referred to as the VO, PB, and BGY/PB theories, respectively. Note that in (83), $c_{3m}(r_{13})$ and $h_{3m}(r_{23})$ are obtained from (25). Thus in both HNC/MS equation (83) and BGY/HNC/MS equation (85), the EDLs of each plate are simply convoluted, although through different recepies.

In Figs. 5 and 6, respectively, TPE HNC/MS and VO counterion and co-ion distributions inside a slit for different plate separations are shown. The plate separation, τ, is given in terms of ionic diameters (see Fig. 2); x' is the distance to one of the plates. Only half of the profile is shown, since the other half is, by symmetry, its mirror image. The counterion concentration increases as τ decreases. The opposite is true for co-ion

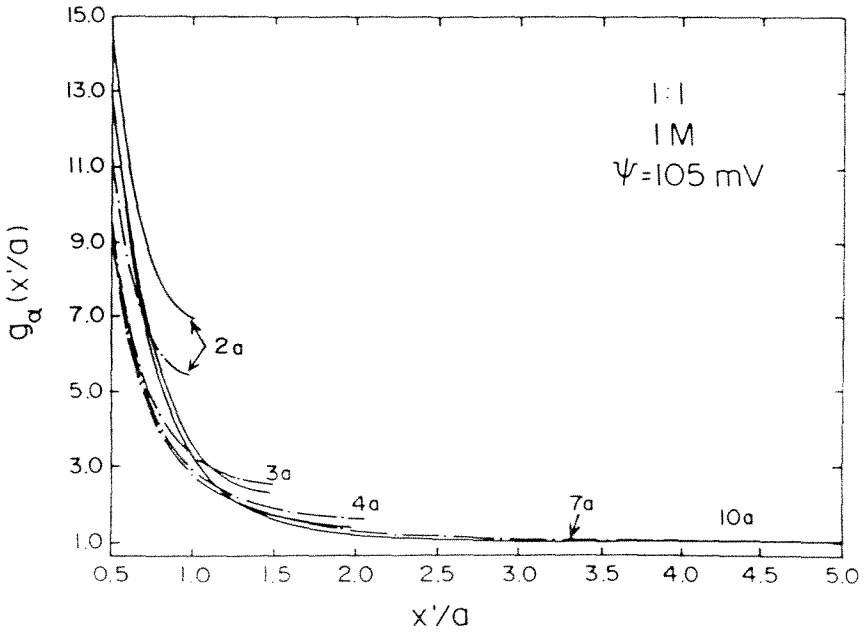


Figure 5 Counterion distribution functions for a model monovalent electrolyte (a = 4.25 Å, $z_+ = |z_-| = 1$, T = 298 K, ϵ = 78.5) confined between two charged plates at ψ_0 = 105 mV for different plate separations (measured in ionic diameters); x' is the distance to one of the plates. The solid and dot–line curves are the TPE HNC/MS and VO results, respectively. Only half of the profile is shown, since the other half is, by symmetry, its mirror image. (From Ref. 37.)

concentration. For plate separation larger than $7a$, the density and potential profiles change very little. In fact, the profiles for the region outside the plates (i.e., for $\tau \to \infty$) are for many purposes indistinguishable from that for $\tau = 10a$. For a sufficiently large plate separation the single EDL profile is reached (i.e., for large separations there is no EDL interaction). The lower the bulk electrolyte concentration, or the higher the surface potential, the larger the plate separation should be to reach the single EDL limit. In general, the calculated surface charge density is less sensitive than the plate-plate force to plate separation (see Figs. 7 and 8). TPE HNC/MS theory predicts a higher counterion concentration and a lower co-ion concentration inside the slit than its point-ion limit, the VO theory. From HNC/MS theory no such profiles can be calculated in a natural way since these profiles are basically three-particle distribution functions and HNC/MS theory is a two-particle distribution function the-

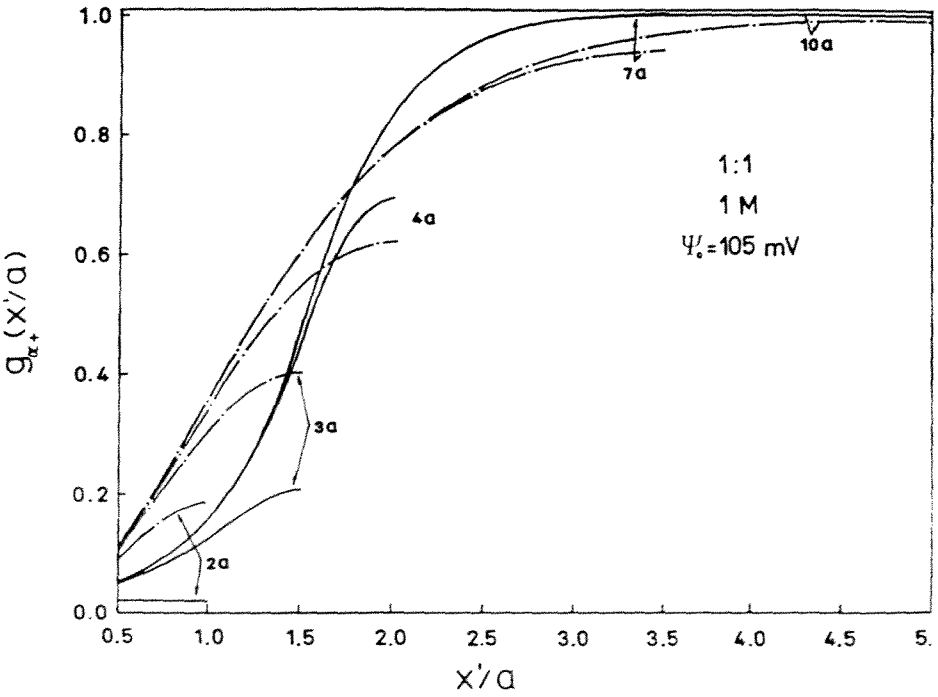


Figure 6 Co-ion distribution functions for different plate separations. The parameters are the same as in Fig. 5, as is the meaning of the symbols. Only half of the profile is shown since the other half is, by symmetry, its mirror image. (From Ref. 37.)

ory. In HNC/MS theory, only the ionic structure around each of the plates can be calculated. As pointed out before, in the point-ion limit the TPE HNC/MS equation becomes the integral equation version of the PB differential equation for the distribution of ions around two plates (i.e., VO theory). In VO theory, the PB differential equation is solved for boundary conditions given by the presence of the two plates. Therefore, the ionic profiles obtained from VO theory are, in fact, three-particle distribution functions. This is also the case for MPB theory. The larger the number of particles in the α-species in (74), the larger the number of particles that must be taken as boundary conditions in the PB and MPB differential equations, to give as a result the same order in the hierarchy of the n-particle distribution functions.

In Fig. 7, the inside surface charge density, σ_{in}, is shown as a function of plate separation. The TPE HNC/MS results are higher than the VO

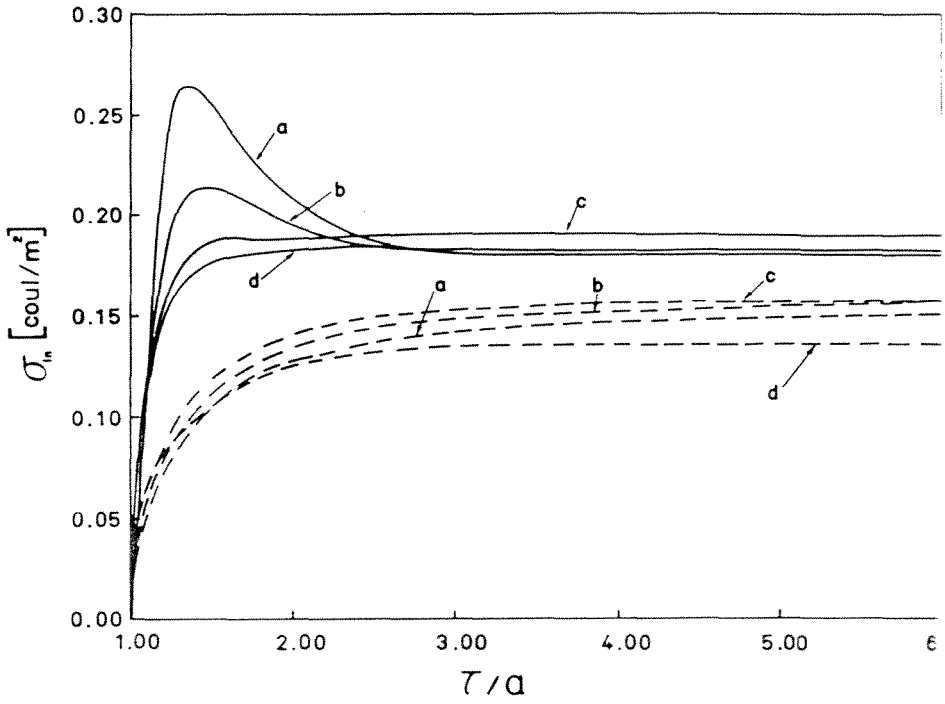


Figure 7 Charge density on the inside surface of the plates $\sigma_{in}(\tau)$ as a function of the distance between the plates. The continuous and dashed curves correspond to the TPE HNC/MS and VO results, respectively. All the curves are for a model monovalent electrolyte (a = 4.25 Å, $z_+ = |z_-| = 1$, T = 298 K, ϵ = 78.5). The bulk electrolyte concentration, the surface potential ψ_0, and the charge density for infinite separation σ_{out} are: (a) 10^{-2} M, 215 mV, and 0.1799 C/m^2 for the TPE HNC/MS and 0.1541 C/m^2 for the VO; (b) 10^{-1} M, 160 mV, and 0.1832 C/m^2, TPE HNC/MS, and 0.1596 C/m^2, VO; (c) 1 M, 105 mV, and 0.1906 C/m^2, TPE HNC/MS, and 0.1578 C/m^2, VO; (d) 2 M, 80 mV, and 0.1804 C/m^2 TPE HNC/MS, and 0.1360 C/m^2, VO. (From Ref. 37.)

results and are nonmonotonic, showing the charge polarization. The outside surface charge density, σ_{out}, is constant. In the limit of large plate separation, $\sigma_{in} \rightarrow \sigma_{out}$. The lower the bulk electrolyte concentration or the higher the surface potential, the larger the polarization effect. This is clearly related to the width of each plate EDL and the strength of their interaction.

In Fig. 8, the net force per unit area of the plates, $P_n(\tau)$, is shown as a function of plate separation. The net pressure is defined as the difference

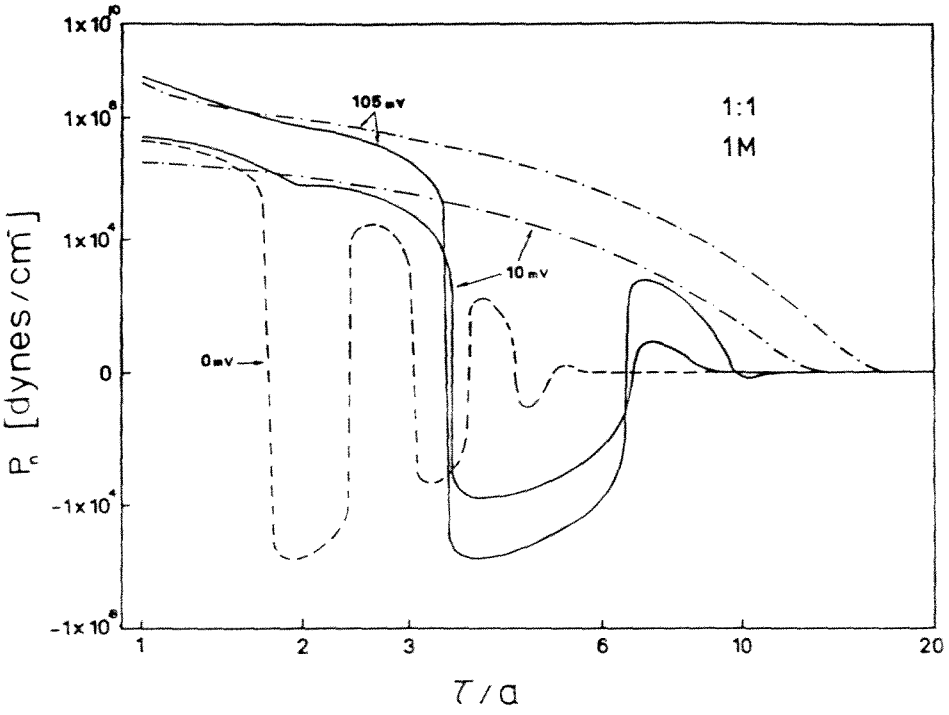


Figure 8 Net pressure between two charged plates as a function of the distance between the plates. The calculation is for a 1:1, 1 *M* bulk electrolyte concentration. The surface potential is indicated in the figure. The solid and dot–line curves are the TPE HNC/MS and VO results, respectively. The dashed curve ψ_0 = 0 mV is equivalent to a hard-sphere fluid of $\rho^* \equiv \rho a^3 = 9.247 \times 10^{-2}$. The scales for P_n and τ/a are $\sinh^{-1}$ scales. (From Ref. 37.)

of pressures on the inside and outside faces of the slit plates and is given by (75). In the figure the O-mV curve corresponds to the TPE HNC/MS solution for a hard-sphere fluid between two hard noncharged walls. A long-range attractive force is predicted by TPE HNC/MS theory, whereas the VO results are always repulsive and monotonic. In a TPE HNC/MS calculation, ionic size is taken into account, whereas in VO theory it is not. Therefore, the attraction has an ionic size origin. However, there does not seem to be a direct correlation between the attractions predicted in a hard-sphere fluid and those for a charged fluid. The pressure on the outside part of the slit is constant. Variations in net pressure are due to variations on the pressure on the inside surfaces of the slit. For a hard-sphere fluid outside the slit, the excluded volume is always constant,

whereas inside the slit this volume changes depending on how many spheres can be accommodated between the walls. In principle a relative minimum of the excluded volume exists for plate separation which is a multiple of the hard-sphere diameter. This implies a maximum in the accessible volume, and therefore a minimum in the inside pressure and in the net pressure. The higher the hard-sphere concentration, the more important this effect should be (see Fig. 9).

In Fig. 9, the net pressure between two hard plates immersed in a hard-sphere fluid with a reduced bulk concentration $\rho^* \equiv \rho a^3 = 0.8$ is shown. The hard-sphere diameter is taken to be 3 Å, to resemble roughly the size of a water molecule. Qualitative agreement with some macroscopic plate interaction experiments [93] is remarkable good. This, of course, shows the relevance of taking into account solvent molecular size in colloidal interaction studies. However, as shown in Fig. 8, size effects on the force

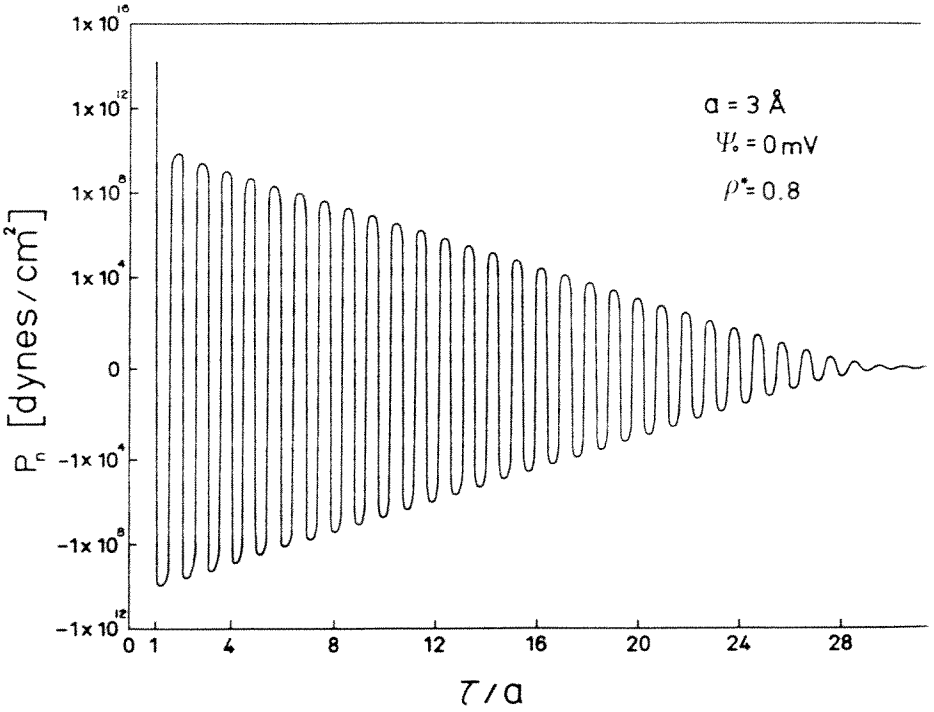


Figure 9 Net pressure between two noncharged hard plates as a function of the distance between the plates. The liquid is a hard-sphere fluid (a = 3 Å, ρ^* = 0.8, T = 298 K). The curve represents TPE HNC/MS results. (From Ref. 37.)

can be greatly modified by the charge, and simple superposition of the solvent origin force between the plates with a force of ionic origin might not be a good approximation.

Since for the force between the two plates in (75) the plate-ion unscreened interaction potential $U_{3m}(r_{13})$ is separated into a hard plate–hard sphere part and a Coulombic part, the net force can be analyzed in terms of a hard-sphere force and an electrostatic force. In the exact equation (75) the total plate–plate force is given in terms of the unscreened plate–plate Coulombic force, $dU_{33}(\tau)/d\tau$, plus, by definition of average quantity, the average hard plate–hard sphere force plus the average charged plate–ion force. The sum of the unscreened plate–plate Coulombic force with the average charged plate–ion force gives the Maxwell stress tensor contribution to the force between the plates. Although the hard-sphere contribution and the Maxwell stress tensor contribution are not really independent, an analysis of the two contributions allows a better understanding of the origin of the long-range attractive force.

The TPE HNC/MS hard-sphere and Maxwell stress tensor components of the net force are shown in Fig. 10 as a function of the plate separation. As can be seen, for some slit separation the Maxwell stress tensor component dominates, and for some others the hard-sphere component dominates. For lower bulk electrolyte concentration, the Maxwell stress tensor component seems to be responsible for the attraction region observed [37]. In general, however, it is perhaps somewhat inaccurate to attribute the origin of the long-range attractions observed in colloidal dispersion forces only to the electrical charge [89–92], since for point ions this force is always repulsive.

Although the TPE HNC/MS is not expected to give its best results for hard-sphere fluids, since the PY approximation is known to be better than the HNC approximation for this type of system, it is encouraging that the TPE HNC/MS density profiles for a hard-sphere fluid between two hard plates have in general good agreement with MD calculation [63]. In Fig. 11, such a comparison is shown for two plate separations, and a given average concentration of the hard-sphere fluid inside the slit, $\rho^* = 0.2$. In this figure a very good agreement of the TPE HNC/MS results with the MD calculations is shown.

In Fig. 12, the TPE HNC/MS [Eq. (75)], the HNC/MS [Eq. 83)], and the BGY/HNC/MS [Eq. (85)] predictions for the net pressure between two charged plates are shown. Also shown are the results of their point-ion limit theories (i.e., the VO, PB, and BGY/PB). The fluid is a 1:1, 10^{-2} M electrolyte and the surface potential on the plates is 50 mV. For this low potential and electrolyte concentration the ionic size effect should be negligible. Therefore, the RPM calculations should agree with their

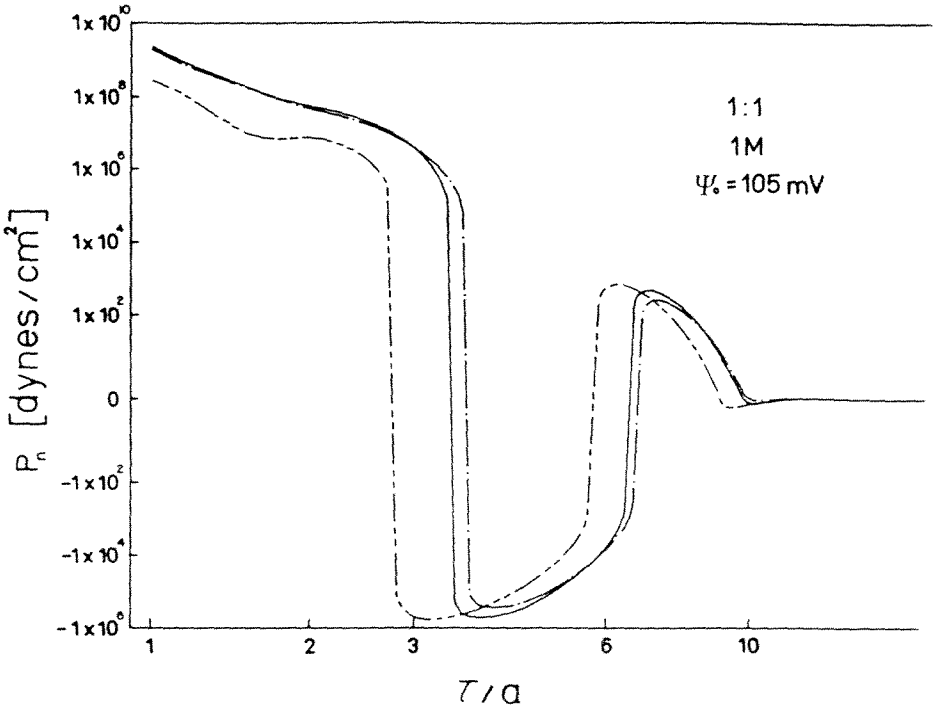


Figure 10 The Maxwell stress tensor (— - - -) and hard-sphere (— · —) net contributions to the net pressure (—) between two charged plates as a function of the distance between the plates. The curves are the TPE HNC/MS results. (From Ref. 37.)

corresponding point-ion limits. This is observed in Fig. 12. The force is always repulsive and of long range (i.e., the force is different from zero for relatively large plate separations). A long-range interaction can be expected since for low electrolyte concentrations the EDL of each plate is thick. The HNC/MS force is in excellent agreement with the TPE HNC/MS force for large plate separations. For small values of τ, an important quantitative disagreement exists. The BGY/HNC/MS results are in poorer agreement with the TPE HNC/MS force even for large values of τ but are in good agreement with the HNC/MS force.

In Fig. 13, results for a higher surface potential are shown. At this higher potential the ionic size effect should be noticeable, at least for small values of τ, since the counterion concentration near the plates is much higher [37]. This effect explains the observed small disagreement

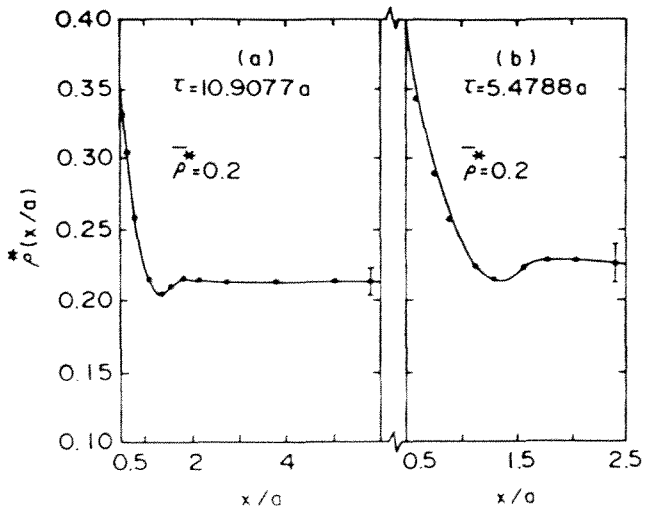


Figure 11 Reduced density profiles for a hard-sphere fluid between two hard walls for two different plate separations. The average reduced density inside the slit is 0.2. The solid lines are TPE HNC/MS results and the dots represent MD data. (From Ref. 63.)

of the RPM theories with their corresponding point-ion theories for $\tau < 3a$. More important is the fact that the HNC/MS, BGY/HNC/MS, and their point-ion limits predict a wide and strong attractive region. This finding is not surprising since the HNC/MS force between two large spheres has also been found to be attractive [86]. The fact that the PB result is also attractive has been used by Teubner [88] to suggest that the attraction predicted by HNC/MS theory is not real. At this point, the TPE HNC/MS results support Teubner's suggestion. Note that although the HNC/MS and BGY/HNC/MS agree qualitatively, the HNC/MS is in excellent agreement with the TPE HNC/MS for sufficiently large values of τ.

In Figs. 14 and 15, results for a higher electrolyte concentration and for two surface potentials are shown. In Fig. 14, for a very low surface potential, $\psi_0 = 10$ mV, almost perfect agreement among the three RPM theories, on the one hand, and among the point-ion theories, on the other hand, is observed. Note that in opposition to the point-ion results shown in Fig. 13, in Fig. 14 the PB and BGY/PB forces are always repulsive. The three RPM theories predict a long-range attractive region. In Fig. 15, results for a much larger surface potential, $\psi_0 = 105$ mV, are shown. The

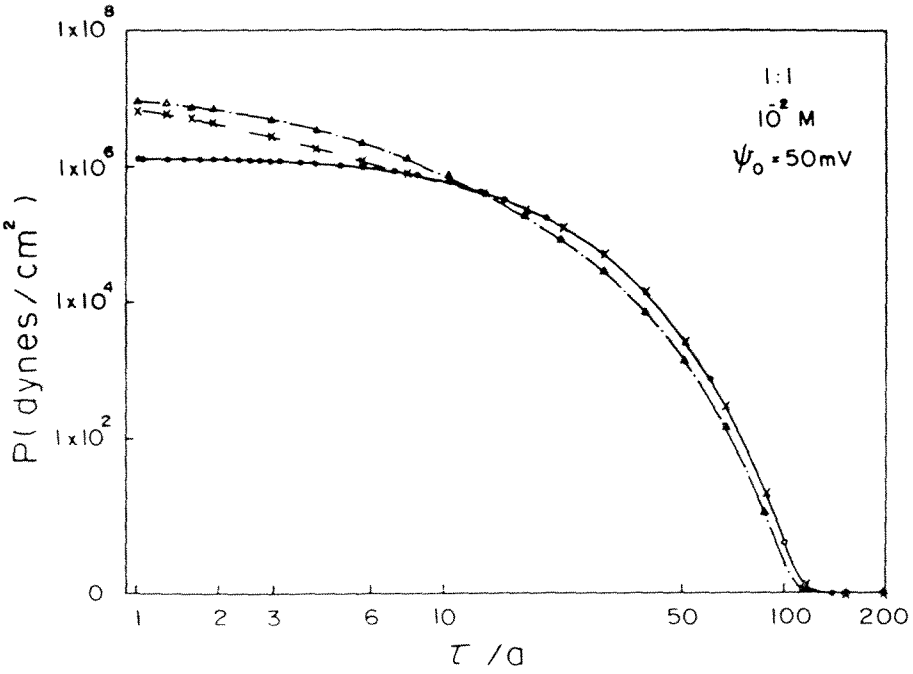


Figure 12 Net pressure between two charged plates as a function of the distance between them. The calculation is for a 1:1, 10^{-2} M electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). The surface potential ψ_0 = 50 mV. The solid, dashed, and dot–line curves are the TPE HNC/MS, HNC/MS, and BGY/HNC/MS results, respectively. The white dots, crosses, and triangles are the VO, PB, and BGY/PB results, respectively. A $\sinh^{-1}$ scale is used for P and τ/a. The width d of the plates for the HNC/MS and PB calculations is taken to be zero. (From Ref. 38.)

three point-ion theories predict a repulsive force for all values of τ. This is in agreement with a Bell and Levine [94] result in which from free-energy consideration it is shown that for point ions the force should always be repulsive. The agreement of the PB force with the VO force is very good. The agreement of the HNC/MS force with that from TPE HNC/MS theory is excellent, but for small values of τ. In the region where they agree, the two theories predict the existence of an attractive force region. However, both the HNC/MS and BGY/HNC/MS predict apparently non-existent attractive regions for small values of τ.

In the HNC/MS and BGY/HNC/MS theories the EDLs of the two plates are simply convoluted [see (83) and (85)], whereas in the TPE HNC/MS the EDL of a dumbbell formed by two plates is used to calculate their interaction force [see (75)]. HNC/MS and BGY/HNC/MS are two-particle

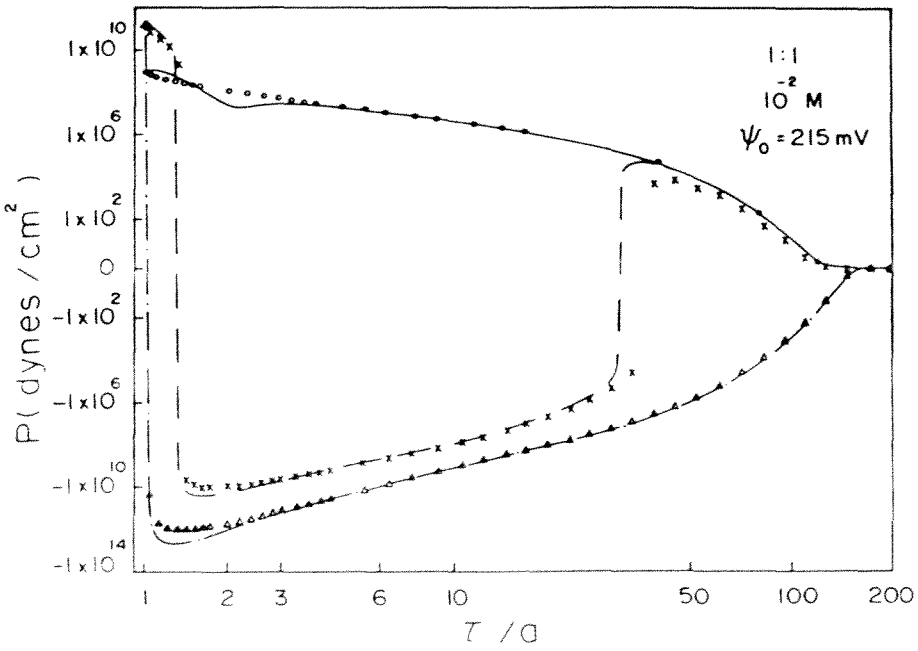


Figure 13 Net pressure between two charged plates as a function of the distance between them. The meaning of the curves is as in Fig. 12. (From Ref. 38.)

distribution function theories, whereas TPE HNC/MS is a three-particle distribution theory. In HNC/MS theory bridge diagrams are not taken into account. In TPE HNC/MS theory many classes of these type of diagrams are included. Both the HNC/MS and BGY/HNC/MS theories use some type of superposition approximation for interaction of the EDLs of each plate or colloidal particle. The lower the electrolyte concentration and/or the higher the surface potential, the wider and/or more concentrated the EDLs, and hence the worse the superposition approximation would be. This is corroborated by the results of Figs. 12 to 15.

In the limit of point ions, the HNC/MS theory reduces to the PB theory. For large values of τ and/or low surface charge or potential, the PB equation for the distribution of ions around a colloidal particle can be linearized. The solution to the linear PB equation is the Yukawa potential [84,85,95]. These Yukawa EDLs have been used extensively to study colloidal dispersions [84,85,95–103]. Even MC and MD dynamics studies that corroborate Yukawa potential theories have been made [104,105]. From the discussion above, it is clear that all these theories neglect ionic

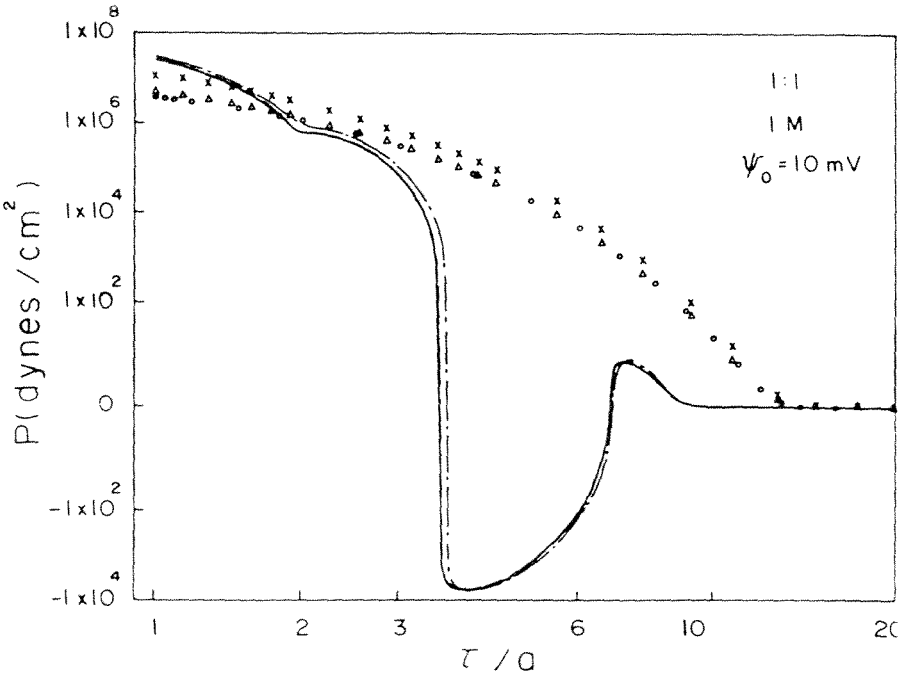


Figure 14 Net pressure between two charged plates as a function of the distance between them. The meaning of the curves is as in Fig. 12. (From Ref. 38.)

size effects and use a strong superposition approximation in the region of small values of τ. Even if the average distance in a dilute colloidal dispersion is sufficiently large to presume the validity of a superposition approximation, in statistical mechanical theories an average over all interparticle distances is implicit. The use of these theories to predict colloidal phase transitions is perhaps too optimistic, since near a phase transition, long-range correlations appear. Thus the use of an effective or renormalized potential that does not take into consideration changes in the correlation of the small particles around the large particles might be a too strong assumption. Therefore, perhaps a more careful analysis of the regions of validity of these widespread theories should be made. Finally, it should be pointed out that predicting the general features of experimental structure factors does not seem to be a strong test to the theory. Apparently, any theory that correctly takes into account the long-range limit behavior of an interaction potential is in reasonable agreement with experimental structure factor measurements [96–102,105]. Theories of renormalized potentials, which include long-range attractive potentials,

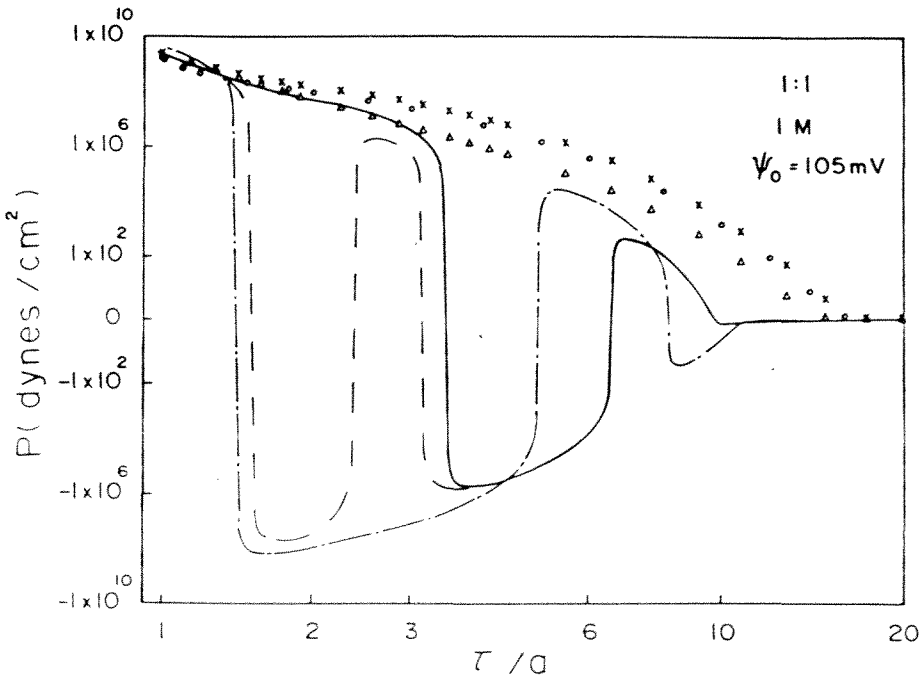


Figure 15 Net pressure between two charged plates as a function of the distance between them. The meaning of the curves is as in Fig. 12. (From Ref. 38.)

might be more realistic [106]. TPE studies could be used to evaluate their adjustable parameters. All the renormalized theories share the advantage of their simplicity, and therefore their value for practical applications. Their validity for low-charge colloidal systems and/or high electrolyte concentrations seem to be supported by the results shown in Figs. 12 to 15.

B. Cylindrical Pore

Through the DM, the equations for an electrolyte inside a charged cylindrical pore can easily be derived from the equations for a homogeneous electrolyte solution [64]. Equations (55) and (62) are equivalent and are the HNC/MS equations for an electrolyte in a charged cylindrical pore. In the limit of point ions the HNC/MS equation reduces to the integral version of the PB differential equation [i.e., in (62), $I_j(x) = 0$]. In the limit of zero charge, the HNC/MS equation for a hard-sphere fluid in a cylindrical pore is obtained. Using the DM, Zhou and Stell [43] have obtained and solved the PY equation for a hard-sphere fluid in a cylindrical pore.

Equation (33) for a slit pore, Eq. (55) for a cylindrical pore, and Eq. (74) for two spheres in a fluid differ only in the geometry of the external field. However, in terms of the homogeneous fluid approach presented in this chapter, (33) and (74) are in fact three-particle distribution function equations, and (55) is a two-particle distribution function equation. Hence (33) and (74) are referred to as TPE HNC/MS equations, whereas (55) will be referred to as a HNC/MS equation. Due to the peculiar geometry of a cylindrical pore the HNC/MS equation for a cylindrical pore has, to some extend, the same level of approximation as the TPE HNC/MS equation for a slit pore, before integration of the third particle through (75) is made.

A numerical solution to (55) has been obtained [64]. In Fig. 3, the geometry for a cylindrical pore is shown. In Figs. 16 and 17, the HNC/MS and PB counterion and co-ion reduced density profiles for an electrolyte inside a charged cylindrical pore are shown for two cylinder diameters, $\tau = 3a$ and $6a$, respectively. The TPE HNC/MS and VO results for a slit pore are included for comparison. The bulk electrolyte concentration and hence the chemical potential of the salt is the same in both the cylindrical pore and the slit pore.

The smaller the value of τ, the larger the counterion concentration. This effect is much more pronounced in the cylindrical pore than for the

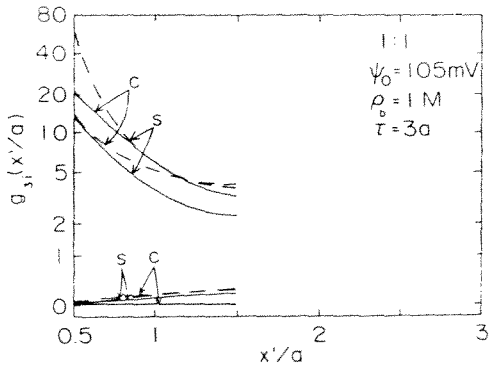


Figure 16 Reduced counterion and co-ion density profiles for a cylindrical pore (c) and a slit pore (s). The diameter of the cylindrical pore and the width of the slit pore have the same value, $\tau = 3a$. The fluid is a 1:1, 1 *M* electrolyte (a = 4.25 Å, T = 298, ϵ = 78.5). The potential on the surfaces of both the cylinder and the slit is ψ_0 = 105 mV. The continuous lines are the HNC/MS (c) or the TPE HNC/MS (s) results. The dashed curves are the PB (c) or VO (s) results. The density profile is given on a $\sinh^{-1}$ scale. Only half of the profile is shown since the other half is, by symmetry, its mirror image.

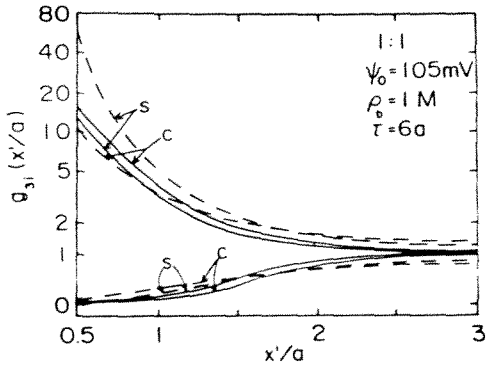


Figure 17 Reduced counterion and co-ion density profiles for a cylindrical pore (c) and a slit pore (s) for a cylinder diameter and slit width of $\tau = 6a$. The symbols are as in Fig. 16. Only half of the profile is shown since the other half is, by symmetry, its mirror image.

slit. The RPM counterion profile for the cylinder is higher than that for the slit. However, due to the geometry of the systems, the average electrolyte concentration in the slit is in general higher than that for the cylinder. This can be seen from the counterion adsorption values, which for the slit are -0.15498 C/m^2 for $\tau = 3a$ and -0.1488 C/m^2 for $\tau = 6a$, whereas for the cylinder they are -0.126 C/m^2 for $\tau = 3a$ and -0.1363 C/m^2 for $\tau = 6a$. For the point-ion electrolyte the counterion adsorption is much higher in the slit than in the hollow cylinder; counterion adsorptions in the slit are -0.4037 C/m^2 for $\tau = 3a$ and -0.3955 C/m^2 for $\tau = 6a$, whereas for the cylinder they are -0.0883 C/m^2 for $\tau = 3a$ and -0.1051 C/m^2 for $\tau = 6a$. Note that while in the cylinder the RPM counterions are more absorbed than the point counterions, in the slit the opposite is true. In Figs. 16 and 17, the RPM counterion profiles are seen to be larger in cylindrical geometry than in planar geometry. The opposite is true for the point counterion profiles. From the results presented here, larger values of τ imply lower counterion adsorption for a slit but larger counterion adsorption for a hollow cylinder. The point ion electrolyte seems to overestimate the counterion adsorption. In all cases presented here, the co-ion adsorption is usually an order of magnitude lower than its corresponding counterion adsorption. These results could be of interest in catalysis studies.

For the slit pore, in the limit of zero charge the HNC/MS equation, (55) becomes the HNC/MS equation for a hard-sphere fluid in a cylindrical pore. The HNC/MS equation density profile for a hard-sphere fluid in a

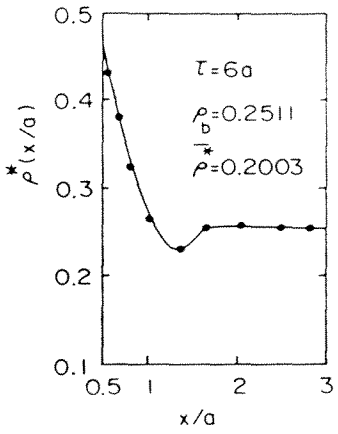


Figure 18 Reduced density profile for a hard-sphere fluid in a cylindrical hard pore of diameter $\tau = 6a$. The average reduced density inside the pore is 0.2. The solid line is the HNC/MS profile and the dots are the MC data. Only half of the profile is shown since the other half is, by symmetry, its mirror image.

cylindrical pore is compared to MC calculations in Fig. 18. Very good agreement is found at this relatively low bulk concentration. The agreement of the HNC/MS results with the MC data shown in Fig. 18 seems to be at least as good as that shown by the PY/PY calculation of Zhou and Stell [43]. However, their calculations are for higher bulk fluid concentrations. Extensive HNC/MS comparisons with MC data will be presented later elsewhere.

IV. TRANSPORT COEFFICIENTS FOR AN ELECTROLYTE IN A SLIT

The application of a pressure gradient P_z and/or an electrical potential gradient E_z in the z direction across a charged slit immersed in an electrolyte solution, [i.e., in a direction parallel to the plates (see Fig. 2)] gives origin to a net volume flux of solution V and the appearance of a net electrical current I. Both of these quantities can, in principle, be measured experimentally. If E_z and P_z are not too large, I and V can be assumed to have a linear dependence and to be coupled through a set of Onsager transport coefficients L_{mn},

$$V = L_{11}P_z + L_{12}E_z \tag{86}$$

$$I = L_{21}P_z + L_{22}E_z \tag{87}$$

For a laminar steady-state flow between the parallel plates, under the stresses E_z and P_z, the one-dimensional Navier–Stokes equation of a viscous fluid is

$$\eta \frac{d^2U(x)}{dx^2} + \rho_c(x)E_z + P_z = 0 \tag{88}$$

where $U(x)$ is the flux velocity profile, x the distance to the midplane, η the viscosity coefficient to the bulk electrolyte solution, and the total charge density $\rho_c(x)$ is defined by (36).

The velocity $U(x)$ inside the pore must satisfy the boundary conditions

$$\frac{dU(x)}{dx} = 0 \qquad \text{at} \quad x = 0 \tag{89}$$

$$U(x) = 0 \qquad \text{for} \qquad t - \delta \leq x \quad \text{or} \quad x \leq \delta - t \tag{90}$$

where $x = 0$ corresponds to the midplane and $x = \pm(t - \delta)$ corresponds to the so-called nonslipping planes. Since the exact location of these hydrodynamic planes is somewhat ambiguous, the simplifying assumption that $\delta = a/2$ is often made [18,107,108] (i.e., the nonslipping planes coincide with the planes of closest approach of the hydrated ions to the walls). As a consequence, the zeta potential will correspond to the mean electrostatic potential at $x = a/2$; that is, the zeta potential is taken as $\zeta \equiv \psi(a/2)$ and is given in terms of $\rho_c(x)$ by (40).

Integrating (88), using the fact that $d\psi/dx = 0$ and $\psi = \zeta$ at $x = (\tau - a)/2$, we get

$$U(x) = U_P(x) + U_E(x) \tag{91}$$

where $U_P(x)$ is the Poiseuille velocity component,

$$U_P(x) = -\frac{P_z}{2\eta}(x^2 - h^2) \tag{92}$$

and $U_E(x)$ is the electro-osmotic velocity component,

$$U_E(x) = -M\left(1 - \frac{\psi(x)}{\zeta}\right)E_z \tag{93}$$

where $h \equiv (\tau - a)/2$ and $M = \epsilon\zeta/(4\pi\eta)$.

The total volume flow per unit width, V, is obtained by integrating $U(x)$ between the plates

$$V = \int_{-h}^{h} U(x)\, dx \tag{94}$$

By comparison with (86), we find that

$$L_{11} = \frac{2}{3\eta} h^3 \tag{95}$$

and

$$L_{12} = L_{12}^{\infty}(1 - G) \tag{96}$$

where

$$L_{12}^{\infty} = -2hM \tag{97}$$

and

$$G = \frac{1}{h\zeta} \int_0^h \psi(x)\, dx \tag{98}$$

L_{12}^{∞} corresponds to the ratio V/E_z for $P_z = 0$ in the Smoluchowski limit (i.e., $h \to \infty$).

Under applied electrical potential and pressure gradients, the velocity of the ions of species j in the electrolyte solution has three components:

$$U_j(x) = U_p(x) + U_E(x) + U_j^T(x) \tag{99}$$

where the first two terms on the right-hand side correspond to the flow velocity of the entire solution, given by (92) and (93). The last term is the transport velocity due to the ionic mobility m_j of the charged species j under the electrical field E_z, namely

$$U_j^T(x) = \frac{z_j}{|z_j|} m_j E_z \tag{100}$$

The total current is then

$$I = 2 \int_0^h \sum_j z_j e \rho_j(x) U_j(x)\, dx \tag{101}$$

$$= I_c^P + I_c^E + I_T \tag{102}$$

where

$$I_c^P = 2 \int_0^h \rho_c(x) U_p(x)\, dx \tag{103}$$

$$I_c^E = 2 \int_0^h \rho_c(x) U_E(x)\, dx \tag{104}$$

$$I_T = 2 \int_0^h \sum_j z_j e \rho_j(x) U_j^T(x)\, dx \tag{105}$$

Using (35), (89), (90), (92), and (103) it is found that

$$I_c^P = L_{21}P_z \tag{106}$$

where L_{21} turns out to be identical to L_{12}, given by (96), as expected from the Onsager reciprocity principle. Similarly, from (93) and (104),

$$I_c^E = L_{12}^{\infty}E_zJ_c \tag{107}$$

where

$$J_c = \frac{1}{h}\int_0^h \rho_c(x)\left[1 - \frac{\psi(x)}{\zeta}\right]dx \tag{108}$$

In terms of the ionic conductivities $\sigma_j = \rho_j \mid z_j \mid m_j e$, the transport current is

$$I_T = 2E_z\int_0^h \sum_j \sigma_j g_j(x)\,dx \tag{109}$$

So far no approximations have been introduced in the theory other than those inherent to the application of hydrodynamics to narrow pores and to the fact that the total electric field in the solution is taken to be simply the sum of the applied electric field E_z and the EDL electric field. For small Reynolds numbers these two assumptions are justified. At this point, however, let us assume that the ionic transport numbers for the ions inside the slit are nonlocal and equal to those for the bulk solution. That is, the ionic transport number of species j, t_j, is defined as

$$t_j = \frac{\sigma_j}{\sigma_T} \tag{110}$$

where $\sigma_T = \sum_j \sigma_j$ is the total salt conductivity of the bulk solution.

Equation (109) can now be written as

$$I_T = 2h\sigma_T E_z(1 + K_T) \tag{111}$$

where

$$K_T = \frac{1}{h}\int_0^h \sum_j t_j h_j(x)\,dx \tag{112}$$

In terms of the Onsager coefficients, the total current I is given by (87), with

$$L_{22} = L_{22}^{\infty}(1 + K) \tag{113}$$

where $L_{22}^{\infty} = 2h\sigma_T$ and K is the capillary conductance coefficient, com-

monly referred to as the surface conductance in the classical colloidal literature. K has convective and ionic transport contributions

$$K = K_c + K_T \tag{114}$$

where K_T is given by (112) and

$$K_c = \frac{L_{12}^{\infty}}{L_{22}^{\infty}} J_c \tag{115}$$

The potential difference E_z produced when a pressure gradient P_z is applied across a capillary under conditions of zero current is referred to as the streaming potential and is obtained from (87) as

$$E_{str} = -\frac{L_{12}}{L_{22}} P_z \tag{116}$$

From (96) and (113) it is found that

$$E_{str} = \frac{\epsilon \zeta F}{4\pi\eta\sigma_T} P_z \tag{117}$$

where $\epsilon\zeta P_z/(4\pi\eta\sigma_T)$ is the Smoluchowski limiting value and

$$F = \frac{1 - G}{1 + K} \tag{118}$$

is the factor that corrects the streaming potential for finite micropores. Defining an apparent zeta potential ζ_a as

$$\zeta_a = \zeta F \tag{119}$$

the form of the Smoluchowski formula for the streaming potential can be preserved.

In a stationary regime under an applied pressure gradient, in the absence of an applied electrical field, an induced streaming potential will be established. Under a zero-current condition, from (86) and (116), the volume flux V is given by

$$\left(\frac{V}{P_z}\right)_{I=0} = L_{11}\left(1 - \frac{L_{12}^2}{L_{11}L_{22}}\right) \tag{120}$$

Equation (120) can be written as

$$V = \frac{2h^3}{3\eta_a} P_z \tag{121}$$

where an apparent viscosity coefficient, η_a, has been defined, in order

to maintain the form of the Poiseuille formula for the volume flux [see (95)].

From (95) to (97), (113), and (120) and (121) it is found that

$$\frac{\eta_a}{\eta} = \left\{1 - \frac{3\beta^*}{1 + K}\left[\frac{\zeta^*(1 - G)}{\kappa h}\right]^2\right\}^{-1} \tag{122}$$

where

$$\zeta^* = e\beta\zeta \tag{123}$$

$$\beta^* = \left(\frac{\epsilon}{4\pi e\beta}\right)^2 \frac{\kappa^2}{\eta\sigma_T} \tag{124}$$

and

$$\kappa^2 = \frac{4\pi\beta e^2}{\epsilon}\sum_j \rho_j z_j^2 \tag{125}$$

In the formulas above for the electrokinetic properties of an electrolyte inside a slit, the local charge density $\rho_c(x)$ must be given. From (33) or (45) the RPM TPE HNC/MS local charge density profile can be calculated, and from (39) the mean electrostatic potential profile can be obtained. In the point-ion limit, (33) and (45) reduce to the integral version of the PB differential equation. Hence, in this limit, from (33) the VO local charge density profile can be obtained. In Figs. 5 and 6, some TPE HNC/MS and VO density profiles are shown for a constant surface potential. In this section, results for the electrokinetic properties of a RPM and a point-ion model electrolytes will be shown within the TPE HNC/MS and VO theories, respectively. The results are for situations in which a constant zeta potential is given.

The TPE HNC/MS and VO (or PB) reduced mean electrostatic potential profiles for a 1:1, 1*M* electrolyte are shown in Fig. 19. Two different separations between the walls and two different values of the ζ potential are considered. The absolute value of the reduced electro-osmotic velocity profile is indicated on the right-hand scale [see (93)]. As a general rule, the higher the ζ potential and/or the plate separation, the deeper the potential well [55]. The TPE HNC/MS potential curve is always deeper than the corresponding PB curve. For sufficiently high ζ potentials and/or electrolyte concentrations, the TPE HNC/MS curves become oscillatory and/or negative, whereas the PB curves are always positive and monotonic.

The deeper the reduced potential profile, the smaller the area under the potential curve. From (98) it is clear that a smaller area under the reduced potential curve implies a smaller value of G and thus a larger

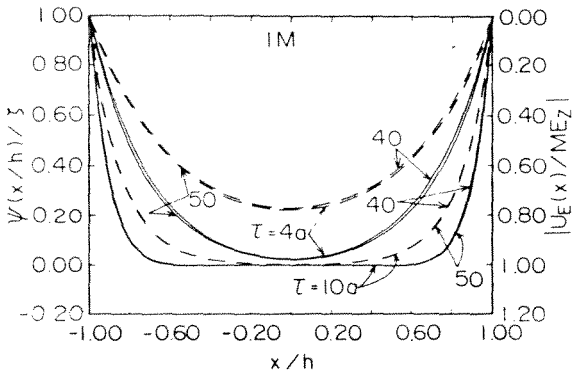


Figure 19 Reduced mean electrostatic potential profile for two slit widths, τ = $4a$ and $10a$, and two ζ potentials, ζ = 40 mV and 50 mV. The fluid is a 1:1, 0.1 M electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). On the right-hand scale the absolute value of the reduced electro-osmotic velocity profile is indicated. The dashed lines are the point-ion model, VO results. The solid lines are the TPE HNC/MS results for the RPM. (From Ref. 55.)

value of $(1 - G)$. Thus larger values of τ and/or ζ imply higher values of $(1 - G)$ (see Fig. 20). Although no results for the potential profile as a function of the electrolyte concentration are presented, higher electrolyte concentrations imply deeper potential wells and hence higher values of $(1 - G)$ [55]. Larger values of $(1 - G)$ produce a larger negative contribution to the net volume flux and net electrical current [Eqs. (86) and (87)] of the so-called electrical double-layer strength Onsager coefficient, L_{12}.

For infinite plates separations, the width of the EDL associated with each plate is defined as the distance from the plate, where the counterion (co-ion) concentration is maximum (minimum), to the point in the solution where the electrolyte concentration becomes equal to that of the bulk electrolyte (i.e., where the electrode's electric field is neutralized by the electrolyte). For narrow slits the electric field is zero at the center of the slit, although the electrolyte concentration at that point is in general different from that of the bulk. Hence for narrow slits the EDL width is given by half the distance between the plates. In every case, a measure of the EDL strength is given by the slope of the potential profile. A steeper $\psi(x)$ curve implies a stronger attraction of the counterions to the walls and therefore a narrower, more concentrated EDL. In Fig. 19, the normalized potential profiles $\psi(x/h)$ are steeper for larger values of h. However, $\psi(x)$ has the opposite behavior. Thus, in general, the larger the

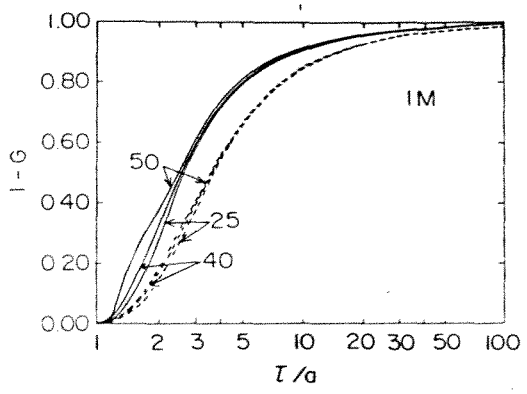


Figure 20 Reduced EDL Onsager coefficient as a function of the slit width for ζ = 25, 40, and 50 mV. The fluid is a 1:1, 1 *M* electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). The solid and dashed lines are the TPE HNC/MS and VO results, respectively. (From Ref. 55.)

electrolyte bulk concentration and/or the higher the ζ potential and/or the narrower the slit, the stronger and more concentrated the EDL. This can be seen in Fig. 21, where the TPE HNC/MS and VO reduced average counterion concentrations are shown as a function of the distance between the plates for ζ = 40 mV and three different bulk electrolyte concentra-

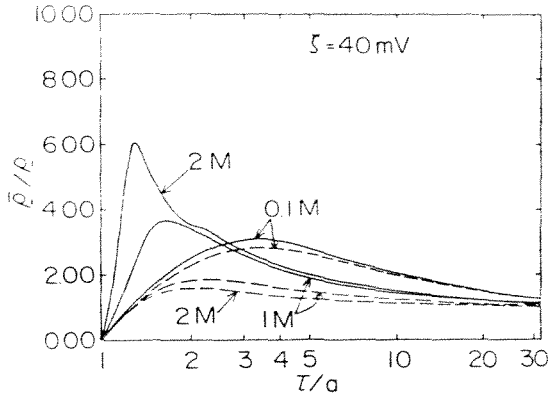


Figure 21 Reduced average counterion concentration inside the slit as a function of the slit width for ζ = 40 mV and 1:1, 0.1 *M*, 1 *M*, and 2 *M* electrolytes (a = 4.25 Å, T = 298 K, ϵ = 78.5). The solid and dashed lines are the TPE HNC/MS and VO results, respectively. (From Ref. 55.)

tions. The average concentration of species i is defined as

$$\overline{\rho_i} = \frac{1}{t}\int_0^h \rho_i(x)\,dx \tag{126}$$

The average co-ion concentration is in general much lower than the counterion concentration, and is not shown. In Fig. 21, important quantitative and qualitative differences can be seen between the TPE HNC/MS and VO results.

The negative contribution of the electro-osmotic velocity to the volume flow per unit width is equal to $2h$ multiplied by the average electro-osmotic velocity [see (86), (93), and (94)]. For very wide slits, the electro-osmotic velocity goes from zero at $x = \pm h$ to a terminal constant velocity at the point where the EDL completely vanishes [i.e., $\psi(x) = 0$]. Therefore, for a sufficiently wide slit, toward the middle of the slit there is a region where the electro-osmotic velocity becomes constant. Hence for very wide slits the average electro-osmotic velocity will be approximately equal to this terminal velocity. For narrower slits, however, the friction with the walls will become important and the average velocity will decrease; that is, the second term on the right-hand side of (93) will become significant (see, e.g., the $\tau = 10a$ curve of Fig. 19). This implies a lower average electro-osmotic velocity with respect to the very wide slit case. However, since the Poiseuille flux decreases with h^3, in fact, the electro-osmotic retardation will become more significant, and a larger flux retardation will be observed. For still narrower slits, the EDLs on each side of the slit interact with each other. As a result, the mean electrostatic potential profile increases and thus the electro-osmotic velocity decreases. This is so because although the counterion concentration increases for narrower slits (see Fig. 21) and this favors an increase in the electro-osmotic velocity, the friction with the walls becomes even more important. The L_{12} Onsager coefficient is proportional to the average electro-osmotic velocity. Thus it will have a similar behavior.

The contribution of the E_z convection current [Eq. (107)] and the contribution of the transport current [Eq. (109)] to the net current [Eq. (87)] are taken into account through the L_{22} Onsager coefficient [Eq. (113)]. Applying the mean value theorem for integrals to (108), it can be shown that the convection current is equal to $\rho_c(x_0)\cdot(1 - G)$, where $0 < x_0 < h$. Hence $\rho_c(x_0)$ is proportional to $\overline{\rho_c}$, defined through (36) and (126). As pointed out before, $\overline{\rho_-}$ dominates over $\overline{\rho_+}$, for small values of τ. Therefore, the value of K_c as a function of τ, given by (115), is expected to be proportional to $\overline{\rho_-}$, shown in Fig. 21. On the other hand, from (112) and (126),

$$K_T = \frac{t}{h} \sum_j t_j \frac{\overline{\rho_j}}{\rho_j} - 1 \tag{127}$$

Since the transport coefficients, $t_+ = t_- = 0.5$ are taken to be constant, K_T as a function of τ is also expected to be proportional to the average number of counterions in the slit (i.e., to $\overline{\rho_-}$). Therefore, K, given by (114), is expected to be a function of τ similar to that of $\overline{\rho_-}$. That this is indeed the case can be seen in Fig. 22. In both the $\overline{\rho_-}$ and $1 + K$ curves a maximum is observed in the region of small values of τ. $L_{22}/2h$ can be taken as a generalized conductivity. Thus the effect of ionic size is to increase the electrolyte conductivity considerably for narrow slits. For macroions one would expect this effect to be even larger.

In Fig. 23 the reduced apparent ζ potential, F, is shown as a function of τ for a 1:1, 1M electrolyte. The behavior of the curves can be inferred from Figs. 20 and 22 [see (118)]. The physically relevant quantities in (117) are the P_z convection current [Eq. (106)] and a generalized resistance $1/L_{22}$. For narrow slits, a higher ζ potential and/or bulk electrolyte concentration increases both the P_z convection current and the L_{22} Onsager coefficient, in different proportions. A higher conductivity, $L_{22}/2h$, implies a lower resistance, $1/L_{22}$. Hence the resistance $1/L_{22}$ and the P_z convection current, I_c^P, compete in (117) to produce a higher or lower streaming potential, with increasing values of the ζ potential and/or bulk

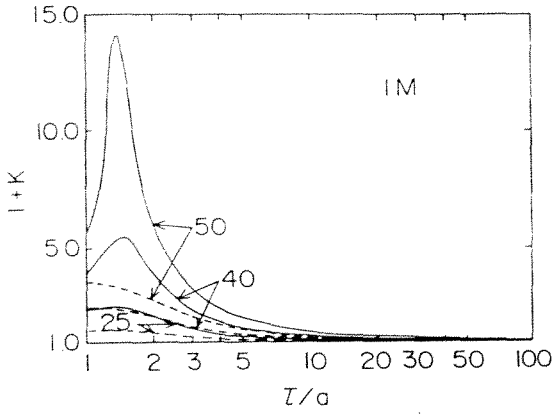


Figure 22 Reduced Onsager coefficient L_{22}/L_{22}^{∞} as a function of the slit width for ζ = 25, 40, and 50 mV. The fluid is a 1:1, 1 M electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). The solid and dashed lines are the TPE HNC/MS and VO results, respectively. (From Ref. 55.)

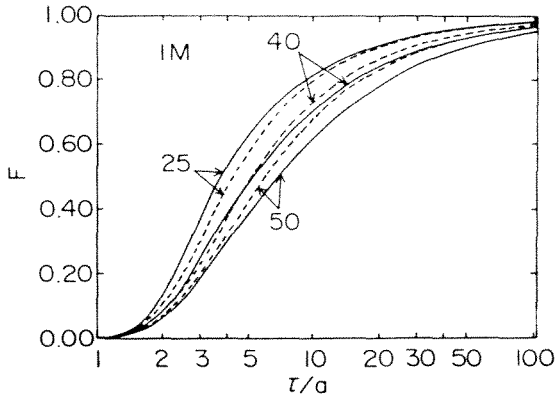


Figure 23 Reduced apparent ζ potential as a function of the slit width for ζ = 25, 40, and 50 mV. The fluid is a 1:1, 1 *M* electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). The solid and dashed curves are the TPE HNC/MS and VO results, respectively. (From Ref. 55.)

electrolyte concentration. The result is that, in general, the streaming potential increases with decreasing bulk electrolyte concentration and can increase or decrease with increasing ζ potential. In Fig. 23, a higher ζ potential implies a higher streaming potential for large values of h, and a lower streaming potential for small values of h.

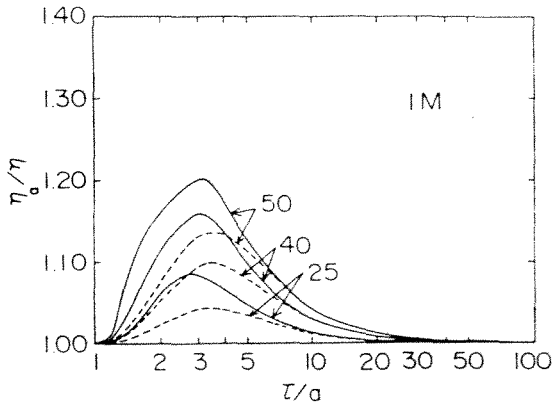


Figure 24 Reduced apparent viscosity coefficients as a function of the slit width for ζ = 25, 40, and 50 mV. The fluid is a 1:1, 1 *M* electrolyte (a = 4.25 Å, T = 298 K, ϵ = 78.5). The solid and dashed curves are the TPE HNC/MS and VO results, respectively. (From Ref. 55.)

Under an applied pressure gradient and zero-current condition, a generalized Poiseuille volume flux can be defined through (121). From (120), this volume flux can be written as

$$V = L_{11}P_z + L_{12}E_{str} \tag{128}$$

The first term on the right-hand side of (128) is the Poiseuille flux, which is retarded by the electro-osmotic flux produced by the induced streaming potential [i.e., the second term on the right-hand side of (128)]. The resulting flux would appear to be that of a more viscous fluid. From the previous discussion on the average electro-osmotic velocity (i.e., the L_{12} Onsager coefficient), $| L_{12} |$ increases with increasing bulk concentration and ζ potential and decreases with decreasing h. On the other hand, E_{str} decreases with increasing bulk concentration and decreasing value of h and does not have a simple dependence with the ζ potential.

For very wide slits, $L_{12}E_{str}$ becomes a constant multiplied by h [see (96) and (117) and Figs. 20 and 23]. Thus the flux retardation becomes negligible compared to the Poiseuille flux, which increases with h^3, and the apparent viscosity tends to the bulk fluid viscosity. For narrower slits, the retardation flux becomes more important and the apparent viscosity increases. Finally, for very narrow slits the apparent viscosity goes to a maximum and decreases as the streaming potential and $| L_{12} |$ decrease faster with decreasing h. This behavior of the reduced apparent viscosity is shown in Fig. 24. In the limit of point ions the results of Levine et al. [51] are obtained.

REFERENCES

1. McQuarrie, D. A. (1976). *Statistical Mechanics*, Harper & Row, New York.
2. Eyring, H., Henderson, D., Stover, B. J., and Eyring, E. M. (1982). *Statistical Mechanics and Dynamics*, 2nd ed., Wiley, New York.
3. Friedman, H. L. (1985). *A Course in Statistical Mechanics*, Prentice Hall, Englewood Cliffs, N.J.
4. Hansen, J. P., and McDonald, I. R. (1986). *Theory of Simple Liquids*, 2nd ed., Academic Press, New York.
5. Lozada-Cassou, M., Saavedra-Barrera, R., and Henderson, D. (1982). The application of the hypernetted chain approximation to the electrical double layer: comparison with Monte Carlo results for symmetric salts, *J. Chem. Phys. 77*: 5150.
6. Lozada-Cassou, M., and Henderson, D. (1983). Application of the hypernetted chain approximation to the electrical double layer: comparison with Monte Carlo results for 2:1 and 1:2 salts, *J. Phys. Chem. 87*: 2821.

7. Henderson, D., and Lozada-Cassou, M. (1983). The application of the HNC/MSA approximation to bulk electrolytes, *J. Chem. Phys. 79*: 3055.
8. Carnie, S. L. (1985). Hypernetted-chain theories of the primitive model double layer: a last look at binary symmetric electrolytes, *Mol. Phys. 54*: 509.
9. Vlachy, V., and Haymet, A. D. J. (1986). A grand canonical Monte Carlo simulation study of polyelectrolyte solutions, *J. Chem. Phys. 84*: 5874.
10. Rossky, P. J., Dudowicz, J. B., Tembe, B. L., and Friedman, H. L. (1980). Ionic association in model 2–2 electrolyte solutions, *J. Chem. Phys. 73*: 3372.
11. Attard, P., and Patey, G. N. (1990). *J. Chem. Phys. 92*: 4970.
12. Outhwaite, C. W. (1969). Extension of the Debye–Hückel theory of electrolyte solutions, *J. Chem. Phys. 50*: 2277.
13. Outhwaite, C. W. (1971). Partial treatment of the fluctuation potential in the Debye–Hückel theory of electrolyte solutions, *Mol. Phys. 20*: 705.
14. Burley, D. M., Hutson, V. C. L., and Outhwaite, C. W. (1974). A treatment of the volume and fluctuation term in Poisson's equation in the Debye–Hückel theory of strong electrolyte solutions, *Mol. Phys. 27*: 225.
15. Lozada-Cassou, M. (1981). A new method of deriving electrical double layer equations from electrolyte theories, *J. Chem. Phys. 75*: 1412.
16. Lozada-Cassou, M. (1982). Comments on "A new method of deriving electrical double layer equations from electrolyte theories," *J. Chem. Phys. 77*: 5258.
17. lozada-Cassou, M. (1983). Hypernetted chain theory for the distribution of ions around a cylindrical electrode, *J. Phys. Chem. 87*: 3729.
18. González-Tovar, E.. Lozada-Cassou, M., and Henderson, D. (1985). Hypernetted chain approximation for the distribution of ions around a cylindrical electrode. II. Numerical solution for a model cylindrical polyelectrolyte, *J. Chem. Phys. 83*: 361.
19. Bacquet, R., and Rossky, P. J. (1984). Ionic atmosphere of rodlike polyelectrolytes: a hypernetted chain study, *J. Phys. Chem. 88*: 2660.
20. González-Tovar, E., and Lozada-Cassou, M. (1989). The spherical double layer: a hypernetted chain mean spherical approximation calculation for a model spherical colloid particle, *J. Phys. Chem. 93*: 3761.
21. Percus, J. K. (1964). The pair distribution function in classical statistical mechanics, in *The Equilibrium Theory of Classical Fluids* (H. L. Frisch and J. L. Lebowitz, eds.), W. A. Benjamin, New York, p. II-33.
22. Henderson, D., Abraham, F. F., and Barker, J. A. (1976). The Ornstein–Zernike equation for a fluid in contact with a surface, *Mol. Phys. 31*: 1291.
23. Henderson, D., and Blum, L. (1978). Some exact results and the application of the mean spherical approximation to charged hard spheres near a charged hard wall, *J. Chem. Phys. 69*: 5441.
24. Henderson, D., Blum, L., and Smith, W. R. (1979). Application of the hypernetted chain approximation to the electric double layer at a charged planar interface, *Chem. Phys. Lett. 63*: 381.

25. Sitillinger, F. H., and Kirkwood, J. G. (1960). Theory of the diffuse double layer, *J. Chem. Phys. 33*: 1282.
26. Kirkwood, J. G., and Poirier, J. C. (1954). The statistical mechanical basis of the Debye–Hückel theory of strong electrolyte, *J. Phys. Chem. 58*: 591.
27. Croxton, T. L., and McQuarrie, D. A. (1979). Numerical solution of the Born–Green–Yvon equation for the restricted primitive model of ionic solutions, *J. Phys. Chem. 83*: 1840.
28. Croxton, T. L., and McQuarrie, D. A. (1979). A theory of the electrical double layer through the Born–Green–Yvon equation, *Chem. Phys. Lett. 68*: 489.
29. Lozada-Cassou, M. (1981). A simple procedure for deriving an electrical double layer equation from the Born–Green–Yvon theory, *Chem. Phys. Lett. 81*: 472.
30. Croxton, T. L., and McQuarrie, T. A. (1981). The electrical double layer in the Born–Green–Yvon equation, *Mol. Phys. 42*: 141.
31. Henderson, D., Blum, L., and Bhuiyan, L. B. (1981). Some comments on the Born–Green–Yvon equation for an electrified interface, *Mol. Phys. 43*: 1185.
32. Outhwaite, C. W. (1970). A modified Poisson–Boltzmann equation in the double layer, *Chem. Phys. Lett. 7*: 636.
33. Burley, D. M., Hutson, V. C. L., and Outhwaite, C. W. (1971). Numerical solution of a modified Poisson–Boltzmann equation in electrolyte solution theory, *Chem. Phys. Lett. 9*: 109.
34. Lozada-Cassou, M. (1984). The force between two planar electrical double layers, *J. Chem. Phys. 80*: 3344.
35. Lozada-Cassou, M., and Henderson, D. (1986). THe force between two planar electrical double layer some numerical results, *Chem. Phys. Lett. 127*: 392.
36. Lozada-Cassou, M., and Díaz-Herrera, E. (1988). The interaction of electrical double layers: superpositon vs. non-superposition, in *Ordering and Organization in Ionic Solutions* (N. Ise and I. Sogam, eds.), World Scientific Publishing Co., Singapore, p. 555.
37. Lozada-Cassou, M., and Díaz-Herrera, E. (1990). Three point extension for hypernetted chain and other integral equation theories: numerical results, *J. Chem. Phys. 92*: 1194.
38. Lozada-Cassou, M., and Díaz-Herrera, E. (1990). Three point extension hypernetted chain, conventional hypernetted chain, and superposition approximations: numerical results for the force between two plates, *J. Chem. Phys. 93*: 1386.
39. Sánchez, J., and Lozada-Cassou, M. (1992). Exact numerical solution to the integral equation version of the Poisson–Boltzmann equation for two interacting spherical colloidal particles, *Chem. Phys. Lett. 190*: 202.
40. Kjellander, R., and Marčelja, S. (1985). Inhomogeneous Coulomb fluids with image interactions between planar surfaces, I, *J. Chem. Phys. 82*: 2122.

41. Attard, P. (1989). Spherically inhomogeneous fluids. I. Percurs–Yevick hard spheres: osmotic coefficients and triplet correlations, *J. Chem. Phys. 91*: 3072.
42. Zhou, Y., and Stell, G. (1989). Fluids inside a pore: an integral-equation approach. I. General formalism and hard spheres inside spherical and slit pores, *Mol. Phys. 66*: 767.
43. Zhou, Y., and Stell, G. (1989). Fluids inside a pore: an integral-equation approach. II. Cylindrical pores, *Mol. Phys. 66*: 791.
44. Verwey, E. J. W., and Overbeek, J. Th. G. (1948). *Theory of the Stability of Lyophobic Colloids*, Elsevier, New York.
45. Wennerström, H., Jönsson, B., and Linse, P. (1982). The cell model for polyelectrolyte systems: exact statistical mechanics relations, Monte Carlo simulations, and the Poisson–Boltzmann approximation, *J. Chem. Phys. 76*: 4665.
46. Linse, P., and Jönsson, B. (1983). A Monte Carlo study of the electrostatic interaction between highly charged aggregates: a test of the cell model applied to micellar systems, *J. Chem. Phys. 78*: 3167.
47. Guldbrand, L., Jönsson, B., Wennerström, H., and Linse, P. (1984). Electrical double layer forces: a Monte Carlo study, *J. Chem. Phys. 80*: 2221.
48. Burgreen, D., and Nakache, F. R. (1964). Electrokinetic flow in ultrafine capillary slits, *J. Phys. Chem. 68*: 1084.
49. Rice, C. L., and Whitehead, R. (1965). Electrokinetic flow in a narrow cylindrical capillary, *J. Phys. Chem. 69*: 4017.
50. Hildreth, D. (1970). Electrokinetic flow in fine capillary channels, *J. Phys. Chem. 74*: 2006.
51. Levine, S., Marriott, J. R., and Robinson, K. (1975). Theory of electrokinetic flow in a narrow parallel-plate channel, *J. Chem. Soc. Faraday Trans. 2 71*: 1.
52. Levine, S., Marriott, J. R., Neale, G., and Epstein, N. (1975). Theory of electrokinetic flow in fine cylindrical capillaries at high zeta-potentials, *J. Colloid Interface Sci. 52*: 136.
53. Brož, Z., and Epstein, N. (1976). Electrokinetic flow through porous media composed of fine cylindrical capillaries, *J. Colloid Interface Sci. 56*: 605.
54. Olivares, W., Croxton, T. L., and McQuarrie, D. A. (1980). Electrokinetic flow in a narrow cylindrical capillary, *J. Phys. Chem. 84*: 867.
55. González-Tovar, E., Lozada-Cassou, M., and Olivares, W. (1991). Electrokinetic flow in ultrafine slits: ionic size effects, *J. Chem. Phys. 94*: 2219.
56. Rasaiah, J. C., and Friedman, H. (1968). Integral equation method in the computation of equilibrium properties of ionic solutions, *J. Chem. Phys. 48*: 2742.
57. Rasaiah, J. C., Card, D. N., and Valleau, J. P. (1972). Calculations on the restricted primitive model for 1–1 electrolyte solutions, *J. Chem. Phys. 56*: 248.
58. Mier y Terán, L., Díaz-Herrera, E., and Lozada-Cassou, M. (1989). A comparison of numerical methods for solving nonlinear integral equations found in liquid theories, *J. Comput. Phys. 84*: 326.

59. Henderson, D., Barker, J. A., and Lozada-Cassou, M. (1985). Comparison of the nonlinear Poisson–Boltzmann approximation with Monte Carlo results for the primitive model of an electrolyte, *J. Chem. Soc. Faraday Trans. 2 81*: 457.
60. Mier y Terán, L., Díaz-Herrera, E., Lozada-Cassou, M., and Henderson, D. (1988). Temperature dependence of the primitive model double layer differential capacitance: a hypernetted chain/mean spherical approximation calculation, *J. Phys. Chem. 92*: 6408.
61. Outhwaite, C. W., and Bhuiyan, L. B. (1982). A further treatment of the exclusion-volumen term in the modified Poisson–Boltzmann theory of the electric double layer, *J. Chem. Soc. Faraday Trans. 2 78*: 775.
62. Outhwaite, C. W. (1983). Comments on the modified Poisson–Boltzmann theory for the electric double layer, *J. Chem. Soc. Faraday Trans. 2 79*: 1315.
63. Alejandre, J., Lozada-Cassou, M., González-Tovar, E., and Chapela, G. A. (1990). Molecular dynamics of a hard-sphere fluid between two walls: a comparison with the three-point extension hypernetted chain approximation, *Chem. Phys. Lett. 175*: 111.
64. Yeomans, L., and Lozada-Cassou, M. Electrolyte structure in a cylindrical pore (in preparation).
65. Hill, T. L. (1956). *Statistical Mechanics: Principles and Selected Applications*, McGraw-Hill, New York.
66. Enos, B., McQuarrie, D., and Colonomos, P. (1975). Force balance in systems of spherical polyelectrolytes, *J. Colloid Interface Sci. 52*: 289.
67. Féat, G. R., and Levine, S. (1976). The double-layer interaction of two charged colloidal spherical particles of a concentrated dispersion in a medium of low dielectric constant. *J. Colloid Interface Sci. 54*: 34.
68. Bell, G. M., Levine, S., and McCartney, L. N. (1970). Approximate methods of determining the double layer free energy of interaction between two charged colloidal spheres, *J. Colloid Interface Sci. 33*: 335.
69. Overbeek, J. Th. G. (1977). Recent developments in the undertaking of colloid stability, *J. Colloid Interface Sci. 58*: 408.
70. Nielaba, P., and Forstmann, F. (1985). Packing of ions near an electrolyte–electrode interface in the HNC/LMSA approximation to the RPM model, *Chem. Phys. Lett. 117*: 46.
71. Plischke, M., and Henderson, D. (1988). Pair and singlet correlation functions of inhomogeneous fluids calculated using the Ornstein–Zernike equation, *J. Phys. Chem. 92*: 7177.
72. Plischke, M., and Henderson, D. (1989). Pair correlation functions and the structure of the electric double layer, *Electrochim. Acta 34*: 1863.
73. Plischke, M., and Henderson, D. (1989). The primitive model of the electric double layer: nonsymmetric electrolytes, *J. Chem. Phys. 90*: 5738.
74. Kjellander, R., and Marčelja, S. (1986). Interaction of charged surfaces in electrolyte solutions, *Chem. Phys. Lett. 127*: 402.
75. Kjellander, R. (1988). Inhomogeneous Coulomb fluids with image inter-

actions between planar surfaces. II. On the anisotropic hypernetted chain approximation, *J. Chem. Phys. 88*: 7129.
76. Mier y Terán, L., Suh, S. H., White, H. S., and Davis, H. T. (1990). A nonlocal free-energy density-functional approximation for the electrical double layer, *J. Chem. Phys. 92*: 5087.
77. Tang, Z., Mier y Terán, L., Davis, H. T., Scriven, L. E., and White, H. S. (1990). Non-local free-energy density-functional theory applied to the electrical double layer. I. Symmetrical electrolytes, *Mol. Phys. 71*: 369.
78. Mier y Terán, L., Tang, Z., Davis, H. T., Scriven, L. E., and White, H. S. (1991). Non-local free-energy density-functional theory applied to the electrical double layer. II. 2:1 electrolytes, *Mol. Phys. 72*: 817.
79. Robledo, A., and Varea, C. (1981). On the relationship between the density functional formalism and the potential distribution theory for non-uniform fluids, *J. Stat. Phys. 26*: 513.
80. Tarazona, P. (1985). Free energy density functional for hard spheres, *Phys. Rev. A 31*: 2672.
81. Freasier, B. C., Woodward, C. E., and Nordholm, S. (1989). Generalized van der Waals theory of hard sphere oscillatory structure, *J. Chem. Phys. 90*: 5657.
82. Percus, J. K. (1976). Model for density variation at a fluid surface, *J. Stat. Phys. 15*: 423.
83. Percus, J. K. (1976). Equilibrium state of a classical fluid of hard rods in an external field, *J. Stat. Phys. 15*: 505.
84. Beresford-Smith, B., and Chan, D. Y. C. (1982). Highly asymmetric electrolytes: a model for strongly interacting colloidal systems, *Chem. Phys. Lett. 92*: 474.
85. Beresford-Smith, B., and Chan, D. Y. C. (1983). Electrical double-layer interactions in concentration colloidal systems, *Faraday Discuss. Chem. Soc. 76*: 65.
86. Patey, G. N. (1980). The interaction of two spherical colloidal particles in electrolyte solution: an application of the hypernetted-chain approximation, *J. Chem. Phys. 72*: 5763.
87. Medina-Noyola, M., and McQuarrie, D. (1980). On the interaction of spherical double layers, *J. Chem. Phys. 73*: 6279.
88. Teubner, M. (1981). On the applicability of the HNC approximation to highly charged polyelectrolyte, *J. Chem. Phys. 75*: 1907.
89. Ise, N., and Okubo, T. (1966). Mean activity coefficient of polyelectrolytes. II. Measurements of sodium salts of polyvinyl alcohols partially acetalized with glyoxylic acid, *J. Phys. Chem. 70*: 1930.
90. Ise, N., Okubo, T., Ito, K., and Dosho, S. (1985). Visible evidence for interparticle attraction in polymer latex dispersions, *Langmuir 1*: 176.
91. Ito, K., Ise, N., and Okubo, T. (1985). Determination of the number of effective charges of colloidal particles by transference measurements, *J. Chem. Phys. 82*: 5732.
92. Matsuoka, H., Ise, N., Okubo, T., Kunugi, S., Tomiyama, H., and Yosh-

ikawa, Y. (1985). Ordered structure in dilute solutions of byopolymers as studied by small-angle x-ray scattering, *J. Chem. Phys. 83*: 378.
93. Israelachvili, J. N. (1985). Measurements of hydration forces between macroscopic surfaces, *Chem. Scr. 25*: 7.
94. Bell, G. M., and Levine, S. (1958). Statistical thermodynamics of concentrated colloidal solutions, *Trans. Faraday Soc. 54*: 785.
95. Arrieta, E., Jedrzejek, C., and Marsh, K. N. (1987). Numerical MSA solution for binary Yukawa mixtures, *J. Chem. Phys. 86*: 3607.
96. Bratko, D., Sheu, E. Y., and Chen, S.-H. (1987). Analysis of intermicellar structure factors with the mean spherical and hypernetted-chain approximations, *Phys. Rev. A 35*: 4359.
97. Bendedouch, D., and Chen, S.-H. (1984). Effect of an attractive potential on the interparticle structure of ionic micelles at high salt concentration, *J. Phys. Chem. 88*: 648.
98. Wu, C.-F., Sheu, E. Y., Bendedouch, D., and Chen, S.-H. (1987). Studies of double layer interactions in micelle and protein solutions by small angle neutron scattering, *Kinam 8A*: 37.
99. Khan, S., Morton, T. L., and Ronis, D. (1987). Static correlations in macroionic suspensions: analytic and numerical results in a hypernetted-chain-mean-spherical approximation, *Phys. Rev. A 35*: 4295.
100. Medina-Noyola, M. (1982). Spatial correlations between charged colloidal particles: the mean spherical (dense point limit) approximation, *J. Chem. Phys. 77*: 1428.
101. Ronis, D. (1984). A variational approach to static correlations in dilute colloidal suspensions, *J. Chem. Phys. 81*: 2749.
102. Alexander, S., Chaikin, P. M., Grant, P., Morales, G. J., and Pincus, P. (1984). Charge renormalization, osmotic pressure, and bulk modulus of colloidal crystals: theory, *J. Chem. Phys. 80*: 5776.
103. Derian, P.-J., Belloni, L., and Drifford, M. (1987). Contribution of small ions to the scattered intensity in the hypernetted chain approximation: application to micellar solutions, *J. Chem. Phys. 86*: 5708.
104. González-Mozuelos, P., and Medina-Noyola, M. (1990). Rescaled mean spherical approximation for wall–particle correlations in colloidal suspensions, *J. Chem. Phys. 93*: 2109.
105. Robins, M. O., Kremer, K., and Grest, G. S. (1988). Phase diagrams and dynamics of Yukawa systems, *J. Chem. Phys. 88*: 3286.
106. Tata, B. V. R., Sood, A. K., and Kesavamoorthy, R. (1990). Structure factor of charged colloidal suspensions using Brownian-dynamics simulations: comparison of Yukawa and Sogami pair potentials, *Pramana J. Phys. 34*: 23.
107. Lyklema, J. (1977). Water at interfaces: a colloid-chemical approach, *J. Colloid Interface Sci. 58*: 242.
108. Van Den Hul, H. J. (1983). Estimation of outer Helmholtz plane potentials from negative adsorption of co-ions, *J. Colloid Interface Sci. 92*: 217.

9

Freezing

A. D. J. Haymet
University of Sydney
N. S. W., Australia

I. THEORIES OF FREEZING

In this chapter we adopt the point of view that a crystal is a *very* inhomogeneous liquid, albeit one with a high degree of symmetry. It may seem surprising at first that this point of view has anything to recommend it. Nevertheless, over the last decade the density functional (DF) theory of classical statistical mechanics [1] has been developed by Ramakrishnan and Yussoff [2] and Haymet and Oxtoby [3,4] to describe successfully the freezing transition from exactly this point of view. The theory exploits advances in both the density functional formalism, reviewed by Evans in Chapter 3 of this volume, and advances since 1960 in describing the structure of bulk homogeneous liquids with workable theories.

The goal of the DF freezing theory is to predict the thermodynamic conditions under which a liquid will freeze and to predict the symmetry and density of the crystal with which the liquid coexists. Central to the theory for both one-component and multicomponent systems is a thermodynamic functional Taylor series expansion of the free energy of the crystal about the free energy of the liquid. Implicit in this expansion is the assumption of a similarity between the local environments of the crystal and the liquid phases along the coexistence line. Quantitatively, the similarities in the local environments can be measured in terms of a co-

ordination number, the number of nearest neighbors around a particle, which is often the same for both phases. For multicomponent systems, the local environment depends not only on the location and number, but also the species, of nearest neighbors.

Starting in 1983 [5], the DF freezing theory has been applied numerically to many one-component systems, such as hard spheres [5–12], the Lennard–Jones liquid [19,20], the hard-core Yukawa fluid [21–25], water [26], mixtures [27], hard ellipsoids and dumbbells [28], and even reformulated for quantum liquids [29–31]. Over the years, a number of review articles have appeared [13–18]. In this chapter we discuss in detail applications of this DF theory to quantum systems, to classical hard- and soft-sphere systems using newly formulated functionals [22], and to the calculation of the vacancy density in the equilibrium crystal at melting [32].

The freezing theory requires as input the correlation functions of the liquid and a choice of crystal symmetry. Freezing theory can thus be split into two independent steps, summarized by

$$v_{ij}(r) \rightarrow c_{ij}(r) \rightarrow \text{phase diagram} \tag{1}$$

The potential energy v_{ij} between each pair of species i and j is used to calculate the direct correlation functions c_{ij}, which in turn are used in the density functional to predict the phase diagram. A wide variety of modern methods of liquid-state chemistry can be used to generate the required input, and in addition, experimental structure factor data can be used directly.

The outline of this chapter is as follows. In Section II a brief guide to density functional methods is presented by picking up results and methods derived in Chapter 3 by Evans. In Section III the quantum theory is derived, and a review of the first application—to the freezing of liquid helium—is presented in Section IV. Modern classical methods, which go beyond a second-order perturbation approach, are summarized in Section V together with applications by several groups to single-component and binary hard-sphere systems, and to the prediction of the vacancy concentration in equilibrium crystals at the melting point. Some prospects for future research are collected in Section VI. Note that this chapter has strong connections with Chapters 3, 10, and 11.

II. BRIEF GUIDE TO THE DF LITERATURE

The first density functional calculation for the freezing of a monatomic liquid using accurate liquid input was presented in 1983 [5]. Following

related ideas from gas–liquid coexistence theory, developed among others by Yang et al. [33] and Saam and Ebner [34], the functional used in this work is

$$\Delta\beta\Omega_V[\rho] = \int_V d\mathbf{r}_1 \left[\rho(\mathbf{r}_1) \ln \frac{\rho(\mathbf{r}_1)}{\rho_L} - \Delta\rho(\mathbf{r}_1) \right] - \frac{1}{2} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2 \, c(|\mathbf{r}_1 - \mathbf{r}_2|) \, \Delta\rho(\mathbf{r}_1) \, \Delta\rho(\mathbf{r}_2) \tag{2}$$

where ρ_L is the density of the liquid, $\rho(\mathbf{r}_1)$ is the spatially varying density of the crystal, $\Delta\rho(\mathbf{r}_1) = [\rho(\mathbf{r}_1) - \rho_L]$, and $c(r_{12})$ is the direct correlation function of the liquid. Upon minimizing the functional $\Delta\beta\Omega_V$, one obtains the grand potential difference $\Delta\beta\Omega$ between the coexisting liquid at crystal phases at temperature T and chemical potential μ.

Density functional methods were introduced into classical statistical mechanics around 1963 [35–37], but were somewhat underutilized until the 1970s. In the context of freezing, calculations by a number of different groups show that the density functional theory of freezing has made substantial progress in the long-standing question of liquid–crystal phase coexistence. Before proceeding to the details of some representative calculations, it is perhaps appropriate to summarize the shortcomings of this approach.

Since the density functional theory is by its very construction a mean-field theory, it suffers the same problems as does any such theory when fluctuation effects become dominant. For example, Evans [1,18] has summarized the problems in the analogous gas–liquid theory near the critical point. In two dimensions, the DF theory in its simple form does not predict the algebraic decay of correlation functions in two-dimensional crystals (although it still has some practical uses). In the isotropic–nematic transition in liquid crystals (see Chapter 11), the DF theory, together with any mean-field theory, cannot predict the small entropy change at this very "weak" first-order transition. It is important to remember that these interesting examples do not negate the central problem of describing the "usual," "strong" first-order transitions found in most materials.

The only approximation in the DF theory is the truncation of a Taylor series expansion of the (grand) free-energy functional. The vexing question of analyticity of the functional at a first-order transition remains. Most researchers agree that the fact that there is a weak singularity does not prevent the application of theory to practical problems. At least one fully nonperturbative freezing theory has been proposed [38], although any such approach involves much more computational work.

Optimization of the free-energy functional is an extremely active and promising area of research [22,23,39–41]. In practical applications one is forced (usually, by ignorance of higher-order correlation functions in the liquid) to truncate the above-mentioned Taylor expansion at second order. Given this constraint, at least two independent methods for improving the DF theory are available. Very recently it has been realized that the ideal system, which underlies the derivation of the DF theory, can be chosen to minimize higher-order contributions. For example, in all the early DF calculations, the ideal system was chosen (almost by default) to be the usual, classical ideal gas of point particles. This is a sensible choice since it implies that the liquid correlation function $c(r)$ in the functional is the usual Ornstein–Zernike direct-correlation function. But the recent work shows that other choices of the ideal system are possible and even preferable. The initial calculations have been made for quantum freezing (e.g., for helium) but the applications to freezing of polymers [43] and other complex liquids are immediate. In simple terms, the procedure is simply to choose the ideal system to minimize higher-order contributions to the functional.

The alternative approach, which has become popular very recently, is the approximate inclusion of higher-order terms in the functional. Most recently, at least four groups [22,41,48,97] have proposed different "infinite-order" functionals, which at the present time appear to possess very different properties. Originally, Tarazona [11] used smoothed density approximations to build into the functional partial inclusion of higher-order terms, involving higher-order liquid correlation functions. This approach is related to an earlier fine-grained generalized van der Waals theory of Nordholm and co-workers [44,45], and has been generalized and improved in important calculations by several groups. Other empirical or semiempirical choices can be made: for example, by choosing a polynomial weighting function such that for hard spheres either the Carnahan–Starling equation is recovered or thermodynamic consistency is enforced. The great value of this approach has been demonstrated by Tarazona, who has presented a functional that displays three-phase coexistence at the triple point.

The activity in the extensions of DF theory indicates that the renaissance of density functional theory in statistical mechanics, especially phase transitions, is far from complete. The most important mathematical properties of the functional remain unknown, and opportunities for extension to dynamical questions and nonsimple molecules (such as polymers [43]) are abundant.

III. PATH INTEGRAL FORMULATION OF GENERAL QUANTUM THEORY

A. Quantum DF Theory

The DF theory for the freezing of quantum liquids at $T \neq 0$ has been developed by McCoy et al. [29–31] to predict the freezing of quantum liquids, such as helium. As summarized above, the freezing of a wide variety of classical liquids has been described by the density functional (DF) theory of statistical mechanics. By choosing a new *ideal* system, a new DF theory is constructed that addresses directly many of the unusual features of the freezing of liquid helium, such as the weakly modulated liquid pair correlation function $g(r)$ at freezing densities. The quantum theory combines DF techniques with the Feynman path integral formulation of quantum mechanics to include dispersion effects. In classical DF theories, the density and external field of the ideal system are connected by a Boltzmann relation $\rho(r) \propto \exp[-\beta V_{ext}(r)]$. In the quantum DF theory, one relates the density and external field of the ideal system through the Feynman path integral representation, in which the quantum particle is represented by a classical ring polymer of P beads. The classical DF theories fail for helium because they employ an ideal system that is too far removed from the interacting system. For a class of simple but interesting problems, the density path integral of the ideal system can be performed in closed form, leading to a compact, physically descriptive theory. The full theory has been applied to the freezing of helium-4 [31], and as summarized in Section III.B, it yields good results.

Although the density functional theory of freezing has been applied successfully to a variety of classical liquids [2,3,5,13,14], it does not work for helium [46], a liquid with significant quantum character. Classical DF freezing theory demonstrates that liquids freeze when they are strongly correlated, as measured, for example, by the pair correlation function or structure factor, in accord with empirical freezing rules [47,85]. When helium freezes, the pair correlation function is not strongly modulated, which implies that factors other than classical intermolecular correlations are important.

The quantum DF theory takes a different approach and treats successfully the quantum swelling (dispersion) of particles, which results in freezing of quantum systems at densities which by classical standards are extremely low. This quantum swelling is quantified by the single-particle self-correlation function discussed below.

The core of the density functional theory of freezing is a functional Taylor series of the excess Helmholtz free energy $\mathcal{F}_{ex}$, an expansion that is truncated after the term second order in the singlet density difference $\Delta\rho(\mathbf{r})$. The convergence (or otherwise) of this expansion is discussed at length below. The point of quantum DF theory is that it chooses a better ideal system, which reduces significantly higher-order contributions to the excess free energy. For a workable theory, this ideal system should be solvable and tractable. To date, the ideal system has been chosen to include (quantum) dispersion in the representation of noninteracting particles, instead of the usual (classical) noninteracting point particles.

The quantum aspects of the problem are treated using the Feynman path integral techniques. Because of the ring-polymer representation of the quantum molecule, the quantum freezing theory naturally has features in common with the freezing theory for polyatomic molecules. The numerical tests of the quantum theory are summarized in Section III.B [29,31].

The theory may be developed as follows. First, the grand potential functional for an arbitrary ideal system is derived. This result is used to determine the free-energy functional for an ideal system that is a noninteracting quantum system with dispersion. One can also recover the classical theory. Finally, the second-order contribution to the quantum free energy is calculated.

B. Free-Energy Density Functionals

The central quantity in the theory is the grand potential function Ω, which is derived here via the total differential of the thermodynamic energy of a system in an external field. Legendre transformation is then used to switch among the various, equivalent thermodynamic potentials. The most important step in the derivation is the functional Taylor expansion of the excess Helmholtz free energy $\mathcal{F}_{ex} = \mathcal{F} - \mathcal{F}_{id}$ about the (reference) liquid state in powers of the singlet density difference. Here $\mathcal{F}$ denotes the total Helmholtz free energy of the actual system, and $\mathcal{F}_{id}$ the free energy of an ideal system, in the notation of Evans in Chapter 3 [1,18]. The remaining steps are (1) the selection of an ideal system, and (2) the approximation of the functional by truncation of the Taylor series after second order. McCoy et al. [30] argue that this approach is superior to an attempt to evaluate higher-order terms in the classical reference system. The final step of the derivation is a Lengendre transform to the grand potential.

The differential form of the thermodynamic energy U is

$$dU = T\,dS - P\,dV + \mu\,dN + \int_V d\mathbf{r}\ \rho(\mathbf{r})\ \delta V_{\text{ext}}(\mathbf{r}) \tag{3}$$

where T is the temperature, S the entropy, P the pressure, V the volume, μ the chemical potential, N the number of particles, $\rho(\mathbf{r})$ the singlet density (possibly spatially varying), $V_{\text{ext}}(\mathbf{r})$ the external field, and $\delta V_{\text{ext}}(\mathbf{r})$ the variation in $V_{\text{ext}}(\mathbf{r})$. Note that N is related to the other variables via $N = \int_V d\mathbf{r}\ \rho(\mathbf{r})$. This expression is consistent with the energy functional used in quantum field theory [49]. The natural variables of U are S, V, N, and $V_{\text{ext}}(\mathbf{r})$. It is convenient to transform to an intermediate thermodynamic energy $\mathcal{U}$ defined to be

$$\mathcal{U} = U - \int_V d\mathbf{r}\ \rho(\mathbf{r}) V_{\text{ext}}(\mathbf{r}) \tag{4}$$

In differential form this is

$$d\mathcal{U} = T\,dS - P\,dV + \int_V d\mathbf{r}\ u(\mathbf{r})\ \delta\rho(\mathbf{r}) \tag{5}$$

where $u(\mathbf{r}) = \mu - V_{\text{ext}}(\mathbf{r})$ and the constraint on $\rho(\mathbf{r})$ is thereby removed. The Helmholtz free energy is

$$\mathcal{F} = \mathcal{U} - TS \tag{6}$$

and the grand potential

$$\Omega = \mathcal{F} - \int_V d\mathbf{r}\ u(\mathbf{r})\rho(\mathbf{r}) \tag{7}$$

Two important relations should be noted at this point. From (5) to (7),

$$\frac{\delta\mathcal{F}}{\delta\rho(\mathbf{r})} = u(\mathbf{r}) \quad \text{and} \quad \frac{\delta\Omega}{\delta u(\mathbf{r})} = -\rho(\mathbf{r}) \tag{8}$$

The excess Helmholtz free energy, truncated at second order in the density difference, becomes

$$\begin{aligned}\beta\mathcal{F}_{\text{ex}} = \beta\mathcal{F}_{\text{ex},L} &+ \int_V d\mathbf{r}\ [\beta u_L(\mathbf{r}) - \beta u_{\text{id},L}(\mathbf{r})]\ \Delta\rho(\mathbf{r}) \\ &- \frac{1}{2}\int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2\ c(\mathbf{r}_1 - \mathbf{r}_2)\ \Delta\rho(\mathbf{r}_1)\ \Delta\rho(\mathbf{r}_2)\end{aligned} \tag{9}$$

where $\beta = 1/kT$, the subscript L indicates the liquid state, the subscript id indicates the ideal state, and $\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_L$. The direct correlation

function $c(\mathbf{r}_1 - \mathbf{r}_2) = c^{(2)}(\mathbf{r}_1,\mathbf{r}_2)$ is one of the family of correlation functions,

$$c^{(n)}(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \frac{-\beta\delta^n(\mathcal{F} - \mathcal{F}_{\mathrm{id}})}{\delta\rho(\mathbf{r}_1)\cdots\delta\rho(\mathbf{r}_n)} \tag{10}$$

For the liquid state, both ρ and u_L are constants. Since $\rho(\mathbf{r})$ is a natural variable of $\mathcal{F}$, the expansion is straightforward, and the transform to the grand thermodynamic potential yields

$$\beta\Delta\Omega = \beta\Delta\mathcal{F}_{\mathrm{id}} + \int_V d\mathbf{r}[\beta u_L - \beta u(\mathbf{r})]\rho(\mathbf{r}) - \int_V d\mathbf{r}\, \beta u_{\mathrm{id},L}\, \Delta\rho(\mathbf{r}) - \frac{1}{2}\int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2\, c(\mathbf{r}_1 - \mathbf{r}_2)\, \Delta\rho(\mathbf{r}_1)\, \Delta\rho(\mathbf{r}_2) \tag{11}$$

where $\Delta\Omega = \Omega - \Omega_L$ and $\Delta\mathcal{F}_{\mathrm{id}} = \mathcal{F}_{\mathrm{id}} - \mathcal{F}_{\mathrm{id},L}$.

Since $u(\mathbf{r})$ is a natural variable of Ω, and $\rho(\mathbf{r})$ is the variable conjugate to $u(\mathbf{r})$, Ω is a minimum with respect to variations in $\rho(\mathbf{r})$. One would usually treat Ω as a functional of T, V, and $u(\mathbf{r})$ with $\rho(\mathbf{r})$ defined by these variables, but one can also consider Ω to be a functional of T, V, $u(\mathbf{r})$, and $\rho(\mathbf{r})$, with an added condition that fixes the value of $\rho(\mathbf{r})$, namely that $\rho(\mathbf{r})$ is determined by Ω being minimized for fixed T, V, and $u(\mathbf{r})$. The quantum DF theory exploits this second approach. The functional which is not yet minimized, namely when $\rho(\mathbf{r})$ is still a free variable, is denoted by a generalization of the notation due to Evans [1], $\Omega_V[\rho]$. With the definition $\Delta\Omega_V = \Omega_V - \Omega_{V,L}$, one obtains from (11)

$$\beta\Delta\Omega_V = \beta\Delta\mathcal{F}_{\mathrm{id}} + \int_V d\mathbf{r}\, [\beta u_L - \beta u(\mathbf{r})]\rho(\mathbf{r}) - \int_V d\mathbf{r}\, \beta u_{\mathrm{id},L}\, \Delta\rho(\mathbf{r}) - \frac{1}{2}\int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2\, c(\mathbf{r}_1 - \mathbf{r}_2)\, \Delta\rho(\mathbf{r}_1)\, \Delta\rho(\mathbf{r}_2) \tag{12}$$

A specific ideal system must be chosen before this equation can be applied to a specific system. Two ideal systems can now be considered: (1) the classical Boltzmann ideal system, which will be delayed until the discussion of nonperturbative functionals, and (2) the Feynman ideal system.

C. Quantum Mechanics: The Feynman Ideal System

For liquid helium, classical DF theory does not predict freezing at *any* density [46,105], even though formally the classical functional is correct to second order in the density. Quantum corrections to the classical func-

tional contribute greatly beyond second order—in fact, so greatly that freezing does not occur if they are omitted.

The resolution of this difficulty is to use an ideal system that is closer to the true nature of the quantum system. One is then able to incorporate the major quantum effects from the beginning and to derive a quantum functional that predicts the helium phase diagram [29,31]. The singlet density is related to the quantity $u_{id}(\mathbf{r})$ for the case of noninteracting quantum particles [50–52].

$$\rho(\mathbf{r}_1^{(1)}) = \lim_{P\to\infty} P^{(3/2)}\left(\frac{\gamma P}{\pi}\right)^{3(P-1)/2} \int_V d\mathbf{r}_1^{(2)} \cdots \int_V d\mathbf{r}_1^{(P)} \exp\left[\frac{\beta}{P}\left(u_{id}(\mathbf{r}_1^{(1)}) + \cdots u_{id}(\mathbf{r}_1^{(P)})\right)\right] e^{-\gamma P(\mathbf{r}_1^{(1)}-\mathbf{r}_1^{(2)})^2} e^{-\gamma P(\mathbf{r}_1^{(2)}-\mathbf{r}_1^{(3)})^2} \cdots e^{-\gamma P(\mathbf{r}_1^{(P)}-\mathbf{r}_1^{(1)})^2} \tag{13}$$

where $\gamma = (2\pi^2 m/\beta h^2)$, m is the mass of a particle, h is Planck's constant, and P is the number of beads in the ring-polymer isomorphism.

To calculate the quantum ideal free energy $\mathcal{F}_{id}$, one can no longer use equation (8) because (14) cannot be inverted to give $u_{id}(\mathbf{r})$ as a functional of $\rho(\mathbf{r})$. Instead, one obtains Ω_{id} through (8) and performs the Legendre transform using (7) to obtain $\beta\mathcal{F}_{id}$. Substitution into (13) yields

$$\beta\Delta\Omega_V = 1 + \beta\mathcal{F}_{id} + \int_V d\mathbf{r}\,[\beta u_L - \beta u(\mathbf{r})]\rho(\mathbf{r}) - \int d\mathbf{r}\,\rho(\mathbf{r}) \ln \rho_L - \frac{1}{2}\int d\mathbf{r}_1 \int d\mathbf{r}_2\, c(|\,\mathbf{r}_1 - \mathbf{r}_2\,|)\,\Delta\rho(\mathbf{r}_1)\,\Delta\rho(\mathbf{r}_2) \tag{14}$$

where $u_{id}(\mathbf{r})$ is given *implicitly* by (14).

The second-order term in this equation is evaluated most readily by exploiting the periodic nature of the singlet density and expanding $\Delta\rho(\mathbf{r})$ in the Fourier series $\Delta\rho(\mathbf{r}) = \sum_{\mathbf{k}}\rho(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}$, where the sum is over the reciprocal lattice vectors (RLVs) of the crystal [30].

To be useful in freezing calculations, one must be able to minimize $\Delta\Omega_V[\rho(\mathbf{r})]$ [Eq. (15)] with respect to $\rho(\mathbf{r})$. However, $u_{id}(\mathbf{r})$ cannot be obtained explicitly. Therefore, it is more convenient to consider $\Delta\Omega_V$ as a functional of both $\rho(\mathbf{r})$ and $u_{id}(\mathbf{r})$. To interpret $\Omega[\rho(\mathbf{r}),u_{id}(\mathbf{r})]$, defined in (15), as the grand potential, the two equations

$$\frac{\delta\Omega_V}{\delta\rho(\mathbf{r})} = 0 \quad \text{and} \quad \frac{\delta\Omega_V}{\delta u_{id}(\mathbf{r})} = 0 \tag{15}$$

must be satisfied. For this choice of ideal system, the function $c(\mathbf{r}_1 - \mathbf{r}_2)$ is related to the direct correlation sometimes used in interaction site models (ISM) of molecular liquids.

D. Quantum Direct Correlation Function

The quantum direct correlation function is defined in (10). This definition is a generalization of the classical definition in which $\mathcal{F}_{id}$ in (10) is the free-energy functional of the Boltzmann system. McCoy et al. [30] discuss in detail the definition of $c(\mathbf{r}_1,\mathbf{r}_2)$, in particular its relation to the measurable particle–particle correlation function $\chi(\mathbf{r}_1,\mathbf{r}_2)$ and the form it takes when $\mathcal{F}_{id}$ is the Feynman free-energy functional.

Following Percus [53], one finds that

$$\beta\chi(\mathbf{r}_1,\mathbf{r}_2) = \frac{\delta\rho(\mathbf{r}_1)}{\delta u(\mathbf{r}_2)} \tag{16}$$

This is the probability [less the mean field contribution of $\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)$] that a particle exists at position $\mathbf{r}_1$ *and* simultaneously a particle, either the same particle as at $\mathbf{r}_1$ or at different one, exists at $\mathbf{r}_2$. If a classical particle exists at $\mathbf{r}_1$, that particle exists only at $\mathbf{r}_1$, and the classical "self" contribution to the particle–particle correlation function is proportional to a delta function. For the ideal Boltzmann system,

$$\chi_B^0(\mathbf{r}_1 - \mathbf{r}_2) = \rho_L\delta(\mathbf{r}_1 - \mathbf{r}_2) \tag{17}$$

A quantum particle is more complicated and interesting. Because of uncertainty in the position, a particle "existing" (in the quantum sense) at $\mathbf{r}_1$ may also exist at $\mathbf{r}_2 \neq \mathbf{r}_1$. McCoy et al. [30] have shown that for the ideal Feynman system in the liquid state

$$\chi_F^0(\mathbf{r}_1 - \mathbf{r}_2) = \rho_L \frac{2\gamma}{\pi} \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} e^{-4\gamma(\mathbf{r}_1 - \mathbf{r}_2)^2} \tag{18}$$

where $\gamma = \pi/\Lambda^2$ and Λ is the de Broglie wavelength. Note that, as expected,

$$\lim_{\gamma\to\infty} \chi_F^0(\mathbf{r}_1 - \mathbf{r}_2) = \chi_B^0(\mathbf{r}_1 - \mathbf{r}_2) \tag{19}$$

It is customary to split $\chi(\mathbf{r}_1,\mathbf{r}_2)$ into self and a distinct particle contributions via

$$\chi(\mathbf{r}_1,\mathbf{r}_2) = \sum(\mathbf{r}_1,\mathbf{r}_2) + \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)h(\mathbf{r}_1,\mathbf{r}_2) \tag{20}$$

where $h(\mathbf{r}_1,\mathbf{r}_2)$ is the so-called total correlation function and $\sum(\mathbf{r}_1,\mathbf{r}_2)$ is the self-correlation function. Note that the relevant correlation functions here *differ* from the correlation functions usually calculated via path integral techniques. Usually, only the diagonal (same imaginary time) terms are calculated, since these terms contribute to the partition function. The correlations that concern us here involve both diagonal and off-diagonal

terms. The off-diagonal terms arise from considerations of linear response theory embodied in (16). In fact, these off-diagonal interactions are central to the success of a quantum theory of freezing [31].

From (16) and (8) one obtains

$$\frac{\delta^2 \beta \mathcal{F}}{\delta\rho(\mathbf{r}_1)\delta\rho(\mathbf{r}_2)} = \chi^{-1}(\mathbf{r}_1,\mathbf{r}_2) \tag{21}$$

where $\chi^{-1}(\mathbf{r}_1,\mathbf{r}_2)$ is the functional inverse of $\chi(\mathbf{r}_1,\mathbf{r}_2)$. A short deviation shows that [30] the Fourier transform of the quantum direct-correlation function is

$$c(\mathbf{k}) = \Delta\sum{}^{-1}(\mathbf{k}) + \frac{\sum^{-1}(\mathbf{k})\rho^2 h(\mathbf{k})}{\sum(\mathbf{k}) + \rho^2 h(\mathbf{k})} \tag{22}$$

The definition of the quantum $c(\mathbf{r})$ has been constructed so that it reduces to the usual classical limit. In the quantum theory $\sum(\mathbf{r} - \mathbf{r}')$ is no longer a delta function, and use of the Boltzmann ideal system in the freezing theory would have caused a huge second-order contribution to the free energy in (13) from the term involving $\Delta\sum^{-1}(\mathbf{k})$. The general principle stated and applied by McCoy et al. is that the ideal system should be chosen so that this $\Delta\sum^{-1}$ contribution to $c(\mathbf{k})$ is eliminated or at least minimized. In numerical work, Rick et al. [31] have explored one method for doing this, namely the choice of an effective mass for the ideal system such that $\sum_{id}(r) = \sum(r)$.

E. Gaussian Approximation: Quantum Systems

The equations above constitute a complete freezing theory which one expects will work well for quantum systems. Solution of these equations is a nontrivial computational problem, and it is worthwhile exploring simplified, yet accurate versions of the general theory. One simplified version is the quantum analog of the Gaussian approximation often used in classical freezing theory [13,14].

For classical systems, the Gaussian approximation is straightforward. One assumes that the crystal singlet density is of Gaussian form (and hence spherically symmetric) about each lattice site, and then searches for the Gaussian with the lowest free energy. In other words, one performs a partial minimization of $\Delta\Omega$ with the constraint that the density is Gaussian. This has proved to be a very good approximation [6,7] for close-packed crystals. One sees that for the quantum case, where the expressions for both $\rho(\mathbf{r})$ and $u_{id}(i\mathbf{r})$ are needed, it is easier to assume an equiv-

alent form for $u_{id}(\mathbf{r})$ and solve (14) for $\rho(\mathbf{r})$ rather to assume a form for $\rho(\mathbf{r})$ and solve for $u_{id}(\mathbf{r})$.

If the function $c(\mathbf{k})$ can be obtained for any liquid density (which is a purely liquid-state problem separate from the minimization of $\Delta\Omega_V$, which is the focus here), the course is clear: One assumes a reasonable form for $u_{id}(\mathbf{r})$ and uses (14) to find the corresponding form of $\rho(\mathbf{r})$. McCoy et al. assumed that $u_{id}(\mathbf{r})$ has a parabolic form in each unit cell $\beta u_{id}(\mathbf{r}) = a - br^2$, where $\mathbf{r}$ is the distance from the nearest lattice point. The normalization of $\rho(\mathbf{r})$ yields parameter a as a function of b, and both parameter b and ρ_S will be obtained by minimization of the free energy. This form for $\beta u_{id}(\mathbf{r})$ makes the system identical to a system of harmonic oscillators, and following Feynman [50], the density takes the form

$$\rho(\mathbf{r}) = e^a \left(\frac{\alpha}{\gamma}\right)^{3/2} e^{-2\alpha(C-1)r^2} \tag{23}$$

where $C = \cosh f$, $f = (b/\gamma)^{1/2}$, and

$$\alpha = \frac{(2bm/\beta)^{1/2}\pi}{h \sin h(f)} \tag{24}$$

Normalization of $\rho(\mathbf{r})$ to one particle per unit cell gives $a = \ln[(2\gamma(C - 1)/\pi)^{3/2}]$ and hence

$$\rho(\mathbf{r}) = \left[\frac{2\alpha(C-1)}{\pi}\right]^{3/2} e^{-2\alpha(C-1)r^2} \tag{25}$$

Clearly, $\rho(\mathbf{r})$ is of Gaussian form, although the relation between $\rho(\mathbf{r})$ and $u_{id}(\mathbf{r})$ is more complicated that it is in the classical case. However, the classical limit (i.e., $\alpha \to \infty$) is recovered correctly since $\lim_{\alpha\to\infty} 2\alpha(C - 1) = b$ and $\lim_{\alpha\to\infty} \rho(\mathbf{r}) = (b/\pi)^{3/2} e^{-br^2}$. The final form of the quantum grand potential $\beta\Delta\Omega$ is

$$\frac{\beta\Delta\Omega}{V} = \rho_L - \rho_S \left[1 + \ln \rho_L - \tfrac{3}{2} \ln \left[\frac{2\gamma(C-1)}{\pi}\right] + \frac{3b}{4\alpha(C-1)}\right] - \tfrac{1}{2}\rho_S^2 \sum_{\mathbf{k}} c(\mathbf{k})\rho^2(\mathbf{k}) \tag{26}$$

where the Gaussians are assumed to be nonoverlapping. Again, (26) has the correct classical limit [5]. The derivation above provides the quantum analog of the classical Gaussian approximation for $\Delta\Omega$, which is known to work well for classical close-packed systems. It turns out [29,31] that this simplified quantum form is also successful in predicting the freezing of liquid helium.

F. Limitations of the $T \neq 0$ Formalism

The quantum freezing theory described above includes the dispersive nature of the particles to all orders in singlet density by including these effects in the ideal system. Application of the old classical freezing theory [5] to a quantum system would include dispersion only to second order and add Boltzmann contributions at higher orders. Exchange is included in both types of freezing theories only to second order: The liquid pair correlation function (if exact) has exchange as well as dispersion effects included in it, so it is only the *extra* exchange effects (if any) that are present in the crystal but *not* in the liquid which are neglected by these theories. In this version of the quantum theory, exchange has not been included in the ideal system because—as a many-body effect—one expects that is can be handled accurately in the same manner as the pair potential interactions. It is possible by dimensional analysis to include approximate exchange effects in the ideal system according to Thomas–Fermi theory [49].

It is interesting to examine the direct correlation function as one passes from a classical system to a quantum system. In the classical system, $h(\mathbf{k}) \to 0$ and $\sum(\mathbf{k}) \to \rho_L$ in the large $\mathbf{k}$ limit, while in the quantum case both $h(\mathbf{k})$ and $\sum(\mathbf{k}) \to 0$. This has a large effect on the second-order contribution to the free energy via the function $c(\mathbf{k})$. For the classical system, one can see from (22) that $c(\mathbf{k})$ goes as $h(\mathbf{k})$ for large $\mathbf{k}$. On the other hand, in the quantum system both the denominator and the numerator approach zero. This produces an amplifying effect on $h(\mathbf{k})$ at large $\mathbf{k}$. Hence a relatively featureless $h(\mathbf{k})$ can have a dramatic impact on the free energy if the self term decreases rapidly with distance. In summary, to second order in perturbation theory, it may be stated that it is this fact which causes the freezing of quantum systems. It is a direct consequence of the quantum self-correlation function.

G. Quantum Freezing at $T = 0$

Subsequent to the development of the $T \neq 0$ quantum theory, at least three independent groups have developed versions of a $T = 0$ theory [54–57]. The most complete of these is due to Ashcroft and co-workers [54], who consider the freezing of the Bose hard-sphere liquid at $T = 0$. Unlike helium-4, Bose hard spheres freeze into a face-centered cubic (fcc) crystal at $T = 0$. Although these authors do not identify the separate contributions of the self and distinct terms to the direct-correlation function [denoted $v(k)$ in their paper], it is easy to show that their $T = 0$ functional is the zero-temperature limit of the quantum functional derived above (with a

slight modification to include the modified weight-density functional discussed below). Most interestingly, there is a deep and important connection between the $T = 0$ quantum freezing problem and *classical* flux-lattice melting, which arises, for example, in theories of high-temperature superconductivity, as discussed by Nelson [58]. This is beyond the scope of this chapter.

IV. APPLICATION OF QUANTUM DF THEORY TO HELIUM

A. Liquid Helium

The path integral density functional theory of quantum freezing has been used to study the freezing of ^{4}He and ^{3}He. Rick et al. [31] examined the crystal–liquid phase diagram over the temperature range 8 to 204.4 K and studied the isotopic shift in the liquid–crystal coexistence line. They concluded that mass effects rather than quantum statistics are important in the freezing transition at these temperatures.

The phase diagram of helium has many unique properties. It is the only element without a triple point; there is no point in phase space where the liquid, crystal, and gas phases are all in equilibrium. Also, unlike any other element, helium remains a liquid at the lowest temperatures and 1 atm pressure. Pressure needs to be applied for helium to form a solid at any known temperature [63].

For the purposes of this calculation, perhaps the most important feature of the helium phase diagram is the significant difference between the liquid–crystal coexistence curves of helium and those of other noble gases. The phase diagrams of neon, argon, krypton, and xenon can be modeled fairly well as a system interacting via the classical, two-parameter Lennard–Jones potential (see below). The two parameters are an energy parameter, ϵ, and a size parameter, σ. The phase diagram of the Lennard–Jones system has been calculated using density functional theory [19,6]. Figure 1 displays the calculated liquid–crystal and gas–liquid coexistence lines (dashed lines). The triple points of neon, argon, krypton, and xenon, determined experimentally [64], are also plotted. The points are all scaled by the appropriate Lennard–Jones parameters [65–67]. The triple point of the Lennard–Jones system is the intersection of the crystal–liquid and gas–liquid coexistence regions, which is in the same region as the noble gas triple points. This is true even for neon, which displays some quantum behavior. However, because quantum effects are much more important for helium, ^{4}He (the solid line) has a vastly different phase diagram, and

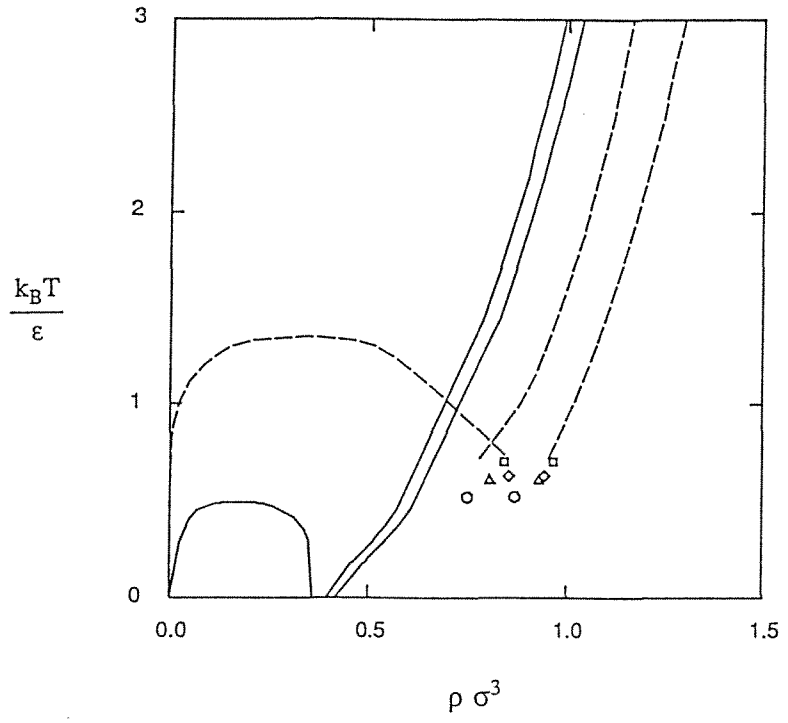


Fig. 1 Phase diagrams for helium, the (classical) Lennard–Jones system, and the noble gases: experimental data for ^{4}He (solid line) [68,69]; density functional theory calculations for the Lennard–Jones system (dashed line) [6]; and experimental data for neon (circles), argon (squares), krypton (triangles), and xenon (diamonds).

two features are apparent: helium freezes at a much lower density, and the density change upon freezing is much smaller [68,69].

Even the phase diagrams of the helium isotopes, ^{3}He and ^{4}He, differ. For example, ^{3}He always freezes at a lower density (although not always at a lower pressure [68]!). Except for very low temperatures, these differences can be attributed to the difference in mass, which affects the quantum dispersion of the atoms. Dispersion is related to the uncertainty principle; it is the reading out of the probability of finding an atom at a particular position. The effects of dispersion decrease slowly with increasing temperature as $1/T$ (as shown below), so they remain important for helium up to relatively high temperatures. These effects are also proportional to the inverse mass of the atoms and will be greater for ^{3}He

than ^{4}He. It is shown below that this causes ^{3}He to freeze at a lower density.

There is also a difference in the type of statistics obeyed by each isotope, since ^{3}He is a fermion and ^{4}He is a boson. As explained in Ref. 30, these effects are not included in the present choice of the ideal system. The ideal system here consists of noninteracting quantum particles (with dispersion), which obey Boltzmann statistics. Exchange effects are important only if the distance over which dispersions cause the atoms to be read is comparable to the typical interparticle distance, when there can be significant overlap between the two atoms. For the temperatures considered here, above 8 K, this is never the case, and the effects of exchange should not be important.

To study lower temperatures, especially the strange behavior of the liquid–crystal phase diagram as the temperature approaches absolute zero, exchange should be included in calculation of the liquid pair correlation function $g(r)$, which is the input for the present freezing theory [30]. A correct treatment of Bose–Einstein or Fermi–Dirac (rather than Boltzmann) statistics is nontrivial. It can be done for bosons with quantum Monte Carlo techniques developed by Pollock and Ceperley [70,71], who have calculated a $g(r)$ value for ^{4}He at 2 K both with and without exchange and find essentially no measurable difference. Since this is true at 2 K, it is safe, at least to second order in the free-energy expansion, to neglect exchange at higher temperatures.

The crystal–liquid coexistence line for ^{4}He and also partially for ^{3}He can now be calculated, proceeding as follows. First, the methods used to calculate the liquid-phase input needed to start the quantum freezing calculation are described. The predicted phase diagrams, for both ^{3}He and ^{4}He, are then compared with experimental results.

B. Liquid Phase Input

To study the phase diagram of helium with the quantum DF theory, the pair correlation function $g(r)$ and the self-correlation function $\sum(r)$ for the liquid are required as input. The first calculation [29] used readily available integral equation methods, coupled with helium pseudopotentials, to examine quickly quantum freezing theory. A more thorough examination using the best available liquid input data was presented subsequently [31]. To calculate these data, one may use quantum Monte Carlo simulations which have been developed and described in detail by others [70]. This entire subsection is concerned with the details of this liquid-phase input rather than with freezing theory per se. However, certain unusual or unfamiliar functions are required by freezing theory.

The quantum Monte Carlo method may be summarized as follows. Neglecting exchange, the partition function for N atoms with Hamiltonian $\mathcal{H}$ is found by taking the trace over a density matrix for each atom,

$$Z = \int d\mathbf{r}_1 \cdots d\mathbf{r}_N \langle \mathbf{r}_1 | e^{-\beta\mathcal{H}} | \mathbf{r}_1 \rangle \cdots \langle \mathbf{r}_N | e^{-\beta\mathcal{H}} | \mathbf{r}_N \rangle \tag{27}$$

where $| \mathbf{r}_n \rangle$ is an eigenfunction of particle n in the position representation, β is $1/k_B T$, T is temperature, and k_B is Boltzmann's constant.

The Hamiltonian may be written as the sum of a free-particle term $\mathcal{H}_0$ and a potential energy $V = \sum_{n=1}^{N} V_n^{(j)}$. Assuming that the potential energy is pairwise additive and of Lennard–Jones form, then

$$\sum_{n=1}^{N} \frac{1}{P} V_n^{(j)} = \sum_{n=1}^{N-1} \sum_{m=n+1}^{N} \frac{1}{P} u_{\mathrm{LJ}}(r_{nm}^{(j)}) \tag{28}$$

where

$$u_{\mathrm{LJ}}(r) = 4\epsilon \left[\left(\frac{r}{\sigma}\right)^{-12} - \left(\frac{r}{\sigma}\right)^{-6} \right] \tag{29}$$

$r_{nm}^{(j)} = | \mathbf{r}_n^{(j)} - \mathbf{r}_m^{(j)} |$, and for helium $\epsilon/k_B = 10.22$ K and $\sigma = 2.556$ Å. The partition function may be simplified [31] to

$$Z = (\gamma P)^{3P/2} \int d\mathbf{r}_1^{(1)} \cdots d\mathbf{r}_N^{(P)} \exp \left\{ -\beta \sum_{n=1}^{N} \sum_{j=1}^{P} \times \left[\frac{\gamma P}{\beta} (\mathbf{r}_n^{(j)} - r_n^{(j+1)})^2 + \frac{1}{P} \sum_{m=n+1}^{N} U_{\mathrm{LJ}}(\mathbf{r}_{nm}^{(j)}) \right] \right\} \tag{30}$$

where $\gamma = mk_B T/2\hbar^2$ and $\mathbf{r}_n^{(P+1)} = \mathbf{r}_n^{(1)}$.

This form of the partition function demonstrates the well-known isomorphism between a system of quantum particles and a classical ring polymer with P monomer units or beads [51]. From this starting point it is straightforward to simulate the liquids. The jth bead on an atom is coupled to beads $j - 1$ and $j + 1$ on the same atom with a spring constant $2P\gamma/\beta$, and it interacts only with the jth bead on other atoms through the interaction potential $V(\mathbf{r})/P$.

Since the movements of the beads are strongly correlated due to the nearest-neighbor springs, it is convenient to calculate the normal modes or diagonal elements of the free-particle contribution to the density matrix. The normal modes are by definition independent of each other, so one can move the beads more efficiently by moving the normal modes [72]. A Monte Carlo move consists of (1) moving the entire atom (i.e., the entire necklace of beads) a random distance uniformly distributed

within a cube of length ΔX, then (2) moving each normal mode uniformly within a cube of a length dependent on the normal mode. The new positions of the atom and its beads are then calculated and the move accepted or rejected according to the usual Metropolis algorithm [73]. The parameters in the simulation are adjusted to give an acceptance ratio of about 30%.

C. Quantum Liquid Pair Correlation Function

The liquid pair correlation function may be calculated from the equation

$$g(r) = \frac{1}{N(N-1)P^2}\left\langle \sum_{n=1}^{N-1}\sum_{m=n+1}^{N}\sum_{i=1}^{P}\sum_{j=1}^{P}\delta[r-(r_n^{(j)}-r_n^{(i)})]\right\rangle \tag{31}$$

and the self-correlation function from

$$\Sigma(r) = \frac{1}{NP^2}\left\langle \sum_{n=1}^{N}\sum_{j=1}^{P}\sum_{i=1}^{P}\delta[r-(r_n^{(j)}-r_n^{(i)})]\right\rangle \tag{32}$$

where $\langle\cdots\rangle$ denotes the ensemble average. These correlation functions include correlations between different imaginary times as well as the same imaginary time. However, for liquid helium the $g(r)$ given by (31) is very similar to the same imaginary time pair correlation function. The pressure, p, is calculated from the virial equation, which for this system is

$$\begin{aligned} p = \rho_L k_B T P - \tfrac{2}{3}\rho_L k_B T\gamma \frac{P}{N}\left\langle \sum_{n=1}^{N}\sum_{j=1}^{P}(r_n^{(j)}-r_n^{(j+1)})^2\right\rangle \\ - \frac{\rho_L}{3NP}\left\langle \sum_{n=1}^{N-1}\sum_{m=n+1}^{N}\sum_{j=1}^{P}\mathbf{r}_{nm}^{(j)}\frac{\partial u_{LJ}(r_{nm}^{(j)})}{\partial \mathbf{r}_{nm}^{(j)}}\right\rangle \end{aligned} \tag{33}$$

where $\rho_L = N/V$ is the number density of the liquid.

To obtain the required input into the quantum DF theory, Rick et al. performed simulations with 500 particles, with $P = 10$ beads per atom for temperatures below 21 K and $P = 3$ at higher temperatures. Additional simulations were run with many more beads to check the convergence of their simulations.

There are two principal computational limitations of path integral simulations. First, there is a finite limit to the number of beads P for each atom. Second, the total number of atoms is finite. These limitations compromise the simulations in different ways and demand the introduction of additional approximations.

The finite bead approximation restricts the accuracy of the calculation of the self-correlation function $\Sigma(r)$, but not the pair correlation function

$g(r)$, since it converges rapidly with increasing P. Unfortunately, $\sum(r)$ converges very slowly. It turns out that the second moment of the distribution, $\langle r^2 \rangle$, converges very slowly as $1/P$ at any nonzero temperature and mass. This demonstrates that dispersion effects go to zero as $1/T$, since γ is proportional to temperature. It is reasonable to expect the self term with interactions to converge in terms of P in a similar manner. This fact prohibits the approximation of the infinite P limit of the self-correlation function $\sum(r)$ with any reasonably sized finite P system. One simple resolution of this difficulty is to approximate $\sum(r)$ with the infinite-P limit. As shown by Rick et al. [31], this method is very effective. One may also approximate this quantity as an infinite-P ideal $\sum(r)$ with an effective mass m^*. This idea has been investigated by calculating $\langle r_p^2 \rangle$ from a simulation at a particular temperature and density for a number of different values of P, and adjusting $m^*(P)$ so that the ideal $\langle r_p^2 \rangle$ equals the simulation $\langle r_p^2 \rangle$. One then extrapolates $m^*(P)$ to infinite P.

The second limitation of the liquid simulations, finite box size, does not affect the short-range $\sum(r)$ but does affect $g(r)$. In particular, long-wavelength information about the system cannot be obtained. As shown by Rick et al., this problem can be ameliorated by using an effective potential, coupled with classical integral equation theory to yield a pair correlation function for all distances, which agrees with the simulated Monte Carlo $g(r)$ at small distances. It is this pair correlation function that is used as input into the quantum DF theory. Analogous problems must be overcome to provide the input for the freezing of Bose hard spheres at $T = 0$ [54].

D. Results: Freezing of Helium

Rick et al. [31] have calculated the crystal–liquid phase existence points of ^{4}He at the temperatures 8, 20, 51.1, 102.2, and 204.4 K. For ^{3}He, a low-temperature point at 20 K, a point in the middle of the temperature range at 102.2 K, and a high-temperature point at 204.4 K have been calculated to show that the theory is capable of investigating isotopic effects on freezing. Comparison of the results for the two isotopes is a good test of the approximations in the theory. By examining isotopic differences, one can judge the importance of features neglected in this theory, such as exchange.

Table I displays the calculated [31] ^{4}He phase coexistence properties, the temperature T, the coexisting liquid and crystal densities ρ_L and ρ_S, the Gaussian width parameter b, the fractional density change on freezing $\eta = (\rho_S - \rho_L)/\rho_L$, and the pressure p. These results were obtained by Rick with the self-correlation function $\sum(r)$ approximated by the ideal

Table 1 Liquid–Crystal Phase Coexistence Properties for ^{4}He

T (K)	ρ_L (Å^{-3})	ρ_S (Å^{-3})	b (Å^{-2})	η	P (kbar)
8.0	0.0409	0.0505	45.3	0.236	0.45
20.0	0.0546	0.0646	60.7	0.183	1.72
51.1	0.0732	0.0864	71.4	0.180	7.74
102.2	0.0898	0.1022	63.0	0.137	21.97
204.4	0.1090	0.1209	58.8	0.109	61.34

function $\sum^0(r)$. The effective mass approximation for $\sum(r)$, investigated at a temperature of 51.1 K, yielded only slightly poorer agreement with experiment than did the full calculation.

Experimentally, the stable crystal in this temperature range is face-centered cubic (fcc) above 14.9 K and hexagonal close-packed (hcp) for lower temperatures [75]. The present (Gaussian) theory always predicts fcc to be slightly more stable than hcp even at 8 K. The resolution of this tiny free-energy difference is a demanding problem even for a classical theory. In Fig. 2 the results are compared with the experimental phase diagram [68,75,76]. The exact placement of the points is sensitive to the values of the Lennard–Jones parameters, and the calculations are sensitive to the form of the pair potential itself. To implement the freezing theory the "standard" literature values for the parameters [70,71] have been used. Nevertheless, the agreement is very good. Less accurate is the predicted crystal density, which is overestimated just as in classical density functional theory [6]. This is probably not the result of neglected quantum effects, since the density change is fairly uniform over the range of temperature. It is more likely a consequence of the truncated free-energy expansion and also, to a lesser degree, the Gaussian approximation for the crystal density.

The ^{3}He results are shown in Table 2. The predicted stable crystal phase at both temperatures is again found to be fcc. Experimentally, the stable phase at 20 K is fcc; the fcc-to-hcp transition occurs at 17.73 K [78]. Apparently, the melting curve for ^{3}He has not been measured above 30 K, but the melting curve for both isotopes has been studied through other theoretical methods [79,80]. One very interesting feature of the ^{3}He and ^{4}He phase diagrams is the isotopic shift in the pressure at the freezing point, $\Delta p = p_f(^3\text{He}) - p_f(^4\text{He})$, at a fixed temperature [68]. At low temperatures, Δp is positive, indicating that although ^{3}He freezes at a lower density, it has a higher pressure than does ^{4}He at its slightly higher freezing density. This is perhaps not surprising, since the lighter mass of

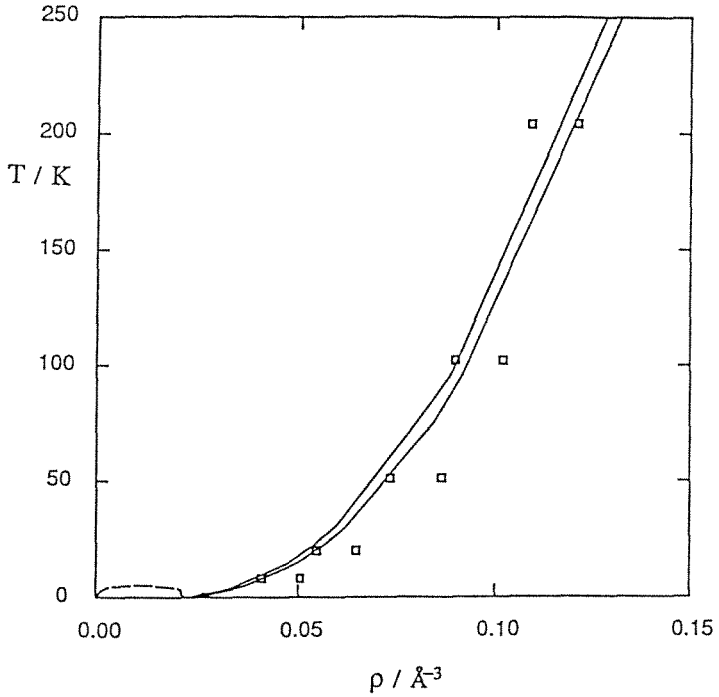


Fig. 2 ^{4}He phase diagram: experimental results for the liquid–crystal coexistence curve (solid line) [68,75–77], the gas–liquid coexistence line (dashed line) [69], and the DF calculations (squares) [31].

^{3}He will cause a higher pressure through the second term in the virial equation (33). The dispersion term of (33), $\langle\sum\sum(r_n^{(j)} - r_n^{(j+1)}\rangle$, will be greater for ^{3}He than for ^{4}He, but when multiplied by γ, which is proportional to the mass, the second term will have a smaller magnitude and the ^{3}He pressure will be greater. On the other hand, at higher temperatures, ^{3}He still freezes at a lower density, dispersion terms diminish, and

Table 2 Liquid–Crystal Phase Coexistence Properties for ^{3}He

T (K)	ρ_L (Å^{-3})	ρ_S (Å^{-3})	b (Å^{-2})	η	P (kbar)
20.0	0.0535	0.0632	61.2	0.181	1.82
102.2	0.0892	0.1005	58.2	0.127	21.93
204.4	0.1086	0.1209	68.7	0.113	60.85

the other terms in the virial equation dominate. This causes the melting pressure of ^{3}He to be less than ^{4}He. Very recent Monte Carlo studies of Barrat and co-workers [79], which calculated the free energies of liquid and crystal using an approximate technique, estimate that the crossover from positive to negative Δp occurs at 100 K.

The quantum DF calculations predict that Δp is 0.10 kbar at 20 K, in good agreement with the experimental value of 0.11 kbar [68]. Also in good agreement is the isotopic shift in the freezing density, -0.0011 Å^{-3} compared to the experimental value -0.0009 Å^{-3}. This is a positive indication that the quantum properties relevant to the freezing transition have been treated correctly in the theory. In this temperature range, the isotopic shift is caused by the differences in mass, not statistics. The theory shows that, more precisely, the isotopic shift is due to differences in the self-correlation functions.

At 102.2 K one finds that the pressure difference Δp is -0.04 kbar, indicating that Δp changes sign at a temperature slightly less than 102.2 K. At 204.4 K, the calculations predict that Δp is -0.51 kbar. The isotopic shift in the freezing density decreases monotonically with increasing temperature. The isotopic shift in the pressure as a fraction of the freezing pressure, $\Delta p/p$, is also decreasing, and eventually both isotopes should freeze at the same pressure, but this will occur at a relatively high temperature. It is interesting that isotopic effects on the phase diagram persist to such high temperatures.

E. Empirical Rules for Quantum Freezing

These results can be used to reflect upon simple, empirical models of melting and freezing. The Sutherland [81]–Lindemann [82] ratio $\mathcal{L}$ is the average root-mean-square deviation of a particle in the crystal from its lattice site, measured in units of the nearest-neighbor distance, d_{nn}. For the Gaussian density given by (3.2) of Ref. 30, the ratio $\mathcal{L}$ is

$$\mathcal{L}^2 d_{nn}^2 = \int_{\text{peak}} d\mathbf{r}\, r^2 \rho(\mathbf{r}) = \frac{3}{4}\frac{1}{\alpha(C-1)} \tag{34}$$

The Sutherland–Lindemann empirical rule [81] states that a crystal will melt when $\mathcal{L}$ exceeds 0.1 and seems to hold for a number of simple crystals. It is certainly not true for helium (Table 3), in which $\mathcal{L}$ is as large as 0.28 for ^{4}He and 0.38 for ^{3}He at low temperatures [83] This is, in fact, a measure of the large zero-point motion of helium. The calculated $\mathcal{L}$ ratios are smaller than the experimental values but (correctly) larger than the values calculated from classical density functional theory for the Lennard–Jones system. The standard version of classical density functional

Table 3 Magnitude of the Sutherland–Lindemann Ratio and Standard Deviation of the Crystal Singlet Density

	T (K)	$\mathscr{L}$	s (Å)
^{4}He	8.0	0.145	0.255
	20.0	0.112	0.189
	51.1	0.0979	0.144
	102.2	0.0915	0.127
	204.4	0.0840	0.114
^{3}He	20.0	0.124	0.202
	102.2	0.0990	0.138
	204.4	0.0838	0.114

theory also predicts Sutherland–Lindemann ratios consistently smaller than those observed [6].

The crystal density profile about a fixed lattice site for ^{4}He at different temperatures can be compared with the classical Lennard–Jones result [6] as shown in Fig. 3. The low-temperature helium density is in fact very broad. The crystal singlet densities display a strong temperature dependence, although this is not apparent from the value of the Gaussian width parameter b. The standard deviation, s, of the density is $s = [\alpha(C - 1)]^{-1/2}/4$, and this does indeed have a strong temperature dependence. This implies that while the effect of the neighboring particles is relatively constant over a wide range of temperatures, the response of the central particle to that effect is strongly temperature dependent.

Another empirical rule, the Verlet [84] or Hansen–Verlet rule [85], states that the first peak of the structure factor of a liquid at its freezing point takes a value from 2.85 to 3.05. This rule is also not obeyed by helium. Table 4 documents the position of the first peak, k_q, and the value of the structure factor at the peak, $S(k_q)$, from the calculations. The low-temperature values demonstrate that the structure factor is much weaker at the freezing point than the empirical rule would claim.

There is an interesting generalization of this rule to the quantum case. When used as input into the freezing theory [29], the structure factor is amplified by the self-correlation function. To study this amplification, it is useful to consider an effective structure factor, $S_e(k) = [\sum(k) + \rho_L{}^2 h(k)]/\sum(k)$, where $\sum(k)$ is ρ_L times the Fourier transform of $\sum(r)$ and $h(k)$ is the Fourier transform of $[g(r) - 1]$. In the classical limit, $S_e(k)$

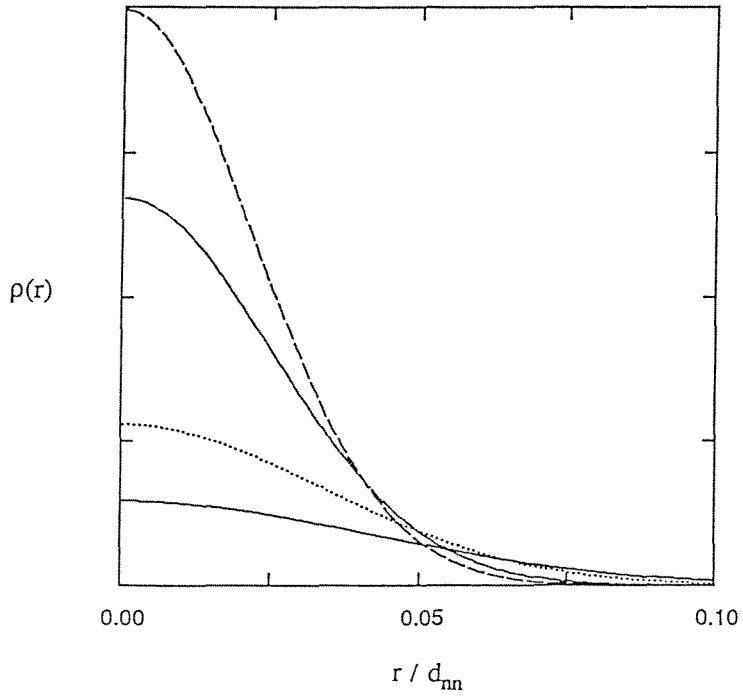


Fig. 3 Crystal singlet density profiles for ^{4}He at $T = 8$ K (solid line), 20 K (dotted line), 204.4 K (dot-dashed line), and classical Lennard–Jones results for $k_B T/\epsilon = 2$ (dashed line), normalized as described in the text.

Table 4 Location and Amplitude of the First Peak of the Structure Factor, $S(k_q)$, and Effective Structure Factor, $S_e(k_q)$, at the Predicted Freezing Point

	T (K)	k_q (Å^{-1})	$S(k_q)$	$S_e(k_q)$
^{4}He	8.0	2.5	1.90	2.68
	20.0	2.7	2.51	3.15
	51.1	3.0	2.70	3.04
	102.2	3.2	3.03	3.24
	204.4	3.4	3.35	3.47
^{3}He	20.0	2.7	2.34	3.13
	102.2	3.2	2.97	3.25
	204.4	3.4	3.32	3.52

reduces to the usual structure factor $S(k)$. The values of the first peak of $S_e(k)$ are also shown in Table 4. The effective structure factor does more or less conform to the empirical rule. It is important to emphasize that even at the higher temperatures, it is the *self*-correlation function that drives the freezing process. Without it, helium remains liquid.

F. Prospects for the Quantum Theory

The spatial extent of the self-correlation function can be measured from the magnitude of $\langle r^2 \rangle$, which averages over the squared distance between each pair of beads on a single atom. This quantity is similar to, but distinct from, the radius of gyration about the center of the atom, R_0, which averages the squared distance between each bead and the center of mass of the atom. In terms of the correlation functions described above, $\langle r^2 \rangle = \int d\mathbf{r}\, r^2 \sum(r)$ and $R_0^2 = \int d\mathbf{r}\, r^2 \sum_{cm}(r)$. As the density is increased isothermally, $\langle r^2 \rangle$ will decrease due to interparticle interactions. This dependence has been shown clearly by Rick et al. [31]. As the temperature is decreased, $\langle r^2 \rangle^{1/2}$ has a more pronounced density dependence, and the approximation of the self term by the ideal self term is better at higher temperatures. A potential cure for this difficulty at low temperatures, the effective mass approximation, reproduces the self-correlation function accurately, but nevertheless does not predict freezing at the lowest temperatures investigated by Rick et al..

It may be anticipated that removal of the Gaussian approximation will improve the results of freezing theory significantly. Due to the very low density of the crystal, especially by comparison with classical crystals, it is extremely unlikely that the crystal density is spherically symmetric, as assumed in the Gaussian approximation. In particular, a more exact parameterization of the crystal, for example by Fourier expansion, should decrease the fractional density change and improve the calculations of other properties as well. Unfortunately, this improvement comes at a high price: Considerable numerical complexity is reintroduced when the Gaussian approximation is removed, especially in performing the path integrals.

The results for ^{4}He and ^{3}He presented above show that the quantum density functional theory is a useful and accurate method for studying the phase diagram of quantum liquids. Calculations in the high-temperature region provide useful information for comparison with both high-pressure experiments and large-scale computer simulations of all possible phases. From a practical perspective, use of a helium pseudopotential produces accurate pair correlation functions $g(r)$ without the need for computer simulation of each phase point. This provides the basis for in-

teresting studies not only of helium but for other quantum liquids an mixtures [86].

V. RECENT FUNCTIONALS FOR CLASSICAL SYSTEMS

A. Original Classical Density Functional

The derivation of quantum DF theory appears to differ considerably fron the earlier classical presentations [2,3,5,13,14], and hence it is now helpfu to recover conventional DF theory. This also provides the starting poin for more recent, alternative generalizations of classical DF theory.

The classical DF theory is recovered by defining the classical *idea* system through the relation between the singlet density $\rho(\mathbf{r})$ and $u_{id}(\mathbf{r})$ For monatomic liquids such as inverse power potentials, hard spheres and Lennard–Jones systems, DF theory uses the Boltzmann relation fo noninteracting particles in an external field, namely,

$$\rho(\mathbf{r}) = \Lambda^{-3} e^{\beta u_{id}(\mathbf{r})} \tag{35}$$

where Λ is the DeBroglie thermal wavelength.

From (35) and (8), the classical ideal free energy is then

$$\beta \mathcal{F}_{id}[\rho] = \int_V d\mathbf{r}\, \rho(\mathbf{r})[\Lambda^3 \ln \rho(\mathbf{r}) - 1] \tag{36}$$

Substitution into (13) leads to

$$\beta \Delta\Omega_V = 1 + \int_V d\mathbf{r}\, \rho(\mathbf{r}) \left[\ln \left(\frac{\rho(\mathbf{r})}{\rho_L} \right) - 1 \right] + \int_V d\mathbf{r}\, [\beta u_L - \beta u(\mathbf{r})]\rho(\mathbf{r}) - \frac{1}{2} \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2\, c(\mathbf{r}_1 - \mathbf{r}_2)\, \Delta\rho(\mathbf{r}_1)\, \Delta\rho(\mathbf{r}_2) \tag{37}$$

By setting $u(\mathbf{r}) = u_L(-\mu)$ in a zero external field, and minimizing $\Delta\Omega_V$ with respect to $\rho(\mathbf{r})$, one obtains the grand potential difference $\Delta\Omega_V$ between the system under study (the crystal) and the liquid reference system at the same chemical potential. The liquid chemical potential is varied (usually via the liquid density) to find a liquid reference system that satisfies $\Delta\Omega = 0$; that is, one finds the thermodynamic conditions under which the liquid and solid are in equilibrium. Note that with this choice of the ideal system, the function $c(\mathbf{r}_1 - \mathbf{r}_2)$ reduces to the usual direct correlation function introduced by Ornstein and Zernike.

It is worth emphasizing that a practical expression for the excess free energy is unknown, in general, for inhomogeneous or *homogeneous* systems. Hence the central task of any statistical mechanical theory may be viewed as the approximation of this quantity. For example, quests for the so-called bridge function in integral equation theories can be viewed in this light, with profit [87].

The family of nonperturbative or weighted-density functionals introduced by Tarazona [10,11] is introduced most conveniently by returning to the canonical ensemble. The total (Helmholtz) free energy $\mathcal{F}[\rho]$ is the sum of the ideal $\mathcal{F}_{id}[\rho]$ above, and a nonideal (or excess) contribution $\mathcal{F}_{ex}[\rho]$, which arises from interparticle interactions: $\mathcal{F}[\rho] = \mathcal{F}_{id}[\rho] + \mathcal{F}_{ex}[\rho]$. The classical theory above yields an excess free energy

$$\beta\mathcal{F}_{ex}[\rho] = \beta\mathcal{F}_{ex}[\rho_L] - c^{(1)}(\rho_L)\int_V d\mathbf{r}_1\,\Delta\rho(\mathbf{r}_1) - \int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2\, c^{(2)}(|\,\mathbf{r}_1 - \mathbf{r}_2\,|;\rho_L)\,\Delta\rho(\mathbf{r}_1)\,\Delta\rho(\mathbf{r}_2) + \cdots \tag{38}$$

This is the free energy that was used in the first full calculation of hard-sphere freezing via the density functional theory [5]. For the Lennard–Jones (LJ) system, this level of theory generates the phase diagram shown in Fig. 1, as calculated by Marshall et al. in 1985 [19]. This theory has the advantages of simplicity and clarity: The only approximation is to truncate the functional Taylor series expansion of $\mathcal{F}_{ex}[\rho]$ (assumed analytic) at second order. Even eight years after the first HS calculation, it is still somewhat hard to accept that such a straightforward theory can generate results such as the LJ phase diagram reproduced above. It is interesting to plot the pressure–density phase diagram from the Lennard–Jones system (here at the reduced temperature $T^* = 1$). Figure 4 displays the pairs of liquid and crystal states, which have the same temperature and chemical potential μ, linked by the dotted lines. Note that the unique pair, which *in addition* share the same pressure, comprise the coexistence point. This plot is the analog of the Maxwell construction in the grand ensemble.

B. Methods for Improving Perturbation Theory

It is now helpful to summarize methods for going beyond second-order perturbation theory while retaining the simplicity of the DF formalism:

1. Choose a "better" ideal system, to minimize the difference between the reference system and the system under study. This is the ap-

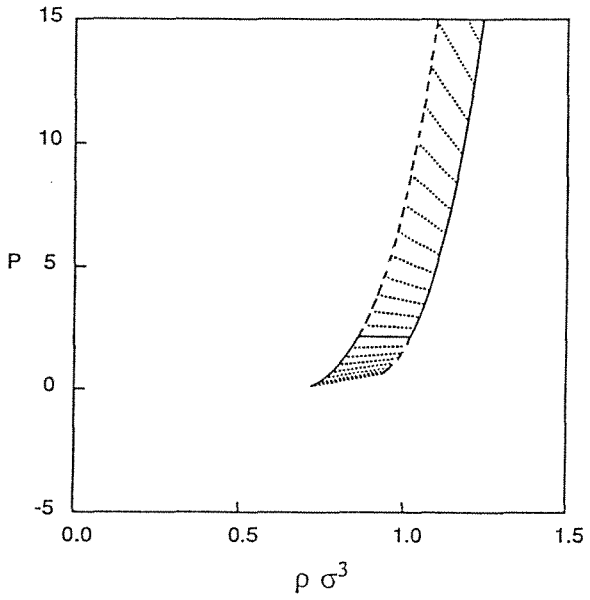


Fig. 4 Pressure versus density phase diagram for the LJ system at $T^* = 1.0$, calculated from density functional theory. Pairs of liquid and crystal points with the same temperature and chemical potential μ are linked with dotted lines.

proach advanced in quantum DF theory, and in the author's view the option with the best long-term prospects.

2. Keep higher-order terms in the straightforward perturbation expansion. This is the most obvious choice, which was investigated partially even in the original papers [5]. For example, the next term in the perturbation expansion for the excess free energy $\beta\mathcal{F}_{ex}$ is

$$\frac{1}{3!}\int_V d\mathbf{r}_1 \int_V d\mathbf{r}_2 \int_V d\mathbf{r}_3\, c^{(3)}(\mathbf{r}_1,\mathbf{r}_2,\mathbf{r}_3)\,\Delta\rho(\mathbf{r}_1)\,\Delta\rho(\mathbf{r}_2)\,\Delta\rho(\mathbf{r}_3) \qquad (39)$$

where $c^{(3)}$ is the next member of the family of n-body direct correlation functions, $c^{(n)}(\mathbf{r}_1,\cdots,\mathbf{r}_n;[\rho])$, defined by functional derivatives of $\mathcal{F}_{ex}[\rho]$ in (10).

For two reasons, this line of development is less promising than the other two options listed here. First, the higher-order correlation function $c^{(n)}$ are unknown even in the bulk liquid, except perhaps for the triplet case $n = 3$ [93,100]. In fact, recent progress in the field of bulk high-order correlation function is based on adopting an approximate free energy de-

scribed in option 3 below and using (10) to approximation correlation functions. Second, and more important, nothing is known about the convergence (or lack of it) of the functional Taylor series expansion. Experience from critical phenomena and quantum mechanics shows that perturbation expansions are unlikely to be useful unless they work at second order, and in fact it is not wise to push these expansions further.

A purely empirical modification of the second-order perturbation theory was proposed by Baus and Colot, who call it the effective-liquid approximation (ELA) [9]. This approximation is to replace the liquid density ρ_L in the excess free-energy functional 39 with an effective density ρ_{eff} chosen (arbitrarily) such that the first peak of the Fourier transform of $c^{(2)}$ coincides with the first nonzero reciprocal-lattice vector of the crystal. We know of no justification for this type of empiricism.

3. Avoid perturbation theory and invent a direct approximation to the full, infinite-order excess free energy. In the context of freezing, this approach was invented by Tarazona [10], although his work is not always cited. Rosenfeld [48] has also pioneered the fundamental understanding of this approach.

C. Approximate Weighted-Density Functionals

There are many versions of the weighted-density approach. All consider a weighted density $\rho_{\text{eff}}(\mathbf{r})$, which is a nonlocal functional of the actual density

$$\rho_{\text{eff}}(\mathbf{r}_1) = \int_V d\mathbf{r}_2\, w(\mathbf{r}_1 - \mathbf{r}_2, \rho_{\text{eff}}(\mathbf{r}_1))\rho(\mathbf{r}_2) \tag{40}$$

where w is a weighting function to be specified. The excess free energy may then be written

$$\mathcal{F}_{\text{ex}}[\rho] = \int d\mathbf{r}_1\, \psi(\rho_{\text{eff}}(\mathbf{r}_1))\rho(\mathbf{r}) \tag{41}$$

where $\psi(\rho)$ is the excess Helmholtz free energy per particle of a homogeneous system of density ρ. Usually, the weighting function w is chosen such that the known (reference) bulk liquid direct correlation function is recovered upon differentiating (41).

Laird and Kroll have shown [22] the systematic relationship between the wide variety of approximate weighting functions. There is a close connection between the freezing theories and earlier local density approximations for inhomogeneous liquids [44,45]. In the context of freezing it is customary to average the local density of a small region of space, as

shown in (40). The goal of this class of approximate theories is to choose a weighting function that describes accurately the crystal phase [88–94].

There are two subclasses of weighted-density functional theories: (1) those with spatially varying weighted density, and (2) those with constant weighted density. Unfortunately, each subtheory has come to be labeled by a set of letters. Class (1) theories, represented by the weighted density approximation (WDA) of Curtin and Ashcroft [8], are more computationally demanding and are not discussed here. This simpler and equally accurate class (2) theories are represented by the modified WDA of Denton and Ashcroft [42], which by a considerable margin is to date the most numerically successful of the weighted-density theories. The MELA [9] and GELA [97] approximations of Baus and co-workers are closely related but have severe limitations documented by Laird and Kroll [22].

In the MWDA of Denton and Ashcroft [42], the weighting function $w(\mathbf{r}_1 - \mathbf{r}_2,\rho_{\text{eff}})$ is chosen such that both the free energy and the two-particle direct correlation function $c^{(2)}$ are recovered in the homogeneous limit. The excess free energy of the crystal phase is approximated by the excess free energy of the liquid, evaluated at a spatially *independent* weighted density:

$$\mathcal{F}_{\text{ex}}[\rho] = N\psi(\rho_{\text{eff}}) \tag{42}$$

where the weighted density is defined by

$$\rho_{\text{eff}}N = \int d\mathbf{r}_1\, \rho(\mathbf{r}_1) \int d\mathbf{r}_2\, \rho(\mathbf{r}_2)\, w(\mathbf{r}_1 - \mathbf{r}_2,\rho_{\text{eff}}) \tag{43}$$

Applying the two constraints used to choose the weighting function leads to an algebraic equation for the weighting function,

$$w(\mathbf{r}_1 - \mathbf{r}_2,\rho_{\text{eff}}) = -\frac{1}{2\beta\psi'(\rho_{\text{eff}})}\left[c(\mathbf{r}_1 - \mathbf{r}_2;\rho_{\text{eff}}) + \frac{\rho_{\text{eff}}\beta\psi''}{V}\right] \tag{44}$$

Note that the class (1) theories with spatially varying weighted densities lead to a complicated differential equation instead of an algebraic equation, but for the examples investigated to date lead to essentially the same numerical predictions [22].

Other variants of the weighted-density approach can be derived by starting from the definition of the direct correlation function in integrated form,

$$\beta\mathcal{F}_{\text{ex}}[\rho] = -\int d\mathbf{r}_1 \int d\mathbf{r}_2 \int_0^1 d\lambda \int_0^\lambda d\lambda'\, \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)\, c^{(2)}(\mathbf{r}_1,\mathbf{r}_2;[\lambda'\rho]) \tag{45}$$

Lutsko and Baus [97] assume that the correlation functions in the crystal can be approximated by those in the liquid at some effective density, and approximate the last factor in (45) by $c^{(2)}(\mathbf{r}_1 - \mathbf{r}_2;\rho_{\text{eff}}[\lambda'\rho])$, where $c^{(2)}(r;\rho_{\text{eff}})$ is now the two-body direct correlation function for a *homogeneous* system with density ρ_{eff}. The form of this functional is such that the sum rules for higher-order direct correlation functions are automatically satisfied. The approximate weighting functional is now determined by requiring that the reference liquid that was used above to approximate the correlation functions *also* determine the thermodynamics, which leads to

$$w(\mathbf{r}_1 - \mathbf{r}_2,\rho_{\text{eff}}) = \frac{-\rho_{\text{eff}}}{\beta\psi(\rho_{\text{eff}})}\int_0^1 d\lambda \int_0^\lambda d\lambda'\, c^{(2)}(\mathbf{r}_1 - \mathbf{r}_2;\rho_{\text{eff}}[\lambda'\rho]) \quad (46)$$

This weighting function is called the GELA; the related MELA approximation replaces the functional $\rho_{\text{eff}}[\lambda\rho]$ in (46) by $\lambda\rho_{\text{eff}}[\rho]$, which leads to a computationally simpler approximation but one that does not recover the correct liquid direct correlation function in the homogeneous limit.

The WDA and MWDA approximations described above were introduced originally by defining a priori the weighting functions. Laird and Kroll [22] have derived these approximations by expanding the exact functional $\mathcal{F}$ about a reference liquid and then choosing a well-defined optimal reference density. Their derivation not only unifies the derivation of all the approximations above (and related ones), but also shows the correspondence with the approximation introduced by Groot and van der Eerden [102] in their modification of the weighted-density functional theory of Meister and Kroll [103].

D. Comparison of Weighted-Density Functionals

Laird and Kroll [22] have compared many versions of the density functional theory for single-component systems. In particular, they investigated purely repulsive, inverse power potentials of the form

$$v_{ij}(r) = \epsilon\left(\frac{\sigma}{r}\right)^n \quad (47)$$

the phase diagrams of which depend on a single dimensionless parameter

$$\gamma_n = \rho\sigma^3\left(\frac{kT}{\epsilon}\right)^{-3/n} \quad (48)$$

The freezing of these inverse power potentials has been investigated ex-

tensively by computer simulations by Hoover and collaborators and hence provide crucial test for the density functional theory.

The set of freezing calculations used a Gaussian approximation for the crystal single-particle density,

$$\rho(\mathbf{r}) = \left(\frac{\pi}{\epsilon}\right)^{3/2} \sum_{\{Rj\}} \exp[-\epsilon \mid \mathbf{r} + \mathbf{R}_j \mid^2] \tag{49}$$

where ε measures the width of the Gaussian peaks, and $\{\mathbf{R}_j\}$ is the set of real-space lattice vectors of the crystal under study, here the face-centered cubic (fcc) structure. [Important considerations concerning the body centered cubic (bcc) phase are discussed by Laird and Kroll.] The accuracy of this approximation has been tested by Laird et al. [6]. With this parameterization, the ideal free energy is given by

$$\beta\mathcal{F}_{\mathrm{id}}(\epsilon) = \tfrac{3}{2} \ln\left(\frac{\epsilon}{\pi}\right) + 3 \ln \Lambda - \tfrac{5}{2} \tag{50}$$

and for each fixed liquid density ρ_L, minimization of the full functional is required only with respect to ϵ.

The values of γ_L and γ_S at phase coexistence, together with the fraction density change and Sutherland–Lindemann parameter $\mathcal{L}$, are collected in Table 5 and indicate that the density functional theory is an extremely useful and accurate theory for predicting phase diagrams. The MWDA theory is probably the "best" theory, although in many aspects, for example the Sutherland-Lindemann ratio $\mathcal{L}$, it does no better than the original second-order perturbation theory [5].

There are many more properties of interest that the DF theory can calculate. For example, Laird has shown [24] that the MWDA, generalized to include a less restrictive Gaussian crystal singlet density, reproduces almost exactly the hard-sphere fcc elastic constants c_{11}, c_{12}, *and* c_{44} measured in computer simulations by Frenkel and Ladd [61]. These calculations also shed light on the absolute limit of mechanical stability of the HS fcc crystal as the density is lowered.

E. Binary Mixtures of Hard Spheres

The phase diagram of a binary mixture of hard spheres has been examined by a number of groups [27,31,86,95,101] and reexamined very recently by two groups using the weighted-density methods described above [40,41]. The second-order perturbation theory density functional is gen-

Table 5 Laird–Kroll Comparison of FCC Freezing Results for Inverse-Power Potentials

n	Method	γ_L	γ_S	$\frac{\Delta\rho}{\rho_s}$	$\mathscr{L}$
∞	MWDA[42]	0.910	1.036	0.13	0.097
	GELA[97]	0.945	1.041	0.10	0.095
	MELA[97]	0.970	1.070	0.084	0.099
	ELA[9]	0.99	1.08	0.09	0.07
	Perturb[6,98]	0.967	1.147	0.18	0.06
	Haymet[5]	0.976	1.035	0.06	0.06
	Simulation[104]	0.939–0.948	1.036–1.045	0.09–0.11	0.126
12	MWDA[22]	1.194	1.252	0.046	0.096
	GELA[22]		No minima		
	MELA[22]		No minima		
	ELA[98]	1.305	1.380	0.06	0.07
	Perturb[98]	1.28	1.37	0.07	0.07
	Simulation[99]	1.15	1.19	0.035	0.15
6	MWDA[22]	2.666	2.720	0.020	0.074
	GELA[22]		No minima		
	MELA[22]		No minima		
	ELA[98]	3.33	3.39	0.02	0.07
	Perturb[98]	3.43	3.52	0.026	0.07
	Simulation[99]	2.18	2.21	0.013	0.17
4	MWDA[22]	8.176	8.238	0.0075	0.07
	GELA[22]		Not attempted		
	MELA[22]		Not attempted		
	ELA[98]	11.34	11.43	0.007	0.07
	Perturb[98]	12.30	12.47	0.014	0.07
	Simulation[99]	5.54	5.57	0.005	0.18

eralized easily to a ν-component system, to give

$$\Delta\beta\Omega_V = \sum_{i=1}^{\nu} \int d\mathbf{r}_1 \left[\rho_i(\mathbf{r}_1) \ln \frac{\rho_i(\mathbf{r}_1)}{\rho_{iL}} - \Delta\rho_i(\mathbf{r}_1) \right] - \frac{1}{2} \sum_{i,j=1}^{\nu} \int d\mathbf{r}_1 \int d\mathbf{r}_2 \, c_{ij}(\mathbf{r}_1 - \mathbf{r}_2) \, \Delta\rho_i(\mathbf{r}_1) \, \Delta\rho_j(\mathbf{r}_2) \tag{51}$$

and the weighted-density methods are generalized similarly.

Binary mixtures display a rich variety of phases, of which five are considered here: (1) the substitutionally disordered fcc crystal, which has

the different atoms of the solid mixture disturbed randomly on a common lattice, in this case a fcc lattice; (2) the ordered cesium chloride symmetry, which consists of two interpenetrating simple cubic lattices, one for each type of atom; (3) the ordered sodium chloride structure, which consists of two symmetrically placed interpenetrating fcc lattices; (4) the ordered zincblende structure, which also consists of two interpenetrating fcc lattices, but with one displaced from the other along the body diagonal of the cubic cell by one-fourth of the length of the diagonal; and (5) the fast sphere phase, which is Smithline and Haymet's terminology for a crystal with the larger atoms fixed on a lattice, in this case a fcc lattice, and the smaller atoms free to flow through the lattice of the larger atoms. This phase is analogous to a fast ion phase, although hard spheres of course have no charge. Denton and Ashcroft call this a sublattice melt.

In Fig. 5 we reproduce the binary hard-sphere phase diagram calculated by second-order perturbation theory [86] and the MWDA theory [41]. Similar results are obtained from the effective liquid free-energy model (ELFEM) of Zeng and Oxtoby [39,40], which in their CA1 approximation reduces to MWDA. The phase diagrams are very similar; the MWDA should be regarded as providing slightly more accurate predictions, although by comparison with recent computer simulations [62], both appear to be deficient in predicting that the disordered fcc crystal is *always* more stable than any substitutionally ordered phase.

Zeng and Oxtoby [40] have calculated the same diagram, and in addition plotted the temperature–comparison (T–x) diagrams for three typical cases: (1) spindle, (2) azeotrope, and (3) eutectic phase diagrams. These are reproduced in Fig. 6 from top to bottom, respectively.

Rick and Haymet have made a complete study of the freezing of Lennard–Jones binary mixtures [86], including comparison with experimental data for argon, krypton, and methane mixtures, and also calculated freezing of charged mixtures in their recent review [60]. These calculations, together with the figures reproduced here, indicate that the DF theory is a simple and workable theory for studying mixture phase diagrams. Ternary and more complex liquids remain to be studied in detail, although freezing of an "infinite" number of components, in the sense of polydisperse liquids, has already been studied successfully [59,96].

F. Vacancy Concentrations in the Equilibrium Crystal

The density functional theory of freezing has recently been extended by McRae et al. [32] to predict the equilibrium concentration of vacancies in the crystal. The vacancy density in the crystal at phase coexistence is

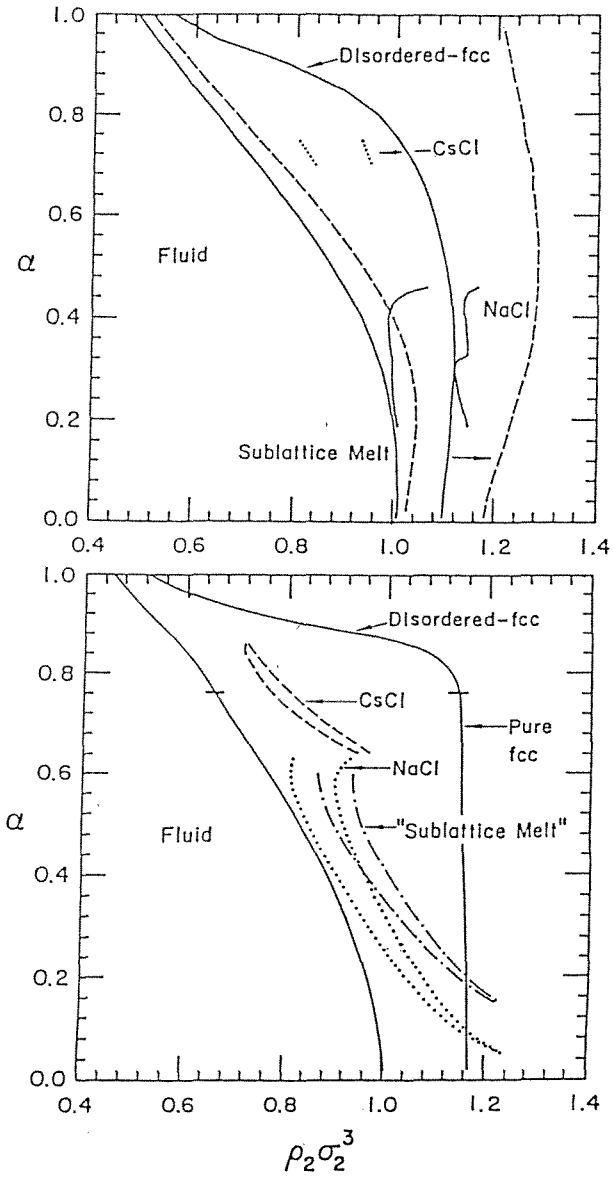


Fig. 5 Binary hard-sphere crystal-phase diagram. Plotted is the crystal density of large spheres, which coexists with a 1:1 liquid, for the radius ratio $\alpha = \sigma_{small}/\sigma_{large}$ as a function of the reduced density of large spheres, as calculated (top) originally by Rick and Haymet [86] and (bottom) from MWDA theory by Denton and Ashcroft [41].

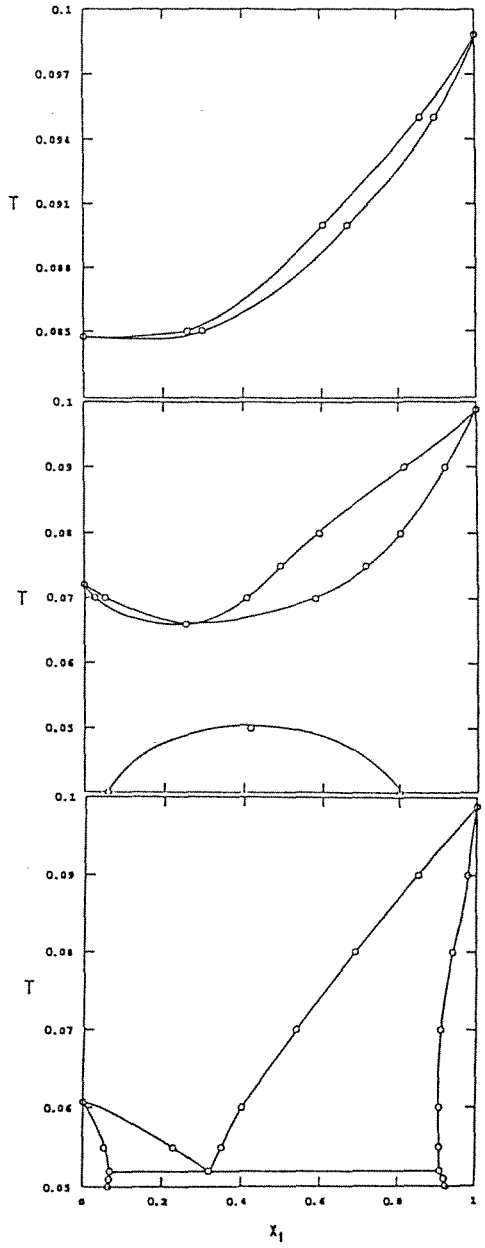


Fig. 6 Binary hard-sphere mixture temperature–composition phase diagrams, calculated by Zeng and Oxtoby [40] for the radius ratios from top to bottom 0.95, 0.90, and 0.85.

predicted to be approximately 5×10^{-5} for the Lennard–Jones model systems near the triple point, which is comparable to concentrations measured in real systems, such as argon and krypton. McRae et al. also investigated the dependence of the vacancy density on the temperature and pressure of the system. They also calculated the concentration of vacancies in the hard-sphere crystal at coexistence and found it to be very small. The perfect crystal approximation used in all the work described above and all previous calculations was found to be very accurate for both the hard-sphere and Lennard–Jones systems.

For close-packed crystals of simple substances, one may assume that vacancies are the primary point defect and that the number of vacancies in the crystal, n, is *much* smaller than the number of particles in the crystal, N, so that $N + n \approx N$. Recent experimental measurements in various noble-gas solids show this to be a valid assumption. Since the vacancies are present in such low relative densities, vacancy–vacancy interactions are assumed to be negligible. A direct consequence of these two assumptions is that any arrangement of the vacancies on the crystal lattice has the same energy and thus the same probability of occurrence. This is not strictly true in real crystals, where a divacancy (for example) certainly has a different energy than that of two isolated monovacancies. However, divacancies are thought to account for less than 2% of the vacancies in noble-gas solids, with larger vacancy clusters accounting for an even smaller fraction. Also neglected are any effects due to interstitials, since the concentration of interstitials in close-packed noble-gas solids is very low relative to the vacancy concentration. Line defects such as dislocations are generally considered to be nonequilibrium (albeit important) defects and are not treated by this theory.

The free-energy functional is generalized by adding the contribution to $\mathcal{F}_{ex}$ due to all possible arrangements of vacancies on the lattice under consideration. The details are straightforward but lengthy [30]. The final expression may be written, in the dilute vacancy limit, as the above-grand potential difference $\Delta\beta\Omega_V^{(pc)}$, now with the superscript (pc) to indicate the perfect crystal, plus additional terms due to the vacancies,

$$\Delta\beta\Omega_V = \Delta\beta\Omega_V^{(pc)} + \rho_{pc}V[a \ln a + (1 - a)\ln(1 - a)] + \rho_v V w_v \quad (52)$$

where $\rho_{pc} = N_{pc}/V$ is the density of the (hypothetical) perfect crystal, $a = \rho/\rho_{pc} \leq 1$, and

$$w_v = -\int_V d\mathbf{r}\, \rho(\mathbf{r},\mathbf{0})\left[\ln\frac{\rho(\mathbf{r},\mathbf{0})}{\rho_L} - 1\right] - \hat{c}(\mathbf{0}) + \rho_{pc}\sum_{\{\mathbf{k}\}}\hat{c}(\mathbf{k})\hat{\rho}^2(\mathbf{k}) - \frac{1}{2}\int_V d\mathbf{r}_1\int_V d\mathbf{r}_2\, c(\mathbf{r}_1 - \mathbf{r}_2)\rho(\mathbf{r}_1,\mathbf{0})\rho(\mathbf{r}_2,\mathbf{0}) \quad (53)$$

where the hats denote Fourier transforms. Neither $\Delta\Omega_V^{(pc)}$ nor w_v depends on ρ_v, and therefore one can minimize with respect to ρ_v to obtain an expression for the (dilute limit) equilibrium vacancy density,

$$\rho_v = \rho_{pc}[e^{\rho_L w_v} + 1]^{-1} \tag{54}$$

This calculation shows that earlier imperfect crystal approximations, which assumed that the singlet density is periodic but with a normalization of $a \neq 1$ particles per unit cell, are incorrect.

Using the structural information about the Lennard–Jones liquid, $\hat{c}(k)$, the equilibrium vacancy density ρ_v may be predicted as a function of temperature and density along the calculated solid–liquid coexistence curve of the LJ liquid. The results of such a calculation at several coexistence points are displayed in Table 6, and the logarithm of the vacancy density along the coexistence line is plotted as a function of temperature in Fig. 7. At the reduced temperature $T^* = 1.0$, the value predicted by the DF theory is 4.7×10^{-5}, which is within the plausible range one would deduce from experimental measurement on real rare-gas solids. Many-body effects, which are neglected in this calculation, are thought to contribute significantly to the energy difference between the perfect crystal and the crystal with vacancies. McRae et al. have also calculated the LJ vacancy density along other paths in the $T\rho$-plane [32].

The extended DF theory of freezing described above provides a simple method for computing the concentration of vacancies in crystalline states at or near liquid-phase coexistence. The formalism for calculating vacancy related properties is capable of further extension, such as the inclusion of lattice relaxation about the vacancies. The first numerical calculations using this DF theory predict a vacancy concentration that decreases rapidly as the density is increased along the phase coexistence line.

Table 6 Vacancy Density Along the Lennard–Jones Freezing Line

T	ρ_L	ϵ	ρ_s	$\log_{10}(\rho_v/\rho_{pc})$
1.00	0.8630	0.0668	1.0226	−4.35
1.25	0.9303	0.0661	1.0715	−6.04
1.50	0.9791	0.0657	1.1140	−7.07
1.75	1.0192	0.0654	1.1524	−7.78
2.00	1.0542	0.0651	1.1878	−8.34
2.50	1.1130	0.0646	1.2500	−9.24
3.00	1.1626	0.0638	1.3048	−9.74

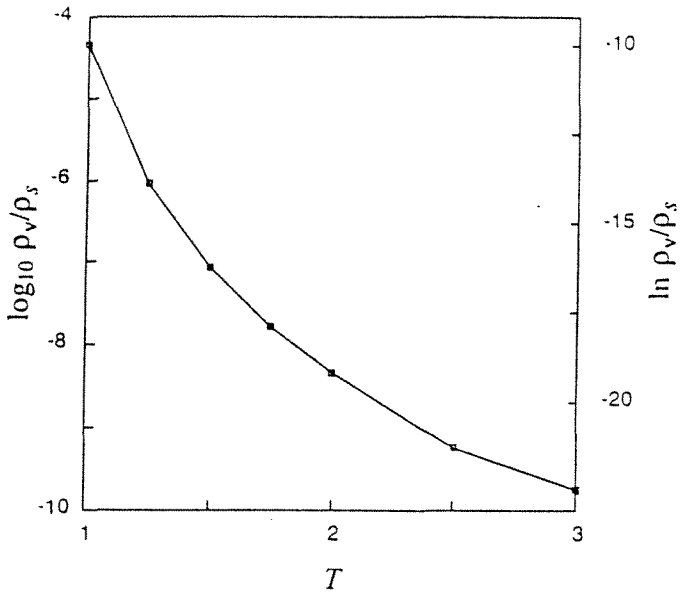


Fig. 7 Logarithm of the relative vacancy concentration versus temperature along the liquid–crystal phase coexistence line for the Lennard–Jones liquid [32]. The scale on the left is $\log_{10}$ and on the right is the natural log.

By taking advantage of the near-equilibrium metastable states located by DF theory, one can investigate the variation of the vacancy concentration on the temperature, and independently the pressure, of the system. The ratio of the vacancy density to the particle number density, ρ_v/ρ_{pc}, is found to be an exponentially decreasing function of pressure (at constant temperature) and an exponentially increasing function of temperature (at constant pressure). For hard spheres, the theory predicts a vacancy concentration that is several orders of magnitude less than the concentration in a Lennard–Jones crystal of comparable reduced density on the phase coexistence line. Finally, the predicted vacancy densities for the hard-sphere system and the Lennard–Jones system are both so low relative to the respective particle densities that the effect on the bulk density and Gaussian widths of the crystal in equilibrium coexistence with the fluid is negligible for both systems. The perfect crystal approximation used in previous freezing calculations is a very accurate approximation for the density functional theory of freezing, at least in simple systems that freeze preferentially into close-packed crystal structures.

VI. CONCLUSIONS AND FUTURE PROSPECTS

The calculations above demonstrate the strength and usefulness of the DF approach. The present limitations of the theory have been summarized in Section I. Certainly, there is a great deal of activity at present concerning non-close-packed crystal structures and the freezing of ever-more-complicated molecules, such as water, polymers, and liquid crystals. At the time of writing there are a number of promising investigations of the freezing of inverse-sixth-power liquids and other liquid models, such as rubidium, which freeze into bcc structures, and it is likely that this long-standing limitation of DF theory will have been removed by the end of 1992. At the same time, *reasons* for the predictive value of weighted density methods for hard-core systems are being uncovered. These developments reinforce the observation that the great value of DF theory is that it can be generalized readily to problems of ever-increasing difficulty. Now under construction are workable dynamical theories designed to describe not only nucleation and spinodal decomposition, but the time dependence of these phenomena and crystal growth. The beginning of this exciting field is reviewed in Chapter 10.

ACKNOWLEDGMENTS

All of our research on freezing has been supported by grants from the Petroleum Research Fund, administered by the American Chemical Society, and NSF and Ford Research through the PYI program. It is also a pleasure to acknowledge many helpful discussions over many years with Professor David Oxtoby and Drs. John McCoy, Shep Smithline, Brian Laird, Robin McRae, and Chris Marshall. Current research is supported by the Australian Research Council (ARC) Grant No. A29131271.

REFERENCES

1. Evans, R. (1979). *Adv. Phys. 28*: 143.
2. Ramakrishnan, T. V., and Yussouff, M. (1979). *Phys. Rev. B 19*: 2775.
3. Haymet, A. D. J., and Oxtoby, D. W. (1981). *J. Chem. Phys. 74*: 2559.
4. Oxtoby, D. W., and Haymet, A. D. J. (1982). *J. Chem. Phys. 76*: 6262.
5. Haymet, A. D. J. (1983). *J. Chem. Phys. 78*: 4641.
6. Laird, B. B., McCoy, J. D., and Haymet, A. D. J. (1987). *J. Chem. Phys. 87*: 5449.
7. Laird, B. B., McCoy, J. D., and Haymet, A. D. J. (1988). *J. Chem. Phys. 88*: 3900.
8. Curtin, W., and Ashcroft, W. (1985). *Phys. Rev. A 32*: 2909.
9. Baus, M., and Colot, J. L. (1985). *Mol. Phys. 55*: 653.

10. Tarazona, P. (1984). *Mol. Phys. 52*: 81.
11. Tarazona, P. (1985). *Phys. Rev. A 31*: 2672. (1985). Erratum, *Phys. Rev. A 32*: 3148.
12. Jones, G., and Mohanty, U. (1985). *Mol. Phys. 54*: 1241.
13. Haymet, A. D. J. (1987). *Annu. Rev. Phys. Chem. 38*: 89.
14. Baus, M. (1987). *J. Stat. Phys. 48*: 1129.
15. Baus, M. (1990). *J. Phys. Condens. Matter 2*: 2111.
16. Oxtoby, D. W. (1990). *Nature 347*: 725.
17. Oxtoby, D. W. (1990). In *Liquids, Freezing and the Glass Transition*, Les Houches, Session L1 (J.-P. Hansen, D. Levesque, and J. Zinn-Justin, eds.), Elsevier, Amsterdam.
18. Evans, R. (1989). In *Les Houches Summer Lectures, Session XLVIII, 1988* (J. Chavolin, J. F. Joanny, and J. Zinn-Justin, eds.), Elsevier, Amsterdam.
19. Marshall, C. H., Laird, B. B., and Haymet, A. D. J. (1985). *Chem. Phys. Lett. 122*: 320.
20. Curtin, W., and Ashcroft, N. W. (1986). *Phys. Rev. Lett 56*: 2775. (1987). Erratum, *Phys. Rev. Lett. 57*: 1192.
21. Kloczkowski, A., and Samborski, A. (1988). *J. Chem. Phys. 88*: 5834.
22. Laird, B. B., and Kroll, D. M. (1990). *Phys. Rev. A 42*: 4810.
23. Laird, B. B., and Kroll, D. M. (1990). *Phys. Rev. A 42*: 4805.
24. Laird, B. B. (1992). *J. Chem. Phys.* (submitted).
25. Renkin, M., and Hafner, J. (1991). *J. Chem. Phys. 94*: 541.
26. Ding, K., Chandler, D., Smithline, S. J., and Haymet, A. D. J. (1987). *Phys. Rev. Lett. 59*: 1698.
27. Smithline, S. J., and Haymet, A. D. J. (1987). *J. Chem. Phys. 86*: 6486. (1988). Erratum, *J. Chem. Phys. 88*: 4104.
28. Smithline, S. J., Rick, S. W., and Haymet, A. D. J. (1988). *J. Chem. Phys. 88*: 2004.
29. McCoy, J. D., Rick, S. W., and Haymet, A. D. J. (1989). *J. Chem. Phys. 90*: 4622. Note that in this reference the self correlation function is denoted $S(r)$, rather than $\sum(r)$ as used here.
30. McCoy, J. D., Rick, S. W., and Haymet, A. D. J. (1990). *J. Chem. Phys. 92*: 3034.
31. Rick, S. W., McCoy, J. D., and Haymet, A. D. J. (1990). *J. Chem. Phys. 92*: 3040.
32. McRae, R., McCoy, J. D., and Haymet, A. D. J. (1990). *J. Chem. Phys. 93*: 4281.
33. Yang, A. J. M., Fleming, P. D., and Gibbs, J. H. (1977). *J. Chem. Phys. 67*: 74.
34. Saam, W. F., and Ebner, C. (1977). *Phys. Rev. A 15*: 3566.
35. Lebowitz, J. L., and Percus, J. K. (1963). *J. Math. Phys. 4*: 116.
36. Stillinger, F. H., and Buff, F. P. (1962). *J. Chem. Phys. 37*: 1.
37. Mermin, D. (1965). *Phys. Rev. 137*: A1441.
38. McCoy, J. D., and Haymet, A. D. J. (1989). *Int. J. Thermophys. 10*: 87.
39. Zeng, X. C., and Oxtoby, D. W. (1990). *J. Chem. Phys. 93*: 2692.
40. Zeng, X. C., and Oxtoby, D. W. (1990). *J. Chem. Phys. 93*: 4357.

41. Denton, A. R., and Ashcroft, N. W. (1990). *Phys. Rev. A 42*: 7312.
42. Denton, A. R., and Ashcroft, N. W. (1989). *Phys. Rev. A 39*: 4709.
43. McCoy, J. D., Honnell, K. G., Schweizer, K. S., and Curro, J. G. (1991). *J. Chem. Phys. 95*: 9348.
44. Nordholm, S., and Haymet, A. D. J. (1980). *Aust. J. Chem. 33*: 2013.
45. Nordholm, S., Johnson, M., and Freasier, B. C. (1980). *Aust. J. Chem. 33*: 2139.
46. Smithline, S. J., and Haymet, A. D. J. (1986). Unpublished.
47. Verlet, L. (1968). *Phys. Rev. 165*: 201.
48. Rosenfeld, Y. *Phys. Rev. A 43*: 5424.
49. March, N. H. (1983). In *Theory of the Inhomogeneous Electron Gas* (S. Lundqvist and N. H. March, eds.), Plenum Press, New York.
50. Feynman, R. P. (1972). *Statistical Mechanics: A Set of Lectures*, W. A. Benjamin, New York.
51. Feynman, R. P., and Hibbs, A. R. (1965). *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York.
52. Chandler, D., and Wolynes, P. G. (1981). *J. Chem. Phys. 74*: 4078.
53. Percus, J. K. (1964). In *The Equilibrium Theory of Classical Fluids* (H. L. Frisch and J. L. Lebowitz, eds.), W. A. Benjamin, New York.
54. Denton, A. R., Nielaba, P., Runge, K. J., and Ashcroft, N. W. (1990). *Phys. Rev. Lett. 64*: 1529.
55. Chui, S. T. (1990). *Phys. Rev. B 41*: 796.
56. Senatore, G., and Pastore, G. (1990). *Phys. Rev. Lett. 64*: 303.
57. Miller, M. D., Mullin, W. J., and Guyer, R. A. (1978). *Phys. Rev. B 18*: 3189.
58. Nelson, D. R., and Seung, S. (1989). *Phys. Rev. B 40*: 6763.
59. McRae, R., and Haymet, A. D. J. (1988). *J. Chem. Phys. 88*: 1114.
60. Rick, S. W., and Haymet, A. D. J. (1990). *J. Phys. Chem. 94*: 5212.
61. Frenkel, D., and Ladd, A. J. C. (1987). *Phys. Rev. Lett. 59*: 1169.
62. Kranendonk, W. G. T., and Frenkel, D. (1989). *J. Phys. Condens. Matter 1*: 7735.
63. Keller, W. E. (1969). *Helium-3 and Helium-4*, Plenum Press, New York.
64. Pollack, G. L. (1964). *Rev. Mod. Phys. 36*: 748.
65. Sherwood, A. E., and Prausnitz, J. M. (1964). *J. Chem. Phys. 41*: 421.
66. Clifford, A. A., Gray, P., and Platts, N. (1977). *J. Chem. Soc. Faraday Trans I 73*: 381.
67. de Boer, J., and Michels, A. (1938). *Physics 5*: 945.
68. Grilly, E. R., and Mills, R. L. (1959). *Ann Phys. (N.Y.) 8*: 1.
69. Roach, P. R., and Douglas, D. H., Jr. (1967). *Phys. Rev. Lett 19*: 287. Roach, P. R. (1968). *Phys. Rev. 170*: 213.
70. Pollock, E. L., and Ceperley, D. M. (1984). *Phys. Rev. B 30*: 2555.
71. Ceperley, D. M., and Pollock, E. L. (1986). *Phys. Rev. Lett. 56*: 351.
72. Runge, K. J., and Chester, G. V. (1988). *Phys. Rev. B 38*: 135.
73. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). *J. Chem. Phys. 21*: 1087.

74. Allen, M. P., and Tildesley, D. J. (1987). *Computer Simulations of Liquids*, Oxford University Press, Oxford.
75. Dugdale, J. S., and Simon, F. E. (1953). *Proc. R. Soc. London A 218*: 291.
76. Mills, R. L., Liebenberg, D. H., and Bronson, J. C. (1980). *Phys. Rev. B 21*: 5137.
77. Loubeyre, P. (1989). In *Simple Molecular Systems at High Density* (A. Polian, P. Loubeyre, and N. Boccara, eds.), Plenum Press, New York.
78. Franck, J. P. (1961). *Phys. Rev. Lett. 7*: 435.
79. Barrat, J. L., Loubeyre, P., and Klein, M. L. (1989). *J. Chem. Phys. 90*: 5644.
80. Young, D. A., McMahan, A. K., and Ross, M. (1981). *Phys. Rev. B 24*: 5119.
81. Sutherland, W. (1890). *Philos. Mag. 30*: 318.
82. Lindemann, F. A. (1910). *Phys. Z. 11*: 609.
83. Daniels, W. B. (1989). In *Simple Molecular Systems at High Density* (A. Polian, P. Loubeyre, and N. Boccara, eds.), Plenum Press, New York.
84. Verlet, L. (1969). *J. Chem. Phys. 7*: 591.
85. Hansen, J. P., and Verlet, L. (1969). *Phys. Rev. 184*: 1501.
86. Rick, S. W., and Haymet, A. D. J. (1989). *J. Chem. Phys. 90*: 1188.
87. Ichiye, T., and Haymet, A. D. J. (1990). *J. Chem. Phys. 93*: 8954.
88. Jacobs, R. L. (1983). *J. Phys. C 16*: 273.
89. Baus, M., and Colot, J. L. (1986). *Mol. Phys. 57*: 809.
90. Iglói, F., and Hafner, J. (1986). *J. Phys. C 19*: 5799.
91. Haymet, A. D. J. (1986). *Prog. Solid State Chem. 17*: 1.
92. Haymet, A. D. J. (1985). *Chem. Phys. Lett. 122*: 324.
93. Haymet, A. D. J. (1985). *J. Phys. Colloq. C9 46*: 27.
94. Sachdev, S., and Nelson, D. R. (1985). *Phys. Rev. B 32*: 1480.
95. Barrat, J. L., Baus, M., and Hansen, J. P. (1986). *Phys. Rev. Lett. 56*: 1063.
96. Barrat, J. L., and Hansen, J. P. (1986). *J. Phys. 47*: 1547.
97. Lutsko, J. F., and Baus, M. (1990). *Phys. Rev. Lett. 64*: 761.
98. Barrat, J. L., Hansen, J. P., Pastore, G., and Waisman, E. M. (1987). *J. Chem. Phys. 86*: 6360.
99. Hoover, W. G., Ross, M., Johnson, K. W., Henderson, D., Barker, J. A., and Brown, B. C. (1970). *J. Chem. Phys. 52*: 4931.
100. Barrat, J. L., Hansen, J. P., and Pastore, G. (1987). *Phys. Rev. Lett. 58*: 2075.
101. Barrat, J. L., Baus, M., and Hansen, J. P. (1987). *J. Phys. C 20*: 1413.
102. Groot, R. D., and van der Eerden, J. P. (1985). *Phys. Rev. A 31*: 4155.
103. Meister, T. F., and Kroll, D. M. (1985). *Phys. Rev. A 31*: 4055.
104. Hoover, W. J., and Ree, F. M. (1968). *J. Chem. Phys. 49*: 3609.
105. de Kuijper, A., Vos, W. L., Barrat, J. L., Hansen, J. P., and Schouten, J. A. (1990). *J. Chem. Phys. 93*: 5187.

10

Nucleation

David W. Oxtoby
James Franck Institute
University of Chicago
Chicago, Illinois

I. INTRODUCTION

Phase transitions in fluids do not always occur under equilibrium conditions. Although the equilibrium freezing point or normal boiling point of a liquid can be measured to high precision, most fluids can readily be undercooled or superheated and held indefinitely in metastable states if they are carefully purified in advance. Water can easily be undercooled to $-15°C$ and in tiny capillary tubes to below $-30°C$; liquid gallium suspended as small droplets in oil can be undercooled by 140°C below its normal freezing point of 30°C. This kinetic stability of states of matter that are unstable thermodynamically arises from the fact that the crystallization and evaporation of a liquid (or the reverse processes of melting of a crystal or condensation of a vapor) are first-order transitions that have kinetic barriers to be overcome. Nucleation is the term given to the first localized appearance of a new stable phase in a metastable state; the rate of nucleation usually controls the rate of appearance of the new phase on a macroscopic level.

Phase transitions are described by order parameters, quantities that characterize the difference between the two states of matter. In the condensation of a vapor, the density is the obvious order parameter; the crystallization of a liquid can be described as the appearance of long-

range periodic order in the average density, and an appropriate order parameter in this case is a Fourier component of the density evaluated at a reciprocal lattice vector of the crystal. The dependence of the grand canonical free energy Ω on an order parameter ρ is shown schematically in Fig. 1. At a given temperature T and chemical potential μ, this free energy shows a double-minimum structure in mean field theory. At the thermodynamic phase transition (Fig. 1a) the values of Ω are equal in the two phases; in Fig. 1b, state A has become metastable relative to B; finally, in Fig. 1c, state A loses even its local stability. Loss of local stability occurs when the spinodal is crossed, so that the system becomes unstable to arbitrarily small fluctuations in the order parameter and the new phase appears rapidly throughout the sample.

How does a metastable state (state A in Fig. 1b) evolve into a stable state? The double-minimum-plus-barrier structure of the free energy suggests that some kind of thermally activated barrier crossing is necessary, and that is true, but the nature of this process is more complicated than is implied by the figure. The figure shows the variation of the free energy for *uniform* systems in which the order parameter has a constant value everywhere in space. Although in principle it is possible for the order parameter to change everywhere at the same time so that the system follows the curve shown in Fig. 1b toward state B, the resulting fluctuation in free energy at the top of the barrier is prohibitively costly in free energy. Instead, the order parameter changes locally in some region of space, forming a small portion of the new phase that can then grow to fill the system. The barrier to nucleation occurs in a state that is inhomogeneous, with a spatially varying order parameter, and cannot be represented on a one-dimensional plot such as in Fig. 1.

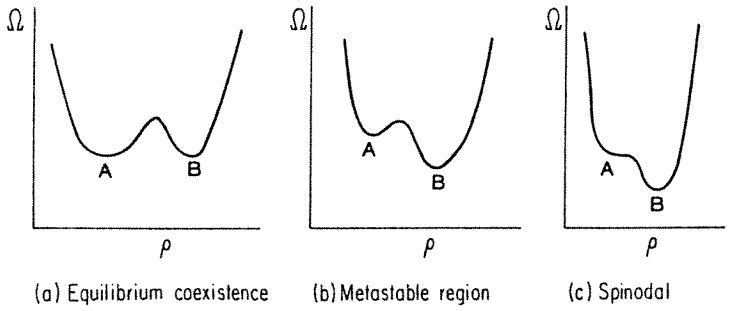


Fig. 1 Variation of the grand potential Ω with an order parameter ρ in three regions.

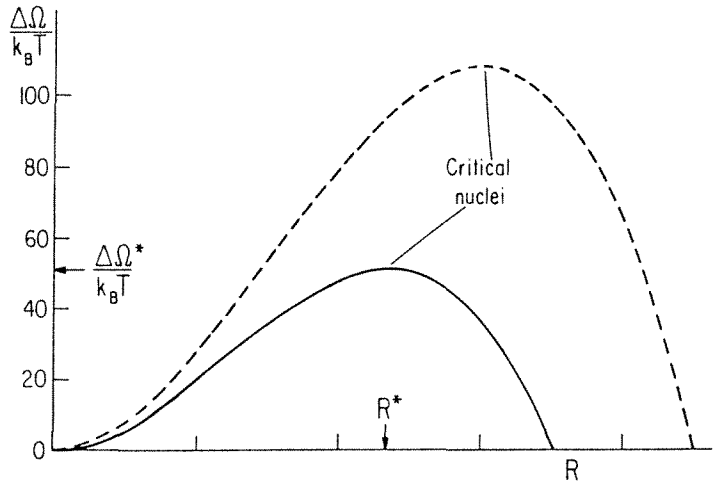


Fig. 2 Variation of the grand potential Ω with cluster radius R. The solid line shows conditions farther from the equilibrium phase transition line than the dashed line.

Consider a large spherical region of radius R of the new phase B surrounded by phase A. The free energy of this state can be estimated as the sum of a bulk term proportional to the volume of the sphere and an interfacial term proportional to its surface area:

$$\Delta\Omega(R) = \tfrac{4}{3}\pi R^3(\Omega_B - \Omega_A) + 4\pi R^2\sigma$$

The first term is negative because creation of new bulk B lowers the free energy, but the second term is positive because it costs an amount σ per unit area to create an interface. A graph of $\Delta\Omega(R)$ (Fig. 2) has a maximum at a critical radius R^*, and the height at this point, $\Delta\Omega^*$, gives the barrier to nucleation under these conditions. The cluster of radius R is referred to as the critical nucleus. Close to the coexistence curve, the coefficient of the R^3 term is small (because Ω_A and Ω_B are nearly equal), the critical nucleus is very large, and the barrier is very high. The barrier height drops and the rate of nucleation increases rapidly as the system penetrates farther into the metastable region.

The calculation just presented of the critical nucleus free energy is referred to as classical nucleation theory or the capillarity approximation [1–3]. It assumes that even a very small nucleus can be described in macroscopic terms as contributing a bulk and a surface term to the excess free energy. Nucleation can be either homogeneous (occurring in the bulk

of a pure phase) or heterogeneous (occurring about an impurity or at a surface). In this chapter we expand on the classical theory of homogeneous nucleation and compare it with experiment, and then examine more microscopic approaches to calculating free energies of nuclei using computer simulations and density functional theory. In Section II we follow this program for the gas–liquid transition and in Section III turn to the liquid–solid transition. In Section IV we examine nucleation in binary systems and in Section V take a brief look at heterogeneous nucleation.

II. VAPORIZATION AND CONDENSATION OF LIQUIDS

A. Classical Nucleation Theory

Statistical mechanical calculations of phase transitions (especially those using density functional methods) are carried out most conveniently in the grand canonical ensemble, for which the grand potential Ω is a natural function of T, V, and μ. For a single-phase system,

$$\Omega = -pV$$

where p is the pressure and V the volume, and for a two-phase system with a planar interface

$$\Omega = -pV + \sigma A$$

where A is the area of the interface. In vapor nucleation experiments, a more convenient potential is the Gibbs free energy, a natural function of pressure p, number N, and temperature T. It is related to Ω through a Legendre transform,

$$G = \Omega + \mu(\rho_v)N + p(\rho_v)V$$

where ρ_v is the number density in the vapor. Rearranging this gives

$$G - \mu(\rho_v)N = \Omega + p(\rho_v)V$$

Because $\mu(\rho_v)N$ is the Gibbs free energy of a uniform vapor of density ρ_v, and $-p(\rho_v)V$ is the grand potential for the same uniform system, the two excess quantities must be equal:

$$\Delta G = \Delta\Omega$$

The nucleation free-energy barrier can be calculated in the grand ensemble.

Now consider a spherical droplet of liquid with radius R surrounded by vapor. The pressure at the center of the liquid droplet p_l will differ

from the pressure of the vapor p_v. In the capillarity approximation the grand potential is

$$\Omega = -p_l V_l - p_v V_v + 4\pi R^2 \sigma$$

where V_l and V_v are the volumes of liquid and vapor:

$$V_l = \frac{4\pi}{3} R^3$$

$$V_v = V - \frac{4\pi}{3} R^3$$

From this the excess grand potential to create the droplet is

$$\Delta\Omega = -\frac{4\pi}{3} R^3 \Delta p + 4\pi R^2 \sigma$$

where $\Delta p = p_l - p_v$ is the pressure difference between the center of the liquid droplet and the bulk vapor. For a planar gas–liquid interface, σ is the surface tension, which is measurable by independent means. The assumption of the capillarity approximation is that the same value of σ can be used for small, curved droplets in nucleation. If the liquid is assumed to be incompressible,

$$\Delta p = \rho_l(\mu - \mu_{\text{sat}})$$

relates the pressure difference to the chemical potential difference between the supersaturated vapor and the saturated vapor. If the vapor is an ideal gas,

$$\mu - \mu_{\text{sat}} = k_B T \ln\left(\frac{p_v}{p_{\text{sat}}}\right) = k_B T \ln S$$

where k_B is Boltzmann's constant, p_{sat} is the saturated vapor pressure at the temperature T of the experiment, and S is the relative supersaturation.

To make contact with more microscopic calculations of free energies, it is convenient to replace the radius R with the discrete variable n (number of molecules in the liquid droplet), which is the product of the liquid number density and the droplet volume. The excess grand potential then has the form

$$\Delta\Omega = -nk_B T \ln S + (36\pi)^{1/3} \rho_l^{-2/3} \sigma n^{2/3}$$

The maximum of $\Delta\Omega$ occurs at

$$n^* = \frac{32\pi\sigma^3}{3\rho_l^2 (k_B T)^3 (\ln S)^3}$$

and the height of the barrier is

$$\Delta\Omega^* = \frac{16\pi\sigma^3}{3\rho_l^2(k_BT)^2(\ln S)^2}$$

As S increases from its value of 1 at coexistence, the height of the barrier falls.

So far, we have calculated only equilibrium quantities, albeit for a metastable state. To predict the nucleation rate, a dynamical theory is needed. In the original Becker–Döring theory, the number of critical nuclei per unit volume (those of size n^*) was taken to be proportional to $\exp(-\Delta\Omega^*/k_BT)$; the rate of nucleation was written as the product of this number density and the rate at which single gas-phase molecules impinge on the surface of the liquid droplet, a quantity that can be estimated from gas kinetic theory. The resulting nucleation rate is

$$J = J_0 \exp\left(\frac{-\Delta\Omega^*}{k_BT}\right)$$

where the preexponential factor J_0 is given by [4]

$$J_0 = a\sqrt{\frac{2}{\pi}}\frac{\rho_v^2}{\rho_l}\sqrt{\frac{\gamma}{m}}$$

Here a is the sticking probability of molecules on the liquid surface (taken to be unity here) and m is the molecular mass.

The simple theory outlined above can be improved in many ways. First, an equilibrium distribution of nuclei up to the critical size has been assumed, whereas in fact the populations of clusters change according to kinetic equations. If clusters are assumed to grow or decay by gaining or losing single molecules, and if each gain or loss is assumed to occur independently of the past history of the cluster, the rate of change of the number N_n of clusters of size n will be

$$\frac{dN_n}{dt} = (k_{n-1,n}N_{n-1} - k_{n,n+1}N_n)N_1 + (k_{n+1,n}N_{n+1} - k_{n,n-1}N_n)$$

The *ratios* of rate constants for gain and loss of molecules, $k_{n-1}/k_{n,n-1}$, are equilibrium constants that can be related to changes in free energy as given by the capillarity approximation or by some other theory. The forward rate constants $k_{n-1,n}$ can be estimated from collision rates of monomer with clusters of size n (at least for larger clusters). The result is a dynamical theory that can be solved for a given initial condition. In the steady state, the rate at which critical clusters react has the same form as the quasi-equilibrium theory, but with an additional quantity Z,

called the Zeldovich factor, multiplying the preexponential J_0 and reducing the rate by one to two orders of magnitude [3]. Another correction to the preexponential has been described by Courtney [5]. Changes in the preexponential by even several orders of magnitude have a much smaller effect than changes in the exponential factor, however, and do not significantly perturb the critical supersaturation S_c at which a given rate of nucleation is achieved.

One of the more extensively discussed corrections to the classical Becker–Döring theory has been proposed by Lothe and Pound [6]. They argued that the overall translational and rotational motion of the critical nucleus gives a contribution that must be added to the free energy, increasing rates by factors on the order of 10^{17}. Such a large effect certainly *could* be detected, but Reiss and co-workers [7] have argued rather convincingly that the Lothe–Pound correction is not justified. Their approach gives a correction to the capillarity approximation, but its effect on the rate is much smaller. A better theory than the classical one must also address such questions as the effect of curvature and finite size on the surface free energy. Because the free energy of the critical nucleus appears in the exponential of the rate, such effects can be very large.

B. Experimental Results

There is space here only to mention a few experimental results; a complete survey will not be attempted. It is first appropriate to describe briefly two types of experimental methods used to obtain quantitative information about nucleation rates. One is the upward thermal diffusion cloud chamber described by Katz [8]. In this apparatus a liquid at the bottom of a container is heated from below to vaporize it partially. The upper surface of the container is held at a lower temperature, so that a temperature gradient is established from top to bottom. The *total* pressure of gas in the container (background carrier gas plus nucleating gas) is approximately uniform, but the partial pressure of nucleating gas falls linearly with height in the chamber. The saturated vapor pressure, however, depends exponentially on temperature and falls much more rapidly with height. The result is that the local supersaturation has a rather sharp peak at a height about three-fourths of the way up in the chamber. The temperatures at the bottom and top of the chamber are then adjusted so that the maximum nucleation rate is on the order of 1 cm^{-3} s^{-1}; drops form only in the narrow band where the supersaturation S passes through a maximum. As droplets form and grow, they fall under the influence of gravity, and vapor-phase molecules are replenished by evaporation of the liquid. This experiment is run in steady state, and supersaturation is calculated by solving heat and mass transport equations for the fluid in the chamber.

A second experimental technique is the fast expansion chamber described by Schmitt [9]. Here a gas is expanded abruptly and adiabatically by a large amount. As it expands, it cools and reaches a supersaturated state with a predetermined value of S. The vapor is held for a short time in this state (typically 0.01 s), allowing critical nuclei to form. Then the gas is compressed to a final state in which S is small enough that further nucleation does not take place, but in which growth still occurs because S is larger than 1. After the droplets have become large enough to be seen, they are counted. This technique is time dependent (not steady state) but has the advantage that the temperature and partial pressures are uniform throughout the chamber during the nucleation process. It also measures a large range of nucleation rates of up to 10^5 cm^{-3} s^{-1} or higher.

One of the most extensively studied substances is *n*-nonane, C_9H_{20}. Figure 3 illustrates the results obtained for its nucleation in a plot of the ratio of measured to classical (Becker–Döring) rate for a series of tem-

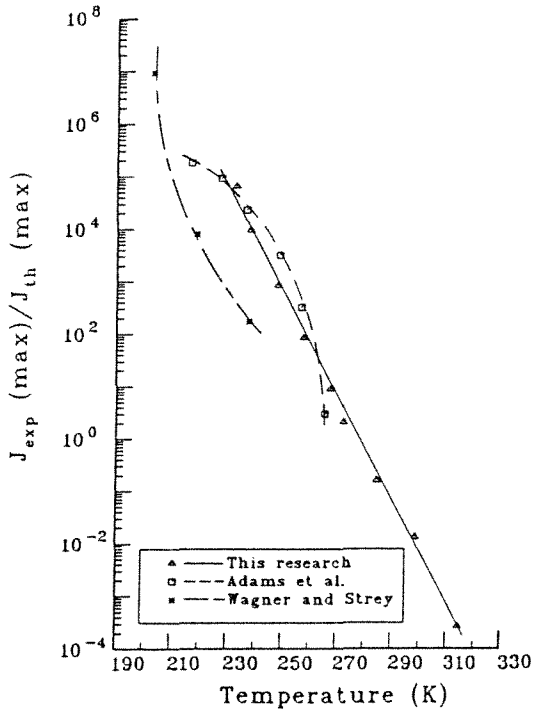


Fig. 3 Ratio of the measured to the classical (theoretical) nucleation rate. Only at one temperature (near 275 K) do the two agree. (From Ref. 10.)

peratures. Several things should be noted in this figure. First, various techniques employed by different groups give results that are consistent with one another over a large range of rates (the actual measured rates range from 10^{-4} to 10^{10} cm^{-3} s^{-1}. Second, the critical supersaturations (those for which $J = 1$ cm^{-3} s^{-1}) are predicted reasonably accurately by the classical theory. Third, however, the classical theory does not give the correct temperature dependence of the nucleation rate. It ranges from being several orders of magnitude too low at low temperatures to being several orders of magnitude too high at high temperatures. Simple modifications of the classical theory give no improvement in this regard. Many other substances have been studied less extensively but agree at least qualitatively with the classical prediction of critical supersaturation. This is true for the condensation of steam to water, for example.

For other substances, the classical prediction of critical supersaturation is significantly in error, leading to rates that are in error by tens of orders of magnitude. In some cases, there are understandable reasons for the discrepancies. Gaseous styrene nucleates far faster than classical theory predicts, a fact that has been explained [11] by the spontaneous thermal gas-phase polymerization of very small numbers of short-chain styrene polymers, which then heterogeneously nucleate the liquid phase at low supersaturations. Gaseous mercury also nucleates much earlier than is expected classically, but this can be explained by noting that small clusters of mercury are distinctly nonmetallic, so that their effective surface free energy will be very different from that of metallic bulk mercury [12].

Leaving aside these explainable exceptions, there remain a number of cases in which large discrepancies with classical theory are seen. Some of these give critical supersaturations 30 to 40% lower than predicted [3]. On the other hand, some dipolar molecules appear to nucleate at significantly higher critical supersaturations than predicted [13]. It seems clear that a better first-principles theory of nucleation is needed.

C. Computer Simulations

The most straightforward way to simulate nucleation of a gas–liquid transition on a computer, at least in principle, would be to start with a uniform supersaturated gaslike configuration and use molecular dynamics to solve Newton's equations of motion. Under favorable circumstances, a liquid drop would appear after a certain time. Averaging over a series of runs would give the average nucleation rate J, and the temperature and initial gas density could then be varied systematically to study their effect. Such simulations have indeed been used to see homogeneous gas-phase nucleation on a computer [14], although the type of systematic study just de-

scribed has never been carried out. There are two ways in which these differ significantly from laboratory experiments, however. First, typical simulation volumes are on the order of 10^{-20} cm^3 and simulation times are on the order of 10^{-10} s, so the nucleation rates that result (if anything is to be seen) must be 10^{30} cm^{-3} s^{-1}, 30 orders of magnitude larger than those typical in the laboratory. Higher supersaturations than normal must therefore be simulated. Second, as a cluster condenses its temperature will tend to rise as the heat of condensation is given off. In the laboratory, this is controlled by having a large excess of a background gas to carry off the heat through collisions with the growing cluster. In the simulation, changes in temperature can be prevented only by artificially scaling velocities, a procedure that may not be equivalent.

Under realistic experimental conditions, critical or near-critical clusters are very rare objects surrounded by a nearly ideal gas of monomer. An attractive idea is therefore to simulate an isolated cluster of a given size to determine its free energy. The variation of that free energy with number of particles can then be used in nucleation theory to relate forward and backward rate constants, or at least to estimate the height of the free-energy barrier to nucleation. The problem is to define a cluster and how to keep it intact long enough to determine its average properties. If the temperature is moderately high and the volume large enough, atoms will evaporate from the surface of the cluster and move away, changing its identity. The problem is especially acute with near-critical clusters, which by definition are unstable: they tend spontaneously to gain or to lose particles to reach the more stable gas or liquid states.

Lee et al. [15] carried out a Monte Carlo simulation of such physical clusters. They defined a cluster of size n by the constraint that all n particles lie within a sphere of radius R_c centered on the center of mass of the cluster [7]. This corresponds to placing the cluster inside a rigid spherical container. Clearly, R_c cannot be made too large, or many configurations of the n-particle system will have widely separated particles that will not look at all like clusters. Nor can R_c be too small, or clusters will be artificially forced to be more compact than they would be in free space. The hope is that there will be a range of values of R_c over which the cluster properties are insensitive to the particular choice made. The Monte Carlo simulations of Ref. 15 focused on such constrained clusters of Lennard–Jones particles, looking most extensively at 87-atom clusters. For low enough temperatures (below about $0.65\epsilon_{LJ}$) a plateau region was seen in the variation of free energy with constraining volume, but this was not true at higher temperatures. For smaller clusters, even lower temperatures were necessary to see a plateau.

Reiss and co-workers [16] have made a significant advance in nucle-

ation theory that has the potential for allowing accurate simulation of cluster free energies. They retain the definition of a physical cluster just described (involving a sphere of volume v about the center of mass) but they make the crucial physical observation that such a sphere not only constrains the n particles to remain inside, but also constrains all the other particles to remain outside. This second aspect contributes an extra term to the work $W(n,v)$ to create a cluster of size n in volume v, in addition to the term from the configurational integral of the cluster itself. W is now a function not of one but of two variables, and the saddle point that determines the nucleation barrier is found by setting the derivatives of W with respect to both n and v equal to zero. The critical cluster is then in unstable "material" equilibrium with respect to the surrounding gas (because the second derivative of W with respect to n is negative, just as in the classical capillarity approximation) but in stable mechanical equilibrium (because the second derivative is positive with respect to v). In Reiss's approach, the shell about the cluster is no longer a constraint but a procedural device for organizing the theory. Quantitative results from this new idea should prove very interesting.

D. Density Functional Theory

In classical nucleation theory, the density at the center of a cluster is assumed to be equal to the bulk liquid density, and the shape and free energy of the surface are taken to be identical to those of a planar interface. There is then only one parameter needed to define a spherical cluster: its radius R (or, equivalently, its particle number n). The critical nucleus is then located by setting the derivative of the grand potential Ω with respect to this single parameter equal to 0:

$$\frac{d\Omega}{dR} = 0 \qquad \text{at } r = R^*$$

In general, of course, there is no reason why the density at the center of a cluster should be constrained to equal the bulk liquid density, nor is there any reason why the density profile should match that at a planar interface. By symmetry, a spherical profile should still give the lowest free energy, but the density $\rho(r)$ in that profile should not be constrained other than to require that it approach the bulk vapor density at large distances. The grand potential is now a *functional* Ω_V of that density, and the nucleation condition is now that the functional derivative be equal to zero:

$$\frac{\delta\Omega_V}{\delta\rho(r)} = 0 \qquad \text{at } \rho(r) = \rho^*(r)$$

A functional relation can be pictured as the dependence of a quantity (here Ω_V) on an infinite number of variables (here the density at all possible points in space). This multidimensional free energy has a minimum at the uniform vapor density and a second, lower minimum at the uniform liquid density. Between these two minima there is a saddle point that can be found by setting the functional derivative to zero. The matrix of second derivatives of the functional evaluated at this saddle point has a single negative eigenvalue, corresponding to the direction of motion over the barrier.

One can define a similar functional derivative to identify the planar equilibrium gas–liquid interface exactly at two-phase coexistence. Here the boundary conditions are that the density approach the vapor or liquid density as z approaches $\pm\infty$. Such an interface is thermodynamically stable (unlike the critical nucleation profile), so the matrix of second derivatives has no negative eigenvalues. It does, however, have a zero eigenvalue, corresponding to uniform translation of the interface in the z-direction, an operation that costs no free energy in the grand ensemble. This zero eigenvalue turns into the negative eigenvalue in supersaturated vapor nucleation. Density functional calculations of equilibrium interfaces were discussed more extensively in Chapter 3.

The grand potential functional is related to the Helmholtz free-energy functional by a Legendre transform:

$$\Omega_V = F - \mu N = F - \mu \int d\mathbf{r}\, \rho(\mathbf{r})$$

The functional derivative equation that gives the density profile in the critical nucleus is then

$$\frac{\delta F}{\delta \rho(r)} = \mu \qquad \text{at } \rho(r) = \rho^*(r)$$

It remains to choose a reasonably accurate Helmholtz free-energy functional and use it to solve this equation. The choice of $F[\rho(\mathbf{r})]$ has been discussed elsewhere in this book (see in particular Chapter 3), so we confine our attention to those few choices that have actually been used in nucleation calculations.

The first density functional calculation of nucleation was published by Cahn and Hilliard [17] in 1959. Their calculation actually applied to nucleation of phase separation in a binary mixture, but the problem is formally identical to gas–liquid nucleation with the substitution of the one-particle density $\rho(r)$ for the concentration profile $c(r)$. We make that substitution in presenting their results below. Cahn and Hilliard took a square-

gradient form for the free energy:

$$F[\rho(\mathbf{r})] = \int d\mathbf{r}[f_u(\rho(\mathbf{r})) + K(\nabla\rho(\mathbf{r}))^2]$$

Here $f_u(\rho)$ is the Helmholtz free energy per unit volume of a uniform system with density ρ everywhere in space; it is a local function rather than a nonlocal functional. A term such as f_u must be evaluated in some kind of mean field theory because, strictly speaking, uniform densities within the coexistence region cannot exist at equilibrium. Nonlocal terms are approximated by the second, square-gradient term, whose coefficient K is allowed to depend on the local density as well. Evaluating the functional derivative then leads to the differential equation

$$\frac{\partial f_u}{\partial \rho} - 2K\nabla^2\rho - \frac{\partial K}{\partial \rho}(\nabla\rho)^2 = \mu$$

which must be solved for the radial geometry of the nucleation problem.

Cahn and Hilliard were interested in nucleation of binary mixtures, and so took f_u from a regular solution theory of mixtures. It has the double minimum structure of Fig. 1 and should apply at least qualitatively to single-component fluids as well. They showed that near the coexistence curve, the results go over to classical nucleation theory, in that the critical nucleus then becomes very large, the density at its center becomes equal to the bulk liquid density, and that density stays almost uniform until the edge of the droplet is reached. The density through the interface is then close to that of a planar interface. As the spinodal is reached, however, the behavior is quite different in this "nonclassical" theory. First, and most important, the barrier to nucleation vanishes as the spinodal is approached, as it must in any correct theory (it does not in classical theory). Second, the nature of the critical droplets changes as the spinodal is approached. As the metastable region is penetrated, the radius of the critical nucleus first decreases (as in classical theory), but then it begins to increase and in fact *diverges* as the spinodal is approached. At the same time, the density at the center of the nucleus approaches the vapor density. Near the spinodal, then, the critical nucleus is large in extent but small in amplitude.

The nature of nucleation near a spinodal has been reexamined more recently by Unger and Klein [18]. They considered the case in which the range of interactions becomes asymptotically large and showed that this limit gives qualitatively different behavior from that seen for nucleation with short-range potentials, as implied by the Cahn–Hilliard model. In particular, for dimensions below 6, the critical nucleus becomes highly

ramified (fractal) near the spinodal and growth occurs preferentially near its center, not at its surface. The authors calculated critical exponents, such as that for the vanishing of the free-energy barrier as the spinodal is approached. Subsequent work verified their theoretical predictions with Monte Carlo simulations of Ising systems [19].

Oxtoby and Evans [4] have investigated another nonclassical density functional model for nucleation. They focused their attention not near the spinodal or near coexistence, but on conditions accessible to experiment, namely those in which the nucleation rate J is on the order of 1 cm^{-3} s^{-1}. They took as their interaction potential the sum of a hard sphere of diameter σ and a Yukawa attractive tail

$$\phi_{att}(r) = \frac{-\alpha\lambda^3 \exp(-\lambda r)}{4\pi\lambda r}$$

The free-energy functional they used was then

$$F[\rho(\mathbf{r})] = \int d\mathbf{r}\, f_h(\rho(\mathbf{r})) + \frac{1}{2}\iint d\mathbf{r}\, d\mathbf{r}'\rho(\mathbf{r})\rho(\mathbf{r}')\phi_{att}(|\,\mathbf{r} - \mathbf{r}'\,|)$$

where $f_h(\rho(\mathbf{r}))$ is the free-energy density of a uniform hard-sphere fluid, treated locally. The effect of the attractive potential was included in perturbation theory but not in the more restrictive square-gradient approximation used by Cahn and Hilliard. For this model the equation of state depends only on the hard-sphere diameter σ and on the potential strength parameter α; the inverse range parameter λ affects only interfacial properties (but, as we shall see, in a very significant way).

The functional derivative equation to be solved for the properties of the critical nucleus then has the form

$$\mu_h[\rho(\mathbf{r})] = \mu - \int d\mathbf{r}'\, \rho(\mathbf{r}')\phi_{att}(|\,\mathbf{r} - \mathbf{r}'\,|)$$

where $\mu_h(\rho)$ is the hard-sphere chemical potential, a known nonlinear function. Note that this is an integral equation, as opposed to the differential equation that arises in the square-gradient theory. Integral equations have some significant advantages, especially in the ease in incorporating the proper boundary conditions. This integral equation was solved by guessing an initial droplet profile (such as a step function with a certain radius R_0) and inserting it into the right side. The function $\mu_h(\rho)$ was then inverted over a grid of distances r from the center of the nucleus to calculate a new density profile, which was then reinserted into the right side and the iteration continued. When applied to a planar interface at equilibrium, this iteration process converges to the stable solution. For nu-

cleation, however, there is an interesting difference because the equilibrium is an unstable one. If R_0 is chosen too small, iteration causes the droplet to disappear, giving a uniform vapor; if it is too large, iteration causes it to grow indefinitely in extent, giving a uniform liquid. For the proper choice of R_0 in between, iteration gives an *almost* convergent density profile (whose free energy can then be calculated) before the nucleus again falls off on one side or the other of its saddle point.

The calculation of Ref. 4 revealed a very strong dependence of nucleation behavior on the range of the attractive potential λ. For the choice $\lambda = 1$ (in units of σ^{-1}) the variation of barrier height with vapor density is compared in Fig. 4 with the result of classical nucleation theory. Note that the classical theory involved exactly the same free-energy functional as that used for the nonclassical theory. The nonclassical theory properly gives a vanishing barrier at the spinodal, and the horizontal line shows the point where the barrier height is $71k_BT$; under these conditions classical theory predicts a rate of 1 cm^{-3} s^{-1}. Clearly, the predicted critical supersaturations are very different in classical and nonclassical theories. In fact, they are *too* different, because, as we have stated, experiment is in at least rough agreement with classical theory. This observation led to an investigation of the role of the range parameter λ. Figure 5 shows the ratio of nonclassical to classical rates as a function of λ, under conditions where the classical rate is 1 cm^{-3} s^{-1}. For $\lambda = 1$, the nonclassical rate is much larger (19 orders of magnitude) than the classical. An increase of λ to 1.5 (which corresponds to a decrease in the range of the potential by only 33%) brings the two into agreement, within experimental error. The Yukawa potential is not a good representation of actual intermolecular interactions, but if the experimental surface tension or its temperature derivative is used to estimate an appropriate value of λ, values of 1.5 to 1.9 result. For these choices the nonclassical and classical theories are in approximate agreement at this temperature. In other words, the success of classical nucleation theory is an accident: if real attractive forces had slightly longer ranges, very large nonclassical effects would be seen.

The reverse transition of bubble formation in a liquid (cavitation) was also studied in Ref. 4. Here, the nonclassical effects were predicted to be even larger and should easily be visible for realistic choices of the parameter λ. Cavitation is studied by exerting a tensile force (negative pressure) on a liquid and measuring the point at which the liquid finally breaks apart through nucleation of bubbles. The prediction is that the liquid will have a lower limiting tensile strength than that from classical theory. There are experiments on liquid tensile strengths [1], some of which do give lower values than classical, but other factors may be contributing as well. Any impurities that can nucleate cavitation heteroge-

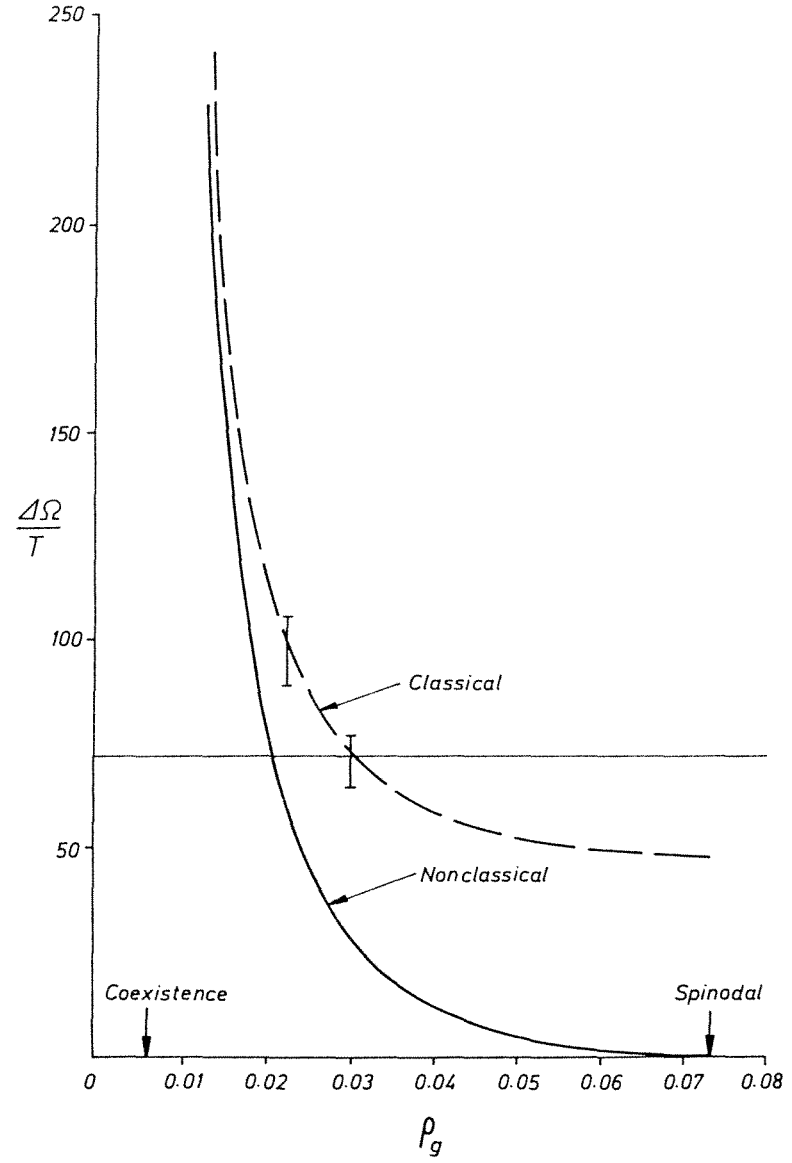


Fig. 4 Dependence of barrier height divided by temperature on vapor density, for $T = 0.6T_c$ and $\lambda = 1$. Densities are expressed in units of σ^{-3}, where σ is the hard-sphere diameter. (From Ref. 4.)

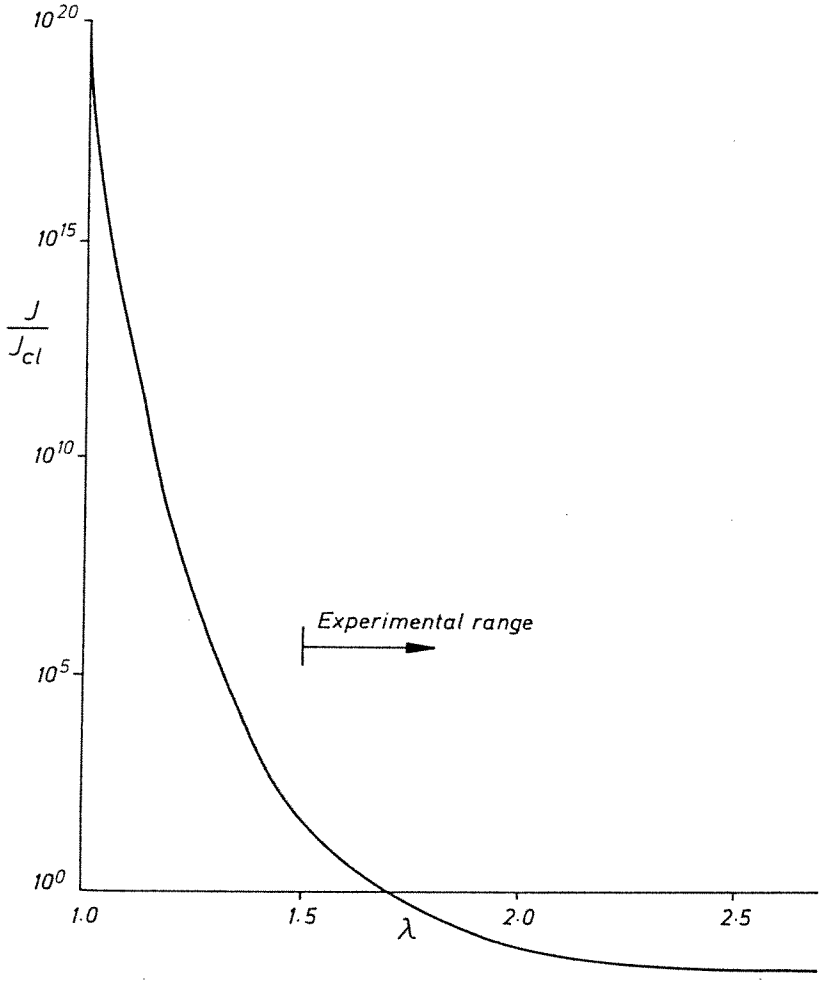


Fig. 5 Ratio of the nonclassical to the classical rate of critical droplet formation, expressed as a function of the range parameter λ at which the classical rate is 1 $cm^{-3}\ s^{-1}$. The experimentally accessible range of values for λ is indicated. $T = 0.6T_c$. (From Ref. 4.).

neously could cause similar trends, and impurities are much harder to eliminate from liquids than from gases.

The free-energy functional employed in Ref. 4 is quite a simple one, but earlier work [20] showed that when a similar functional was applied to stable clusters, it gave rather good agreement with molecular dynamics

simulation. For a quantitative comparison of the nonclassical nucleation theory of condensation with experiment, a better interaction potential and density functional are needed. For example, a Lennard–Jones potential gives a better representation than a Yukawa potential and has been treated in hard-sphere perturbation theory [20a], matching the temperature dependence shown in Fig. 3. The local approximation to the hard-sphere free energy could be replaced by a nonlocal weighted density approximation like those discussed in Chapter 3. Finally, the effect of the attractive tail could be treated more accurately by multiplying the product of densities at two points $\mathbf{r}$ and $\mathbf{r}'$ by the hard-sphere pair distribution function $g_{\text{hs}}(|\mathbf{r} - \mathbf{r}'|)$ evaluated at a suitable average density.

The approach of Ref. 4 focused on the free-energy barrier and used the simple Becker–Döring preexponential to find actual rates. A more complete and consistent theory would use a time-dependent generalization of density functional theory to calculate the preexponential as well. Such a theory has been proposed by Langer and Turski [21] for use near the critical point. They show that the nucleation rate has the form (in our notation)

$$J = \frac{\kappa}{2\pi} P_0 \exp\left(\frac{-\Delta\Omega^*}{k_B T}\right)$$

Here $\Delta\Omega^*$ is the barrier height that we have been calculating, and the prefactor has been divided into two parts. P_0 is a statistical factor, a measure of the phase-space volume of the saddle-point region; it is a generalization of the Zeldovich factor discussed earlier for classical nucleation. It can be related to the eigenvalues of the second functional derivative matrix of Ω with respect to $\rho(\mathbf{r})$, evaluated in the metastable vapor state (where all are positive) and at the saddle point (where one is negative and three are zero, corresponding to uniform translation of the droplet). This factor can be evaluated for a given model of the free-energy functional, although not without difficulty. The dynamical prefactor κ gives the rate of exponential growth of the unstable mode of the critical nucleus and depends on transport processes in the liquid, not just on the free energy. Langer and Turski suggest some approximate ways to estimate this factor near the critical point, but a general theory is lacking. It will also be of interest to calculate the dynamical prefactor for the case of cavitation. In this case there is no simple kinetic model for the growth of cavities in a liquid, in contrast to the situation for accretion of particles by droplets in the gas phase, which can be approximated by gas kinetic theory. A major challenge is to develop an accurate and consistent dynamical extension of density functional theory for nucleation.

III. CRYSTALLIZATION OF LIQUIDS

A. Classical Nucleation: Theory and Experiment

The classical nucleation theory of Becker and Döring described in Section II.A has been extended to the liquid-to-crystal transition by Turnbull and Fisher [22]. The nucleation rate is again taken to have the form

$$J = J_0 \exp\left(\frac{-\Delta\Omega^*}{k_B T}\right)$$

and the variation of barrier height with number of particles n in the crystal nucleus has the form

$$\Delta\Omega(n) = n\,\Delta\mu + (36\pi)^{1/3}\rho_s^{-2/3}\sigma_{sl}n^{2/3}$$

where $\Delta\mu$ is the difference in chemical potential (or Gibbs free energy per particle) between bulk solid and bulk liquid, ρ_s the particle density in the bulk solid, and σ_{sl} the solid–liquid surface free energy.

There are several important differences between the gas–liquid and the liquid–solid transition that make the latter more difficult to study:

1. It is much less obvious how to identify the size n of a crystalline cluster. Even in the case of a liquid droplet forming from a supersaturated vapor, we have seen in Section II.C that there is some uncertainty in how to identify which particles form part of a cluster; for a crystalline cluster in a liquid the distinction between the two phases is even smaller and there is no simple criterion for saying which particular particles are crystalline.
2. The prefactor J_0 is rather different from that in the gas–liquid transition. In the simplest picture, the dynamics involve not molecules moving in nearly ideal gases, but rather "jumping" across the interface to attach. It is not clear whether this jumping occurs at a characteristic phonon frequency or at a significantly smaller frequency given by the rate of diffusion of molecules over intermolecular distances. Some rearrangement of configurations occurs at the interface, but actual transport of material is not necessary in a single-component fluid.
3. $\Delta\mu$ is less accessible to experimental measurements than the corresponding quantity for the gas–liquid transition, namely the supersaturation S. Near equilibrium freezing,

$$\Delta\mu = -\Delta h + T\,\Delta s = -\Delta s(T_m - T)$$

where T_m is the equilibrium melting temperature and Δh and Δs are the enthalpy and entropy of fusion per molecule at T_m. Frequently, nucleation occurs hundreds of degrees below T_m, however, and it is not evident whether this expression can be extrapolated so far. To correct it, heat capacity measurements on the undercooled liquid are needed.

4. The solid–liquid surface free energy is much more difficult to measure than the gas–liquid surface tension. Reference 23 identifies a small number of liquids for which σ_{sl} has been directly measured and discusses some different theoretical models that have been employed to estimate it. Because it is much smaller in magnitude than the corresponding gas–liquid or gas–solid surface free energies, the uncertainties in its value are much larger.

Combining all the factors just discussed and including a Zeldovich factor to take into account the difference between an equilibrium and a steady-state distribution of nucleus sizes leads to the following result in the classical model for the nucleation rate per unit volume [24]:

$$J = 2\rho_s^{-2/3}\sqrt{\frac{\sigma_{sl}}{k_BT}}\frac{k_BT}{h}\exp\left(-\frac{\Delta G_a}{RT}\right)\rho_l V\exp\left[\frac{-16\pi\sigma^3}{3\rho_s^2 k_BT(\Delta s)^2(T_m-T)^2}\right]$$

Here h is Planck's constant and R the gas constant; $(k_BT/h)\exp(-\Delta G_a/RT)$ is an estimate of the jump rate of particles across an activated barrier as they join the crystalline nucleus. As the temperature is brought below T_m, the nucleation barrier height $\Delta\Omega^*$ drops rapidly from infinity. For most simple liquids, this second exponential dominates the temperature dependence, and because it varies so sharply, the nucleation rate changes from being negligibly slow to being unmeasurably fast over a small temperature range. In this case a well-defined undercooling limit can be identified. From a rough estimate of the preexponential factor, the temperature T_n at which the nucleation rate becomes 1 cm^{-3} s^{-1} is

$$\frac{\Delta\Omega^*(T_n)}{k_BT_n} \simeq 76$$

In some cases, the temperature dependence of the first exponential, $\exp(-\Delta G_a/RT)$, becomes important as the liquid is undercooled. This is particularly true near a glass transition. In this case the nucleation rate can pass through a maximum as the temperature decreases; if this maximum rate is small enough, crystallization will be bypassed and a glass will form.

In nucleation experiments, extraordinary care must be taken to exclude impurities that may catalyze heterogeneous nucleation. It is not enough

merely to purify the samples studied because catalytic impurities occur at concentrations well below the limits of detection. The method developed by Turnbull [25] is to form an emulsion of the undercooled liquid in oil by using a high-speed blender to break the liquid into tiny drops. If the drops are small enough, many of them will not contain nucleating impurities and may nucleate homogeneously. As the emulsion is cooled, the process of crystallization is followed by measuring the volume of the suspension or by detecting the heat evolved during crystallization. Some droplets crystallize at higher temperature about impurities, but the temperature at which all the remaining droplets crystallize is identified as the homogeneous nucleation temperature T_n.

As mentioned above, the surface free energy σ_{sl} is known for rather few systems. As a result, rather than test classical nucleation theory, what is usually done is to assume that it is correct and use the measured undercooling limit T_n to calculate σ_{sl}. Turnbull studied a variety of liquid metals in this way [25,26] and proposed the following correlation between σ_{sl} and the enthalpy of fusion Δh:

$$\sigma_{sl} = \frac{\alpha \, \Delta h \rho_s^{2/3}}{f}$$

Here f is a factor that relates the area A per surface atom to the density in the bulk crystal

$$A = \frac{f}{\rho_s^{2/3}}$$

It is equal to 1.12 for the bcc (110) face and 1.09 for the fcc (111) face. Turnbull found that values for the proportionality constant α clustered around 0.45 for metals and around 0.32 for water and semimetals. His observed relative undercoolings $(T_m - T_n)/T_m$ ranged up to 0.25 and clustered around 0.18.

More recent work has succeeded in obtaining larger undercoolings before nucleation takes place. For example, Bosio and co-workers [27] undercooled gallium droplets suspended in ethanol by 140 K, nearly 50% of T_m. Since 1980, Perepezko and co-workers have studied a variety of metals. In the case of bismuth, for example, they extended the measured undercooling limit from 90 K (16% of T_m) to 227 K (41% of T_m) [28]. These results show that at least some of the early measurements were not of homogeneous nucleation. Rather, it is likely that the nucleation was catalyzed by the surface of the droplet where the metal makes contact with the oil and where some oxidation takes place. Different emulsion compositions can give different undercooling limits. In fact, Perepezko

argues that most, if not all, observed undercooling limits represent such surface-catalyzed heterogeneous nucleation. One additional intriguing observation is that deep undercooling in gallium can lead to the formation of metastable crystalline phases instead of the thermodynamically stable phase [28,29]. Reference 23 discussed additional experimental work in this area.

As mentioned earlier, independent measurements of σ_{sl} are available for only a few substances. One of these is water, for which $\sigma_{sl} = 0.029$ J m^{-2} [30]. A study by Wood and Walton [31] was able to determine the temperature dependence of the nucleation rate over a range of 2 to 3°C in the vicinity of −36°C. A two-parameter fit of the preexponential factor and of σ_{sl} gave $\sigma_{sl} = 0.020$ J m^{-2} and a preexponential factor larger than classical theory by a factor of 10^{15}. Because the barrier height is proportional to the surface free energy raised to the third power, it is clear that routine use of classical nucleation theory disagrees with experiment in this case [23].

The temperature dependence of nucleation can be studied most directly in glass-forming liquids. Most of these are liquid mixtures rather than one-component liquids, but we nonetheless discuss their behavior here. Angell et al. [32] studied nucleation in solutions of LiCl in water with a differential scanning calorimeter. This work directly determined the temperature of maximum nucleation rate and measured the variation of the rate about that temperature. They found that the observed maximum rate was limited by the rate of growth of critical nuclei to fill the droplet, so that nuclei that form at lower temperatures grow too slowly to be observed. Near and below the glass-transition temperature, it may take a long time for a steady-state cluster distribution to be established. In this case, a time-dependent theory of nucleation is needed [33], and slow crystallization may arise not from a small value of J (steady state) but from a long transient time. In favorable cases the number of nuclei forming as a function of time can be measured and classical nucleation theory tested [34]. Weinberg and Zanotto [35] have pointed out that if the transient time and steady-state rate are both measured, a clean test of classical nucleation theory is possible that does not rely on assumptions about the form of the preexponential factor. Their study showed that the temperature dependence of nucleation in certain silicate glasses was not well explained by classical theory.

B. Computer Simulations

Computer simulation of crystal-to-liquid nucleation faces a different set of problems from those discussed in Section III.C for gas–liquid nucle-

ation. Because the number of particles in a crystalline nucleus seems impossible to define, simulations that calculate the free energies of constrained clusters of particular sizes appear unlikely (see, however, Ref. 35a). Nucleation of crystals is readily observed in molecular dynamics simulations of undercooled liquids, although of course the extent of undercooling is much larger than that typical for laboratory experiments. The first observation of nucleation was a simulation by Mandell et al. [36] of a Lennard–Jones liquid that nucleated into a (metastable) bcc crystal instead of an fcc crystal. They estimated the critical nucleus to contain 40 to 70 atoms and identified the time of its appearance by randomizing the velocities of the particles at a series of times and determining whether crystallization still took place. A number of other simulations of nucleation are summarized in Ref. 23. Nucleation seems to occur so readily on the computer that those investigating glass transitions have to go to some pains to avoid crystallization.

A crucial question, however, is whether computer-simulated nucleation is truly homogeneous or whether it is determined by the periodic boundary conditions used on the computer. This question was raised in a study by Honeycutt and Andersen [37]. The rate of homogeneous nucleation should be proportional to the size of the sample and should increase with the number of particles. Instead, these authors found that the nucleation rate continued to decrease as the system size increased, even out to the largest sizes studied (1500 atoms). This indicates that the boundary conditions have a large effect on the crystallization. Many of the earlier results must be reinterpreted in the light of this observation.

C. Density Functional Theories of the Crystal–Melt Interface

Before we use density functional theories to look at nucleation of crystals from the melt, we consider a slightly simpler inhomogeneous liquid–crystal system: the equilibrium planar interface between a crystal and its melt. Recall from Chapter 9 that in density functional theory a crystal is considered a highly nonuniform liquid, and that its free energy is estimated from properties of the liquid. In the case of a planar interface there is now an additional complication: the order parameters themselves change in the direction perpendicular to the interface. We describe here two theoretical calculations of the structure and free energy of a crystal–liquid interface. One [38,39] is based on a fairly drastic approximation to the free-energy functional, but within that functional is a rather accurate calculation of the surface properties; the second [40] employs a more accurate functional but is an approximate variational calculation within that

functional. Reference 41 gives a fuller treatment of density functional theories of liquid–solid systems.

The grand potential functional has the form (Chapter 3)

$$\beta\Omega_V[\rho] = \beta\mathcal{F}[\rho] - \mu N = \beta\mathcal{F}_{id}[\rho] + \beta\mathcal{F}_{ex}[\rho] - \beta\mu \int d\mathbf{r}\, \rho(\mathbf{r})$$

where $\mathcal{F}_{id}[\rho]$ is the Helmholtz free-energy functional for an ideal (noninteracting) fluid with density $\rho(\mathbf{r})$, given by

$$\beta\mathcal{F}_{id}[\rho] = \int d\mathbf{r}\, \rho(\mathbf{r})[\ln(\lambda^3\rho(\mathbf{r})) - 1]$$

where λ is the thermal de Broglie wavelength of the particles. $\mathcal{F}_{ex}$ is the difference between the actual Helmholtz free energy and the ideal part. In studying crystallization, we are only interested in the difference between the grand potential of the (nonuniform) solid and that of the uniform liquid at density ρ_l. At constant chemical potential, this difference is

$$\beta\Delta\Omega_V = \beta\mathcal{F}_{id}[\rho] - \beta\mathcal{F}_{id}[\rho_l] + \beta\mathcal{F}_{ex}[\rho] - \beta\mathcal{F}_{ex}[\rho_l] - \beta\mu \int d\mathbf{r}\, (\rho(r) - \rho_l)$$

In Refs. 38 and 39, the excess free energy was written as a functional perturbation expansion about the liquid density. Recall from Chapter 3 that the second functional derivative evaluated in the liquid state has the form

$$-\beta \frac{\delta^2\mathcal{F}_{ex}}{\delta\rho(\mathbf{r}_1)\,\delta\rho(\mathbf{r}_2)} = c^{(2)}(r_{12})$$

in terms of the pair direct correlation function of the liquid. If the expansion of $\mathcal{F}_{ex}$ is truncated at second order (ignoring $c^{(3)}$ and higher-order terms), the grand potential difference becomes

$$\beta\Delta\Omega_V = \int d\mathbf{r}_1\, \rho(\mathbf{r}_1) \ln\frac{\rho(\mathbf{r}_1)}{\rho_l} - \int d\mathbf{r}_1\, \Delta\rho(\mathbf{r}_1) - \frac{1}{2}\int d\mathbf{r}_1\, d\mathbf{r}_2\, c(r_{12})\, \Delta\rho(\mathbf{r}_1)\, \Delta\rho(\mathbf{r}_2)$$

where

$$\Delta\rho(\mathbf{r}) = \rho(\mathbf{r}) - \rho_l$$

For a uniform solid, the density can be written as a Fourier expansion:

$$\rho(\mathbf{r}) = \rho_l\left(1 + \sum_{i=0} \mu_i e^{i\mathbf{k}_i\cdot\mathbf{r}}\right)$$

Here the $\mathbf{k}_i$ are the reciprocal lattice vectors of the solid (plus $\mathbf{k}_0 = 0$) and the μ_i are the crystalline order parameters, equal to zero in the liquid state. For a solid–liquid interface, the order parameters depend on the z-coordinate perpendicular to the interface, so that

$$\Delta\rho(\mathbf{r}) = \rho_l \sum_{i=0} \mu_i(z) e^{i\mathbf{k}_i\cdot\mathbf{r}}$$

Although the density itself varies extremely rapidly through space, it is a plausible assumption that the order parameters vary more slowly. Thus we make a square-gradient approximation for the nonlocal part of the free energy, writing

$$\Delta\rho(\mathbf{r}_2) \simeq \rho_l \sum_{i=0} \left[\mu_i(z_1) + \frac{\partial\mu_i(z_1)}{\partial z_1}(z_2 - z_1) + \frac{1}{2}\frac{\partial^2\mu_i(z_1)}{\partial {z_1}^2}(z_2 - z_1)^2 + \cdots \right] e^{i\mathbf{k}_i\cdot\mathbf{r}_2}$$

and inserting this into the expression for $\beta\Delta\Omega$. The resulting equations include terms such as

$$\int d\mathbf{r}_1\, d\mathbf{r}_2\, c(r_{12}) \left(\sum_i \mu_i e^{i\mathbf{k}_i\cdot\mathbf{r}_1}\right) \left[\sum_j \frac{\partial^2\mu_j}{\partial {z_1}^2}(z_2 - z_1)^2\, e^{i\mathbf{k}_j\cdot\mathbf{r}_2}\right]$$

If μ_i changes on a distance scale large relative to $\exp(i\mathbf{k}_i\cdot\mathbf{r}_1)$—an assumption implicit in the square-gradient approximation—then the only important contributions come from $\mathbf{k}_j = -\mathbf{k}_i$. After several further transformations [41], the final result is

$$\Delta\Omega_V = \Delta\Omega_V^{(\text{local})}[\rho(\mathbf{r})] - \frac{1}{4}\int d\mathbf{r}_1 \sum_i (\hat{k}_i\cdot\hat{z}_1)^2 c''(k_i) \left|\frac{\partial\mu_i}{\partial z_1}\right|^2$$

This expression has the usual square-gradient form, in which the "local" terms are evaluated for a crystal with the local values of the order parameter throughout space. This first term is always positive, because the free energy of the liquid and the crystal are minima, and any order parameters in between must give higher free energies. Near a peak in the structure factor, the second derivative of the Fourier-transformed direct correlation function, $c''(k)$, is negative, making an additional positive contribution to the excess free energy. Together, the two terms give the excess free energy to create a planar interface. The first term is minimized by the sharpest possible order parameter profile (to minimize the amount of space in which $\Delta\Omega_V^{(\text{local})}$ is nonzero), and the second by the broadest possible one (to make the gradients of the order parameters as small as possible). The resulting compromise between the two terms gives the

actual interfacial width and free energy. It is evident that if terms for which $c''(k)$ is *positive* are to be included, the square-gradient approximation can no longer be made. To find the order parameter profiles, the functional derivative of the free energy with respect to each order parameter is set to zero. The result is a set of coupled differential equations:

$$\frac{\partial \Delta \Omega_V^{(\text{local})}}{\partial \mu_i} = -\tfrac{1}{2}(\hat{k}_i \cdot \hat{z})^2 c''(k_i) \frac{d^2 \mu_i}{dz^2}$$

Analogies to coupled equations of motion for interacting particles are discussed in Ref. 41.

The expression just developed for the free energy of an interface involves an infinite sum over reciprocal lattice vectors. In Ref. 39 this sum was approximated further by including only contributions from $\mathbf{k}_0 = 0$ and from the first set of reciprocal lattice vectors $\mathbf{k}_1$. This is a qualitatively reasonable approximation for a body-centered-cubic-forming liquid, in which the first peak in the structure factor is much larger than the subsequent peaks, because it consists of setting $c(k)$ equal to zero beyond the first peak. Furthermore, the second derivative of $c(k)$ at $k = 0$, which appears to have very little effect on the free energy, was set to zero. After making these approximations, the three resulting coupled algebraic and differential equations were solved numerically. The order parameter profile through the interface and the interfacial free energy were calculated for the (100) and (111) surfaces of sodium and potassium, using experimental structure factors for these liquids to calculate $c(k)$. The interface was found to be rather broad (on the order of 16 to 20 Å), suggesting that the interface between the two phases consists of partially ordered liquid and disordered crystalline layers. The two surfaces showed only a small difference (3%) in free energy, and the surface free energy obtained correlated reasonably well with trends predicted on the basis of the density and enthalpy of fusion of a series of metals.

The approximations to the free-energy functional just described were rather drastic: both the perturbation expansion and the sum over reciprocal lattice vectors were truncated at low order. A more accurate functional was used by Curtin in Ref. 40 to study the surface free energy of a hard-sphere and a Lennard–Jones interface. He used a weighted density approximation of the type described in Chapters 3 and 9 and parametrized the uniform solid as a sum of Gaussians:

$$\rho_s(\mathbf{r}) = (\pi\epsilon^2)^{-3/2} \sum_{\mathbf{R}_i} \exp\left[\frac{-(\mathbf{r} - \mathbf{R}_i)^2}{\epsilon^2}\right]$$

where ϵ is a width parameter and $\mathbf{R}_i$ denotes a lattice site in the crystal.

This corresponds to a Fourier expansion

$$\rho_s(\mathbf{r}) = \rho_l \sum_{\mathbf{k}_i} \mu_i e^{i\mathbf{k}_i \cdot \mathbf{r}}$$

in which the values of μ_i in the uniform crystal, μ_i^c, are determined by the value of the single parameter ϵ, which is in turn determined variationally. Through the interface, the density was assumed to have the form

$$\rho_s(\mathbf{r}) = \rho_l \sum_{\mathbf{k}_i} \mu_i^c f_i(z) e^{i\mathbf{k}_i \cdot \mathbf{r}}$$

where the "switching functions" $f_i(z)$ were parametrized as

$$f_i(z) = \begin{cases} 1 & z < z_0 \\ \frac{1}{2}\left[1 + \cos\left(\pi \frac{z - z_0}{\Delta z_i}\right)\right] & z_0 < z < z_0 + \Delta z_i \\ 0 & z_0 + \Delta z_i < z \end{cases}$$

The widths over which each order parameter changes from solid to liquid were taken to be

$$\Delta z_i = \left(\frac{k_l}{k}\right)^{\nu} \Delta z_0 \qquad 0 \leq \nu \leq 1$$

Once the equilibrium crystal structure is calculated, if z_0 is held fixed, there are two parameters that define the density profile: ν and Δz_0. These were determined variationally, and the resulting interface was somewhat narrower than that found in Ref. 39, having a thickness of four layers. The surface free energies for the fcc (100) and (111) interfaces differed by less than 5%. The calculated values were in reasonable agreement with simulation results.

D. Density Functional Theory of Crystal Nucleation

The only density functional calculation of crystal nucleation was carried out by Harrowell and Oxtoby [42]. It was based on the same truncated density functional described above for the planar interface and introduced in Refs. 38 and 39. The coupled differential equations that result from the square-gradient approximation are now radial ones and have the form

$$\frac{\partial \Delta \Omega_V^{(\text{local})}}{\partial \mu_i} = -\tfrac{1}{2} c''(k_i) \left\{ (\hat{k}_i \cdot \hat{r})^2 \frac{d^2 \mu_i}{dr^2} + [1 - (\hat{k}_i \cdot \hat{r})^2] \frac{1}{r} \frac{d\mu_i}{dr} \right\}$$

Here $\Delta\Omega_V^{(\text{local})}$ is the uniform grand potential for an undercooled liquid, in which the solid minimum is now deeper than that of the liquid. These equations, which describe the critical nucleus at the saddle point in function space, can be solved by a "shooting" method in which boundary conditions at $r = 0$ are guessed and the profiles integrated out. The boundary conditions at $r = 0$ are then modified until the correct asymptotic boundary conditions are recovered.

As noted in Section III.C, this free-energy functional is such a drastic approximation that quantitative results cannot be expected. However, it is still of considerable interest to compare the nonclassical nucleation that results from the density functional theory with the classical predictions calculated consistently using the *same* free-energy functional (i.e., the thermodynamic properties and the surface free energy are calculated as in Ref. 39). The nonclassical critical nucleus turns out to be qualitatively similar to the classical one; in particular, the order parameters at its center are quite close to those of the bulk crystal, and the interfacial profile is close to a planar one. This similarity appears to arise from the absence of a spinodal in the undercooled liquid: There is no evidence for an instability in the liquid, which would be signaled by a diverging structure factor at some value of k. There are important quantitative differences between the nonclassical and classical predictions, however, as illustrated in Fig. 6 by the variation with undercooling of the height of the nucleation barrier. The nonclassical theory predicts that deeper undercooling will

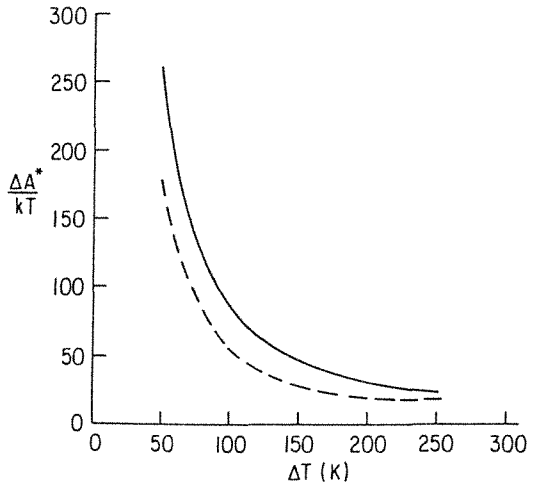


Fig. 6 Calculated free energy of the nucleus in nonclassical theory (solid line) compared with the result from the classical theory (dashed line).

be possible than is given in the classical theory. These results raise the question of whether it is correct to use classical nucleation theory in fitting observed undercooling limits for liquids.

The theory of Ref. 42 calculated only the free-energy barrier to nucleation and estimated the preexponential from classical theory. Grant and Gunton [43] used a generalization of the Langer–Turski [21] theory to estimate this preexponential factor using time-dependent density functional theory. The result was close to that from classical theory.

IV. NUCLEATION IN BINARY MIXTURES

The single-component nucleation theory discussed so far must be generalized to treat binary mixtures. In this section we concentrate on extensions of classical nucleation theory to the binary case and some comparisons with experiment.

The pioneering theory of binary nucleation in the gas–liquid transition is due to Reiss [44]. He used both kinetic and thermodynamic arguments to show that the nucleation rate is determined by passage over a saddle point in a space of droplet compositions. Recall that a liquid droplet is now characterized by the number of particles n_i that it contains of each of two types. In classical theory, the work to create the droplet from the vapor is

$$\Delta\Omega = n_1 \Delta\mu_1 + n_2 \Delta\mu_2 + \sigma(n_1,n_2)A$$

Here $\Delta\mu_1$ and $\Delta\mu_2$ are the differences between the chemical potentials of bulk liquid and gas for the two components and can be calculated using an ideal solution model, σ is the surface tension, and A is the area of the interface. For a spherical nucleus in which the partial molar volumes v_1 and v_2 are independent of pressure, this area can be rewritten as

$$A = (36\pi)^{1/3}(n_1 v_1 + n_2 v_2)^{2/3}$$

Reiss also generalized the classical theory calculation of the preexponential factor to the binary case [44].

The major controversy in binary nucleation theory has revolved about the proper way to find the particle numbers n_1^* and n_2^* that define the saddle point in passing from gas to liquid. The natural expectation is that one must set the derivatives of $\Delta\Omega$ with respect to n_1 and n_2 equal to zero. In Reiss's original work these derivatives were evaluated at constant surface tension; in other words, the composition dependence of σ was not included in evaluating the two partial derivatives. This gave a pair of equations of the form

$$\Delta\mu_i + \frac{2\sigma v_i}{R} = 0$$

where R is the droplet radius. Subsequently, Doyle [45] used Reiss's approach but *added* the extra terms involving $d\sigma/dn_1$ and $d\sigma/dn_2$. The resulting "classical" theory apparently dominated the field for some 20 years; it gave good agreement with some experiments but not with certain water–alcohol mixtures.

In 1981, Renninger et al. [46] pointed out that including the terms with the compositional derivative of the surface tensions is thermodynamically inconsistent. Wilemski [47] has confirmed this in a clear presentation, which shows that if one takes into account surface adsorption, the effect in a thermodynamically consistent theory is to eliminate the additional terms. His revised classical theory, which is thus a return to the original approach of Reiss, given better agreement with experiment in cases where the surface tension varies significantly with composition. Flageollet-Daniel et al. [48] have presented another thermodynamically consistent theory in which the extent of surface adsorption is calculated using a lattice model. The theory gives good agreement with experiment in cases in which the Doyle theory failed (Fig. 7). There is still a need for a first-principles microscopic approach to go beyond the classical model, however. Zeng and Oxtoby [48a] recently extended the nonclassical density functional theory to binary systems and applied it to Lennard-Jones mixtures.

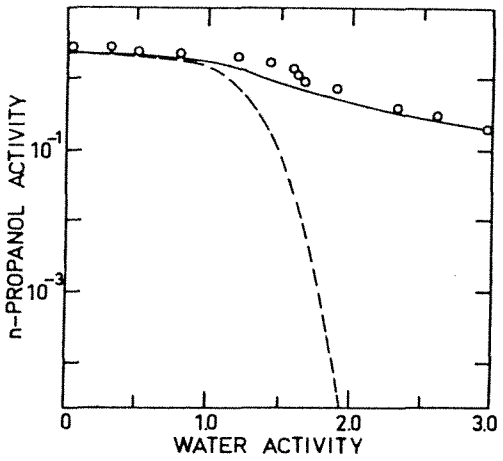


Fig. 7 Comparison of theory and experiment (circles) for the binary mixture water/*n*-propanol at 298 K. The solid line shows the results of the theory of Ref. 48, including surface enrichment, and the dashed line is the result of a theory like that of Ref. 45, which is not consistent thermodynamically. (From Ref. 48.)

Binary nucleation of crystals from the melt has been studied by metallurgists interested in the rapid solidification of alloys. Perepezko and co-workers have studied a large number of low-melting alloys of lead, tin, bismuth, gallium, and other metals; a list of references is given in Ref. 23. They find a general correlation between the dependences of nucleation temperature and equilibrium liquidus temperature on composition. However, they do see metastable phases forming in some deeply undercooled melts, and they also caution that the nucleation they observe may be heterogeneous, occurring at the surface of the liquid metal droplets. In the Cu–Te alloy, which has a deep eutectic, the nucleation temperature falls faster than the liquidus temperature and glass formation occurs over a limited composition range [49]. Thompson and Spaepen [50] used classical nucleation theory for alloys to analyze the observed behavior. They assumed that the composition of the critical nucleus was the same as that of the bulk stable phase and took the surface free energy to be a linear combination of the two single-component surface free energies. Both of these assumptions are open to question [23]. An important aspect that has not been addressed is under what conditions the thermodynamically stable crystal forms as opposed to a metastable solid solution (with a composition close to that of the melt). Such "partitionless solidification" has been seen for Bi–Sn alloys [29]. It seem likely that nucleation may play a role in determining the types of phases seen under rapid solidification (nonequilibrium) conditions, but this problem has not been explored.

V. HETEROGENEOUS NUCLEATION

In the real world, most nucleation is heterogeneous. It is this fact that causes phase transitions to occur close to thermodynamic equilibrium and prevents highly metastable states from forming. Controlled experiments on heterogeneous nucleation are much more difficult to do than on homogeneous nucleation, however, and for this reason our understanding of such processes is much less advanced.

One type of heterogeneous nucleation involves formation of a new phase about a single impurity ion or molecule. This could be described as a limiting case of binary nucleation in which the amount of one of the components is reduced to a very small value, but it is probably more useful to describe it as the formation of a new phase under the influence of a fixed "external" potential from the impurity. This additional potential can significantly lower the free-energy barrier to nucleation. One area of obvious interest is the effect of ions on nucleation of water and ice from

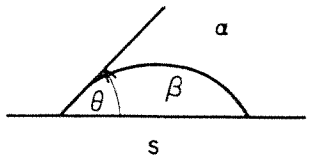


Fig. 8 Contact angle of two phases α and β with substrate.

the vapor present in the atmosphere [1]. Turnbull [51] has shown that impurities can significantly change the preexponential factor in nucleation even in cases where they have no effect on the magnitude of the barrier. Another type of "impurity"-induced nucleation has been mentioned earlier in this chapter: the role of short-chain gas-phase polymers in inducing droplets of monomer to condense [11].

A second type of heterogeneous nucleation occurs when a new phase forms on the surface of a foreign body, such as on the wall of the container or on an aerosol particle. Such surfaces can be flat or can have indentations. Porous media can induce nucleation inside thick or thin channels. Gas-phase aerosol particles that induce nucleation can be large, or they can be small and highly curved, so that the critical nucleus actually forms by enveloping them with a thin layer of material. They can even be highly ramified (fractal) objects. The geometry of the nucleating surface is almost never known, however, so most treatments assume the limiting case of a flat surface. This is the only case we consider here.

Turnbull [52] extended classical nucleation theory to heterogeneous processes occurring on rigid solid substrates. Consider the geometry shown in Fig. 8. Two phases of a substance (α and β) are in contact with a substrate s. These two phases can be vapor and liquid, or liquid and crystal. At equilibrium, there will be a contact angle θ where they meet the substrate. Because there are now two interfaces present, the change in free energy to create this droplet (containing n particles) on the surface will be

$$\Delta\Omega = n\,\Delta\mu - A_{s\beta}(\sigma_{s\beta} - \sigma_{s\alpha}) + A_{\alpha\beta}\sigma_{\alpha\beta}$$

where $\Delta\mu$ is the bulk chemical potential difference between phases α and β, $A_{s\beta}$ and $A_{\alpha\beta}$ are the areas of contact of phase β with substrate and with phase α, and the σ_{ij} are the respective interfacial free energies. For a given value of n, the shape that minimizes $\Delta\Omega$ is a spherical sector, with the contact angle θ given by

$$\cos\theta = \frac{\sigma_{s\alpha} - \sigma_{s\beta}}{\sigma_{\alpha\beta}}$$

The height of the barrier to nucleation is

$$\Delta\Omega^* = \frac{16\pi\sigma^3 f(\theta)}{3\rho_\beta{}^2 \Delta\mu^3}$$

where the only difference from homogeneous nucleation theory is the factor $f(\theta)$, defined by

$$f(\theta) = \frac{(2 + \cos\theta)(1 - \cos\theta)^2}{4}$$

Let us examine the form of this result for the case in which α represents the vapor phase and β the liquid. Consider first the situation when

$$\sigma_{s\alpha} - \sigma_{s\beta} < -\sigma_{\alpha\beta}$$

In this case, a naive application of the equations above would give $\cos\theta < -1$, which is impossible. This simply means that under these conditions a droplet attached to the surface will not form, because the substrate actually causes its free energy to be *higher* than that of a droplet in free space (this can also be seen by noting that as $\theta \to 180°$, the critical nucleus simply detaches itself from the substrate). In this case the situation returns to one of homogeneous nucleation. The second case is defined by the conditions

$$-\sigma_{\alpha\beta} < \sigma_{s\alpha} - \sigma_{s\beta} < \sigma_{\alpha\beta}$$

Over this range a finite contact angle will be seen and the barrier height to nucleation will be reduced by the factor $f(\theta)$, which is less than 1. The third and final case occurs when

$$\sigma_{\alpha\beta} < \sigma_{s\alpha} - \sigma_{s\beta}$$

Here the contact angle has reached 0° and the "droplet" becomes a thin layer spread on the substrate. This is the situation where the liquid "wets" the substrate (see Chapter 3) and the barrier to nucleation disappears.

The contact angle θ is a macroscopic surface property of a three-phase equilibrium, and the same questions about the validity of making macroscopic approximations for droplets of molecular size arise as in the case of homogeneous nucleation. Unfortunately, nucleation experiments have not generally been carried out in cases where all the parameters in the classical theory are known, in order to test the theory. In most cases, θ is used simply as a fitting parameter to bring the theory into accord with experiment. It is clear that there is much theoretical and experimental work to be done in order to understand heterogeneous nucleation.

REFERENCES

1. Hirth, J. P., and Pound, G. M. (1963). *Condensation and Evaporation: Nucleation and Growth Kinetics*, Pergamon Press, London.
2. Zettlemoyer, A. C., ed. (1969). *Nucleation*, Marcel Dekker, New York.
3. Abraham, F. F. (1974). *Homogeneous Nucleation Theory*, Academic Press, New York.
4. Oxtoby, D. W., and Evans, R. (1988). Nonclassical nucleation theory for the gas–liquid transition, *J. Chem. Phys. 89*: 7521.
5. Courtney, W. G. (1961). Remarks on homogeneous nucleation, *J. Chem. Phys. 35*: 2249.
6. Lothe, J., and Pound, G. M. (1962). Reconsiderations of nucleation theory, *J. Chem. Phys. 36*: 2080.
7. Reiss, H., Katz, J. L., and Cohen, E. R. (1968). Translation–rotation paradox in the theory of nucleation, *J. Chem. Phys. 48*: 5553.
8. Katz, J. L., Scoppa, C. J., Ganesh Kumar, N., and Mirabel, P. (1975). Condensation of a supersaturated vapor. II. The homogeneous nucleation of the *n*-alkyl benzenes, *J. Chem. Phys. 62*: 448.
9. Schmitt, J. L. (1981). Precision expansion cloud chamber for homogeneous nucleation studies, *Rev. Sci. Instrum. 52*: 1749.
10. Katz, J. L., Hung, C., and Krasnopoler, M. (1988). The homogeneous nucleation of nonane, in *Atmospheric Aerosols and Nucleation* (P. E. Wagner and G. Vali, eds.), Springer-Verlag, Berlin, p. 356.
11. Reiss, H., Rabeony, H., El-Shall, M. S., and Bahta, A. (1987). Analysis of cloud chamber measurements of gas phase polymerization, *J. Chem. Phys. 87*: 1315.
12. Hensel, F., and Uchtmann, H. (1989). The metal-insulator transition in expanded fluid metals, *Annu. Rev. Phys. Chem. 40*: 61.
13. Wright, D., Caldwell, R., and El-Shall, M. S. (1991). Vapor phase homogeneous nucleation of acetonitrile: the effect of dipole-dipole interaction, *Chem. Phys. Lett. 176*: 46.
14. Rao, M., Berne, B. J., and Kalos, M. H. (1978). Computer simulation of the nucleation and thermodynamics of microclusters, *J. Chem. Phys. 68*: 1325.
15. Lee, J. K., Barker, J. A., and Abraham, F. F. (1973). Theory and Monte Carlo simulation of physical clusters in the imperfect vapor, *J. Chem. Phys. 58*: 3166.
16. Reiss, H., Tabazadeh, A., and Talbot, J. (1990). Molecular theory of vapor phase nucleation: the physically consistent cluster, *J. Chem. Phys. 92*: 1266.
17. Cahn, J. W., and Hilliard, J. E. (1959). Free energy of a nonuniform system. III. Nucleation in a two-component incompressible fluid, *J. Chem. Phys. 31*: 688.
18. Unger, C., and Klein, W. (1984). Nucleation theory near the classical spinodal, *Phys. Rev. B 29*: 2698.
19. Monette, L., Klein, W., Zuckerman, M., Khadir, A., and Harris, R. (1988). Monte Carlo simulation of nonclassical nucleation, *Phys. Rev. B 38*: 11607.

20. Lee, D. J., Telo da Gama, M. M., and Gubbins, K. E. (1986). A microscopic theory for spherical interfaces: liquid drops in the canonical ensemble, *J. Chem. Phys. 85*: 490.
20a. Zeng, X. C., and Oxtoby, D. W. (1991). Gas-liquid nucleation in Lennard-Jones fluids. *J. Chem. Phys. 94*: 4472.
21. Langer, J. S., and Turski, L. A. (1973). Hydrodynamic model of the condensation of a vapor near its critical point, *Phys. Rev. A 8*: 3230.
22. Turnbull, D., and Fisher, J. C. (1949). Rate of nucleation in condensed systems, *J. Chem. Phys. 17*: 71.
23. Oxtoby, D. W. (1988). Nucleation of crystals form the melt, *Adv. Chem. Phys. 70*(2): 263.
24. Walton, A. G. (1969). In *Nucleation* (A. C. Zettlemoyer, ed.), Marcel Dekker, New York.
25. Turnbull, D. (1952). Kinetics of solidification of supercooled liquid mercury droplets, *J. Chem. Phys. 20*: 411.
26. Turnbull, D. (1950). Formation of crystal nuclei in liquid metals, *J. Appl. Phys. 21*: 1022.
27. Bosio, L., Schedler, E., and Windsor, C. E. (1976). Quasielastic neutron scattering from liquid gallium over the temperature range 163 to 333 K, *J. Phys. 37*: 747.
28. Yoon, W., Paik, J. S., LaCourt, D., and Perepezko, J. H. (1986). The effect of pressure on phase selection during nucleation in undercooled bismuth, *J. Chem. Phys. 60*: 3489.
29. Perepezko, J. H., and Anderson, I. E. (1980). Metastable phase formation in undercooled liquids, in *Synthesis and Properties of Metastable Phases* (E. S. Machlin and T. J. Rowland, eds.), TMS-AIME, Warrendale, Pa., p. 31.
30. Hardy, S. C. (1977). A grain boundary groove measurement of the surface tension between ice and water, *Philos. Mag. 35*: 471.
31. Wood, G. R., and Walton, A. G. (1970). Homogeneous nucleation kinetics of ice from water, *J. Appl. Phys. 41*: 3027.
32. MacFarlane, D. R., Kadiyala, R. K., and Angell, C. A. (1983). Homogeneous nucleation and growth of ice from solutions: TTT curves, the nucleation rate, and the stable glass criterion, *J. Chem. Phys. 79*: 3921.
33. Kelton, K. F., Greer, A. L., and Thompson, C. V. (1983). Transient nucleation in condensed systems, *J. Chem. Phys. 79*: 6261.
34. Kelton, K. F., and Greer, A. L. (1988). Test of classical nucleation theory in a condensed system, *Phys. Rev. B 38*: 10089.
35. Weinberg, M. C., and Zanotto, E. D. (1989). Re-examination of the temperature dependence of the classical nucleation rate: homogeneous crystal nucleation in glass, *J. Non-Cryst. Solids 108*: 99.
35a. Van Duijneveldt, J. S., and Frenkel, D. (1992). Computer simulation study of free energy barriers in crystal nucleation, *J. Chem. Phys.*, to be published.
36. Mandell, M. J., McTague, J. P., and Rahman, A. (1977). Crystal nucleation in a three dimensional Lennard–Jones system. II. Nucleation kinetics for 256 and 500 particles, *J. Chem. Phys. 66*: 3070.

37. Honeycutt, J. D., and Andersen, H. C. (1986). Small system size artifacts in the molecular dynamics simulation of homogeneous crystal nucleation in supercooled atomic liquids, *J. Phys. Chem. 90*: 1585.
38. Haymet, A. D. J., and Oxtoby, D. W. (1981). A molecular theory for the solid–liquid interface, *J. Chem. Phys. 74*: 2559.
39. Oxtoby, D. W., and Haymet, A. D. J. (1982). A molecular theory for the solid–liquid interface. II. Study of bcc crystal-melt interfaces, *J. Chem. Phys. 76*: 6262.
40. Curtin, W. A. (1989). Density-functional theory of crystal-melt interfaces, *Phys. Rev. B 39*: 6775.
41. Oxtoby, D. W. (1990). Crystallization of liquids: a density functional approach, in *Liquids, Freezing and Glass Transition*, Les Houches Session LI (J. P. Hansen, D. Levesque, and J. Zinn-Justin, eds.), Elsevier, Amsterdam.
42. Harrowell, P., and Oxtoby, D. W. (1984). A molecular theory of crystal nucleation from the melt, *J. Chem. Phys. 80*: 1639.
43. Grant, M., and Gunton, J. D. (1985). Theory for the nucleation of a crystalline droplet from the melt, *Phys. Rev. B 32*: 7299.
44. Reiss, H. (1950). The kinetics of phase transition in binary systems, *J. Chem. Phys. 18*: 840.
45. Doyle, G. J. (1961). Self-nucleation in the sulfuric acid–water system, *J. Chem. Phys. 35*: 795.
46. Renninger, R. G., Hiller, F. C., and Bone, R. C. (1981). Comment on "Self-nucleation in the sulfuric acid–water system," *J. Chem. Phys. 75*: 1584.
47. Wilemski, G. (1984). Composition of the critical nucleus in multicomponent vapor nucleation, *J. Chem. Phys. 80*: 1370.
48. Flageollet-Daniel, C., Garnier, J. P., and Mirabel, P. (1983). Microscopic surface tension and binary nucleation, *J. Chem. Phys. 78*: 2600.
48a. Zeng, X. C., and Oxtoby, D. W. (1991). Binary homogeneous nucleation theory for the gas-liquid transition: a nonclassical approach, *J. Chem. Phys. 95*: 5940.
49. Perepezko, J. H., and Smith, J. S. (1981). Glass formation and crystallization in highly undercooled Te–Cu alloys, *J. Non-Cryst. Solids 44*: 65.
50. Thompson, C. V., and Spaepen, F. (1983). Homogeneous crystal nucleation in binary metallic melts, *Acta Metall. 31*: 2021.
51. Turnbull, D. (1981). On anomalous prefactors from analysis of nucleation rates, in *Progress in Materials Science* (J. W. Christian, P. Haasen, and T. B. Massalski, eds.), Pergamon Press, London, p. 269.
52. Turnbull, D. (1950). Kinetics of heterogeneous nucleation, *J. Chem. Phys. 18*: 198.

11

Liquid Crystals

J. F. Marko
James Franck Institute
University of Chicago
Chicago, Illinois

I. INTRODUCTION

A principal objective of the description of fluids through the use of statistical mechanics is to understand the relationship between microscopic intermolecular interactions and the macroscopic properties of the stable ordered phases. In this chapter I review recent progress that has been made toward this goal using results from the theory of inhomogeneous fluids. A chief result of the last few years, and the one that will be dealt with in this chapter, is the demonstration that for a variety of model fluids, approximations to the Helmholtz free energy for the inhomogeneous fluid can be used to compute the stability and structure of ordered phases, in quite reasonable agreement with results from computer simulations.

An important class of models (indeed, the class that will be discussed almost exclusively here) incorporates only "hard-core" interactions (i.e., all nonoverlapping configurations of particles have equal statistical weight); those with overlapping molecules are disallowed. A fundamental property of such a fluid of spherical hard cores of diameter D is that it undergoes a first-order freezing transition at a density $N/V \approx 0.9D^{-3}$ [1]. The structural details (e.g., the Lindeman ratio at freezing, lattice constant at melting in terms of the hard-core diameter D) of this freezing-phase

transition are remarkably similar to structural details of freezing in liquids composed of spherical particles with short-range interactions. For nonspherical particles, the analogous problem is to understand the relationship between the shape of constituent molecules and the stability of various ordered phases with various combinations of broken orientational and translational symmetries.

The modern version of the density functional approach [2] to this problem offers a description of freezing by means of functional expansion of the thermodynamical potentials in powers of the one-body density about the isotropic fluid phase. This constructive approach is a natural route to a free-energy functional that can then be minimized with respect to the expectation value of the one-body density to obtain the canonical mean-field description of the ordered phases. In principle, the dispersion of the free energy around this mean-field description gives a Landau–Ginzberg–Wilson field theory that can be used to study elastic properties, and ultimately the effect of fluctuations.

In this chapter this program is outlined, and particular attention will be paid to recent applications of the approach to the study of liquid–crystal phases of fluids composed of hard cores of anisotropic shape. Section II begins with a discussion of these models and reviews the available exact and simulation results. In Section III, the density functional approach to theories for liquid–crystalline ordering will be discussed, with a major objective being to relate the Landau parameters of the usual theories for simple liquid–crystal phase transitions explicitly to static liquid structure. In this discussion we introduce the basic ideas and emphasize the crucial requirement of accurate liquid structure information. We also present an approach that allows one to include contributions to the density functional expansion from higher orders in perturbation theory.

In Section IV we focus on summarizing what we know about the liquid structure of anisotropic hard cores. These pieces are put together in Section V, where the results that have been obtained for various types of phase transitions are discussed. Where possible, comparison is made with computer experiments. Finally, in Section VI we summarize the current state of the theory.

II. MODELS, EXACT RESULTS, AND SIMULATIONS

The freezing of a liquid into a solid is a ubiquitous physical phenomenon. A model that displays all of the essential qualitative properties of this

phase transition is that of spherical particles that interact through a purely exculsion interaction. These "hard spheres" were first observed to crystallize in simulations of Alder and Wainright [3] and have since been the focus of much simulation work [1]. For liquids composed of anisotropic particles, transitions that break various combinations of translational and orientational order are possible, including transitions to liquid–crystal phases, which have lower symmetry than the isotropic liquid but higher symmetry than the high-density crystalline states.

Liquid–crystal mesophases have anisotropic elastic and optical properties which are distinct from the isotropic liquid and crystal phases. We describe these phases in terms of their symmetries: An isotropic liquid is invariant under any rotation or translation. The crystal phase is typically invariant under a smaller set of transformations: translations by a set of lattice vectors and some subgroup of the rotation group consistent with the lattice symmetries [4]. The nematic liquid–crystal phase is translationally invariant but differs from the isotropic phase in not being invariant under rotations, due to the orientation of the constituent molecules along some direction. Typically, nematic phases are observed at lower temperatures or at higher pressures than the isotropic liquid.

Before the formation of the crystal at low temperatures or at high pressures, smectic phases may be observed: the simplest of these, the smectic *A* phase, has in addition to the broken orientational symmetry a peculiarly broken translational symmetry. The smectic *A* is invariant under translations in the plane perpendicular to the direction of molecular orientation (the "director"), but only under translations out of this plane that are multiples of some distance, the smectic wavelength. Many types of smectics are possible, depending on, e.g., the relationship between the molecular orientation and the orientation of the smectic "layers"; a rather complete discussion of the classification of these phases is that of Pershan [5]. No complete review of all of the liquid–crystal phases that have been observed, let alone the entire field of liquid crystals, can be attempted here; the reader is urged to consult one of the excellent standard texts [6,7].

A central problem is that of identifying suitable model systems that will display the essential features of simple liquid–crystal phases. Onsager discovered that there could be a stable nematic state in a system of elongated particles interacting via strictly hard-core interactions [8], and over the past few years it has become apparent that such systems also possess smectic *A* phases [9]. This discovery suggests that one should try to understand the behavior of hard-core models and then take account of longer-range interactions of finite strength perturbatively, an approach that has been successful in the description of dense simple liquids [10].

A. Models and Ordered Phases

We will consider N molecules confined inside a three-dimensional box of volume V, which in addition to a translational degree of freedom $\mathbf{r}$ describing the location of their centers of mass, possess orientation degrees of freedom. The orientation of a molecule can be expressed using a vector $\hat{\mathbf{e}}$ on the surface of the unit sphere to specify the orientation of one of the body axes, and an additional angle χ to specify rotation around that axis. The vector $\hat{\mathbf{e}}$ will often be described in terms of the polar and azimuthal angles θ and ϕ with respect to a Cartesian coordinate system: the z-axis will generally be taken to be parallel to the director of orientationally ordered phases. For axially symmetric particles, the axial angle χ can be ignored. The combination of degrees of freedom for a molecule $(\mathbf{r},\hat{\mathbf{e}},\chi, \ldots)$ will be referred to with the notation X. The orientation of the positional variable $\mathbf{r}$ can be described by the unit vector $\hat{\mathbf{r}} = \mathbf{r}/|\mathbf{r}|$. Finally, we note that integration over all configurations will be denoted by

$$\int dX \equiv \int_V d^3r \int_0^{2\pi} d\chi \int_{-1}^{1} d(\cos\theta) \int_0^{2\pi} \frac{d\phi}{8\pi^2} \tag{1}$$

The total integral of configurations is $\int dX = V$.

The shape of any convex hard core (and many nonconvex hard cores) can be specified using the function $D(X,X')$, which specifies the distance of closest approach of the centers of masses for two hard cores at coordinates X and X'. This distance is the value of $|\mathbf{r} - \mathbf{r}'|$ at which the hard cores touch, given the orientations $(\hat{\mathbf{r}},\hat{\mathbf{e}},\chi)$ and $(\hat{\mathbf{r}}',\hat{\mathbf{e}}',\chi')$. For $|\mathbf{r} - \mathbf{r}'| < D(X,X')$, the two cores overlap. If we consider the intermolecular potential to be the sum of pair interactions of the form

$$\begin{aligned} U(X,X') &= 0, |\mathbf{r} - \mathbf{r}'| > D(X,X') \\ &\quad \infty, |\mathbf{r} - \mathbf{r}'| < D(X,X') \end{aligned} \tag{2}$$

then any overlapping configuration has infinite potential, while any state free of overlapping hard cores has zero potential. This hard-core interaction can of course be supplemented by additional finite-strength pairwise or multiple-particle interactions.

A central approximation is to disregard any momentum dependence to the interactions between particles, which results in an N-particle Hamiltonian of the form

$$H_N(\{X_i,P_i\}) = \sum_{i=1}^{N} T(P_i) + \sum_{j<i} U(X_i,X_j) \tag{3}$$

where $T(P)$ is the kinetic energy of the momenta P conjugate to X (here we have included only the pair interactions described above). For the

degrees of freedom described above,

$$T(P) = \frac{\mathbf{p}^2}{2m} + \frac{\mathbf{L}\mathbf{I}^{-1}\mathbf{L}}{2} \tag{4}$$

where m is the particle mass, $\mathbf{p}$ the linear momentum, $\mathbf{I}$ the moment of inertia tensor, and $\mathbf{L}$ the angular momentum conjugate to the orientational degrees of freedom. The tensor $\mathbf{I}$ is taken to have eigenvalues I_1, I_2, and I_3; we define the geometrical average of these as $I = (I_1 I_2 I_3)^{1/3}$.

The grand canonical partition function, including a one-body potential (essentially a chemical potential that is a function of X) $\mu(X)$, is thus

$$e^{-W} = \sum_{N=0}^{\infty} \frac{1}{N!} \prod_{i=1}^{N} \left(\int \frac{dP_i dX_i}{h^6} \right) \exp[-\beta H_N(\{X_i, P_i\}) + \beta \int dX\, \mu(X)\hat{\rho}(X)] \tag{5}$$

where h is Planck's constant, $\beta = (kT)^{-1}$, $-W = \beta pV$, p is pressure, and the classical one-body density operator is defined as $\hat{\rho}(X) = \sum_{i=1}^{N} \delta(X - X_i)$. The momentum integration is unrestricted, while the spatial integration over $\mathbf{r}_i$ is over the system volume V, and the orientation integration is over all the angles (of total angular volume $8\pi^2$). After transforming the angular momentum to components along the principal axes of the inertia tensor I, we note that all the momenta can be integrated, giving

$$e^{-W} = \sum_{N} \frac{1}{N!} \prod_{i} \left\{ \int dX_i \exp\left[- \int dX\, u(X)\hat{\rho}(X) - \beta \sum_{j<i} U(X_i, X_j) \right] \right\} \tag{6}$$

where the effective one-body potential is $u(X) = -\beta\mu(X) + 3 \log \Lambda + 3 \log \Lambda_a$, $\Lambda = [h^2/(2\pi kTm)]^{1/2}$ is the thermal wavelength, and $\Lambda_a = [h^2/(2\pi kTI)]^{1/2}$ is the thermal angle. Thus the partition function reduces to a sum over configurations of the particles.

Since the quantities Λ and Λ_a contribute only additively to the total grand potential W, any expectation values concerning static (configurational) structure are independent of m and I. We will ignore these quantities for the remainder of this chapter. Additionally, if the pair potential U is of the hard-core form above, we can ignore the factor β in front of it. This removes any dependence on temperature from the configurational integral and vastly simplifies the problem. All phase transitions, and more generally, all of the thermodynamics, are independent (except for trivial details) of temperature. Our strategy in this chapter is to consider these as purely entropic models, which apparently will have only lyotropic phase transitions (the stability of ordered phases depends only on the volume fractions of the constituent particles). Typically, we will present

phase diagrams in terms of dimensionless parameters describing the shape of the particles, and the number density, in units of inverse particle volume (density in these units is often referred to as the packing fraction).

Examples of the models of this sort that we will focus on are motivated by particular ordered phases. The idea is that by breaking the rotational symmetry of the underlying molecular interactions, we can induce the spontaneous breaking of a global symmetry (i.e., induce a phase transition to an ordered phase). We introduce several models that have been studied along with the relevant ordered phases. Figure 1 shows the hard-core shapes, while Fig. 2 schematically shows the structure of the ordered phases.

1. *Spheres.* This simple case has a constant $D(X,X')$ equal to the sphere diameter. In three dimensions, the fluid phase freezes and becomes a close-packed (cubic) solid (Fig. 2a). That this should happen is not immediately obvious; what is obvious is that there is no coupling of orientational degrees of freedom of the molecules, and there are never orientational correlations in the system. The hard-core shape can be described by a single parameter, the sphere diameter D.

2. *Ellipsoids of revolution.* If we elongate or compress a spherical hard core we obtain either prolate or oblate ellipsoidal hard cores. This change in symmetry induces interactions between the orientational and translational molecular degrees of freedom. Thus it is possible to have a fluid (no orientational or translational order), a nematic liquid crystal (no translational order, but orientational order, shown in Fig. 2b), a plastic

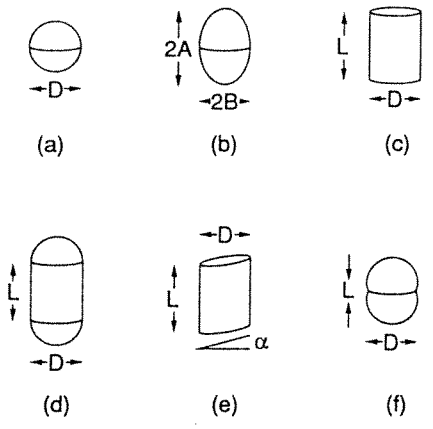


Fig. 1 Hard-core shapes: (a) sphere; (b) prolate ellipsoid of revolution; (c) cylinder; (d) spherocylinder; (e) oblique cylinder; (f) dumbbell.

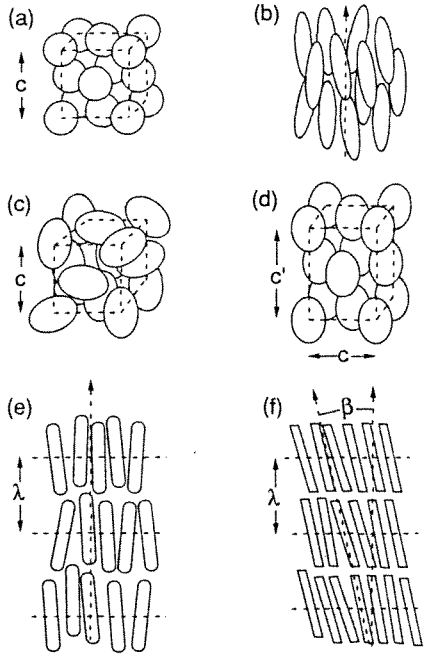


Fig. 2 Structure of various liquid-crystal and crystal phases. (a) Close-packed crystal. The structure shown is a face-centered cubic (fcc) crystal of spheres, with long-range translational order. (b) Nematic liquid crystal. There is long-range orientational order but no positional order. (c) Plastic crystal. The structure shown is an fcc crystal of hard ellipsoids with translational order and with orientational order consistent with the rotational symmetries of the crystal. (d) Oriented solid. Again, hard ellipsoids are shown; however, now the lattice is tetragonally distorted fcc, and the orientational distribution consequently has lower symmetry than in the plastic phase. (e) Smectic *A* liquid crystal. There is orientational order, and cylinders are organized into regularly spaced planes orthogonal to the director axis. There is no long-range translational order in the planes. (f) Smectic *C* liquid crystal. There is again only one-dimensional translational order, but now the director axis makes an angle β with the normal to the smectic planes. Elliptocylinders are shown.

solid (cubic translational order but no orientational order beyond the cubic point group as in Fig. 2c), and an oriented solid (e.g., tetragonal translational order with corresponding uniaxial orientational order, as in Fig. 2d). The hard core is described by the axial diameter, $2A$, and the diameter perpendicular to the axis of rotational symmetry, $2B$. The single dimen-

sionless parameter A/B describes the shape: The core is prolate or oblate if $A > B$ or $B > A$, respectively.

3. *Cylinders and spherocylinders.* A cylindrical hard core has the same symmetry as a prolate ellipsoid, but it is of interest because for strong elongation, there may be a smectic A phase, consisting of layers of orientationally aligned molecules (i.e., translational order along one direction) but with no translational order in the plane of these layers (see Fig. 2e). Again two parameters describe the shape, the diameter D and the length of the cylindrical section L; again the single parameter L/D describes the shape. Often the cylinder is topped with a hemisphere of diameter D as shown in Fig. 1d (a shape referred to as a spherocylinder).

4. *Elliptocylinders.* We can imagine breaking the rotational symmetry of a cylinder by making its cross section elliptical; this might prompt the appearance of a phase with biaxial orientational symmetry (i.e., there might be a transition to a nematic or smectic phase where particles have long-range correlations in their axial angles χ). The parameters describing this core are the two diameters of the elliptical cross section, $2A$ and $2B$, and the length L. We might choose to use dimensionless parameters $2A/L$ and A/B to describe the system.

5. *Oblique cylinders.* By cutting the ends of a cylinder with planes not perpendicular to the axis, we introduce the possibility of tilted phases. The example important here is the smectic C phase, which is identical to the smectic A phase except that the molecules are tilted some angle with respect to the smectic layers (Fig. 2f). The parameters describing the molecules are the cylinder length L and diameter D, as well as the angle χ shown in Fig. 1e.

6. *Dumbbells and other shapes.* Spheres may be fused together to form shapes with low symmetries; for example, two spheres fused together form a dumbbell with an aspect ratio between 1 and 2. Such an object is attractive because the $D(X,X')$ can be constructed exactly from that for two spheres. For dumbbells, we specify the shape with the sphere diameter D and the distance L between the sphere centers (Fig. 1f).

7. *Models with constrained degrees of freedom.* Often, calculation or simulation of some system of molecules with low symmetry is facilitated by the constraint of some of the degrees of freedom. For example, one might study the nematic to smectic A transition by constraint of the orientational degree of freedom along the z-axis. The justification for such a study would be that near the N–S_A transition, the nematic order would be well established, and the orientational fluctuations might not strongly affect the transition. Such models might also be directly relevant, for example, in the case where transitions are studied in strong fields that strongly align the molecules.

8. *Softer interactions.* At this point the rather more complicated world of softer interactions has been studied to a limited extent. We note that work has generally proceeded using potentials that depend on the separation of the centers of mass of the molecules (e.g., a r^{-6} attraction at long ranges), or on some combination of electric multipole interactions.

These models, although somewhat arbitrarily chosen, represent the geometrically simple hard-core shapes that lead to various ordered phases. One should not take the point of view that these models represent actual materials that show these phases. The models are meant only to be tools to be used to develop techniques for studying phase transitions in dense molecular systems; they must be generalized to make them represent particular materials more faithfully.

B. Exact Results

There are a limited number of exact results concerning these sorts of phase transitions, but the ones that exist are informative. An important piece of information was provided by Onsager regarding the isotropic–nematic transition in a system of long, hard spherocylinders [8]. By considering the density expansion for the free energy of such a system, Onsager showed that as the length-to-diameter L/D ratio of fluid of hard spherocylinders is taken to infinity, there is an instability toward nematic ordering, and that this I–N transition density (in units of the inverse of the molecular volume) tends to zero as D/L. This transition is strongly first order; the difference in densities of the nematic and isotropic phases at the transition is about 27% of the density of the isotropic phase. The cause of this transition in such a dilute gas of hard cores is purely due to entropy: In a dilute gas of highly elongated particles, a given molecule enjoys a greater number of allowed configurations in a system with nematic order than it would in an orientationally disordered environment. In Section III the Onsager theory is constructed.

If one imposes perfect orientational order on a system of hard ellipsoids of revolution (prolate or oblate), one might imagine that there would be some interesting mesophases (i.e., smectic A). Such a system consists of hard ellipsoids with their axis of rotational symmetry aligned along some preferred direction; Lebowitz and Perram [11] noted that by rescaling space along this direction, one can reduce the partition function to that for a system of hard spheres. Since hard spheres have only a freezing transition, the only transition in a system of aligned ellipsoids is therefore also a freezing transition to an oriented solid phase, at a packing fraction of 0.7. This is rather surprising in light of the fact that this rescaling

argument applies to all elongations! Evidently, the rounded shape of the molecules precludes any possibility of a smectic A or other mesophase (since the ellipsoids in such a system are parallel, the low-density, disordered phase is a perfectly aligned nematic).

This result may be understood in the light of a recent argument by Wen and Meyer [12], suggesting that there is a stable smectic A phase in a fluid of orientationally aligned hard rods. For definiteness, consider long cylinders with their axes aligned along the z-axis. Wen and Meyer start by considering the fully ordered smectic, which consists of planes separated by the smectic wavelength d, on which molecules have their centers of mass. The distribution of molecules in each plane is considered to be uniform, and molecules in different planes are presumed not to interfere with each other. Now consider the possibility that a molecule (of length $L < d$) is allowed to fluctuate out of a layer; if it moves a distance vertically larger than $d - L$ but less than L, it will interfere with the configurations of molecules in an adjacent layer, reducing the overall configurational entropy of the system. The conclusion is that there may be a transition to the smectic because the two-dimensional fluid layers have larger configurational entropy than does the disordered (nematic) state for some packing fractions. Although this argument has not been proven with an exact calculation, Wen and Meyer present an estimate of the free energy based on this idea, and find that indeed, the smectic phase becomes more stable than the nematic at a packing fraction of 0.223. For the case of prolate ellipsoids, the tapering of the ends of the ellipsoids causes a less drastic loss of configurational entropy in neighboring planes; the result is that the smectic is never stable, in agreement with the exact result.

A final class of interesting exact results are those for models of fluids in one dimension. In one dimension, many problems can be solved due to the limited number of ways that particles can interact; this advantage is offset by the fact that for short-range interactions, there are never phase transitions in one dimension [13]. Nevertheless, the results have been useful, particularly in building intuition about the behavior of correlation functions in the dense liquid phase. Tonks [14] first solved the problem of hard rods in one dimension; these results were expanded by Wertheim [15], who noted that the Percus–Yevick closure of the Ornstein–Zernike equation was exact for any mixture of hard cores of different lengths in the case of one dimension. He thereby obtained not only the free energy but also the pair distribution. Marko [16] noted that this result could easily be modified to solve certain one-dimensional problems involving molecules with both orientational and translational degrees of freedom, and to obtain the direct correlation function in closed form. The thermodynamics of similar systems have been studied by transfer-matrix techniques

[17,18] and via this route, much useful information concerning the equation of state and one-particle distributions has been obtained; in principle, this approach can be applied to calculation of the correlation functions as well. Some of these results are discussed in Section IV.

C. Simulations

With the explosion of computer power over the last decade we have seen a similar explosion of work on simulational approaches to the statistical theory of fluids. Here, we restrict the discussion to a summary of the main results obtained from simulations of hard-core models presented in Section II.A. We skip any detailed discussion of the seminal work concerning freezing of the hard-sphere fluid [1] as it is discussed elsewhere in this volume. However, we note [19] that the transition occurs from the isotropic liquid phase at a density of $0.94D^{-3}$ (D is the sphere diameter) and is first order. The crystal phase is likely a (close-packed) face-centered cubic crystal; at coexistence, the crystal has a density of $1.04D^{-3}$. The entropy difference per particle between the two coexisting phases is $1.16k$, and the Lindemann ratio (the ratio of the rms fluctuations of particles about lattice sites to the nearest-neighbor spacing) is about 0.13. A striking fact about the crystal is that the one-particle fluctuations are well described by Gaussian distributions centered on sites of a fcc lattice for densities near the melting point [20]; this is surprising in light of the highly anharmonic interparticle interaction.

As noted above, elongation or compression of spherical hard cores results in prolate or oblate hard ellipsoids of revolution. Simulation of such a system of particles is difficult due to the large number of degrees of freedom. However, Frenkel, Mulder, and McTague have been able to compute the phase diagram and isotherms (i.e., the pressure versus density) for the hard ellipsoid fluid via constant-pressure Monte Carlo simulation [21]. Their phase diagram is pictured in Fig. 3; the ordinate is the aspect ratio, the ratio of the major semiaxis A to the minor semiaxis B of the ellipsoids. Spheres have $A/B = 1$, needles have $A/B \gg 1$, and platelets have $A/B \approx 0$. The abscissas of Fig. 3 is the number density ρ in units of $8AB^2$; the quantity $8AB^2\rho$ is $6/\pi$ times the packing fraction η. The maximum packing fraction that can be attained for such a system is believed to be that for a close-packed oriented crystal (similar to the close-packed version of Fig. 1d), for which $8AB^2\rho = \sqrt{2}$.

First-order transitions (involving coexisting phases of different densities) are indicated as shaded regions in Fig. 3. For small anisotropies ($A/B \approx 1$), the first transition is from the isotropic liquid (I) to a plastic crystal phase (P). The density at which the isotropic liquid freezes in-

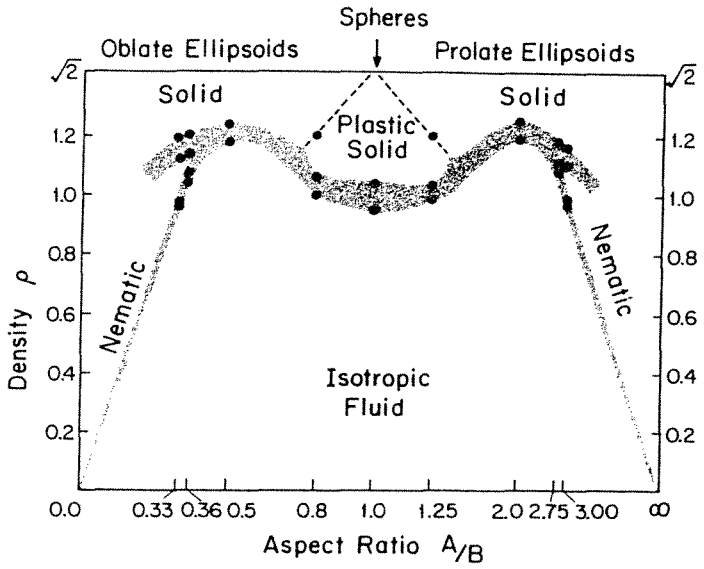


Fig. 3 Phase diagram of hard ellipsoids of revolution. Transition densities in molecular units $8AB^2\rho = 6\eta/\pi$ is shown versus anisotropy A/B. Coexistence regions are indicated with shading; the plastic–solid transition is indicated by a dashed line. Points are Monte Carlo results. (Adapted from Ref. 21.)

creases as A/B is made either slightly larger or smaller than 1. It is important to realize that this crystal phase, although often described as orientationally disordered, consists of particles whose orientational one-particle distribution function is not invariant under arbitrary rotations. The crystal lattice introduces a cubic field, which results in the orientational distribution being invariant only under symmetry operations of the underlying lattice (i.e., cubic rotations). The one-particle orientational distribution becomes fully rotationally invariant only for the case of spheres ($A/B \rightarrow 1$).

As the plastic crystal phase is compressed, a point is reached at which there must be a transition to a crystal phase with ellipsoids aligned along one direction. Such an oriented solid phase (S) has the structure indicated in Fig. 1d; the lattice is tetragonally distorted fcc, with x and y lattice constants c and a z lattice constant of c'. The ellipsoids have their axis of rotational symmetry preferentially aligned with the z axis. For the case of hard spheres, the P–S transition never occurs, of course, but for small elongations it must occur at some density less than $\sqrt{2}$ (we presume that the state of highest density is a perfectly aligned close-packed configu-

ration and that there is no state with the symmetry of the plastic phase that can achieve this density). In fact, the simulations show transitions from the plastic to the solid phase at a packing fraction that drops as anisotropy is made different from $A/B = 1$. Interestingly, this transition is observed to be either weakly first-order, or perhaps second-order; the phase boundary is shown as a dashed line in Fig. 3. The results are not conclusive concerning this point, as the number of particles simulated was small (≈ 100), and thus the orientational correlations could not be observed to become gradually long-ranged, as would be the case for a second-order transition.

At an aspect ratio of about $A/B = 1.25$ on the prolate side, and at about $A/B = 0.8$ on the oblate side, the plastic–solid transition line meets the isotropic–plastic coexistence region. For anisotropies slightly larger than $A/B = 1.25$ or slightly less than $A/B = 0.8$, there is a first-order transition from the isotropic liquid directly to the oriented solid. Exactly what the nature of the I–P–S coexistence point is, is not clear; possibilities include a triple point (if the P–S transition is first order, or if there is a tricritical point along the P–S transition line) and a critical endpoint (if P–S is continuous).

At yet more prolate or oblate aspect ratios, there occurs a triple point between the isotropic, solid, and the nematic (N) phases. Beyond these points (at about $A/B = 2.7$ on the prolate side and $A/B = 0.35$ on the oblate side), the sequence of transitions that is seen as density is increased is I–N–S, with both transitions first order. Although no simulations have been done for $A/B > 3$, Onsager's results [8] indicate that as $A/B \to \infty$, the transition is first order, and that the transition packing fraction tends to zero as B/A. It seems likely that the I–N transition line for prolate ellipsoids can be smoothly extrapolated to $A/B = \infty$ as shown.

In the extreme oblate limit $A/B \to 0$, we obtain platelets. Frenkel and Eppenga determined [22] that for $A = 0$, zero-volume hard platelets undergo a first-order transition from the isotropic liquid at density $8B^3\rho = 3.78$ to a nematic of density $8B^3\rho = 4.07$. These results are for zero density; we conclude that if A is slightly larger than zero, the I–N coexistence region will be between densities $8AB^2\rho = 3.78A/B$ and $4.07A/B$. Reasoning thus, Frenkel extended the I–N line to zero density for $A/B \to 0$ as in Fig. 3.

The phase diagram (more exactly the densities demarcating the coexistence regions for various anisotropies) comprise one component of the MC information. However, we note that the MC results also include accurate isotherms (grand potential, or pressure versus density for different A/B) for all of the phases, and some information concerning orientational correlations. A remarkable result concerning the data as a

whole is the apparent symmetry between ellipsoids of aspect ratio A/B and B/A (Fig. 3 has as abscissas A/B for $A/B < 1$ and B/A for $A/B > 1$), a result that is not present in the viral series beyond second order [21]. This symmetry is apparent not only in the phase diagram but also in the isotherm data. If the pressure is scaled with the molecular volume and temperature, it is found that $8AB^2P/kT$ versus $8AB^2\rho$ differs for ellipsoids with aspect ratios A/B and B/A by at most 7% for all cases considered. This surprising (albeit approximate) symmetry has no explanation. Finally, we note that no smectic phases have been observed in computer simulations of ellipsoids of any anisotropy.

Similar large-scale simulation work has been done on spherocylinders. The computationally simpler case of up to 1080 parallel spherocylinders of cross-sectional diameter D and cylindrical section length L was studied by Stroobants et al. [23]; in such a system, the low-density phase is of course the nematic, as all of the axes of the molecules are lined up by *fiat* along the z-axis. At high densities, we expect a freezing transition to a close-packed crystal; the case $L/D = 0$ is, of course, again that of hard spheres, which has only this phase transition. The maximum density possible in the system is presumed to be $\rho_{cp} = \sqrt{2}/[D^3(1 + \sqrt{3/2}L/D)]$, that of the tetragonally distorted fcc lattice.

The phase diagram that resulted from this study is shown in Fig. 4; indeed, the nematic appears at low density and there is a freezing transition (to a phase again designated solid, or S) to an oriented close-packed

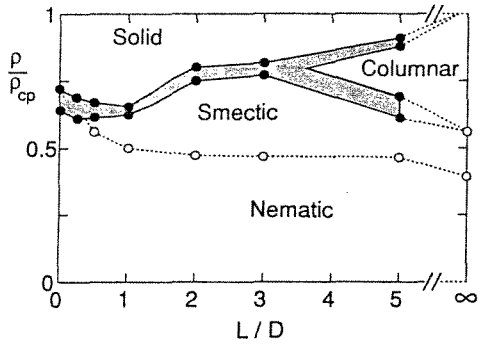


Fig. 4 Phase diagram of hard parallel spherocylinders. Transition densities in units relative to the maximum density ρ_c (of the oriented solid) are shown versus the ratio of the length to the diameter of the cylindrical section L/D. Coexistence regions are shaded, and simulation results are indicated with points. (Adapted from Ref. 23.)

crystal for small L/D. However, for $L/D > 0.5$, a layered phase is thermodynamically stable; this phase is a hard-core smectic A. The transition appears to be continuous, which suggests that the N–A–S point in this system is a critical endpoint. The N–A transition at $L/D \approx 0.5$ is at a reduced density $\rho/\rho_{cp} \approx 0.6$ and drops to near 0.45 for $L/D = 5$. The case $L/D = \infty$ can be mapped onto a system of parallel cylinders; the N–A transition is seen at $\rho/\rho_{cp} = 0.39$. For aspect ratios above $L/D = 3$, a columnar phase has also been observed; for finite L/D the smectic-columnar phase transition is first order, but for $L/D = \infty$, it appears to be continuous. The crystalline solid phase has been observed at high densities for $L/D < \infty$; the columnar-solid transition is first order.

The phase diagram of spherocylinders with both orientational and translational degrees of freedom has also been studied. Simulation of 576 spherocylinders with $L/D = 5$ led Frenkel et al. to the conclusion [9,24] that a sequence of transitions I–N–A occurred as density was increased. Additional work by Veerman and Frenkel [25] has resulted in the phase diagram shown in Fig. 5. Again, the density is in reduced units ρ/ρ_{cp}; for low densities the isotropic liquid is the stable phase. For $L/D < 2$, there is a first-order freezing transition from the isotropic to the close-packed solid phase. Near $L/D = 3$, the simulations suggest, in succession as anisotropy increases, a I–A–S coexistence point and a I–N–A coexistence point. The I–A and A–S transitions appear to be first order; thus the I–A–S point is a triple point. The only N–A transition studied is that at $L/D = 5$; so far it is not clear whether the transition is weakly first order or if it is continuous. The orientationally constrained and free spherocylinder phase diagrams are notably different; in the former (Fig. 4) the liquid–crystal (N,A) phases are found for all $L/D > 0.5$. This apparently is a result of the strong alignment constraint since in the unconstrained case (Fig. 5) the A and N phases do not appear for $L/D < 4$ or so.

We conclude this section by mentioning that there have also been pioneering efforts by Allen and Frenkel [26] to compute the Frank elastic constants of the nematic liquid–crystal phases of hard ellipsoids and hard spherocylinders via computer simulation. We also note that anisotropic hard-core systems in two dimensions may show intriguing behavior. Frenkel and Eppenga have simulated a system of hard "needles" (lines) of length L in two dimensions and report evidence that the isotropic–nematic transition is near $L^2\rho = 7$ (as in the case of platelets, the transition formally is at zero packing fraction since the lines have zero volume) [27]. They also find that the nematic has algebraically decaying orientational correlations, suggesting that the I–N transition for needles in two dimensions may be of Kosterlitz–Thouless type. Analysis of the behavior of topological defects near the transition supports this picture.

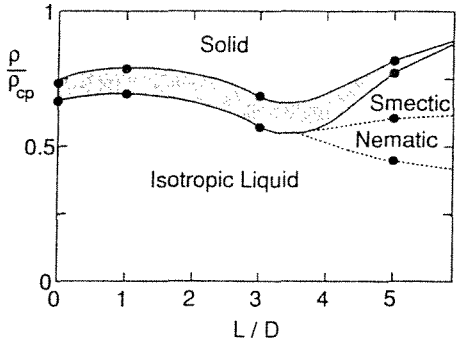


Fig. 5 Phase diagram of hard spherocylinders with orientational and translational degrees of freedom; data are presented as in Fig. 5. (Adapted from Ref. 25.)

III. DENSITY FUNCTIONAL THEORY OF LIQUID CRYSTALS

An alternative, analytical approach to understanding the systems described in Section II is available through a functional expansion of the thermodynamical potential about the isotropic liquid state, in powers of the deviation of the one-body density from the constant density of the isotropic phase. In the remainder of this chapter we will be concerned with the application of this method, commonly termed the density functional approach, to models described in Section II.A. In Section III.A we summarize the basic density functional formalism (treated in greater detail elsewhere in this volume). In Section III.B the problem of carrying out calculations to high order in perturbation theory is addressed; we present a description of a weighted-density approximation appropriate for liquid–crystal systems. The parametrization of ordered phases are discussed in Section III.C. Finally, in Section III.D we present application of these methods to the description of transitions to the nematic and smectic *A* liquid–crystal phases. This will include a derivation of Onsager's theory of the isotropic–nematic phase transition [8] and will indicate the microscopic origin of the Landau coefficients in the mean-field theory of nematic and smectic *A* liquid crystals.

A. Density Functional Expansion for Anisotropic Particles

We begin by considering the grand-canonical potential (6), which can be written as

$$-W = \beta pV = \log\left(\sum_N \frac{1}{N!}\prod_i \left\{\int dX_i \exp[-u(X_i) - \beta \sum_{j<i} U(X_i,X_j)]\right\}\right) \tag{7}$$

The potential W is intrinsically a functional of the dimensionless one-body potential $u(X)$, which is the chemical potential associated with a particle with position and orientation described by X. Since u is coupled linearly to the density operator, it is seen to generate the density distribution functions of the system:

$$\rho(X) = \langle\hat{\rho}(X)\rangle = \frac{\delta W}{\delta u(X)}$$

$$\langle\hat{\rho}(X)\rangle\langle\hat{\rho}(X')\rangle[g(X,X') - 1] \tag{8}$$

$$= \langle\hat{\rho}(X)\hat{\rho}(X')\rangle - \langle\hat{\rho}(X)\rangle\langle\hat{\rho}(X')\rangle = -\frac{\delta^2 W}{\delta u(X)\,\delta u(X')}$$

and so on.

The potential associated with W that is an intrinsic functional of the one-body density $\rho(X)$ is obtained by Legendre transformation:

$$\beta F = \max_{u(X)}[W - \int dX\, u(X)\rho(X)] \tag{9}$$

where, of course, F is the usual extensive Helmholtz free energy. If there were no interactions, namely if $U(X,X') = 0$, the Helmholtz free energy thus defined would be just the entropy of an inhomogeneous ideal gas:

$$\beta\mathcal{F}|_{U=0}[\rho] = \beta\mathcal{F}_{\text{id}}[\rho] = \int dX\, \rho(X)[\log \rho(X) - 1] \tag{10}$$

We define the remainder of the Helmholtz free energy as the excess potential via $\mathcal{F} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{ex}}$ and examine the functional derivatives of the excess Helmholtz potential with respect to the one-body density:

$$\frac{\delta^s}{\delta\rho(X_1)\cdots\delta\rho(X_s)}\beta\mathcal{F}_{\text{ex}}[\rho] \equiv -C^{(s)}(X_1, \ldots, X_s;\rho) \tag{11}$$

A bit of computation shows that for the isotropic liquid with $\mu(X) = \mu$ and $\rho(X) = \rho_0$, $C^{(1)} = \log \rho_0 - \beta\mu$, and that $C^{(2)} \equiv c$ is related to the pair distribution function defined above by the Ornstein–Zernike equation:

$$g(X,X') - 1 = c(X,X') + \rho_0 \int dX''[g(X,X') - 1]c(X'',X') \tag{12}$$

This allows us to identify the functions $C^{(s)}$ as the direct correlation functions of the isotropic liquid. We pause to remark that the Mayer expan-

sions of these objects involve only one-particle irreducible graphs and are thus rather short range. It should also be noted that the $C^{(s)}$ can be defined via (11) in an inhomogeneous system.

Noting the inverse of the transformation (9)

$$W = \min_{\rho(X)}[\beta\mathcal{F} + \int dX\, u(X)\rho(X)] \tag{13}$$

and assuming that the direct correlation functions are known in the uniform state of density ρ_0, we can functionally expand W. In this expansion, $\rho(X)$ and $u(X)$ are treated as independent functional degrees of freedom: They are later linked through the functional minimization in (13). For the purposes of this work, there will be no external potential and equal chemical potentials for all configurations of the one-body distribution. This leads to the expansion for the grand potential:

$$\Delta w = \frac{W - W_0}{\rho_0 V} = 1 + \frac{1}{\rho_0 V}\int dX\, \rho(X)\left[\log\frac{\rho(X)}{\rho_0} - 1\right] - \sum_{s=2}^{\infty}\frac{1}{\rho_0 V s!}\int dX_1 \cdots dX_s\, C^{(s)}(X_1, \ldots, X_s; \rho_0)[\rho(X_1) - \rho_0]\cdots[\rho(X_s) - \rho_0] \tag{14}$$

where W_0 is the grand potential of the isotropic fluid of density ρ_0. The $\rho(X)$ that minimizes this expression is the stable equilibrium configuration of the system, and when two local minima switch global stability, there is a (generally first-order) phase transition.

The role that the various terms of (14) play is reasonably clear; the ideal-gas entropy always favors the disordered state $\rho(X) = \rho_0$. The sum of direct correlation function integrals measures the effect of interactions; for hard-core systems, this will strongly discourage the overlap of hard-core regions [$C^{(s)}$ in the dense fluid regime is large and negative when its s arguments correspond to overlapping configurations of cores]. This tendency to avoid overlapping configurations drives phase transitions by making the uniform state $\rho(X) = \rho_0$ have higher grand potential than some ordered state which (generally) satisfies $[\rho(X) - \rho_0][\rho(X') - \rho_0] < 0$ for $|\mathbf{r} - \mathbf{r}'| < D(X,X')$.

It should be noted that the formal density functional expansion of the grand potential (14) is strictly valid only when the isotropic liquid is the stable thermodynamical state, since the existence of isotropic liquid direct correlations is assumed. An additional limitation is our knowledge of direct correlation functions; for the hard-sphere liquid, arguably the best characterized of simple liquids, we know little about $C^{(3)}$, and basically nothing about higher-order direct correlations. However, the existence

of a functional $W[\rho,u]$ from which the stable state ρ is computed by global functional minimization is guaranteed by the density functional theorem [28]; the idea of the density functional approach is to attempt to find as good an approximation to this (unknown) functional as possible.

One approach to approximate the free-energy functional is simply to truncate the expansion (14) after the $C^{(2)}$ term; unfortunately, for crystallization transitions, there is evidence that the higher-order terms in the expansion are not small [29]. We shall refer to the expansion (14) truncated at second order as the RY theory, after Ramakrishnan and Yussouff. In Section III.B we discuss a nonperturbative approach to this problem which results in a free-energy functional that makes use of the functional relationships between the different direct-correlation functions to avoid these problems.

B. Weighted-Density Approximation for Anisotropic Particles

In Section III.A we discussed the basic functional expansion of the free energy around the isotropic liquid; via truncation of this expansion, an approximation for the free-energy functional for inhomogeneous states is obtained. Unfortunately, we possess such limited information about $C^{(3)}$ and higher-order direct correlations that it is difficult to improve on the second-order RY theory by including higher-order terms. However, it is possible to formulate a nonperturbative approach to the problem: Here we describe such an approach related to that introduced by Curtin and Ashcroft [19] and independently by Tarazona [30].

We focus on the description of the fluid in the canonical ensemble, where the number of particles $\int dX\ \rho(X) = N$ is fixed. The approach begins with the partition of the Helmholtz free energy into ideal and excess portions: $\mathcal{F} = \mathcal{F}_{id} + \mathcal{F}_{ex}$, with $\mathcal{F}_{id}$ the free energy of an inhomogeneous noninteracting gas, defined as in (10). We can express the excess free energy as an integral over a local excess free energy f:

$$\mathcal{F}_{ex} = \int dX\ \rho(X) f(X;[\rho]) \tag{15}$$

In analogy to the local density approximation used in the theory of electronic structure, the local density approximation (LDA) for f is

$$f(X;[\rho]) = f_0(\rho(X)) \tag{16}$$

where $f_0(\rho)$ is the (assumed known) excess Helmholtz free energy per particle of the isotropic fluid state of density ρ. Although this approximation may be suitable to study inhomogeneities induced in a fluid by a

weak external potential, the LDA obviously cannot lead to a theory for ordering in a fluid due to intermolecular interactions.

If we instead introduce a new effective density $\bar{\rho}(X)$, we can rewrite (15) as

$$\mathcal{F}_{ex}^{WDA}[\rho] = \int dX\, \rho(X) f_0(\bar{\rho}(X)) \tag{17}$$

In principle, the right choice of $\bar{\rho}$ will recover the exact excess potential (15). In practice, Curtin and Ashcroft [19] studied the case where the effective density is written as a weighted average:

$$\bar{\rho}(X) = \int dX'\, \rho(X') w(X,X';\bar{\rho}(X)) \tag{18}$$

This theory is generally referred to as the weighted-density approximation (WDA).

The form of w is set by requiring that in the homogeneous limit $\rho(X) \to \rho$, the known properties of the uniform liquid of density ρ are recovered. This limit applied to (18) reveals that

$$\int dX'\, w(X,X';\bar{\rho}(X)) = 1 \tag{19}$$

Imposing that the second functional derivative of the excess potential with respect to the density $\rho(X)$ in the homogeneous limit gives the second direct correlation function of the liquid, that is, that

$$C^{(2)}(X,X';\rho) = -\beta \lim_{\rho(X)\to\rho} \frac{\delta^2 \mathcal{F}_{ex}([\rho])}{\delta\rho(X)\, \delta\rho(X')} \tag{20}$$

leads to a set of integrodifferential equations that determine w uniquely.

This approach has been extremely successful in the study of a variety of phase transitions in fluid systems where the particles interact via spherically symmetric interactions [19,31]. The WDA does a better job in predicting the localization of particles near lattice sites at the freezing transition (the RY theory typically predicts Lindemann ratios that are a factor of 4 too small for hard-sphere freezing) and gives a better account of the entropy difference between the coexisting liquid and solid phases at the freezing point. The WDA also has the virtue of being a suitable technique for studying the ordered phase [31]. The practical implementation of the WDA is aided by the fact that for crystallization transitions, the effective density (18) is everywhere significantly lower (>30%) than that the ordered phase average density, which are densities where liquid structure

is well understood. In contrast, the RY theory requires knowledge of the liquid structure at the freezing point, where it may not be well defined.

An attractive feature of the WDA is that it is an approximate summation to all orders of the density functional expansion (14). This comes about due to the feedback of $\bar{\rho}(X)$ in the definition of the effective density (18). In the homogeneous limit, $\rho(X) \to \rho$, which leads to $\bar{\rho}(X) \to \rho$, the WDA satisfies the following infinite set of sum rules [19]:

$$-\beta \int dX_3 \cdots dX_s \lim_{\rho(x)\to\rho} \frac{\delta^s \mathcal{F}_{ex}^{WDA}([\rho])}{\delta\rho(X_1)\cdots\delta\rho(X_s)} = \frac{\partial^{s-2}}{\partial\rho^{s-2}} C^{(2)}(X_1,X_2;\rho) = \int dX_3 \cdots dX_s C^{(s)}(X_1, \ldots, X_s;\rho) \tag{21}$$

for $s = 2, 3, 4, \ldots$. Thus, for small inhomogeneities, the WDA has as its limiting form the RY theory to second order in $\rho(X) - \rho$, plus higher-order terms that use approximate liquid structure information which satisfy all the sum rules (21). This situation contrasts markedly with the situation for the RY theory, where after two functional differentiations, the remaining excess free energy is independent of the ordered phase density and vanishes after further functional differentiation.

Unfortunately, extension of the WDA to phase transitions involving nonspherical particles is impractical because of the necessity of solving five-dimensional nonlinear integrodifferential equations for the weight function. One possibility is simply to specify the weight function [32]. Alternatively, one might replace the inhomogeneous effective density with a homogeneous effective density. This kind of approach has been used to study nonspherical particles by two groups [33,34]. Here we describe a theory similar to the WDA due to Denton and Ashcroft [35] which retains many of the desirable features of the WDA without leading to a large numerical calculation. It uses a homogeneous effective density and can be used to study anisotropic particles [36].

This modified WDA (MWDA) considers the excess Helmholtz free energy to be derived from that of the liquid at a single effective density $\bar{\rho}$:

$$\mathcal{F}_{ex}^{MWDA}[\rho] = \int dX\, \rho(X) f_0(\bar{\rho}) = N f_0(\bar{\rho}) \tag{22}$$

This effective density is written as a complete contraction of the physical density over a weight function w:

$$\bar{\rho} = N^{-1} \int dX\, \rho(X) \int dX'\, \rho(X') w(X,X';\bar{\rho}) \tag{23}$$

In the homogeneous limit, we recover the normalization condition (19), but now the weight function w is determined by (20) without any trouble as

$$w(X,X';\rho) = -\frac{\rho f''(\rho)}{2Vf'(\rho)} - \frac{C^{(2)}(X,X';\rho)}{2\beta f'(\rho)} \tag{24}$$

The form of the MWDA is much simpler than the WDA, but again, due to the feedback of $\bar{\rho}$ in the definition of the effective density, all the sum rules (21) are exactly satisfied. For the freezing of hard spheres, the MWDA gives results that are essentially identical to those obtained by the WDA with a large reduction in the complexity of the numerical calculations [35]. This agreement is due to the fact that in addition to having the same second-order density functional expansion and satisfying the same infinite set of sum rules, the MWDA and WDA are identical in two limits, the homogeneous limit and the limit of infinitely localized particles (delta functional distributions).

Calculation of phase transitions using the MWDA follows a route similar to that for the RY theory; the main difference is that the WDA and MWDA are concerned with the Helmholtz free energy; one proceeds by thus minimizing $\mathcal{F}[\rho(X)]$ over $\rho(X)$ subject to the constraint that $\int dX\, \rho(X) = \rho$. This reduces the number of free parameters in the variation, which is desirable. The MWDA calculations require the two-point direct correlation function and also the excess free energy per particle of the liquid as a function of density. The latter requirement is simply obtained from calculation of the zero-momentum, zero-angular-momentum part of $C^{(2)}$; from the definition of the direct correlations, this is essentially the inverse compressibility, which can be integrated to yield the pressure and free energy (see Section IV).

In these calculations it is convenient to carry out the minimization in the canonical ensemble and then to calculate phase coexistence by computing the grand potential for the ordered phase and solving for when it equals the grand potential of the liquid at a chemical potential equal to that corresponding to the ordered phase. For reference, the relationship between the Helmholtz free energy $\mathcal{F}$, the chemical potential μ, and the grand potential W is

$$W = \beta\mathcal{F} - \beta\mu N = -V\rho^2 \frac{\partial}{\partial\rho}\frac{\beta\mathcal{F}}{N} \tag{25}$$

where $\rho = N/V$ is the average density. We discuss in Section V some results of calculations for phase transitions of anisotropic particles using the RY and the MWDA theories.

C. Parametrization of Ordered Phases

Having described approaches to construction of the free-energy functional, we now discuss the problem of parametrizing its functional argument, the one-body density $\rho(X)$. The one-body density is the order parameter in a density functional theory of phase transitions. One must suitably parametrize $\rho(X)$ to obtain a tractable numerical calculation; the choice of parametrization must obviously be motivated by the structure of the ordered phases one is concerned with.

A general expression for a real-valued $\rho(X)$ in terms of complex-valued expansion coefficients $a_{\mathbf{k}l}^{mn}$ is

$$\rho(X) = \sum_{\mathbf{k}} \sum_{l=0}^{\infty} \sum_{mn} a_{\mathbf{k}l}^{mn} \, e^{i\mathbf{k}\cdot\mathbf{r}} D_{mn}^{l}(\theta,\phi,\chi) \tag{26}$$

where $\mathbf{k}$ are a set of momenta suitable for describing the inhomogeneous state, $a_{\mathbf{k}l}^{mn} = (-1)^{n+m} a^{*}_{-\mathbf{k}l}{}^{m-n}$, the D_{mn}^{l} are the rotation matrices as discussed by Gray and Gubbins [37], the summations over m and n are from $-l$ to l, and $*$ denotes complex conjugation. Evidently, we can compute the $a_{\mathbf{k}l}^{mn}$ from a given $\rho(X)$ using the orthogonality of the D's:

$$a_{\mathbf{k}l}^{mn} = \frac{2l+1}{V} \int dX \, e^{-i\mathbf{k}\cdot\mathbf{r}} D_{mn}^{*l}(\theta,\phi,\chi)\rho(X) \tag{27}$$

Our approach will be to parametrize $\rho(X)$ in terms of some small set of variational parameters (four or five is a practical number of parameters). The local ideal gas term $\mathcal{F}_{\mathrm{id}}$ is generally easily computed directly; however, it is useful to resolve ρ into its components $a_{\mathbf{k}l}^{mn}$ in order to perform the multiple integrals of the excess free energy.

The two-point contribution to the excess free energy of the RY theory (14) and the MWDA effective density (23) involve integrals of the form

$$V^{-1} \int dX \, dX' \, \rho(X)\rho(X')c(X,X') \tag{28}$$

where c is the pair DCF of the isotropic liquid and is invariant under translations and rotations of the coordinates. The pair DCF c can be expanded as

$$c(X_1,X_2) = \sum_{l_1 m_1 n_1} \sum_{l_2 m_2 n_2} \sum_{lm} (2l_1+1)(2l_2+1)(l_1 l_2 l \mid m_1 m_2 m) \times D_{m_1 n_1}^{l_1}(\theta_1,\phi_1,\chi_1) D_{m_2 n_2}^{l_2}(\theta_2,\phi_2,\chi_2) Y_{lm}(\theta,\phi) c_{l_1 l_2 l}^{n_1 n_2}(|\mathbf{r}_1 - \mathbf{r}_2|) \tag{29}$$

where $(l_1 l_2 l \mid m_1 m_2 m)$ are Clebsch–Gordan coefficients, the Y_{lm} is a spherical harmonic, and the angles θ and ϕ are those describing the vector

$\mathbf{r}_1 - \mathbf{r}_2$. The $c^{n_1n_2}_{l_1l_2l}$ may be extracted from $c(X_1,X_2)$ via

$$c^{n_1n_2}_{l_1l_2l}(|\mathbf{r}_1 - \mathbf{r}_2|) = \int \frac{d(\cos\theta)\, d(\cos\theta_1)\, d(\cos\theta_2)\, d\phi\, d\phi_1\, d\phi_2 d\chi_1\, d\chi_2}{64\pi^4(l_1l_2l \mid 000)} \times D^{*l_1}_{0n_1}(\theta_1,\phi_1,\chi_1)D^{*l_2}_{0n_2}(\theta_2,\phi_2,\chi_2)Y^*_{l0}(\theta,\phi)c(X_1,X_2) \tag{30}$$

Using the decomposition of a plane wave into spherical harmonics,

$$e^{i\mathbf{k}\cdot\mathbf{r}} = 4\pi \sum_{l=0}^{\infty} \sum_{m=-l}^{l} i^l j_l(kr) Y^*_{lm}(\hat{\mathbf{k}}) Y_{lm}(\hat{\mathbf{r}}) \tag{31}$$

we can write (28) as

$$\sum_{\mathbf{k}lm} \sum_{l_1m_1n_1} \sum_{l_2m_2n_2} a^{m_1n_1}_{\mathbf{k}l_1} a^{m_2n_2}_{-\mathbf{k}l_2} Y_{lm}(\hat{\mathbf{k}})(l_1l_2l \mid m_1m_2m)\tilde{c}^{n_1n_2}_{l_1l_2l}(k) \tag{32}$$

where the unit vector $\hat{\mathbf{k}}$ denotes the angles corresponding to the vector $\mathbf{k}$, $r = |\mathbf{r}|$, $k = |\mathbf{k}|$, and where

$$\tilde{c}^{n_1n_2}_{l_1l_2l}(k) = 4\pi(-i)^l \int_0^\infty dr\, r^2 j_l(kr) c^{n_1n_2}_{l_1l_2l}(r) \tag{33}$$

It is often possible to make an approximation that simplifies the calculations. If we are studying an inhomogeneous state $\rho(X)$ where particles are relatively localized in space, often it is a good approximation to decouple the translational and orientational and distributions:

$$\rho(X) = \rho_{\text{tr}}(\mathbf{r})\rho_{\text{or}}(\theta,\phi,\chi) \tag{34}$$

This is a good approximation for many states such as crystalline solids, where the particles are strongly localized to lattice sites, or for simple smectics, where the orientational distribution does not vary much between the smectic layers. The benefits of this decoupling are large. The coefficients in (27) decouple:

$$a^{mn}_{\mathbf{k}l} = t_{\mathbf{k}} u_{lmn} \tag{35}$$

and the ideal gas free energy (10) also decouples additively into contributions from the translational and orientational distributions [38]. We adopt the convention that when such a decoupling is used, the normalizations will be $\int dX\, \rho_{\text{tr}}(X) = N$, $\int dX\, \rho_{\text{or}}(X) = V$.

We now discuss parametrization schemes suitable for various ordered phases.

1. Positional Distribution for Crystalline Phases

To characterize a crystalline phase, we need to specify the lattice, lattice constants, and information concerning the localization of particles to lat-

tice sites. For many purposes, a one-particle distribution consisting of a lattice of Gaussian distributions is suitable. Even dense hard-core solids possess nearly Gaussian lattice-site distributions [1], and the form is convenient for these calculations. Such a distribution might be written as

$$\rho_{tr}(\mathbf{r}) = \rho_1 \Delta \pi^{-3/2} \sigma^{-3} \sum_{\mathbf{s}\in\mathbf{T}} \exp(-\sigma^{-2}[\mathbf{r} - \mathbf{s}]^2) \tag{36}$$

where ρ_1 is the average density of the solid, Δ the volume per lattice site, σ the localization length, and $\mathbf{T}$ the real-space lattice. This can be written in terms of Fourier components as

$$\rho_{tr}(\mathbf{r}) = \rho_1 \sum_{\mathbf{k}\in\mathbf{U}} e^{-(\sigma\mathbf{k}/2)^2} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{37}$$

where now $\mathbf{U}$ is the inverse lattice of $\mathbf{T}$. Distorted unit cells can easily be obtained by performing the transformation

$$\begin{aligned} \rho_{tr}(\mathbf{r}) &\to \det M \, \rho_{tr}(M\mathbf{r}) \\ \mathbf{r} &\to M\mathbf{r} \\ \mathbf{s} &\to M\mathbf{s} \\ \mathbf{k} &\to M^{-1}\mathbf{k} \end{aligned} \tag{38}$$

where M is a matrix that performs the desired transformation. We note that such a distortion does not affect the ideal gas free energy (10) since we can change integration variables back to the undistorted ones.

2. *Positional Distribution for Smectics*

We shall discuss smectic A and C phases, which are described in terms of a single mass density wave, which we will take along the $\hat{\mathbf{z}}$-axis. The positional distribution is invariant under translation in the xy-plane, and thus we write

$$\begin{aligned} \rho_{tr}(\mathbf{r}) &= \rho_1 \frac{c}{\sigma\pi^{1/2}} \sum_{s=-\infty}^{\infty} \exp[-\sigma^{-2}(z - sc)^2] \\ &= \rho_1 \sum_{k=-\infty}^{\infty} e^{-(\pi\sigma k/c)^2} e^{2\pi i k z/c} \end{aligned} \tag{39}$$

which is useful in both the strongly localized ($\sigma \ll c$) and the delocalized ($\sigma \gg c$) regimes. For the latter case the distribution reduces to

$$\rho_{tr}(\mathbf{r}) = \rho_1 \left[1 + b \cos \frac{2\pi z}{c} + \mathcal{O}(b^4) \right] \tag{40}$$

where the smectic order parameter $b = e^{-(\pi\sigma/c)^2}$.

3. *Uniaxial Orientational Distribution*

For the nematic and smectic A phases, the orientational order is uniaxia with the director along the z-axis. The state $\rho(X)$ is independent of th angles θ and ϕ in this case, and a suitable orientational distribution is

$$\rho_{or}(\theta) = \exp(a_0 + a_2\mu^2 + a_4\mu^4 + \cdots) \quad (41$$

where $\mu = \cos\theta$, and where the normalization a_0 is chosen so that $\int d$ $\rho_{or}(\theta) = V$. Oriented crystal phases have strongly uniaxial orientation distributions that are well approximated by such a distribution. Such distribution can of course be written as

$$\rho_{or}(\theta) = 1 - \sum_{l=2,4,6,\cdots} b_l P_l(\mu) \quad (42$$

but the exponential form is useful for cases where the orientational or dering becomes strong, as is the case in hard-core smectics or oriente crystals.

4. *Tilted Orientational Distribution*

We note that for the smectic C, the director $\hat{\mathbf{c}}$ makes a nonzero angle with the normal to (or smectic planes). In this case this tilt will be du to a lack of symmetry of the molecule under rotation about the χ-axis Without lack of generality, we suppose that $\hat{\mathbf{c}}$ is in the xz-plane. In thi case a suitable orientational distribution is

$$\rho_{or}(\theta,\phi,\chi) = \exp(a_0 + a_2\mu_c^2 + a_4\mu_c^4 + \cdots)\exp(b_1\cos\chi + d_1\sin\chi + b_2\cos 2\chi + d_2\sin 2\chi + \cdots \quad (43$$

Here $\mu_c = \sin\alpha\sin\theta\cos\phi + \cos\alpha\cos\theta$ is the dot product of the uni vector described by the angles θ and ϕ and the director axis. The χ de pendence has been decoupled from that on θ and ϕ. We note that if th molecule is symmetric under $\chi \to -\chi$, the coefficients $d_1 = d_2 = \cdots = 0$.

5. *Orientational Distribution for Crystalline Phases*

We close by noting that in any phase with translational symmetry, th orientational distribution will be invariant under a set of orientationa symmetry operations that form a subgroup of the point group of the lattice For example, in the plastic solid phase of a system of hard ellipsoids, th lattice is face-centered cubic, and thus is invariant under the cubic (oc tahedral) point group. The orientational distribution $\rho_{or}(\theta,\phi)$ should als be cubic invariant.

D. Onsager Theory of Isotropic–Nematic Phase Transition

An important requirement of any theory is that it must reproduce known exact results. How can we make contact with the Onsager theory of the I–N transition in a dilute gas of elongated, hard particles? The link between the theories is straightforward and was first pointed out by Sluckin and Shukla [39].

The functional order parameter for a uniaxial nematic is of the form

$$\rho(X) = \rho_0 n(\mu) \tag{44}$$

where $-1 < \mu < 1$ is the cosine of the polar angle describing the molecular orientation. The density of the ordered phase is ρ_N if we use the normalization

$$\int_{-1}^{1} \frac{d\mu}{2}\, n(\mu) = \frac{\rho_N}{\rho_0} \tag{45}$$

The terms involving $C^{(s)}$ are of $(s-1)$th order in ρ_N; thus we may ignore all but the one involving $C^{(2)}$. In this limit of low density, the pair direct correlation function reduces to its lowest-order (zeroth-order) graph, which is just the Mayer function f:

$$C^{(2)}(X,X';\rho_0) = f(X,X') = \exp[-\beta U(X,X')] - 1 \tag{46}$$

For hard-core interactions, $f(X,X')$ is -1 for forbidden configurations [i.e., $|\mathbf{r} - \mathbf{r}'| < D(X,X')$] and vanishes for all other configurations. Thus we rewrite (14) in the form

$$\Delta w = \frac{\rho_0 - \rho_N}{\rho_0} + \int_{-1}^{1} \frac{d\mu}{2}\, n(\mu) \log n(\mu) + \frac{\rho_0}{2V} \int_{|\mathbf{r}-\mathbf{r}|<D(X,X')} dX\, dX'\, n(\cos\theta) n(\cos\theta') \tag{47}$$

This functional for the grand potential can be transformed to the canonical ensemble via (9); we are interested in the case where the chemical potential per particle is the same in either phase, and also where $\rho_0 = \rho_N$. We are led to the difference in Helmholtz free energies of the two phases,

$$\Delta\beta f = \int_{-1}^{1} \frac{d\mu}{2}\, n(\mu) \log n(\mu) + \frac{\rho_0}{2V} \int_{|\mathbf{r}-\mathbf{r}'|<D(X,X')} dX\, dX'\, n(\cos\theta) n(\cos\theta') \tag{48}$$

where now n has the normalization $\int_{-1}^{1}(d\mu/2)n(\mu) = 1$, which is the result due to Onsager [8,6].

E. Isotropic–Nematic–Smectic Phase Transitions

In this section the connection between the density functional and Landau theories of phase transitions is made explicit. We then describe the application of the theory to the transitions between the isotropic liquid, nematic, and smectic A phases [39–41]. Returning to the general case (26), we note that the functional order parameter is $\rho(X) - \rho_0$, where ρ_0 is the density of the isotropic phase. We introduce the dimensionless order parameters

$$b_{\mathbf{k}l}^{mn} = \rho_0^{-1} a_{\mathbf{k}l}^{mn} - \delta_{\mathbf{k}0}\delta_{l0}\delta_{m0}\delta_{n0} \tag{49}$$

The grand potential difference between the isotropic and ordered phases (14) can then be expanded in these order parameters as

$$\Delta w = \sum_{s=2}^{\infty} \frac{1}{s!} \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_s} \sum_{l_1,\ldots,l_s} \sum_{m_1,\ldots,m_s} \sum_{n_1,\ldots,n_s} b_{\mathbf{k}_1 l_1}^{m_1 n_1} \cdots b_{\mathbf{k}_s l_s}^{m_s n_s} W_{\mathbf{k}_1\cdots\mathbf{k}_s l_1\cdots l_s}^{(s) m_1\cdots m_s n_1\cdots n_s}(\rho_0) \tag{50}$$

where the $W^{(s)}$'s are given by

$$\begin{aligned} W_{\mathbf{k}_1\cdots\mathbf{k}_s l_1\cdots l_s}^{(s) m_1\cdots m_s n_1\cdots n_s}(\rho_0) &= \delta_{\mathbf{k}_1+\cdots+\mathbf{k}_s,0} \\ &\times \left[f_s \int \frac{d\cos\theta\, d\phi\, d\chi}{8\pi^2} D_{m_1 n_1}^{l_1}(\theta,\phi,\chi)\cdots D_{m_s n_s}^{l_s}(\theta,\phi,\chi) \right. \\ &\quad - \frac{\rho_0^{s-1}}{V} \int dX_1 \cdots dX_s C^{(s)}(X_1, \ldots, X_s;\rho_0) \\ &\left. \times e^{i(\mathbf{k}_1\cdot\mathbf{r}_1+\cdots+\mathbf{k}_s\cdot\mathbf{r}_s)} D_{m_1 n_1}^{l_1}(\theta_1,\phi_1,\chi_1)\cdots D_{m_1 n_s}^{l_s}(\theta_s,\phi_s,\chi_s) \right] \end{aligned} \tag{51}$$

and where

$$f_s = \frac{d^s}{dx^s}[(1+x)\log(1+x) - x]_{x=0}$$

$W^{(2)}$ can be obtained rather explicitly:

$$\begin{aligned} W_{\mathbf{k}_1\mathbf{k}_2 l_1 l_2}^{(2) m_1 m_2 n_1 n_2}(\rho_0) &= \delta_{\mathbf{k}_1,-\mathbf{k}_2} \left[\frac{(-1)^{m_1+n_1}\delta_{l_1 l_2}\delta_{m_1,-m_2}\delta_{n_1,-n_2}}{2l_1+1} \right. \\ &\left. - \rho_0 \sum_{lm} Y_{lm}(\hat{\mathbf{k}}_1)(l_1 l_2 l \mid m_1 m_2 m)\tilde{c}_{l_1 l_2 l}^{n_1 n_2}(k_1;\rho_0) \right] \end{aligned} \tag{52}$$

We note the explicit form of $W^{(2)}$ since it is the only W involving DCFs in the second-order RY theory; a similar expression is required to compute the effective density in the MWDA theory. The second-order term also determines the stability limit of the isotropic phase, since the isotropic phase becomes unstable when an eigenvalue of $W^{(2)}$ crosses zero. At $\rho_0 = 0$, all eigenvalues of $W^{(2)}$ are positive. In the case of the instability toward the nematic, we consider only $\mathbf{k}_1,\mathbf{k}_2 = \mathbf{0}$ and $n_1,n_2,m_1,m_2,m = 0$; the Clebsch–Gordan coefficient causes $W^{(2)}$ to be explicitly diagonal:

$$W^{(2)00000}_{00ll'}(\rho_0) = \frac{\delta_{ll'}}{2l+1}\left[1 - \rho_0(-1)^l\sqrt{\frac{2l+1}{4\pi}}\,\tilde{c}^{00}_{ll0}\right] = \frac{\delta_{ll'}}{2l+1}\,k_l^{-1} \tag{53}$$

Here the k_l are compressibilities associated with distortions of the system with zero momentum, but with nonzero angular momentum l; k_0 is the compressibility $kT\,d\rho/dp$ and k_2 is the static Kerr constant [42]. These constants can also be written as

$$k_l^{-1} = 1 - \rho_0\int\frac{dX_1\,dX_2}{V}\,C^{(2)}(X_1,X_2;\rho_0)P_l(\hat{\mathbf{e}}_1\cdot\hat{\mathbf{e}}_2) \tag{54}$$

When one of these k_l^{-1} becomes zero, there is an instability to the nematic phase.

For axially symmetric molecules and the uniaxial orientational symmetry of the I–N–A phases, all of the components of c and ρ are zero except for those with $m = n = 0$. In addition, if the molecules have inversion symmetry, only terms with l's even appear. For the I and N phases, only the $\mathbf{k} = \mathbf{0}$ component appears; for smectic A, only $\mathbf{k} = 2\pi n\hat{\mathbf{z}}/\lambda$ for $n = 0, \pm1, \pm2, \ldots$ need be considered, where λ is the smectic wavelength. For the case where only $n = -1, 0, 1$ and angular momenta $l = 0, 2$ are included, a rather complete Landau theory can be developed, with all coefficients determined in terms of the direct correlations. Connections with specific mean-field theories have been discussed by Sluckin and Shukla [39], Singh [43], and Lipkin and Oxtoby [41].

IV. HOMOGENEOUS-PHASE LIQUID STRUCTURE

As we have seen in Section III a density functional theory for liquid crystals requires knowledge of static liquid structure. From a practical standpoint, this means knowledge of at least the two-point direct correlation function (DCF) of the dense isotropic phase over a range of densities

near the ordered phases. In principle this requires the solution, at least approximately, of a strongly interacting many-body problem and is a topic whose even partial treatment is totally beyond the scope of this section. Here we endeavor merely to give the reader an overview of what is known about the DCFs of isotropic liquids composed of anisotropic particles.

Much attention has been paid to integral equations for the correlation functions, and in particular to the Ornstein–Zernike equation (12), which relates the pair distribution g to the DCF c. Approximate closures of this equation are of great interest, as they allow a nonperturbative solution of the problem of determining the correlation functions [44]. One such closure is the Percus–Yevick approximation, which is easily stated in the case of a hard-core potential:

$$c(X,X') = 0, |\mathbf{r} - \mathbf{r}'| > D(X,X') \tag{55}$$

Once this approximation is made (which is reasonable due to the fact that it is known from simulations that the magnitude of c is much smaller outside the hard-core overlap region than inside) we can write the OZ equation in terms of a single continuous function $\tau = g - c$. This reduces the problem to the solution of a quadratic integral equation. (See Chapter 4 for more on this topic.)

We note that for dense fluids composed of particles interacting via spherically symmetric interactions, a good deal is known [10] about the static correlation functions. In particular, we are fortunate to possess a closed-form solution to the Percus–Yevick (PY) closure of the Ornstein–Zernike equation (12) for the case of hard spheres in three dimensions [45]. This is important because the solution is relatively simple in structure and leads to correlation functions and thermodynamics in good agreement with existing simulation data. The resulting DCF has been used successfully to treat freezing of hard spheres [46,47] as well as freezing in systems with longer-range interactions (see Chapter 9).

A. Density Expansions

An important result is simply the density expansion for the pair DCF [48], $c(X,X')$. This expansion is shown graphically in Fig. 6, where a bond between vertices i and j represents the Mayer function $f_{ij} = e^{-v(i,j)/kT} - 1$, where v is the pair interaction potential. The form of this expansion is not dependent on the types of degrees of freedom or the dimension as long as all of the interactions are additive pair potentials. We use the usual notation where filled vertices represent points that are integrated over all configurations and where each graph is multiplied by a factor $\rho = N/V$

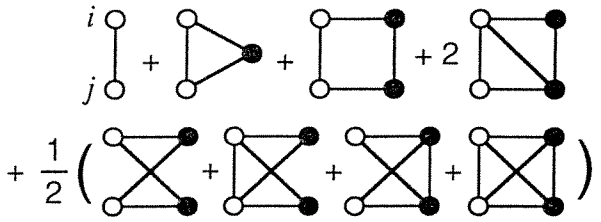


Fig. 6 Density expansion for the direct correlation function $c(1,2)$. A bond between vertices i and j represents a term $\exp[-v(i,j)/kT] - 1$, where v is the pair interaction potential. Filled vertices are integrated over all configurations, and each graph is multiplied by a factor ρ for each filled vertex.

for each filled vertex. The open vertices are of course the external points corresponding to the two configuration arguments of the two-point DCF.

For hard-core potentials the bonds are -1 if the particles at the vertices overlap, and are 0 otherwise. We see that the first term (of order 1) merely contributes a jump from 0 to -1 when the two external points begin to overlap. The second term also is zero when the external points do not overlap, but when they do, the contribution (of order ρ) is obtained by integrating over all mutually overlapping configurations of a third particle. This results in a negative contribution that, for spheres, grows in magnitude as the external points are brought together. These two terms allow us to understand the behavior of integral equation and simulation results for the DCF for hard spheres; the DCF is nearly zero if the two spheres do not overlap, there is a negative jump at contact, and then as the spheres increase their overlap, the DCF monotonically becomes more negative. The DCF for overlapping configurations becomes steadily more negative as the density is increased.

For anisotropic particles we must consider the dependence of the DCF on the orientations as well as the positions of the two external points. The zero-order term in the expansion again is simply -1 or 0 for overlapping or nonoverlapping configurations, respectively. The second term is now more complicated. Suppose that the particles are extremely elongated (e.g., cylinders with $L/D = 10$). If the two external points have parallel orientations, we find that the volume of the overlap region increases as the particles are brought together. This is approximately proportional to the contribution of the first-order graph since for most orientations of the integrated particle, its long axis is not parallel to the external points. Conversely, for perpendicular configurations, the volume of the region of overlap of the external particles increases only while one particle is "pushing through" the other. Thus the overlap increases to

some maximum value, and then is roughly constant for all configurations beyond the point where one cylinder has pushed completely through the other. The value of the graph is again proportional to the overlap volume since most orientations of the integrated particle are out of the plane of the external points.

This argument suggests that at least for prolate particles, the behavior of the two-point DCF at low densities will be quite different for parallel and antiparallel configurations of the arguments. We note that precisely the behavior described above is observed in a simple, exactly soluble one-dimensional model for a fluid of anisotropic hard cores [16]. This model consists of rectangular hard cores constrained to have one of their body axes on a line, as indicated in Fig. 7; the cores can thus be translated along the line and can take on either of two orientations. The model is solved by noting that the graphs that are omitted from the expansion of

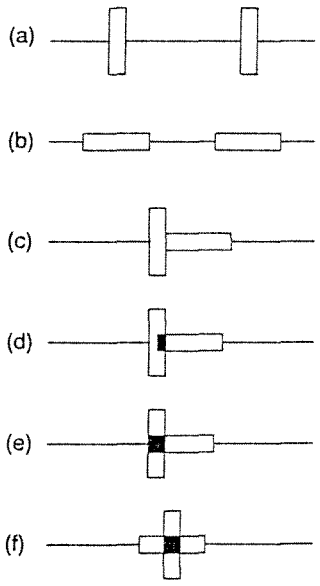


Fig. 7 Orientational configurations and different regimes of DCF behavior. (a), (b) Parallel configurations. DCF jumps from zero at core contact and becomes more negative as the overlap of the cores increases (as particle separation decreases). (c)–(f) Perpendicular configuration. (c) Hard-core contact. DCF jumps from zero to a finite negative value. (d) Partial overlap. DCF becomes steadily more negative as overlap increases. (e) Maximum overlap configuration. All particle separations smaller than this [e.g., (f)] have the same direct correlation.

the correlation functions by the PY approximation all vanish; the solution to the PY equation is therefore the exact solution.

The result is that for all densities (there are, of course, no phase transitions) the two-point DCF jumps from zero to some negative constant at hard-core contact, and for the parallel cases (Fig. 7a and b) becomes steadily more negative as the particle centers are brought together. For the perpendicular case, there is again a jump at contact (Fig. 7c) and an initial increase in the strength of the DCF as the overlap increases (Fig. 7d). At the point where the overlap stops increasing (Fig. 7e), the DCF stops increasing in strength also; it is constant for all smaller particle separations (as in Fig. 7f). For all nonoverlapping configurations, the DCF is zero. This result can be extended to certain cases involving continuous orientational degrees of freedom [16], and the results are similar.

This is interesting as it verifies the argument above for the behavior of the DCF for a one-dimensional case; we note that for strongly prolate particles, the arguments for the behavior of the first-order term presented above become more believable in high dimensions (since there are more orientations out of the plane formed by the long axes of the two external particles) and can be applied to the higher-order terms in the DCF density expansion. These sorts of results might be useful for studying problems such as the Onsager transition, which can be pushed to arbitrarily low densities by elongating the particles. However, apart from cases that permit exact summations, the density expansion cannot adequately address the problem of liquid structure for a dense fluid.

B. Geometrical Approximations

A heuristic approach to the problem of obtaining liquid structure for anisotropic hard cores is to attempt to define scaling parameters in order to find a mapping to a hard-sphere problem. One such prescription was proposed by Pynn and later studied by Wulf [49]; on the basis of the agreement with the results of a variational study of the PY-OZ equation, it was proposed that the hard-sphere DCF, with the sphere diameter D replaced by the orientation-dependent distance-of-closest approach $D(X,X')$, and the sphere packing fraction $\eta = \pi D^3\rho/6$ replaced by the packing fraction $\eta = \rho v$, where v is the molecular volume, would be a good approximate DCF for anisotropic particles. Pynn's original calculations were done with an approximate $D(X,X')$ for ellipsoids of revolution due to Berne and Pechukas [50]:

$$D(X,X') = 2B\left[1 - \chi\frac{(\hat{\mathbf{r}}\cdot\hat{\mathbf{e}})^2 + (\hat{\mathbf{r}}\cdot\hat{\mathbf{e}})^2 - 2\chi(\hat{\mathbf{r}}\cdot\mathbf{e})(\hat{\mathbf{r}}\cdot\hat{\mathbf{e}}')(\hat{\mathbf{e}}\cdot\hat{\mathbf{e}}')}{1 - \chi^2(\hat{\mathbf{e}}\cdot\hat{\mathbf{e}}')^2}\right]^{-1/2} \tag{56}$$

where $\chi = (A^2 - B^2)/(A^2 + B^2)$ is a parameter describing the anisotropy of the ellipsoids of minor axes $2A$ and rotational symmetry axis $2B$. Pynn's calculations agree well with this form for the case that he studied, $A/B = 1.4$.

The zeroth-order term of the DCF density expansion is, of course, given exactly by this *ansatz*. However, this approximation does not have the property (see Section IV.A) of the DCF being different in behavior for parallel and perpendicular configurations, as the scaling makes no reference to the overlap volume. Nevertheless, if one computes the pressure as a function of density by integrating the compressibility, one finds good agreement with Monte Carlo results for relatively high densities ($\eta \approx 0.4$) even up to an anisotropy of $A/B = 3$ [40]. The pressure of the isotropic liquid obtained from the Pynn–Wulf DCF as a function of packing fraction $\eta = \rho v$ can be obtained in closed form:

$$\beta P v = \eta + \frac{4\eta - 2\eta^2 + \eta^3}{(1 - \eta)^3} F(\chi) \quad , \tag{57}$$

The excess free energy per particle can also be obtained:

$$\beta f = [\tfrac{3}{2}(1 - \eta)^{-2} - \tfrac{3}{2} - \log(1 - \eta)]F(\chi) \tag{58}$$

where $v = 4\pi A^2 B/3$ is the molecular volume, and where

$$F(\chi) = \frac{1}{2}\left[1 + \frac{\sin^{-1}\chi}{\chi(1 - \chi^2)^{1/2}}\right]$$

Rosenfeld has discussed the behavior of hard-particle DCFs [51] and has recently proposed a scaled-particle theory for hard-core fluids [52]. The latter work attempts to construct an approximate DCF from the fundamental geometric measures needed to describe higher-order terms in the density expansion. Unfortunately, there has not been a comparison of the results of this method to simulation data.

C. Numerical Solution of Closures of the OZ Equation

No analytical solutions exist for any of the commonly studied closures of the Ornstein–Zernike equation (12) for the case of anisotropic hard cores, but increasing computational speed has made possible the numerical solution of these nonlinear integral equations. We refer the reader to Chapter 5 for more details on the integral equation approach to dense liquid structure; here we merely survey the current results and discuss their implications.

The first efforts to solve such problems seem to be those of Chen and Steele [53], who studied the hard-core PY equation. The quadratic nonlinearity present in this equation greatly simplifies the problem and allowed some information to be obtained concerning the case of hard dumbbells. Low densities and low anisotropies (center-to-center distances of 0.3 and 0.6 sphere diameters) were studied. Spherical harmonics of angular momenta larger than $l = 2$ were ignored: reasonable agreement was obtained, however, with both computer simulations and virial expansion results.

A more general prescription for solving nonlinear integral equations for liquid structure of anisotropic particles was presented by Lado [54], who applied his technique to the dumbbell problem. In this work, the reference-hypernetted-chain closure was used, with angular momenta up to $l = 4$. Lado investigated center-to-center distances of 0.2, 0.4, 0.6, and 0.8, at densities $\rho d^3 = 0.2$, 0.4, 0.6, and 0.8, where d is the diameter of a sphere with the same volume as the hard dumbbell. Lado found his results for thermodynamical quantities as well as the pair distribution to be in fairly good agreement with MC results.

A later paper by Lado [55] applied the technique to solve the PY equation for dumbbells. In this work, a comparison was made of the resulting DCF with the Pynn–Wulf approximation, and surprisingly good agreement was found for center-to-center distance-to-sphere diameter ratios up to 0.6 and densities up to $\rho d^3 = 0.8$. This work, combined with the original papers of Pynn [49], tells us the geometrical approximations discussed in Section IV.B are reasonable for dense fluids of prolate particles with aspect ratios less than 1.5 or so.

Unfortunately, from the review of simulation results in Section II, we are led to conclude that transitions to nematic or smectic liquid crystals occur for dense fluids composed of particles with aspect ratios larger than 2, precisely where the techniques mentioned above begin to break down. In the last few years, some progress has been made toward obtaining liquid structure for highly anisotropic hard cores. Perera, Kusalik, and Patey have solved the Percus–Yevick and hypernetted-chain closures for the hard ellipsoid [42] fluid; Perera and Patey went on to study the hard spherocylinder [56] fluid similarly. Spherical harmonic expansions up to angular momentum $l = 8$ were used.

These expansions appear to converge well even for aspect ratios of 5 at densities (for ellipsoids with symmetry axis length $2A$ and width $2B$) of $8\rho AB^2 = 0.45$. We comment that the PY closure generally works better than the HNC for the case of hard spheres. What is found for prolate ellipsoids is that for anisotropies up to $A/B = 3$, the equation of state obtained from the PY compressibility is in somewhat better agreement

with Monte Carlo data than either the virial PY equation of state or either type of equation of state obtained from the HNC equation. However, it seems that the orientation-integration radial distribution given by the HNC equation is in better agreement with the MC data. In particular, we note that the inverse Kerr constant obtained from the HNC solutions is more reasonable than that obtained from the PY solutions at densities approaching the isotropic–nematic transition. One expects the inverse Kerr constant to go to zero as the instability toward the nematic is approached (see Section III); the HNC (extrapolated) data appear to do this, while the PY data do not. In Fig. 8 we show the inverse Kerr constant for the PY and HNC solutions as a function of density for $A/B = 3$, and we have indicated the location of the isotropic–nematic transition observed in MC simulations. We also display the Pynn–Wulf result for the

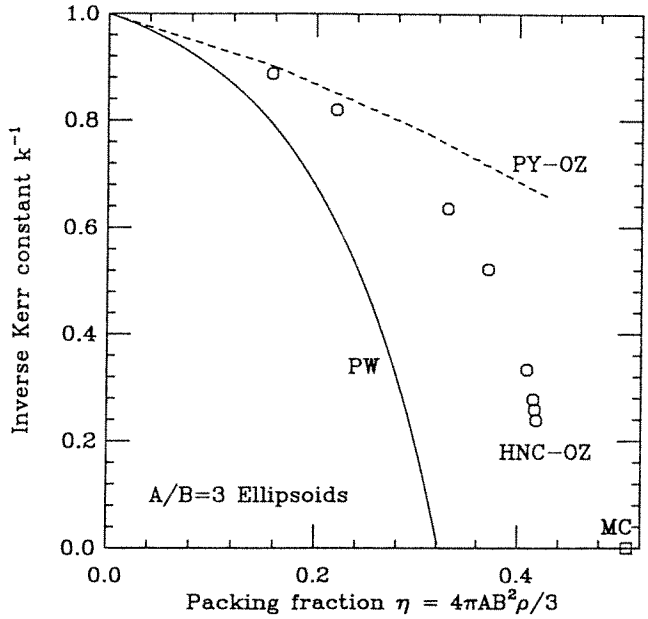


Fig. 8 Inverse Kerr constant for prolate hard ellipsoids of revolution of aspect ratio $A/B = 3$, as a function of density ρAB^2. Results from numerical solution of the HNC-OZ equation (filled points), numerical solution of the PY-OZ equation (dashed line), along with the result obtained from the Pynn–Wulf approximation (solid line). The isotropic–nematic transition density for $A/B = 3$ obtained from MC simulation is indicated on the density axis with an open box. (Adapted from Ref. 42.)

inverse Kerr constant; we note that like the PY result, it does not tend to zero at a reasonable density for the $A/B = 3$ case.

Perera and Patey's results for the spherocylinder case are quite similar, leading to the conclusion that although the PY compressibility equation of state is more accurate than HNC or PY virial results for relatively low anisotropies ($L/D = 1$), the PY virial, and HNC compressibility results agree better at higher anisotropies (up to $L/D = 5$ was studied). The larger volume of data available for spherocylinders allowed a better comparison of the correlation functions; it seems that the HNC results are in better agreement with the MC data. Again, the inverse Kerr constant given by the HNC solution is much more reasonable (again tending to zero near the known isotropic–nematic transitions) than the PY result.

A study of the parallel spherocylinder fluid was undertaken by Caillol and Weis [57], again comparing the HNC and PY equations. In this case it was found that the amplitude of the first peak of the HNC structure factor (along the axis of the spherocylinder) tended to diverge as the density where the transition to the smectic was observed in MC simulations. In the PY case, no such divergence was observed.

The DCF data obtained from these numerical studies is rarely presented as a function of radial separation for various orientations. However, this has been done by Lago and Sevilla [58], who studied the PY closure for spherocylinders with $L/D = 1$ for packing fractions ρv up to 0.39. (We note the agreement between the HNC and PY correlation functions at these densities in the study of Perera and Patey [56] for the case $L/D = 1$.) The most striking feature of these results is the apparent dependence of the DCF on the overlap volume. To quote, "the most outstanding feature is seen in the T orientation, where a large plateau is noticeable at higher densities" [58]. For parallel particle orientations, the DCF decreases to a much more negative value as the particle separation is reduced. This effect is reminiscent of the plateau in the DCF discussed in Section IV.A for configurations of prolate cores such that one punches through the other as the radial separation is decreased. It would be interesting to see similar plots of the real-space dependence of the DCF for the HNC and PY results discussed previously.

The preliminary conclusion is that although the PY approximation gives good thermodynamics and good liquid structure for weakly anisotropic prolate hard cores (aspect ratio < 2), for the more anisotropic particles likely to form nematics and smectics (aspect ratios of 3 or larger), the HNC equation appears to be preferable. In particular, the behavior of the inverse Kerr constant appears to be a useful way of examining the behavior of the anisotropic part of the liquid structure. The inverse Kerr constant is the $l = 2$ analog of the $l = 0$ part of the zero-momentum

component of the DCF, the inverse compressibility, which has long been used as a guide for determining the suitability of integral equations for liquid structure. What is required is a way of systematically improving approximate integral equations for structure of dense liquids, which is an old, unsolved problem.

V. DENSITY FUNCTIONAL STUDIES OF LIQUID–CRYSTAL PHASE TRANSITIONS

The following survey of recent density functional results obtained for various sorts of phase transitions is not complete, due to the volume of work that has been published on this topic in the past few years. It seems that the first application of the modern formulation of the density functional approach to a liquid–crystal phase transition was the study of the isotropic–nematic phase transition by Sluckin and Shukla [30]. As mentioned in Section III.D, this study developed the grand potential as a functional of the one-body density. An order-parameter theory was obtained by expanding the one-body density $\rho(\theta)$ in Legendre polynomials, which allows contact to be made with various types of theories of the I–N transition (i.e., Maier-Saupe, Onsager, de Gennes). Similar ideas were presented by Singh [43]. Lipkin and Oxtoby [41], who applied the density functional expansion to isotropic–nematic–smectic A phase transitions, making contact with the theory of McMillan.

A. Isotropic–Nematic Transition

Although the early studies emphasized that one could construct a mean-field theory for transitions to mesophases from microscopic information, no explicit calculation for a specific molecular model was carried out until Singh and Singh studied the isotropic–nematic transition in the hard ellipsoid system [40]. The one-body density was again expanded in Legendre polynomials as discussed in Section III.C. The Pynn–Wulf DCF (Section IV.B) was used, and all higher-order direct correlation functions were ignored (RY theory). At large anisotropies, the isotropic–nematic transition densities obtained from this calculation qualitatively agree with the simulation results of Frenkel et al. [21], although the density functional theory predicts the transition at densities 40% below the simulation result for aspect ratios of 2.75 and 3. This is presumably due to the inaccuracy of the Pynn–Wulf DCF (see Section IV). Singh and Singh also noted that their choice of DCF leads to a density-functional theory that is invariant under the transformation $A/B \rightarrow B/A$, $AB^2 \rightarrow 1/(AB^2)$; this is an approx-

imate symmetry plainly apparent in the MC phase diagram [21] (see Fig. 3). This calculation has been studied by other authors; similar results have been obtained [38,59–61].

The hard-ellipsoid nematic was also considered by Baus et al. [33], who considered an exact expression of the difference between the excess Helmholtz free energy of states with one-body densities $\rho(X)$ and ρ_0:

$$\beta\mathcal{F}_{ex}[\rho] - \beta\mathcal{F}_{ex}[\rho_0] = -\int dX \int dX' \int_0^1 d\lambda(1-\lambda) \times c^{(2)}(X,X';[\rho_\lambda])[\rho(X)-\rho_0][\rho(X')-\rho_0] \tag{59}$$

where $\rho_\lambda(X) = \rho_0 + \lambda(\rho(X) - \rho_0)$. Although formally exact, this expression requires the pair DCF of the inhomogeneous system; Baus et al. propose an *ansatz* for this object. The resulting theory is equivalent to the Onsager theory in the limit of high anisotropy and also possesses the $A/B \rightarrow B/A$ symmetry mentioned above. For less extreme anisotropies, the transition properties are in good agreement with the MC results.

More recently, computations have been carried out using liquid structure data obtained from the solution of nonlinear integral equations (see Section IV.C). As described previously, it seems that the HNC equation is most suited to the calculation of liquid structure of prolate hard cores. Perera, Patey, and Weis have applied solutions to the HNC equation to the RY theory of the isotropic–nematic transition of hard ellipsoids and hard spherocylinders [62]; they find that the transition densities are somewhat below those observed in MC simulations, but in much better agreement than those obtained using the Pynn–Wulf DCF.

Marko and Curtin have also carried out similar calculations [36] using the MWDA and RY theories and have obtained remarkably similar results from the two theories. The MWDA theory predicts slightly higher transition densities and slightly less sharp orientational distributions, but the differences are much less than those encountered in the comparison of MWDA and RY calculations of freezing [19]. This suggests that the higher-order DCF contributions are small enough to be ignored in the case of the isotropic–nematic transition.

In Table 1 are listed the properties of the coexisting isotropic and nematic phases of hard ellipsoids of various anisotropies. Results from a variety of calculations, including simulations, are listed. The orientational components of the nematic phase listed in the table are defined as $\langle P_l \rangle = \rho_N^{-1} \int dX\, \rho(X) P_l(\cos\theta)$, where P_l are Legendre polynomials, ρ_N is the density of the nematic, and θ is the angle between the nematic director and the symmetry axis of the ellipsoids.

Table 1 Properties of Coexisting Isotropic and Nematic Phases of Hard Ellipsoid Liquid[a]

A/B	ρ_I	ρ_N	$\beta P v$	$\beta\mu$	$\langle P_2 \rangle$	$\langle P_4 \rangle$	Calc.
1/3	0.952	0.972	17.47	24.03			MC [21]
	0.590	0.630			0.547	0.197	RY [40]
	0.901	0.925	14.82	22.3	0.561		DF [33]
	0.790	0.820			0.638	0.350	RY [62]
1/2.75	1.040	1.066	25.69	41.60			MC [21]
	0.628	0.663			0.532	0.186	RY [40]
	0.957	0.978	18.37	25.7	0.548		DF [33]
2.75	1.072	1.089	30.00	35.68			MC [21]
	0.628	0.663			0.532	0.186	RY [40]
	0.957	0.978	18.37	25.7	0.548		DF [33]
3	0.969	0.988	18.69	25.15			MC [21]
	0.590	0.630			0.547	0.197	RY [40]
	0.901	0.925	14.82	22.3	0.561		DF [33]
	0.693	0.721			0.657	0.358	RY [62]
	0.842	0.850	11.14	17.10	0.480	0.158	RY [36]
	0.843	0.875	11.21	17.18	0.451	0.139	MWDA [36]
5	0.613	0.654	5.31	12.2	0.64		DF [33]
	0.562	0.602			0.649	0.356	RY [62]
	0.599	0.615	4.94	10.5	0.500	0.172	RY [36]
	0.602	0.639	5.02	10.7	0.476	0.160	MWDA [36]

[a] The symmetry axis is of length $2A$, the other axes are of length $2B$, densities are in units of $(8AB^2)^{-1}$, and the molecular volume is $v = 4\pi AB^2/3$. Listed are the densities of the isotropic and nematic phases at coexistence, the pressure P, and the chemical potential μ. The $l = 2$ and $l = 4$ components of the orientational distribution are defined in the text. The results from Monte Carlo (MC) simulation, Ramakrishnan–Yusouff (RY) and MWDA density functional theories, and other density functional (DF) formulations are listed.

B. Crystallization of Weakly Anisotropic Particles

The freezing of anisotropic particles into orientationally ordered or disordered states with broken translational symmetry is a familiar phenomenon and can be studied using the density functional approach. The first such calculation was that of Singh and Singh [40], who studied the isotropic–plastic freezing transition for hard ellipsoids with aspect ratios up to $A/B = 2.25$. Unfortunately, there are errors in their calculations that qualitatively change the results.

This transition was correctly treated by Marko [60,38] and Smithline et al. [63]. Both groups used the RY theory and the Pynn–Wulf DCF with

the $D(X,X')$ due to Berne and Pechukas [50] known to be accurate for low anisotropies, and both used a translational distribution that was a sum of Gaussian distributions. The main difference between the two calculations seems to be that Smithline et al. used a one-body density constrained to have one molecule per lattice site, while Marko carried out an unrestricted minimization. In any case, the two sets of results are quite similar.

In opposition to the computations of Singh and Singh, it was found that the isotropic–face-centered-cubic (fcc) plastic transition density increased (in units of inverse molecular volume) as anisotropy A/B increased, in agreement with the MC result [21] (see Fig. 3 and Section II.C). However, it was also found by both groups that the coexistence density increased rapidly near $A/B = 1.1$. No stable plastic crystal states were found beyond this point.

Marko and Curtin have carried out MWDA calculations for the I–P ellipsoid transition and have found that as in the freezing of spherical particles [19,35], the MWDA (with the Pynn–Wulf DCF) leads to a doubling of the localizations (widths of the Gaussian distributions), but otherwise, the MWDA results are close to those obtained via the RY theory. In particular, an upturn in the transition density near $A/B = 1.1$ is seen in the MWDA calculations, suggesting that the inclusion of higher-order DCF contributions would not change this result. The same effect was seen in the freezing of hard dumbbells [63]; in this calculation, Smithline et al. used the approximate DCF due to Pynn [49].

Singh et al. [59] have carried out essentially the same calculations but with a DCF obtained by numerical solution of the PY-OZ equation. They report that the orientationally disordered fcc plastic phase of hard ellipsoids is stable from $A/B = 1$ to $A/B = 1.25$, which is the range in which this phase is seen in MC simulations [21]. They have studied hard dumbbells using the same approach; the plastic phase is observed for anisotropies L/D up to 1.26. They remark that "the location of coexistence parameters depends sensitively on the accuracy of the harmonics of the direct correlation function" [59]. This last calculation is in the best agreement with the simulation results of all of the density functional calculations, which is encouraging since it incorporates the most realistic liquid structure data. As anisotropy increases, the I–P freezing density in this calculation becomes rather large compared to the simulation results, but this is rather a minor defect compared to the lack of stability of the P phase predicted by the calculations that use the Pynn–Wulf DCF.

In none of these calculations has the case of orientational distributions with the (cubic) symmetry of the lattice been studied. It is possible that cubic anisotropy of the one-body orientational distribution may play an

important role in reducing the free energy of the plastic phase. This effect must become more important as A/B increases, and in particular in the vicinity of the I–P–S coexistence point. In Table 2 are listed properties of the coexisting liquid and plastic phases of hard ellipsoids obtained from the calculations discussed above.

Smithline et al. [63] have searched for orientationally aligned crystal phases of dumbbells and ellipsoids, but they have restricted their search to cubic and hexagonal close-packed lattices. The only study to date of an oriented tetragonal crystal was carried out by Marko [38], who found that for hard ellipsoids of aspect ratio $A/B > 1.05$, the tetragonal oriented solid had a lower free energy value than that of the fcc plastic. The Pynn–Wulf DCF was used in an RY calculation; a triple-point phase diagram was found with a weakly first-order plastic–solid phase transition at higher densities than the first-order liquid–plastic transition. Although the topology of the phase diagram is in agreement with simulation results, the triple point is seen in the MC at $A/B \approx 1.25$. The ratio of lattice constants (the distortion of the lattice away from cubic) was found to be nearly A/B, reflecting the fact that the solid is a close-packed lattice of ellipsoidal hard cores. Unfortunately, the minima found were boundary minima, as

Table 2 Properties of Coexisting Isotropic and Plastic Crystal Phases of Hard Ellipsoid Liquid[a]

A/B	ρ_I	ρ_P	βPv	$\beta\mu$	$\sigma/2B$	Calc.
1	0.94	1.04	11.7	16.6	0.11	MC [19]
	0.946	1.052	12.81	17.56	0.042	RY [36]
	0.877	1.019	9.263	13.02	0.095	MWDA [36]
1.050	0.961	1.057	13.77	18.57	0.044	RY [36]
	0.924	1.027	11.56	15.58	0.087	MWDA [36]
	0.949	1.052			0.0436	RY [59]
1.060	0.990	1.146			0.0477	RY [63]
1.095	1.050	1.168			0.0446	RY [63]
1.100	1.028	1.097	19.14	23.98	0.045	RY [36]
	1.041	1.076	20.40	24.55	0.064	MWDA [36]
	0.983	1.056			0.0498	RY [59]
1.250	0.983	1.039	14.34	18.44		MC [21]
	1.030	1.078			0.0585	RY [59]
1/1.250	0.998	1.060	15.51	19.54		MC [21]

[a] Symbols are as in Table 1. Listed are the densities of the isotropic and plastic phases at coexistence, the pressure P, the chemical potential μ, and the localization parameter σ. Results from Monte Carlo (MC) simulation and from Ramakrishnan–Yusouff (RY) and MWDA density functional theories are listed.

only the orientational $l = 2$ Legendre amplitude was used. It would be of interest to carry out RY and MWDA calculations of this system, using PY-OZ liquid structure data and using an orientational distribution capable of producing sharper peaks.

C. Smectic Liquid Crystals

The treatment of transitions to smectic phases has been a recent application of density functional theories. Most transitions to smectics are observed to occur from a strongly orientationally ordered nematic phase rather than from the disordered isotropic phase. Thus use of the RY theory is questionable (the isotropic phase is probably not metastable at *N–A* transitions, hence an isotropic phase DCF is probably not well defined). One might still use an RY theory if direct correlations of the nematic phase were employed instead of those of the isotropic liquid. To date, this approach has not been used.

Historically, the first simulations of hard cores that indicated stable smectic phases were done without orientational fluctuations; particles were constrained to be parallel [23]. This idealization not only facilitates simulations, but also simplifies the density functional theory. It should be noted that the partition functions of aligned hard cylinders of different aspect ratios L/D can be mapped onto one another by anisotropically scaling space; in terms of the packing fraction (or equivalently, the density in units of the close-packing density), the phase diagram is the same for all L/D. This property obviously does not hold for spherocylinders or for nonaligned cylinders.

For parallel hard particles, the disordered phase is a perfectly aligned nematic, and the direct correlation function, although anisotropic, is only a function of relative position. Mulder carried out a density functional calculation for aligned hard cylinders [64], using a density expansion to approximate the excess free energy. An interesting feature of his calculation is the factorization of the diagrams into $d = 2$ hard disk and $d = 1$ hard rod contributions. Mulder finds that the transition is of second order (mean-field critical behavior) and that the critical density and smectic wavelength tend toward values in agreement with MC simulations, as higher-order density contributions are added to the excess free energy. In this system, the maximum density is $\rho_{cp} = 2/(\sqrt{3}\, LD^2)$. The critical density was found to be $\rho^*/\rho_{cp} = 0.41$, and the smectic wavelength at the transition is $\lambda^*/L = 1.34$ (MC results of $\rho^*/\rho_{cp} = 0.39$ and $\lambda^*/L = 1.27$ are quoted, which are results similar to those obtained from aligned spherocylinders, which have a *N–A* transition at about $\rho/\rho_{cp} = 0.45$ to 0.5).

Aligned hard spherocylinders were also studied by Somoza and Tar-

azona [34]. They note the success of a method of Lee [65], who approximated the excess free energy of a nematic by rescaling the free energy of a hard-sphere system at the same packing fraction by the ratio of the second virial coefficients of the two systems. This approach obviously generates the correct second virial coefficient and appears to lead to an accurate expression for the free energy of the nematic. Somoza and Tarazona adopt a similar approach, but use as their reference not hard spheres but parallel hard ellipsoids. The geometry of the equivalent ellipsoids is obtained by matching the ratios of the moments of inertia of the ellipsoids and the hard particles under study. The inhomogeneous system excess free energy is then written as an integral of the free energy of parallel hard ellipsoids (known as well as that of hard spheres), scaled by a ratio of integrals of Mayer functions chosen to give the correct second virial coefficient in the homogeneous limit. As in weighted density calculations, the reference system free energy is evaluated at an effective density, set by a relation similar to (18). The authors choose to use an effective density that is only a function of position; the weight function w is set by scaling a weight function applied by the authors to hard-sphere freezing.

This approach was applied to the N–A transition of parallel hard spherocylinders, and a phase diagram in good agreement with computer simu-

Table 3 Nematic–Smectic A transition of Parallel Hard Spherocylinders[a]

L/D	ρ_{N-A}/σ_{cp}	βPv	$\beta\mu$	λ/L	Calc.
0.5	0.56	3.56	9.98	3.0	MC [23]
	0.54–0.60			3.0–3.2	DF [34]
1	0.50	2.62	7.29	2.2	MC [23]
	0.46–0.52			2.2	DF [34]
2	0.48	2.31	5.85	1.7	MC [23]
	0.41–0.48			1.7	DF [34]
3	0.47	2.22	5.27	1.5	MC [23]
	0.40–0.46			1.6	DF [34]
5	0.46	2.11	4.56	1.4	MC [23]
	0.40–0.46			1.5–1.6	DF [34]
∞ (cylinders)	0.39	1.50	5.42	1.3	MC [23]
	0.39–0.46			1.3–1.4	DF [34]
	0.41			1.34	DF [64]

[a] The cylinder length is L, the hemispherical caps are of diameter D, densities are in units of the close-packing density ρ_{cp}, and the molecular volumes is $v = \pi D^2(L/4 + D/6)$. Listed is the nematic–smectic transition density, the pressure P, the chemical potential μ, and the smectic wavelength λ. Results from Monte Carlo (MC) simulations and from density functional (DF) theories are listed. In the limit $L/D \to \infty$, parallel hard cylinders are obtained.

lations [23] was obtained. It was demonstrated that for a variety of different choices of the effective density, no qualitative change in the theory occurred. All of the results indicate an abrupt increase in the transition density at low aspect ratios below $L/D = 2$, and a critical density of $\rho^*/\rho_{cp} \approx 0.4$ to 0.5 for larger aspect ratios. The smectic wavelength is seen to be only weakly affected by different effective densities and is in good agreement with the MC results [23]. The properties of the N–A transition in hard cylinder systems is indicated in Table 3; density functional and MC results are listed.

In the same work, the phase diagram of parallel cylinders with oblique ends was also calculated. Smectic C and biaxial nematic phases as well as smectic A and the disordered nematic are shown in Fig. 9. For slightly oblique cylinders, the N–A transition is shifted to somewhat higher packing fractions, while for more oblique cylinders, a first-order transition from the nematic to smectic C is seen. For still more oblique cylinders, the transition to C is precluded by a second-order transition to a biaxial nematic phase. Unfortunately, no simulation results are available for this system.

Poniewierski and Holyst applied a weighted density theory to the problem of the I–N–A transitions in hard spherocylinders with orientational

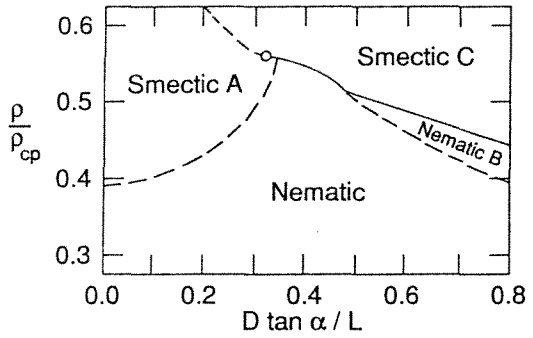


Fig. 9 Phase diagram for parallel oblique cylinders obtained by Somoza and Tarazona [34]. The density is measured in units of the close-packing density, and the skew is expressed as D tan α/L, for diameter D, length L, and skew angle α (see Fig. 1e). For no skew, the N–A transition is seen, but for large enough skew, there is a transition to a tilted smectic C. If the molecules are skewed enough, a biaxial nematic ("nematic B") appears between N and smectic C. Second-order transitions are indicated with dashed lines, first-order with solid lines. The circle indicates the tricritical point along the A–C transition line. (Adapted from Ref. 34.)

degrees of freedom [32]. The effective density use in the free energy per particle is obtained using a weighting similar to (18); the weight function w is chosen so that the virial expansion of the excess free energy is exact to second order in density; the Onsager limit is thus recovered from the theory, so that at low density and low anisotropy ($L/D \to 0$) the theory reduces to a WDA theory due to Tarazona for hard spheres. The authors comment, however, that at low L/D their theory may not be valid, as it predicts hard-sphere crystallization at too low a density. In this theory there is an I–N transition at low densities ($\rho/\rho_{cp} = 0.52$ at $L/D = 3$ dropping to $\rho/\rho_{cp} = 0.34$ by $L/D = 10$) of first order. At higher densities, a line of N–A transitions is reported, which are first order for $L/D < 3.3$ and are second order for $L/D > 3.3$. The N–A transition packing fraction rises from $\rho/\rho_{cp} = 0.57$ at $L/D = 3$ to $\rho/\rho_{cp} = 0.67$ near $L/D = 10$. The I–N and N–A coexistence regions meet at (presumably a triple point) near $L/D = 2.5$. The smectic wavelength is between $\lambda/L = 1.3$ and 1.42 over the range $2 < L/D < 10$.

Somoza and Tarazona have recently applied their approach to the hard sphereocylinder system with unconstrained orientational degrees of freedom [66]. In this calculation the decoupling $\rho(X) = \rho(\mathbf{r})\rho(\hat{\mathbf{e}})$ was used, and again the theory reduces to that of Lee [65] for the isotropic and nematic phases. The N–A line is found to be second order only for large aspect ratios ($L/D > 50$); the transition is first order for shorter molecules. The N–A and I–N coexistence regions meet near $L/D = 3$ at a I–N–A

Table 4 Nematic–Smectic A Transition of Hard Spherocylinders[a]

L/D	ρ_I/ρ_{cp}	ρ_N/ρ_{cp}	ρ_A/ρ_{cp}	ρ_S/ρ_{cp}	λ/L	$\langle P_2 \rangle$	Calc.
1	0.695			0.792			MC/MD [25]
3	0.574			0.688			MC/MD [25]
		0.55	0.55		1.40	0.69	DF [32]
		0.58	0.63		1.37		DF [66]
4		0.57	0.57		1.40	0.81	DF [32]
		0.57	0.61		1.31		DF [66]
5		0.58	0.58		1.34	0.90	MC/MD [24]
		0.58	0.58		1.40	0.88	DF [32]
		0.56	0.61		1.26	0.91	DF [66]

[a] Units are as in Table 3. Listed are the densities of the coexisting nematic and smectic phases (ρ_N and ρ_A). In the cases of simulations with $L/D = 1$, 3, and 4, the N–A transition is preempted by the isotropic–solid freezing transition; we list the coexisting isotropic and solid densities ρ_I and ρ_S for these cases. Also tabulated are the smectic wavelength λ and the quadrupole order pararmeter b_2. Results of Monte Carlo/molecular dynamics (MC/MD) simulations and density functional (DF) calculations are listed.

triple point. The N–A transition packing fraction drops from $\rho/\rho_{cp} = 0.61$ at $L/D = 3$ to about $\rho/\rho_{cp} = 0.45$ as $L/D \to \infty$. Again, the smectic wavelength is between $\lambda = 1.2$ to 1.4 over the range of anisotropies studied.

In Table 4 are summarized these results for the N–A transition in the unconstrained spherocylinder system, along with relevant simulation results. The orientational component listed is defined as in Table 1.

It seems that the behavior of parallel prolate hard particles can be reasonably well described by the density functional approaches described above. The agreement of different calculations with simulation results is quite encouraging, and the mechanisms suggested to stabilize smectic C phases are interesting. In contrast, the question of the structure of the phase diagram of the unaligned spherocylinder system is unsettled. The order of the N–A transitions and the location of the N–A tricritical point has not been pinned down by simulations, and current density functional calculations are not in good agreement with one another. The role of the solid phase expected at high densities has not been investigated via the density functional approach; the simulations suggest that there may be two triple points (I–S–A and I–N–A) in close proximity to one another.

VI. CONCLUSION

In this chapter we have focused on the problem of understanding the macroscopic properties of liquid crystals based on consideration of the interactions between the anisotropic constituent particles. We have discussed only hard-core particles for two reasons: first, the theory is simplified for hard-core interactions, and second, the bulk of simulation data are for hard-core systems. The thermodynamics and phase transitions in these systems are quite similar to those of real liquid crystals.

In Section III it was shown how one can construct a theory for phase transitions in such systems. The microscopic information required is static liquid structure: in particular, the direct correlation functions. The most straightforward approach is to use them to construct a functional perturbation theory for ordered phases; unfortunately, the fact that we have little knowledge even about three-point direct correlations forces us to ignore contributions of other than the pair direct correlations (RY theory). However, we have seen that this crude approximation leads to results in quantitative agreement with simulation results.

The fact that the direct correlation functions are successive functional derivatives of the excess Helmholtz potential suggests that one should construct nonperturbative free-energy functionals that contain contributions at all orders of perturbation theory; it is possible also to satisfy all sum rules derived from the two-point direct correlations. Density func-

tional approaches along these lines have been successful in treating freezing of spherical particles, and weighted-density approaches have also been used to study liquid–crystal phase transitions in hard-particle systems (Section V.C).

It is apparent, however, that in addition to improving the free-energy functional through methods such as weighted-density approximations, the use of accurate liquid structure information is of crucial importance in obtaining accurate predictions. This is particularly apparent in the cases of the isotropic–nematic transition of prolate cores (Section V.A) and in the isotropic–plastic transition in hard ellipsoids (Section V.B). For the I–P transition, Singh et al. [59] showed that the lack of stability of the plastic phase observed by other authors [60,63] for $A/B > 1.1$ was due to the use of inaccurate liquid structure rather than being due to higher-order DCF contributions or inadequate parametrizations of the crystal. Similar improvements have been realized in the description of the I–N transition using liquid structure obtained from solution of integral equations for liquid structure. One hopes that the rapid progress that we have seen in the last few years in simulation and solution of integral equations for liquid structure will continue, as these two sources of information are crucial to constructing and evaluating the theory.

As discussed in Section V, density functional approaches have been used to study a variety of phase transitions in hard-core systems. In addition to freezing transitions, transitions to nematic and smectic liquid crystals have been studied, and in the case of elongated particles constrained to be parallel, these theories predict transition properties in good agreement with simulation results. The application of this approach to particles not constrained to be parallel is a topic of current research, but preliminary results tell us that we can expect accurate predictions in this case as well.

There are a variety of directions in which the theory can be developed. Perhaps the most obvious is the inclusion of finite-energy interactions in models. Some work has been done by Singh and Singh [67] on particles with long-range dispersion interactions treated as a perturbation on the hard-core interactions. The structure of the theory presented in Section III does not change; the interactions are of course taken into account by the direct-correlation functions (which for finite-energy interactions, depend on temperature). The accuracy of the theory will hinge, as before, on the accuracy of the liquid structure used.

Another direction of interest is that of determining elastic properties of ordered phases. In finding a particular ordered phase we have found a minimum of the free-energy functional. One might ask what the quadratic free-energy response is to long-wavelength distortions of such an

ordered phase: This can be summarized with a set of elastic constants that describe the free-energy cost of distortions of various symmetries [5,6]. Some effort has been made to develop a formalism to study this problem based on the RY theory [67,68]. There exist simulation results for the elastic constants for hard ellipsoid and hard spherocylinder nematics due to Allen and Frenkel [69] that do not agree well with existing density functional predictions. No calculations of elastic constants using WDA techniques or liquid structure obtained from integral equation methods have been done to date.

Calculation of the quadratic dispersion of the free energy around a particular ordered state would also allow the computation of corrections to the free energy due to long-wavelength fluctuations. Such corrections will typically increase the free energy of the ordered phase and will thus move isotropic–nematic or isotropic–solid transitions to somewhat higher densities or lower temperatures. In the case of transitions to a smectic (where the free-energy corrections due to elastic fluctuations diverge [7]), it would be useful to use density functional techniques to study the energy of defects that may be fundamental to understanding the nature of the smectic phase [70]. Finally, we note that techniques for handling inhomogeneous systems are useful for studying problems other than bulk properties of condensed phases, such as the behavior of smectics near substrates or free surfaces [71].

ACKNOWLEDGMENTS

I thank Nihat Berker, Bill Curtin, Alan Denton, and Mehran Kardar for many rewarding discussions and insightful suggestions regarding work done at MIT that is discussed in this paper. Research done at MIT was supported by the National Science Foundation through Grants DMR-84-18718 and DMR-87-19217, by the Joint Services Electronics Program through Contracts DAAL 03-83-K0002 and DAAL 03-89-C0001, and by the Natural Science and Engineering Research Council of Canada through a Postgraduate Scholarship. Work done at the University of Chicago was partially supported by the NSF through Grant DMR-88-19860, and by NSERC (Canada) through a Postdoctoral Fellowship.

REFERENCES

1. Frenkel, D., and McTague, J. P. (1980). Computer simulations of freezing and supercooled liquids, *Annu. Rev. Phys. Chem. 31*: 491.
2. Ramakrishnan, T. V., and Yussouff, M. (1981). First-principles order parameter theory of freezing, *Phys. Rev. B 23*: 5871. See also Haymet, A. D.

J. (1987). Theory of the equilibrium liquid–solid transition, *Annu. Rev. Phys. Chem. 38*: 89.
3. Alder, B. J., and Wainwright, T. E. (1959). Studies in molecular dynamics. I. General method, *J. Chem. Phys. 31*: 459.
4. Ashcroft, N. W., and Mermin, N. D. (1976). *Solid State Physics*, W. B. Saunders, Philadelphia, Chapters 4 to 7.
5. Pershan, P. S. (1988). *Structure of Liquid Crystal Phases*, World Scientific, N.J.
6. de Gennes, P.-G. (1974). *The Physics of Liquid Crystals*, Oxford University Press, New York.
7. Chandrasekhar, S. (1977). *Liquid Crystals*, Cambridge University Press, Cambridge.
8. Onsager, L. (1949). The effects of shape on the interaction of colloidal particles, *Ann. N.Y. Acad. Sci. 51*: 627.
9. Frenkel, D., Lekkerkerker, H. N. W., and Stroobants, A. (1988). Thermodynamic stability of a smectic phase in a system of hard rods, *Nature 332*: 822.
10. Barker, J. A., and Henderson, D. (1976). What is "liquid"? Understanding the states of matter, *Rev. Mod. Phys. 48*: 587.
11. Lebowitz, J. L., and Perram, J. W. (1983). Correlation functions for nematic liquid crystals, *Mol. Phys. 50*: 1207.
12. Wen, X., and Meyer, R. B. (1987). Model for smectic-*A* ordering of parallel hard rods, *Phys. Rev. Lett. 59*: 1325.
13. Landau, L. D., and Lifshitz, E. M. (1978). *Statistical Physics*, Pergamon Press, Elmsford, N.Y., p. 478.
14. Tonks, L. (1936). The complete equation of state of one, two and three dimensional gases of hard elastic spheres, *Phys. Rev. 50*: 955.
15. Wertheim, M. S. (1964). Further comments on the preceding paper, in *The Equilibrium Theory of Classical Fluids* (H. L. Frisch and J. L. Lebowitz, eds.), W. A. Benjamin, New York, p. II-276.
16. Marko, J. F. (1989). Exact pair correlations in a one-dimensional fluid of hard cores with orientational and translational degrees of freedom, *Phys. Rev. Lett. 62*: 543.
17. Casey, L. M., and Runnels, L. K. (1969). Models for correlated molecular rotation, *J. Chem. Phys. 51*: 5070.
18. Lebowitz, J. L., Percus, J. K., and Talbot, J. (1987). On the orientational properties of some one-dimensional model systems, *J. Stat. Phys. 49*: 1221.
19. Curtin, W. A., and Ashcroft, N. W. (1985). Weighted-density-functional theory of inhomogenous liquids and the freezing transition, *Phys. Rev. A 32*: 2909. Curtin, W. A., and Ashcroft, N. W. (1986). *Phys. Rev. Lett. 56*: 2775.
20. Young, D. A., and Alder, B. J. (1974). Studies in molecular dynamics. XIII. Singlet and pair distribution functions for hard-disk and hard-sphere solids, *J. Chem. Phys. 60*: 1254.
21. Frenkel, D., Mulder, B. M., and McTague, J. P. (1984). Phase diagram of a system of hard ellipsoids, *Phys. Rev. Lett. 52*: 287. Frenkel, D., and

Mulder, B. M. (1985). The hard ellipsoid-of-revolution fluid. I. Monte Carlo simulations, *Mol. Phys. 55*: 1171.
22. Frenkel, D., and Eppenga, R. (1982). Monte Carlo study of the isotropic–nematic transition in a fluid of thin hard disks, *Phys. Rev. Lett. 49*: 1089.
23. Stroobants, A., Lekkerkerker, H. N. W., and Frenkel, D. (1986). Evidence for smectic order in a fluid of hard parallel spherocylinders, *Phys. Rev. Lett. 57*: 1452. Stroobants, A., Lekkerkerker, H. N. W., and Frenkel, D. (1987). Evidence for one-, two-, and three-dimensional order in a system of hard parallel spherocylinders, *Phys. Rev. A 36*: 2929. Frenkel, D. (1987). Onsager's spherocylinders revisited, *J. Phys. Chem. 91*: 4912.
24. Frenkel, D. (1988). Structure of hard-core models for liquid crystals, *J. Phys. Chem. 92*: 3280.
25. Veerman, J. A. C., and Frenkel, D. (1990). Phase diagram of a system of hard spherocylinders by computer simulation, *Phys. Rev. A 41*: 3237.
26. Allen, M. P., and Frenkel, D. (1988). Calculation of liquid-crystal Frank constants by computer simulation, *Phys. Rev. A 37*: 1813.
27. Frenkel, D., and Eppenga, R. (1985). Evidence for algebraic orientational order in a two-dimensional hard-core nematic, *Phys. Rev. A 31*: 1776.
28. Mermin, D. (1965). Thermal properties of the inhomogenous electron gas, *Phys. Rev. 137*: A1441.
29. Curtin, W. A. (1988). Freezing in the density-functional approach: effect of third-order contributions, *J. Chem. Phys. 88*: 7050.
30. Tarazona, P. (1985). Free-energy density functional for hard spheres, *Phys. Rev. A 31*: 2672.
31. Curtin, W. A., and Runge, K. (1987). Weighted-density-functional and simulation studies of the bcc hard-sphere solid, *Phys. Rev. A 35*: 4755.
32. Poniewierski, A., and Holyst, R. (1988). Density-functional theory for nematic and smectic-A ordering of hard spherocylinders, *Phys. Rev. Lett. 61*: 2461.
33. Baus, M., Colot, J.-L., Wu, X.-G., and Xu, H. (1987). Finite-density Onsager-type theory for the isotropic–nematic transition of hard ellipsoids, *Phys. Rev. Lett. 59*: 2184. Colot, J.-L., Wu, X.-G., Xu, H., and Baus, M. Density-functional, Landau, and Onsager theories of the isotropic–nematic transition of hard ellipsoids, *Phys. Rev. A 38*: 2022.
34. Somoza, A. M., and Tarazona, P. (1988). Nematic-smectic-A-smectic-C transitions in systems of parallel hard molecules, *Phys. Rev. Lett. 61*: 2566. Somoza, A. M., and Tarazona, P. (1989). Density functional approximation for hard-body liquid crystals, *J. Chem. Phys. 91*: 517.
35. Denton, A. R., and Ashcroft, N. W. (1989). Modified weighted-density-functional theory of nonuniform classical liquids, *Phys. Rev. A 39*: 4701.
36. Marko, J. F., and Curtin, W. A. (1990). Weighted-density theory of phase transitions in fluids composed of anisotropic particles, in *Macromolecular Liquids* (C. R. Safinya, S. A. Safran, and P. A. Pincus, eds.), Materials Research Society, Pittsburgh (*Mater. Res. Soc. Proc. 177*: 329).
37. Gray, C. G., and Gubbins, K. E. (1984). *Theory of Molecular Fluids*, Oxford University Press, New York, Vol. 1, Appendix A.

38. Marko, J. F. (1989). First-order phase transitions in the hard-ellipsoid fluid from variationally optimized direct pair correlations, *Phys. Rev. A 39*: 2050.
39. Sluckin, T. J., and Shukla, P. (1983). Molecular field theory of nematics: density functional approach. I. Bulk effects, *J. Phys. A 16*: 1539.
40. Singh, U. P., and Singh, Y. (1986). Molecular theory for freezing of a system of hard ellipsoids: properties of isotropic–plastic and isotropic–nematic transitions, *Phys. Rev. A 33*: 2725.
41. Lipkin, M. D., and Oxtoby, D. W. (1983). A systematic density functional approach to the mean field theory of smectics, *J. Chem. Phys. 79*: 1939.
42. Perera, A., Kusalik, P. G., and Patey, G. N. (1987). The solution of the hypernetted chain and Percus–Yevick approximations for fluids of hard nonspherical particles: results for hard ellipsoids of revolution, *J. Chem. Phys. 87*: 1295. Perera, A., Kusalik, P. G., and Patey, G. N. (1988). Erratum, *J. Chem. Phys. 89*: 5969.
43. Singh, Y. (1984). Molecular theory of liquid crystals: application to the nematic phase, *Phys. Rev. A 30*: 583.
44. Rice, S. A., and Gray, P. (1965). *The Statistical Mechanics of Simple Liquids*, Wiley, New York, Chapter 2.
45. Wertheim, M. S. (1963). Exact solution of the Percus–Yevick integral equation for hard spheres, *Phys. Rev. Lett. 10*: 321.
46. Haymet, A. D. J. (1987). Theory of the equilibrium liquid–solid transition, *Annu. Rev. Phys. Chem. 38*: 89.
47. Jones, G. L., and Mohanty, U. (1985). A density functional-variational treatment of the hard sphere transition, *Mol. Phys. 54*: 1241.
48. Morita, T., and Hiroke, K. (1960). A new approach to the theory of classical fluids, I, *Prog. Theor. Phys. 23*: 1003.
49. Pynn, R. (1974). Theory of static correlations in a fluid of linear molecules, *Solid State Commun. 14*: 29. Wulf, A. (1977). *J. Chem. Phys. 67*: 2254.
50. Berne, B. J., and Pechukas, P. (1972). Gaussian model potentials for molecular interactions, *J. Chem. Phys. 56*: 4213.
51. Rosenfeld, Y. (1985). High-density properties of liquid-state theories: physically intuitive meaning for the direct correlation functions, *Phys. Rev. A 32*: 1834.
52. Rosenfeld, Y. (1988). Scaled field particle theory of the structure and the thermodynamics of isotropic hard particle fluids, *J. Chem. Phys. 89*: 4272.
53. Chen, Y.-D., and Steele, W. A. (1971). Statistical mechanics of linear molecules. VI. Solutions of the Percus–Yevick integral equation for a hard-core model, *J. Chem. Phys. 54*: 703.
54. Lado, F. (1982). Integral equations for fluids of linear molecules. I. General formulation, *Mol. Phys. 47*: 283. Lado, F. (1982). Integral equations for fluids of linear molecules. II. Hard dumbbell solutions, *Mol. Phys. 47*: 299. Lado, F. (1984). Integral equations for fluids of linear molecules. III. Orientational ordering, *Mol. Phys. 47*: 313.
55. Lado, F. (1985). Test of a simple analytic model for fluids of hard linear molecules, *Mol. Phys. 54*: 407.

56. Perera, A., and Patey, G. N. (1988). The solution of the hypernetted-chain and Percus–Yevick approximations for fluids of hard spherocylinders, *J. Chem. Phys. 89*: 5861.
57. Caillol, J. M., and Weis, J. J. (1989). Integral equation study of parallel hard spherocylinders, *J. Chem. Phys. 90*: 7403.
58. Lago, S., and Sevilla, P. (1988). Solution of the Percus–Yevick equation for hard spherocylinders. I. The entire pair correlation function, *J. Chem. Phys. 89*: 4349.
59. Singh, U. P., Mohanty, U., and Singh, Y. (1988). Molecular theory for freezing transition of hard ellipsoid and hard dumbbell molecules, *Phys. Rev. A 38*: 4377.
60. Marko, J. F. (1988). Accurate calculation of isotropic–plastic and isotropic–nematic transitions in the isotropic fluid, *Phys. Rev. Lett. 60*: 325.
61. Mahato, M. C., Lakshmi, M. R., Pandit, R., and Krishnamurthy, H. R. (1988). Liquid–mesophase–solid phase transitions: systematics of a density-wave theory, *Phys. Rev. A 38*: 1049.
62. Perera, A., Patey, G. N., and Weis, J. J. (1988). Density-functional theory applied to the isotropic–nematic transition in model liquid crystals, *J. Chem. Phys. 89*: 6941.
63. Smithline, S. J., Rick, S. W., and Haymet, A. D. J. (1988). Density functional theory of freezing for molecular liquids, *J. Chem. Phys. 88*: 2004.
64. Mulder, B. (1987). Density-functional approach to smectic order in an aligned hard-rod fluid, *Phys. Rev. A 35*: 3095.
65. Lee, S.-D. (1987). *J. Chem. Phys. 87*: 4972.
66. Somoza, A. M., and Tarazona, P. (1990). Nematic and smectic liquid crystals of hard spherocylinders, *Phys. Rev. A 41*: 965.
67. Singh, Y., and Singh, K. (1986). Density-functional theory of curvature elasticity in nematic liquids, I, *Phys. Rev. A 33*: 3481. Singh, K., and Singh, Y. (1986). Density-functional theory of curvature elasticity in nematic liquids, II, Effect of long-range dispersion interactions, *Phys. Rev. A 34*: 548.
68. Lipkin, M. D., Rice, S. A., and Mohanty, U. (1985). The elastic constants of condensed matter: a direct-correlation function approach, *J. Chem. Phys. 82*: 472.
69. Allen, M. P., and Frenkel, D. (1988). Calculation of liquid-crystal Frank constants by computer simulation, *Phys. Rev. A 37*: 1813.
70. Nelson, D. R., and Toner, J. (1981). Bond-orientational order, dislocation loops, and melting of solids and smectic-A liquid crystals, *Phys. Rev. B 24*: 363.
71. Selinger, J. V., and Nelson, D. R. (1988). Density funcitonal theory of nematic and smectic-A order near surfaces, *Phys. Rev. A 37*: 1736.

12

Nature of Microemulsion

K. A. Dawson and Christopher J. Mundy
University of California
Berkeley, California

I. INTRODUCTION

Upon mixing oil, water, and a little surfactant* one often finds a macroscopic thermodynamically stable phase termed *microemulsion* that possess some quite remarkable properties. Two recent reviews of microemulsion contain most of the earlier references and discuss most of their basic properties as described in this section [1] (see also Refs. 2 to 4). Microemulsion has found application in a number of developing technologies, such as enhanced oil recovery and cleanup, catalytic systems, and synthetic blood substitutes. Some of the prominent experimental observations associated with these fluids are as follows.

As a function of the concentration of amphiphile, temperature, and brine one finds a phase equilibrium between microemulsion, oil-in-water, and water-in-oil micellar phases. In some cases the middle microemulsion phase in this equilibrium is replaced by a lamellar liquid–crystalline phase. As one proceeds to the oil-rich or water-rich regions of the phase diagram, the three-phase equilibrium collapses to an oil–microemulsion or water–

* It is often necessary to add a little cosurfactant or brine to obtain microemulsion. There do exist some systems, notably oil–water–C_nE_m mixtures, that are genuinely three-component. We may also note that for the system we discuss, the term *amphiphile* rather than *surfactant* is favored by many. This convention has been chosen here.

microemulsion two-phase equilibrium via critical endpoints of, respectively, the water and microemulsion or oil and microemulsion. This three-phase coexistence, often named the Winsor III [5] state, is of particular interest both from the viewpoint of basic statistical mechanics† and for its technological importance [6].

Thus in this region of the phase diagram the middle microemulsion phase is often opalescent and its interfacial tensions with the coexisting phases are ultralow,‡ leading many to suppose that the liquid is near-critical. However, the volume fractions of the various components in the three phases may differ significantly (differences as high as 80% w/v have been observed) and, characteristically, the microemulsion does not wet the oil–water interface. That the interface is typically nonwet is indicated by, for example, Kunieda and Shinoda [7]. Both of these observations seem to indicate that the properties of the middle phase may not be ascribed solely to a proximate critical point. Furthermore, the interfacial tension data in the three-phase region exhibit a rather characteristic and interesting pattern between the two critical endpoints [8].§ The tension between a pair of incipient critical phases naturally decreases monotonically, while that between the other pair rises monotonically. On the other hand, the oil–water tension actually exhibits a minimum between the two critical endpoints, a phenomenon that has long been used as the criterion of an optimal microemulsion. In that region of three-phase equilibrium where the oil–water tension is a minimum, the microemulsion contains almost equal volume fractions of oil and water and a fairly small amount (between 10 and 20 vol %) of amphiphile. It is bicontinuous in the sense that it conducts electricity and the apolar molecular diffusivity is high [9].

A number of scattering studies of this bicontinuous phase have been undertaken. These include small-angle neutron scattering (SANS) [10–

† The origin of this three-phase equilibrium has long been a matter of debate. In this chapter we offer the opinion that it stems primarily from a remarkable confluence of the various energy scales. This reasoning is to be distinguished from the notion that it is merely an extension of the conventional symmetric three-phase equilibrium of a three (or more)-component mixture.

‡ The true origins of ultralow interfacial tension have also long been a matter of debate. Some have considered the phenomenon to arise as a consequence of near-criticality of the mixture. We shall see that the present model predicts the phenomenon to arise as a consequence of two types on cancellation. Thus the surface pressure of the amphiphilic film almost compensates the bare interaction energy of a flat interface between oil and water. Furthermore, a delicate balance of the chemical potentials of the components means that it is favorable to produce large amounts of free interface.

§ A path between the two critical endpoints may require variation of both temperature and the amount of brine.

12], light scattering [13], and freeze-fracture electron microscopy [14]. In particular, the SANS studies produce a peak in the scattering intensity at small wave numbers, an observation that is consistent with the presence of two relatively long length scales in the microemulsion. The ratio between these lengths appears to be nearly constant in the bicontinuous region* and is approximately equal to 4 [15]. It is also observed that with increasing amphiphile concentration, the peak in the SANS intensity shifts to shorter wavelengths.

Finally, we note that the phase diagram also possesses a one-phase region spanning the oil-in-water, bicontinuous, and water-in-oil phases. It is believed that, typically, this single-phase corridor permits one to accomplish phase inversion without undergoing a phase transition. These, then, are some of the prominent aspects of the phase diagrams of mixtures of amphiphile, oil, and water where one has relatively small amounts of amphiphile. We note in passing that at higher amphiphile concentrations, one finds various liquid–crystalline phases.

The purpose of the present chapter is to show that all of these features are present in a relatively simple lattice model. Within this model it is also possible to explain the origins of these phenomena and to see that the puzzling experimental observations are, indeed, entirely consistent with one another.

A number of theoretical studies of microemulsion have been undertaken, and some of them are now providing quite promising results [16–26]. Before beginning such a description it is worth pointing out that there are few truly universal features in these systems. Rather, they exhibit interesting and important generic properties that one seeks to understand in a unified and consistent manner. Many of the important observations originate in long but not diverging length scales. For this reason it is convenient to construct a model that is computationally tractable, thereby facilitating the study of these various trends, but that is reducible to a more compact-coarse grained Hamiltonian whenever one seeks to study universal features via renormalization-group calculations.

There are some further considerations to be taken into account when formulating the model. Thus, if one is to understand a range of amphiphilic behavior, there should be sufficient flexibility to represent the effect of geometric fluctuations of self-assembled structures, as well as fluctuations causing the breakup of the phase. This indicates that the simplest possible formulation may be that of a lattice model.

* This observation is a reflection of the conditions (referred to in an earlier footnote) that ensure the formation of much amphiphilic film.

Finally, one should also be aware that although mean-field studies of the present lattice model have been shown to elucidate the qualitative features of the lattice Hamiltonian, fluctuations certainly produce significant quantitative changes in the predictions and, in some cases, it has been shown that the method gives qualitatively incorrect results. The present chapter is based largely on mean-field analyses and simulations, but, where possible, these have been checked using other methods. The main purpose of the present chapter is to present the lattice model as the basis for a qualitative description of microemulsion. Thus many technical details are suppressed and references are given to other studies where these may be found.

II. ISING HAMILTONIAN

We begin by systematically constructing the Ising spin model, starting from a few simple hypotheses about the microscopic molecular interactions. The Hamiltonian that we shall present is closely related to that of Widom [16,17], a model that had itself grown from earlier research by Wheeler and Widom [16]. It differs only in the generalization of some of the interactions. As in Widom's original treatment, we divide configuration space into cubes of side a and this, the microscopic distance in the formulation, is chosen to be the length of an amphiphile molecule. Arbitrary configurations of the oil (AA), water (BB), and amphiphile molecules (AB) are then assigned to this lattice, subject to the constraint that only like ends of different molecules be permitted to lie within any cube. Thus each cube of configuration space is composed only of hydrophilic or hydrophobic material, and any molecular configuration may therefore be represented by values of Ising variables at their centers. Evidently there exist no isolated amphiphiles in this model, so to each local configuration of amphiphile molecules we may assign an additional energy that is chosen with respect to a pair of parallel amphiphile molecules that lie side by side. Ultimately, this leads to an unfavorable energy both for bending of an isolated amphiphilic film and for the touching of two such layers. However, these terms, being based only on pairs of amphiphiles, necessarily treat edges and corners of the amphiphilic film on equal terms in that a corner energy is calculated as the sum of edge energies. Since we conceive the microscopic origins of this energy scale to lie in the partial free energies of proximate amphiphile molecules that are surrounded by oil, water, and possibly cosurfactant, one expects the interaction Hamiltonian to possess many-body terms. In particular, the bending of an amphiphilic film certainly has a potential energy contribution, but there

is also penetration of water (or oil) into the film that must be accounted for by the microscopic interaction parameters. It is therefore evident that such contributions would be different for edges and corners of the amphiphilic film, and thus, in principle, require independent energy parameters. We therefore choose as our basic interactions the parameters V_1–V_5, these corresponding to the energies of the local configurations that are given in Fig. 1.

Note that the essential content of the foregoing arguments is that small regions or aggregates of amphiphile have only three basic significant energy scales: the energies due to bends of an isolated film, these being distinct for edges and corners, and the energy of two proximate flat pieces of the film. For generality we must also differentiate between the ends of the amphiphiles, and this is reflected in the fact that parameters V_1 and V_3 are different from V_2 and V_4. Note that, in principal, two independent three-body interactions for the amphiphiles at a corner have been assigned the same energy. Thus the configuration where the hydrophilic heads form a corner has the same energy (V_5) as the configuration shown in Fig. 1. Had we chosen to resolve between these local configurations, the form of the Hamiltonian would remain unchanged, although the four-body spin coupling parameter is modified. The essential features of the mean-field

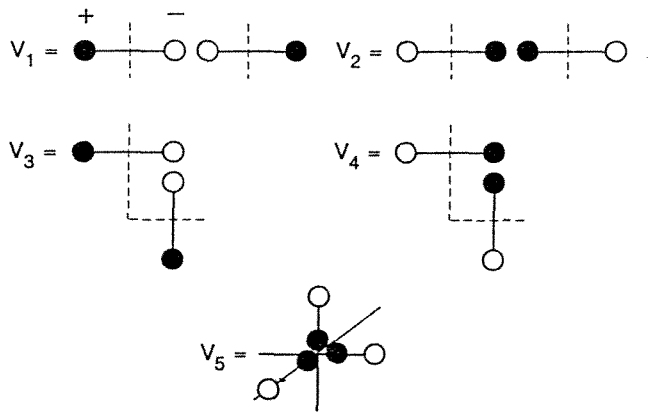


Fig. 1 Molecular configurations of amphiphile that correspond to the interaction energies V_1–V_5. These energies are measured relative to that of a pair of parallel amphiphiles. The amphiphiles (AB) are represented by a hydrophobic (●) and a hydrophilic (○) region and are considered to lie along the bonds of a simple cubic lattice. The midpoints of every bond occupied by an amphiphile molecule collectively represent the amphiphilic film.

predictions are also unchanged. However, for the moment we consider the spontaneous three-body curvature effect to be smaller than the two-body spontaneous curvature. The retention of the three-body curvature energy causes the ordinarily second-order micellar–bicontinuous phase transition to become first order. The retention of the three-body spontaneous curvature introduces no qualitatively new feature.

The centers of all cubes that compose configuration space define a regular cubic lattice. Since the interactions of Fig. 1 require knowledge only of a central and six surrounding cubes, one may construct the total energy from octahedral clusters of Ising spins. Each octahedron involves seven Ising variables, and we may represent the local energy as a function of these Ising spins. The energy of each of the 20 such octahedral fragments (see Fig. 2) may be compiled from the rules in Fig. 1. We then require that the chosen function of the Ising variables correctly reproduce the energies of all 20 local-spin configurations given in Fig. 2.

It transpires that the local energy contribution requires only one- to four-body spin interactions and the Hamiltonian may be written

$$\mathcal{H} = H \sum_{n} \sigma_n + \sum_{n,n'} J_{n,n'} \sigma_n \sigma_{n'} + \sum_{n,n',n''} L_{n,n',n''} \sigma_n \sigma_{n'} \sigma_{n''} + \sum_{n,n',n'',n'''} P_{n,n',n'',n'''} \sigma_n \sigma_{n'} \sigma_{n''} \sigma_{n'''} \quad (1)$$

The spin-interaction coupling constants refer to the interactions for isolated spins as well as those between pairs, triples, and quadruples of spins. The two-body spin interactions $\{J_{nn'}\}$ are nearest-neighbor (J), diagonal neighbor (M_1) and linear next-nearest neighbor (M_2), and their definitions in terms of the elementary energies V_1–V_5 are given in Table 1. Similarly, the three-body interactions $\{L_{nn'n''}\}$ are of two types reflecting the coupling between bent triples (L_1) and linear triples (L_2) of connected Ising spins. Finally, the four-body spin interactions $\{P_{nn'n''n'''}\}$ are of one type only (P), with four connected spins forming the corner drawn in Fig. 5. All of these coupling constants are given in terms of the basic interaction energies in Table 1.

Previous studies of this model [16,25] have been confined to the choice of parameters $M_1 = M$, $M_2 = 2M$, $L_1 = L_2 = P = 0$, and more detailed information is available only for the cut $H = 0$. If one makes the choice $H = L_1 = L_2 = 0$, corresponding to zero spontaneous curvature of the amphiphile film and equality of the chemical potentials of oil and water, one would expect to find the bicontinuous microemulsion to be one of

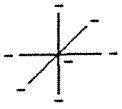
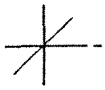
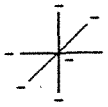
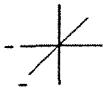
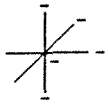
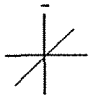
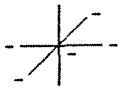
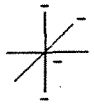
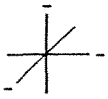
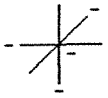
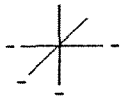
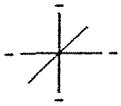
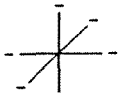
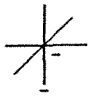
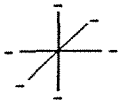
Cluster	Symbol	Local Energy	Cluster	Symbol	Local Energy
	C_0^0			C_0^{0*}	
	C_1^0			C_1^{0*}	
	C_2^0	V_4		C_2^{0*}	V_3
	C_2^1	V_2		C_2^{1*}	V_1
	C_3^1	$V_2 + 2V_4$		C_3^{1*}	$V_1 + 2V_3$
	C_3^0	$3V_4 + V_5$		C_3^{0*}	$3V_4 + V_5$
	C_4^1	$5V_4 + 2V_5 + V_2$		C_4^{1*}	$5V_3 + V_1 + 2V_5$
	C_4^2	$4V_4 + 2V_2$		C_4^{2*}	$4V_3 + 2V_1$
	C_5^2	$2V_2 + 4V_5 + 8V_4$		C_5^{2*}	$2V_1 + 4V_5 + 8V_3$
	C_6^2	$3V_2 + 12V_4 + 8V_5$		C_6^{2*}	$3V_1 + 12V_3 + 8V_5$

Fig. 2 Energy of every local-octahedral Ising spin configuration. These values may be calculated using the definitions in Fig. 1.

Table 1 Transcription from Ising Variables to Interaction Energies and Chemical Potentials in the Solution[a]

Ising variables	Microemulsion model
H	$3/2(\mu_{BB} - \mu_{AA}) + 9/8V_1 - 9/8V_2 + 3/2V_3 - 3/2V_4$
J	$1/4V_1 + 1/4V_2 - 1/2V_3 - 1/2V_4 - 1/4[\mu_{AB} - (\mu_{BB} + \mu_{AA})/2] + 1/2V_5$
M_1	$-1/8(V_3 + V_4) + 1/4V_5$
M_2	$-1/4(V_1 + V_2)$
L_1	$-1/8(V_1 - V_2)$
L_2	$-1/8(V_3 - V_4)$
P	$1/8V_5$

[a] The molecular configurations corresponding to the interaction energies V_1–V_5 are given in Fig. 1.

the prominent phases. The Hamiltonian then becomes

$$\mathcal{H} = \frac{1}{2}\sum_n \sigma_n O_n \sigma_n + P \sum_{n,n',n'',n'''}^{4B} \sigma_n\sigma_{n'}\sigma_{n''}\sigma_{n'''} \tag{2}$$

$$O_n = a_4\Delta_n^4 + a_2\Delta_n^2 + a_0 \tag{3}$$

$$\Delta_x^2 f_x = f_{x-1} + f_{x+1} - 2f_x \tag{4}$$

$$\Delta_x^2 + \Delta_y^2 + \Delta_z^2 = \Delta_n^2 \tag{5}$$

$$a_4 = -M \tag{6a}$$

$$a_2 = -(J + 12M) \tag{6b}$$

$$a_0 = -6(J + 5M) \tag{6c}$$

where, to make connection with earlier research, we set $M_1 = 2M$ ($M_2 = M$) and thus

$$J = \tfrac{1}{2}[\mu_{AB} \tfrac{1}{2}(\mu_{AA} + \mu_{BB})] - \tfrac{5}{2}K \tag{7}$$

$$M = -\frac{K}{4}$$

where K is the bending energy of the amphiphilic film. For this set of restrictions one finds that the present model is equivalent to that of Widom, except for the presence of a four-body term that reflects the selection of an independent corner energy in Fig. 1.

Finally, if one generalizes the definition of the linear operator to

$$O_n = M_2(\Delta_x^4 + \Delta_y^4 + \Delta_z^4) - M_2(\Delta_x^2\Delta_y^2 + \Delta_x^2\Delta_z^2 + \Delta_y^2\Delta_z^2) + (J + 4M_1 + 4M_2)\Delta_n^2 - 6(J + M_2 + 2M_1) \tag{8}$$

one may conveniently write the general model as

$$H = H\sum_n \sigma_n + \frac{1}{2}\sum_n \sigma_n O_n \sigma_n + \sum_{n,n',n''} L_{n,n',n''}\sigma_n\sigma_{n'}\sigma_{n''} + \sum_{n,n',n'',n'''} P_{n,n',n'',n'''}\sigma_n\sigma_{n'}\sigma_{n''}\sigma_{n'''} \quad (9)$$

In this chapter we study selected portions of the global phase diagram of the full Hamiltonian and evolve a more complete picture for the restricted case (2). In many cases our study is facilitated by the results presented in earlier papers on Widom's model [16,25], and such information will therefore not be discussed in much detail. We also note that the zero-temperature states of the Hamiltonian are known exactly for the extended Hamiltonian [16,26,27] and that low-temperature analyses have been used to determine the structure of the phase diagram for the ordered phases [27].

The important observations from the zero- and low-temperature analysis that will be qualitatively important for our understanding of amphiphilic systems are outlined below. We first note that if $J > 0$ and $M < 0$, the lattice model Hamiltonian is spatially frustrated, and this results in the stabilization of many structured phases. It transpires (see Section III) that these ordered phases may be put in correspondence with the liquid–crystal and crystalline phases observed in multicomponent mixtures of oil, water, and amphiphile. However, even for the disordered phase that corresponds to microemulsion, the presence of these competing interactions is most important, and as we shall see (Section IV), actually implies the presence of the two long length scales observed in the SANS experiments. It is therefore worth reflecting a little on the origins and consequences of this spatial frustration.

The fact that the underlying spin model is spatially frustrated is hardly surprising given that such models are known to produce many flat domain walls between plus and minus spins [27,28] and that, in the solution picture, such domain walls represent amphiphilic film. Nevertheless, the equivalence of the solution model to a spatially frustrated lattice model with only a few coupling constants is already a nontrivial and interesting aspect of the study. When the chemical potential term $\mu_{AB} - 1/2(\mu_{AA} + \mu_{BB})$ is large and positive, so is the coupling constant, J (small amphiphile concentration), and the spin model has a strong ferromagnetic coupling and therefore disfavors the formation of magnetic domain walls or, what is equivalent in the liquid model, amphiphile layers. Those that do form are encouraged to remain apart by the linear next-nearest-neighbor antiferromagnetic coupling and to remain flat by the diagonal-neighbor in-

teraction, the latter two effects reflecting the underlying molecular interactions.* At higher temperatures, layer fluctuations will force the lamellae even further apart because of the favorable contributions that arise from entropy of wandering of the surfaces and the unfavorable effects from their collision. It is important to note, however, that above the layer-roughening temperature, the principal contributions to the effective forces between the domain walls arise not from simple entropy of wandering, but from a convolution of the layer fluctuations over the energy of contact and bending of the amphiphilic film. This will ultimately mean that the length scales in the liquid–crystal and microemulsion phases are set by these effects rather than by Helfrich forces.†

As mentioned earlier, the zero-temperature states of the Hamiltonian can be determined exactly. This is accomplished by determining those local octahedra that minimize the energy for fixed values of the coupling constants. Since for $H = L_1 = L_2 = 0$ there is inversion symmetry and every local octahedron is self-tiling or tiles with its inversion image, one can construct all the zero-temperature states. There are, however, some regions of parameter space (termed *multistate surfaces*) where two or more octahedra are energetically degenerate. Such degeneracies reflect special choices of the values of the microscopic energies, as well as the relative chemical potentials of the oil, water, and amphiphile. They may occur between octahedra with either different numbers of amphiphile molecules or with the same number but different geometric arrangements. In either case zero-temperature states may then be constructed from arbitrary mixtures of them, and this results in the formation of layered, rippled, tubular, and cubic ordered phases. At nonzero temperatures, fluctuations break the degeneracy existing on the zero-temperature multistate

* This is an important point. Even at low temperatures where bilayer fluctuations are negligible, the characteristic layer separations in this model are set by two factors. The first is the amphiphile end-to-end energetic contribution that appears in Fig. 1; relative to the side-by-side contribution, this is unfavorable. However, in principle one should have further-neighbor end-to-end interactions that reflect the attractive dispersion interactions between amphiphiles, and this causes some modification of the swelling progression predicted by the model. The other more important factor affecting the low-temperature layer separations is that of the relative chemical potential of oil, water, and amphiphile. Evidently, if insufficient oil or water is present, the amphiphile layers cannot swell. Large chemical potentials of oil and water produce large (ferromagnetic) nearest-neighbor coupling, and this results in plus-minus domain walls being well separated.

† This matter has also been discussed for continuum models of interacting bilayers (A. Parsegian, preprint and private communications). The convolution of the layer fluctuations with interaction potentials produces effective forces that differs from the entropic interactions of Helfrich that are discussed in Ref. 29. Such considerations are all accounted for in the formulation of a lattice model.

surfaces, providing the thermodynamic ordered phases of the lattice model. For example, it may be shown that upon the surface

$$J + 2M_2 + 4M_1 + 4P = 0 \tag{10}$$

the local octahedra C_0^0 and C_0^1 (see Fig. 2) are energetically degenerate and that only the lamellar and the two ferromagnetic phases are stable. The effects of this zero-temperature multistate region in stabilizing long-period lamellar order are, as we shall see, ultimately profoundly important in understanding the properties of microemulsion. It is worth noting that since an amphiphile lies along every bond connecting a plus and minus spin, Eq. (10) is, from the viewpoint of the underlying solution model, a surface that divides parameter space into a region where it is unfavorable ($J + 2M_2 + 4M_1 + 4P > 0$) and favorable ($J + 2M_2 + 4M_1 + 4P \leq 0$) to form large amounts of (flat) amphiphilic film. This is in accord with the observation that at finite temperature and upon the surface defined by (10), zero work is required for the insertion of amphiphile into the amphiphilic film, the underlying reason for degeneracy of the clusters C_0^0 and C_0^1. One may also describe this phenomenon qualitatively by saying that the mesoscopic coarse-grained tension of an amphiphilic film is zero, even though the curvature energy is large and positive. However, the definition based on the lattice model is regarded as canonical and yields a connection to microscopic parameters and chemical potentials of the system. Such connections could, in principal, be checked experimentally and one expects them to be obeyed, at least qualitatively.*

At higher temperatures, in the vicinity of the surface (10), the flat lamellar sheets crumple and a disordered phase results. However, the structural properties of this disordered phase still reflect many of the features of the zero-temperature multistate sheet, a theme that will arise frequently in our description of microemulsion.

It is worth noting that although the sheet defined in (10) is central to our understanding of microemulsion, there exist others that presage the emergence of liquid–crystalline phases. Inasmuch as the model reflects such near-cancellations in continuum systems, one might expect a range

* There are, of course, different levels at which the conclusions drawn from this model may be evaluated experimentally. At one level one may simply locate the various phase equilibria by adjusting the parameters of the model. Various quantities may then be computed along these phase equilibria and compared with the corresponding equilibria in experiment. One may also determine the parameters of the lattice model using SANS data. These are the types of strategy pursued in the present chapter. It would be interesting, however, to use experimental data to fit the evolution of phase boundaries and, subject to these fixed values, compute the observables.

of interesting experimental phenomena near these surfaces. We are unaware of careful experiments in these regions of the phase diagram.

At higher temperatures the low-temperature expansions referred to above must be replaced by other types of calculation. The most basic analysis is that based on local-density mean-field theory. It is useful partly because it yields a qualitatively correct description of the global phase diagram* and partly because it is analytically tractable, an important feature given the complexity of the model.

In the next section we construct and study the mean-field theory of the Hamiltonian [Eq. (2)]. Some results for the full Hamiltonian (9) are also presented.

III. LOCAL DENSITY MEAN-FIELD THEORY

To derive the mean-field equation, we apply the Gibbs–Bogoliubov variational principal, which establishes an upper bound on the free-energy density. Thus we find that

$$G \leq G_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle \tag{11}$$

Here, G is the true free energy of the system, G_0 is the free energy with respect to a reference Hamiltonian,

$$\mathcal{H}_0 = -\sum_n h_n \sigma_n \tag{12}$$

and $\langle X \rangle_0$ means that the average of X is taken with respect to $\mathcal{H}_0$. Using (11) and (12) one can show that the free-energy density may be written in the thermodynamic limit as

$$\mathcal{G} = \lim_{l \to \infty} \frac{1}{l} \left(\begin{array}{l} \tfrac{1}{2} \sum_n S_n O_n S_n + \tfrac{1}{2} kT \sum_n (1 + S_n) \ln \tfrac{1}{2}(1 + S_n) \\ \qquad + (1 - S_n) \ln \tfrac{1}{2}(1 - S_n) + P \sum_n S_n \sum_{n',n'',n'''}{}' S_{n'} S_{n''} S_{n'''} \end{array} \right) \tag{13}$$

where $S_n = \langle \sigma_n \rangle_0$ is the spin density at site n and the primed sum means that the sum is restricted to the appropriate connected four-body terms (see Fig. 1). To proceed we now expand around the minimal spin density S_n,

$$\mathcal{G}[S_n + \delta S_n] = \mathcal{G}_0 + \sum_n \frac{\delta \mathcal{G}}{\delta S_n} \delta S_n + \tfrac{1}{2} \sum_{n,n'} \frac{\delta^2 \mathcal{G}}{\delta S_n \delta S_{n'}} \delta S_n \delta S_{n'} + \cdots \tag{14}$$

* This point has been to a large degree established by checking the conclusions from mean-field theory with low-temperature, high-temperature, and loop expansions in addition to simulation and analytical renormalization-group calculations. There remain some open questions, however, such as those discussed in Ref. 30 (also, Y. Levin and K. A. Dawson, unpublished).

The Euler–Lagrange equations are

$$\frac{\delta \mathcal{G}}{\delta S_n} = 0 = O_n S_n + kT \tanh^{-1} S_n + P \sum_{n',n'',n'''}' S_{n'} S_{n''} S_{n'''} \qquad (15)$$

and these ensure that the resulting spin density correspond to a local minimum or maximum of the free energy. Again, the prime implies that the sum is restricted to run only over all appropriate three-spin terms that are connected to the central site (n). For a local minimum, the second derivative term must be greater than zero for all variations in the spin density. The second-order derivative matrix evaluated in the disordered phase ($S_n = 0$) is

$$K_{n,n'} = (O_n + kT)\delta_{n,n'} \qquad (16)$$

Within local density mean-field theory, we shall be interested in constructing analytically the phase diagram in the vicinity of the disordered phase (microemulsion) without a priori knowledge of the symmetry of the states. We must, therefore, first locate that surface in parameter space where the matrix (16) has one or more zero values for $S_n = 0$. This will define a critical surface* in terms of the lattice-model parameters J, M, and P, beneath which the disordered phase is no longer stable. The second derivative matrix (16) is diagonalized by Fourier transformation and has eigenvalues ($M_1 = 2M$, $M_2 = M$)

$$K_q = -4Me_q^2 - 2Je_q + 6M + kT \qquad (17)$$

where

$$e_q = \sum_{i=1}^{3} \cos q_i \qquad \mathbf{q} = (q_1, q_2, q_3) \qquad (18)$$

When one or more of the eigenvalues becomes zero (say, $K_{\mathbf{q}_c}$) this is an indication that the paramagnetic phase is unstable with respect to infinitismal sinusoidal spin-density variations with momentum $\mathbf{q}_c$. To enforce convexity on the spectrum of eigenvalues we require that

$$K_{\mathbf{q}_c} = 0 \qquad (19)$$

and

$$K_{\mathbf{q}_c} = \min_{\mathbf{q}} \{K_{\mathbf{q}_c}\} \Rightarrow \nabla_{\mathbf{q}} K_{\mathbf{q}} \big|_{\mathbf{q}=\mathbf{q}_c} = 0 \qquad (20)$$

* This defines a true critical surface only if at the paramagnetic phase is a global minimum of the free energy as it becomes unstable. It is possible that for some values of the parameters, another minimum is then a global free-energy minimum and the critical transition has then been preempted by a first-order transition. One can prove (see Ref. 25) that for $P = 0$ the paramagnetic phase is the unique minimum until the multifurcations occur. This comment is valid up to a value of P, say P_c, that is a function of J and M. Note that $P_c(J,M)$ defines a curve of tricritical points.

To study the region below the critical surface ($T < T_c$) systematically, a perturbation theory must be developed to solve the nonlinear mean-field equations.

Reparameterization of the mean-field equations at the critical surface is achieved upon division of (15) by kT_c ($m = M/kT$, $j = J/kT$, $\mathcal{P} = P/kT$, $\mathcal{O}_n = O_n/kT$).

$$\mathcal{O}_n S_n + (1 - \epsilon) \tanh^{-1} S_n + \mathcal{P} \sum_{n',n'',n'''}{}' S_{n'} S_{n''} S_{n'''} = 0 \tag{21}$$

where the parameter ϵ ($= 1 - T/T_c$) will be the small parameter in the theory. The solution to Eq. (21) may be written as an expansion in ϵ,

$$S_n(\epsilon) = \sum_{p=0} \epsilon^{\beta+p} S^{(p)} \tag{22}$$

We emphasize that the spin density may still possess arbitrary spatial variations and that, as yet, no judgments have been made about the symmetry of the solutions. The leading order term, $S^{(0)}$, is premultiplied by ϵ^β, where β is the mean-field critical exponent of $\frac{1}{2}$. Then, by expanding the nonlinearity in (22) and equating equal powers of ϵ, an infinite set of coupled equations is obtained and these define the perturbation theory. To $\mathcal{O}(\epsilon^{5/2})$ one obtains

$$\epsilon^{1/2}\colon (\mathcal{O}_n + 1)S^{(0)} = K_q S^{(0)} = 0 \tag{23a}$$

$$\epsilon^{3/2}\colon (\mathcal{O}_n + 1)S^{(1)} = S^{(0)} - \tfrac{1}{3}S^{(0)3} - \mathcal{P}(S^{(0)}_{x,y,z+1}S^{(0)}_{x,y+1,z}S^{(0)}_{x+1,y,z} + \text{terms related by symmetry}) = f^{(1)} \tag{23b}$$

$$\epsilon^{5/2}\colon (\mathcal{O}_n + 1)S^{(2)} = S^{(0)2}S^{(1)} + S^{(1)} + \tfrac{1}{5}S^{(0)5} - \tfrac{1}{3}S^{(0)3} + \mathcal{P}(S^{(0)}_{x,y,z+1}S^{(0)}_{x,y+1,z}S^{(1)}_{x+1,y,z} + \text{terms related by symmetry}) = f^{(2)} \tag{23c}$$

The remaining terms in (23b) are composed of all combinations of three-spin terms that are connected to the central site. In (23c) the same is also true, but one permutes the index of $S^{(1)}$ through every position that is connected to the central site.

The solutions, $S^{(0)}$, of (23a) are plane waves so that one may write

$$S^{(0)} = A_q \prod_{i=1}^{3} \cos(q_i x_i + \alpha_i) = A_q \varphi_q \tag{24}$$

Here the q_i satisfy (19) and x_1, x_2, x_3, correspond to x, y, z, respectively. At this order in perturbation theory the only nondegenerate solutions that are consistent with the convexity relations are $\mathbf{q}_c = (0,0,0)$ and $\mathbf{q}_c = (\pi,\pi,\pi)$. However, these solutions are valid only in the vicinity of their

respective critical lines,

$$j_c + 5m_c = \tfrac{1}{6} \tag{25}$$

and

$$j_c - 5m_c = -\tfrac{1}{6} \tag{26}$$

Note that these equations are true for arbitrary values of $\mathcal{P}$, provided that the order–disorder transition is second-order. They correspond, respectively, to the paramagnetic–ferromagnetic critical line and the paramagnetic–antiferromagnetic critical lines. The third region of parameter space that satisfies the convexity relations is given by the relation

$$-4m_c e_{\mathbf{q}_c}^2 - 2j_c e_{\mathbf{q}_c} + 6m_c + 1 = 0 \qquad -3 \leq e_{\mathbf{q}_c} \leq 3 \tag{27}$$

again, valid for arbitrary $\mathcal{P}$ provided that the transition is continuous. There is also a constraint on the sum of the cosines of the critical modes that is derived from the convexity relation (20),

$$e_{\mathbf{q}_c} = \frac{j_c}{4m_c} \tag{28}$$

Since (25) to (27) have no explicit $\mathcal{P}$ dependence, we conclude that the order–disorder transition consists of two flat sheets that cap a segment of an elliptical cylinder. Beneath this ellipsoidal surface ($T < T_c$) one is not yet able to resolve the degeneracy of states having many different periodicities. One therefore uses the Fredholm alternative to solve the equations to $O(\epsilon^{3/2})$. The Fredholm alternative requires that for there to be a solution to (23b), all homogeneous solutions must be orthogonal to the inhomogeneous term. For example,

$$\langle S^{(0)} \mid f^{(1)} \rangle = 0 \tag{29}$$

where the brackets denote the sum over all lattice spacings,

$$\langle X_n \rangle = \lim_{l \to \infty} \left(\frac{1}{(1 + l)^3} \sum_n X_n \right) \tag{30}$$

Application of the Fredholm alternative leaves us with a number of special cases for the spatial modulations of phases that minimize the free energy to $O(\epsilon^3)$. The free-energy density to $O(\epsilon^3)$ is

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{4} \langle (S^{(0)})^2 \rangle \tag{31}$$

Performing the sums we find that to $O(\epsilon^3)$, there are nine special cases to consider in the free-energy density. These are:

1. $\mathbf{q} = (\pi,0,\cos^{-1}(-j_c/4m_c))$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{6}\left\{\frac{1}{(\frac{1}{3} + 32(-j_c/4m_c - 2)\mathcal{P})}\right\} \tag{32}$$

2. $\mathbf{q} = (0,0,\cos^{-1}(-j_c/4m_c - 2))$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{6}\left\{\frac{1}{(\frac{1}{3} + 32(-j_c/4m_c - 2)\mathcal{P})}\right\} \tag{33}$$

3. $\mathbf{q} = (\pi,\pi,\cos^{-1}(-j_c/4m_c + 2))$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{6}\left\{\frac{1}{(\frac{1}{3} + 32(-j_c/4m_c - 2)\mathcal{P})}\right\} \tag{34}$$

4. $\mathbf{q} = (0,\cos^{-1}(-j_c/8m_c - \tfrac{1}{2}),\cos^{-1}(-j_c/8m_c - \tfrac{1}{2}))$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{4}\left\{\frac{1}{(\frac{3}{4} - 18(-j_c/4m_c - 1)^2\mathcal{P})}\right\} \tag{35}$$

5. $\mathbf{q} = (\pi,\cos^{-1}(-j_c/8m_c + \tfrac{1}{2}),\cos^{-1}(-j_c/8m_c + \tfrac{1}{2}))$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{4}\left\{\frac{1}{(\frac{3}{4} + 18(-j_c/4m_c + 1)^2\mathcal{P})}\right\} \tag{36}$$

6. $\mathbf{q} = (\cos^{-1}(-j_c/4m_c - q_2 - q_3),q_2,q_3)$

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{32}\left\{\frac{1}{(\frac{9}{64} + \frac{27}{2}(-j_c/4m_c \cos q_2 \cos q_3 - \cos^2 q_2 \cos q_3 - \cos^2 q_3 \cos q_2)^2\mathcal{P})}\right\} \tag{37}$$

The remaining three cases (7 to 9) have wave vectors that have value $\pi/2$ in at least one direction [$\mathbf{q} = (\pi/2,\pi,q_3)$, $\mathbf{q} = (\pi/2,0,q_3)$, $\mathbf{q} = (\pi/2,\pi/2,q_3)$]. For these cases the resulting free energies are degenerate with value

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{2} \tag{38a}$$

In the limit $\mathcal{P} = 0$, the free energies reduce to the values obtained for Widom's model. In that case, one finds that if two special values are chosen (0, $\pi/2$, or π), the third value direction of $\mathbf{q}$ is not yet fixed by (27) and (28). Higher orders of perturbation theory must be used to break the remaining degeneracy. Here we are concerned with finite $\mathcal{P}$ and at $O(\epsilon^{3/2})$ most of the degeneracy is broken. However, there are some regions in parameter space where one must go to second-order perturbation theory [$O(\epsilon^{5/2})$] to break the remaining degeneracy. It transpires that second-order perturbation theory is needed only for those cases where at least one of the special values is $\pi/2$ (cases 7 to 9). It may also be shown

that for the special cases 7 to 9, the problem is independent of $\mathcal{P}$ up to and including $O(\epsilon^{5/2})$. The free energy to $O(\epsilon^4)$ for cases 7 to 9 is then determined to be

$$\mathcal{G} = -\ln 2 - \frac{\epsilon^2}{2} - \epsilon^3\left(\frac{1}{9K_{113}} + \frac{11}{24}\right) + O(\epsilon^4) \tag{38b}$$

Here, K_{113} corresponds to the eigenvalue $K_{q_1q_2q_3}$. One then minimizes the free energy (38) with respect to q_1, q_2, q_3 subject to the constraint (28).

Since we have a free parameter, $\mathcal{P}$, premultiplying a four-spin term, we must also solve for values of $\mathcal{P}$ where a surface of second-order order–disorder phase transitions terminates in a tricritical line, beyond which a surface of first-order order–disorder phase transitions develops. To see how this might arise, we consider expansion of the free-energy density to $O(S_n^6)$.

$$\mathcal{G} = \tfrac{1}{2}\langle S_n \mathbb{O}_n S_n\rangle + \tfrac{1}{2}(1-\epsilon)\langle S_n + S_n^2 + S_n^6 + O(S_n^8)\rangle \tag{39}$$
$$\mathcal{P}\langle S_n S_{x,y+1,z} S_{x+1,y,z} S_{x,y,z+1} + \text{terms related by symmetry}\rangle$$

We know that $S_n = \epsilon^{1/2} A_q \varphi_q^{(0)} + \epsilon^{3/2} B_q \varphi_q^{(1)}$. By inserting this result into (39) and collecting powers of ϵ we have

$$\epsilon\colon \langle A_q^2(\varphi_q^{(0)} \mathbb{O}_n \varphi_q^{(0)} + \tfrac{1}{2}\varphi_q^{(0)2})\rangle \tag{40}$$

$$\epsilon^2\colon \left\langle \begin{array}{c} \mathcal{P} A_q^4 \varphi_q^{(0)} \varphi_{x+1,y,z}^{(0)} \varphi_{x,y+1,z}^{(0)} \varphi_{x,y,z+1}^{(0)} \\ + \text{terms related by symmetry} + A_q^4 \tfrac{1}{12} \varphi_q^{(0)4} + B_q \varphi_q^{(0)} \varphi_q^{(1)} \end{array} \right\rangle \tag{41}$$

We may also examine the Landau–Ginzburg free-energy functional

$$\mathcal{G} = aS^2 + bS^4 + cS^6 + \cdots \tag{42}$$

At a second-order phase transition $a = 0$, $b > 0$, $c > 0$. One can see a clear analogy between (39)–(41) and (42) along with the conditions on a, b, and c. Thus at the critical point the $O(\epsilon)$ term is zero from (23a), and for stability the $O(\epsilon^2)$ must be positive. This condition leads to the requirement that

$$\left\langle \begin{array}{c} \mathcal{P} A_q^4 \varphi_q^{(0)} \varphi_{x+1,y,z}^{(0)} \varphi_{x,y+1,z}^{(0)} \varphi_{x,y,z+1}^{(0)} \\ + \text{terms related by symmetry} + A_q^4 \tfrac{1}{12} \varphi_q^{(0)4} + B_q \varphi_q^{(0)} \varphi_q^{(1)} \end{array} \right\rangle \geq 0 \tag{43}$$

Performing the sums in (43) in the infinite volume limit [Eq. (30)], the following condition for stability is given:

$$\mathcal{P} > -\frac{1}{96(\cos q_1 \cos q_2 \cos q_3)} \tag{44}$$

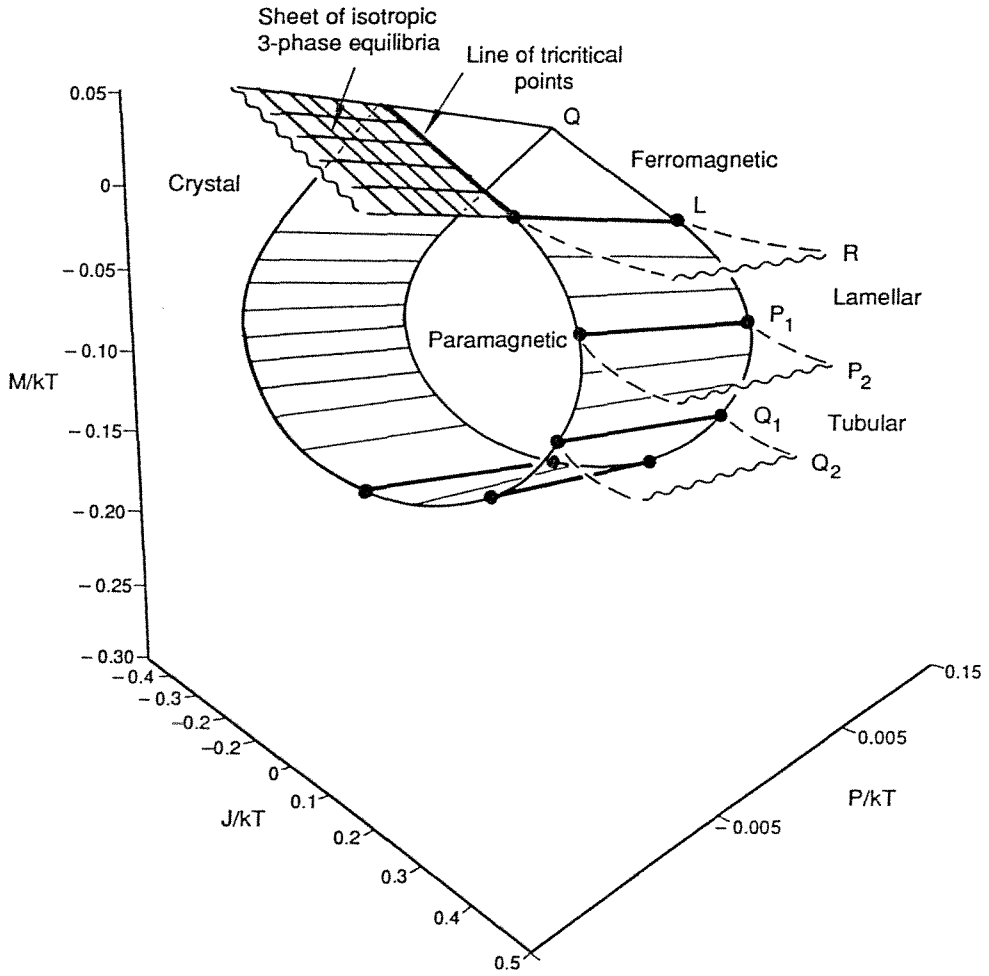


Fig. 3 Phase diagram in space of Ising model couplings, j ($= J/kT$), m ($= M/kT$), $\mathcal{P}$ ($= P/kT$), for the range $-1/96 \leq \mathcal{P} \leq 0$. The points Q, L, R, P_1, P_2, Q, Q_2 have been marked to provide a relation to the more detailed phase diagram of Fig. 5. The paramagnetic phase is identified with isotropic mixtures of oil, water, and amphiphile. In the vicinity of curve L and in the interior of the capped elliptical cylinder the disordered phase takes on the properties of microemulsion.

This constraint permits us to determine whether a given portion of the order–disorder transition surface is first or second order.

The foregoing analysis is useful because it permits us to determine analytically the structure of the mean-field phase diagram in the vicinity of the order–disorder transitions. This is in itself a highly important region

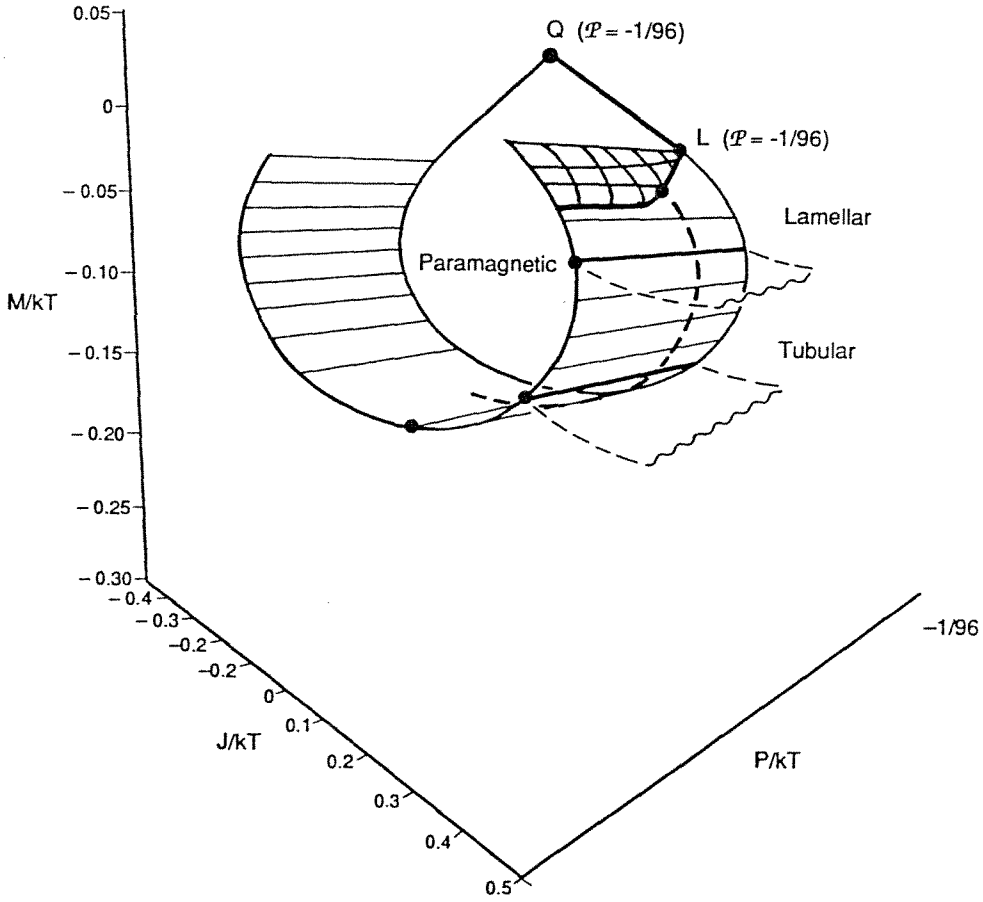


Fig. 4 Mean-field phase diagram for $\mathcal{P} \leq -1/96$. The sheet of first-order phase transitions between paramagnetic and ferromagnetic phases that grows out of the plane from line QL has not been marked. It is, however, described in more detail in Fig. 16. Note that the hatched portion of the lamellar-disordered phase boundary is first order; the remainder is second order.

of the phase diagram and provides a useful check on numerical studies of the mean-field equations (15). The basic structure of the resulting phase diagram is laid out in Figs. 3 and 4. Thus in Fig. 3 we present the results for $j,m,\mathcal{P}$ (for $-1/96 \leq \mathcal{P} \leq 0$). Note that many of the essential features found for the cut $\mathcal{P} = 0$ (see Fig. 5) are also present for nonzero $\mathcal{P}$. Thus the phase boundaries for the ordered phases shift, but the states remain, while the ferromagnetic–paramagnetic phase transition remains second order.

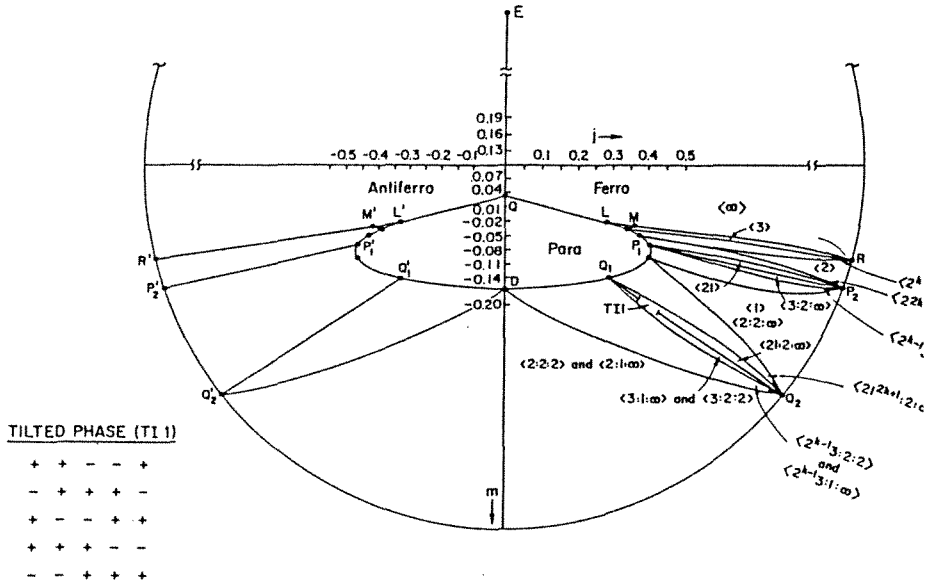


Fig. 5 Cut ($\mathcal{P} = 0$) of the phase diagram in Fig. 3. This phase diagram is described in more detail in Ref. 25, as is the notation used to describe the layered (lamellar), tubular, and cubic phases.

At $\mathcal{P} = -1/96$ the uniform phases disorder via a tricritical transition, and beyond this value there is a surface of three-phase equilibria between the ferromagnetic and uniform phases. This line of tricritical points that separates second- and first-order transitions of the uniform phases extends into the periodic-phases region, where, however, it becomes a curve parameterized by the wavelength of the critical mode. Beyond this one may construct another perturbation theory that is also parameterized by the distance from the tricritical curve. The resulting analysis is lengthy and will not be presented here. To clarify the situation further, in Fig. 6 we have presented a solution variable cut of the phase diagram (Fig. 4) at $\mathcal{P} = -0.15$. The variables we have used are the chemical potential of amphiphile relative to that of oil and water, and the parameter $K = -M/4$.

The results presented in Figs. 3 to 6 comprise the full mean-field phase diagram of Hamiltonian (3). We now turn to its interpretation and discussion as a model of microemulsion. To do this we will rely on the results from the mean-field theory and from extensive Monte Carlo calculation [31]. These two types of study of the Hamiltonian are, to a significant degree, in agreement. Some important differences are discussed in Ref. 31. However, besides quantitative Monte Carlo results one can use the

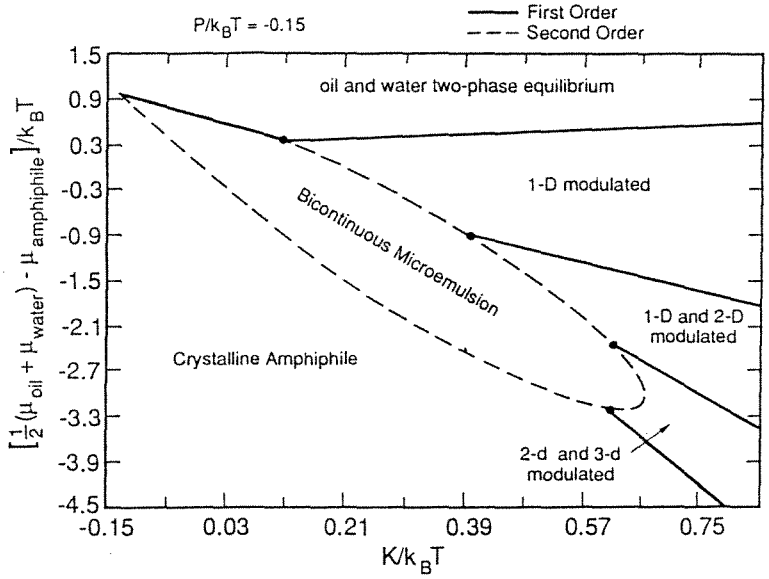


Fig. 6 Cut ($\mathcal{P} = -0.15$) of the phase-diagram shown in Fig. 4 in terms of the solution variables. The abscissa and ordinate axes represent, respectively, the chemical potential and curvature parameter defined in Table 1.

instantaneous configurations to produce quantitative descriptions of the various phases.

We identify the disordered phase of the spin model as the amphiphile-rich isotropic phase that is found in mixtures of oil, water, and amphiphile or alcohol. The ferromagnetic phases correspond to the oil-rich and water-rich phases that are typically found to be in coexistence with the middle bicontinuous microemulsion. There are also many ordered phases of the spin model and these are considered to represent the liquid-crystalline and crystalline materials that are found at high concentration of amphiphile. In particular, the one-dimensional periodic structures correspond to the lamellar liquid–crystalline phases that are almost always found near the bicontinuous microemulsion phase. These lamellae may be characterized by a repeat distance that increases as they are swollen by the addition of oil or water. Near the limits of swelling that, in this model, correspond to the intersection of the flat sheet [Eq. (25)] with the ellipsoidal cylinder (curve marked L of Figs. 3 and 4) one finds that there is a transition to the disordered phase. Upon further examination by Monte Carlo simulation this disordered phase is found to consist of fairly well separated crumpled amphiphilic sheets, an observation that is further elu-

cidated in Section IV. We now contend that the region of the disordered phase that lies close to where the interior of the ellipsoidal surface and the flat sheet join corresponds to the bicontinuous microemulsion phase. One would consider this classification to persist within the ellipsoidal surface up to the vicinity of its intersection with the multistate sheet (10), the latter being defined from the condition that the insertion of an amphiphile at the expense of half an oil and water cube costs zero work. This is an important conclusion that will be argued in more detail later in the chapter. It will ultimately permit us to establish relations between the microscopic parameters of the model in order that the bicontinuous phase be present. These relations may ultimately be expressed in terms of ratios of various experimental variables that should then be nearly universal. An example is given in (59). Concisely, then, we expect to find the bicontinuous microemulsion in the vicinity of the order–disorder transition within the sector defined by the relations

$$J + 4M_1 + 4M_2 \leq 0 \text{ (all } P) \tag{45}$$

and (10), that is,

$$J + 4M_1 + 4M_2 - 8P \geq 0 \tag{46}$$

The origin of the first of these two equations is discussed further just above (54).

For our study of the Winsor III equilibrium, the most important changes on going from zero to finite $\mathcal{P}$ are twofold. First, the surface of transitions from the ferromagnetic to the disordered phase becomes first-order, and part of it becomes identifiable with the Winsor III equilibrium. The second is more subtle and its importance will become clear only when we discuss (see Section IV) the correlation functions of the theory. In essence, though, it will transpire that for certain choices of $\mathcal{P}$, the model predicts a nonmonotonic small-wavelength scattering intensity for a disordered microemulsion phase that is in equilibrium with the micellar phases. This phenomenon is indeed sometimes observed for Winsor III middle-phase microemulsions.

The fact that for sufficiently large values of $\mathcal{P}$ (and therefore of V_5) the micellar and microemulsion phases are in first-order equilibrium is hardly surprising. The four-body term admits the assignment of independent edge and corner energies. This is, in turn, reflected in the emergence of a discontinuous transition between the ferromagnetic and disordered phases, the former containing numerous corners, the latter consisting mainly of flat domain walls with occasional edges. It is clear, therefore,

that we have implicitly asserted that the principal cause of the Winsor III three-phase equilibrium lies in the difference between these different curvature–energy scales. That there exist other physical effects, not accounted for by the original two-body spin Hamiltonian is clear. For example, the configurational entropy for a three-state Potts model is different from that of the Ising model, and this results in a symmetric three-phase equilibrium. Within such a formulation, one might therefore assume the Winsor III state to arise as a consequence of the three-component nature of the mixture.* While such effects are surely present in real solutions, it is unlikely that they are relevant to the Winsor equilibrium. For them to be so, one would require many of the amphiphile molecules to be isolated in solution, and this is known not to be the case for moderately long amphiphiles. Now in the case of mixtures of short-chain alcohols, oil, and water, one also often finds a three-phase equilibrium between an oil-rich, water-rich, and isotropic alcohol-rich phase. Substantial portions of the alcohol exist as solvated isolated molecules. For this reason one believes that the origin of these first-order transitions lies in large measure in translational entropy effects due to three distinguishable components. This case should therefore be well described by the symmetric equilibrium in the Potts model. It has sometimes been claimed that the Winsor equilibrium is a natural extension of this type of three-phase equilibria associated with short-chain alcohols. This is probably not true. In essence we have argued that the Winsor III state is much more remarkable and is largely a reflection of the different bending energies of the amphiphilic film. If this is indeed the case, the Winsor III transitions should become increasingly first-order as the amphiphile length and therefore the curvature energy are increased. The converse would be true if the origin lay in translational entropy effects, since as the amphiphiles become longer it has a much greater tendency to remain within the amphiphilic film and the system becomes increasingly two-component in nature. Undoubtedly, there exist experimental data that would resolve this issue.

IV. STRUCTURE FACTOR IN THE MICROEMULSION PHASE

Some of the most revealing insights into the structure of microemulsion have been offered by fairly recent small-angle neutron scattering (SANS)

* This effect would in some measure account for the symmetric three-phase equilibrium found in the models discussed in Ref. 18.

experiments. However, as we shall see, such measurements are of additional importance to our theoretical study because, from them, one may uniquely determine the microscopic coupling parameters of the lattice model.

Teubner and Strey [41] analyzed some of the earlier experimental data and pointed out that the appearance of a pronounced peak in the scattering intensity for the bicontinuous microemulsion regime could not be understood in terms of a classical droplet model. They suggested a simple Ornstein–Zernike form for the intensity of scattered radiation,

$$I(q) = \frac{I(0)}{1 - bq^2 + cq^4} \tag{47}$$

where b and c are positive quantities. The presence of the $O(q^2)$ and $O(q^4)$ terms of opposite sign means that this function has a peak at finite values of q. It was also pointed out that this scattering behavior is consistent with the long-distance behavior of the correlation function,

$$h(r) = \frac{A}{r} e^{-r/\xi} \sin \frac{2\pi r}{d} \tag{48}$$

where ξ and d are given by

$$\frac{d}{2\pi} = \left(\frac{1}{2c^{1/2}} + \frac{b}{4c} \right)^{-1/2} \tag{49}$$

$$\xi = \left(\frac{1}{2c^{1/2}} - \frac{b}{4c} \right)^{-1/2} \tag{50}$$

When the experimental data are fitted to (47), the two putative lengths, ξ and d, may be determined as a function of the field variables. On that region of the phase diagram typically considered to be bicontinuous microemulsion, both lengths are typically found to be on the scale of hundreds of angstroms. One concludes, therefore, that there are at least two significant lengths that are important in understanding the bicontinuous microemulsion phase. A clear understanding of their origin is central to the elucidation of the microstructure of microemulsions. The results of this section are derived from a Gaussian fluctuation calculation [15] and Monte Carlo [31] simulations and the reader should consult the original references for technical details.

In fact, it was also possible to show that the isotropic phase of the lattice model of microemulsion exhibits such behavior [15,31]. The reason is straightforward. From the mean-field free-energy functional [15], one can determine an approximate susceptibility as the second-derivative matrix with eigenvalues given by [17]. The Fourier-transformed lattice-dif-

ference operator may, in the long-wavelength limit, be replaced by its small-momentum expansion, and one thereby obtains a structure factor and scattering intensity with coefficients of opposite sign for the terms of order q^2 and q^4. Thus, using (17), one obtains

$$b = \frac{j + 12m}{-6(j + 5m) + 1} \tag{51}$$

$$c = \frac{-m}{-6(j + 5m) + 1} \tag{52}$$

$$I(0) = \frac{1}{-6(j + 5m) + 1} \tag{53}$$

and also a nonspherically symmetric term of form $(q_1^4 + q_2^4 + q_3^4)$. Clearly, the latter term is present because of the underlying lattice. In the microemulsion regime its contribution is, however, numerically small and need not be considered at present. Indeed, for this region of parameter space it may further be shown that the coefficient of the nonspherically symmetric term is an irrelevant variable for the renormalization group flow. Thus, for long lengths d and ξ this term does not contribute much beyond naive fluctuation theory. For shorter lengths, ξ and d, the structure factors are still affected by the presence of the underlying lattice, although these remain numerically small.

It is, however, also possible to approximate the scattering data for the full Hamiltonian [Eq. (9)] in a slightly different manner. Thus, as for (17), we determine the eigenvalues of the second-derivative matrix and spherically average to obtain generalizations of (51) and (52). The results are

$$b' = \frac{J + 4M_1 + 4M_2}{6(J + M_2 + 2M_1) - 1} \tag{54}$$

$$c' = \frac{\frac{53}{64}\left(\frac{J}{6} + \frac{8}{3}M_2 + M_1\right)}{6(J + M_2 + 2M_1) - 1} \tag{55}$$

The most important point to be deduced from these formulas is that the characteristics of the scattering curve [Eq. (47)] are actually determined not just by the concentrations of the components and some curvature energy, but also by the interlayer interaction energy scale set by the parameter M_1. Thus although the form of (47) is quite simple, variations with the parameters reflect a number of system-dependent variables. This presumably accounts for the remarkably common pattern of the scattering from microemulsion.

We should also note a few further general features of the predictions

for b and c. For example, both diverge [as does $I(0)$] on approach to the sheet in parameter space given by (25). Part of that surface corresponds to the ferromagnetic to disordered phases disordering transition. Elsewhere on the surface the periodic phases are stable and it no longer has the significance of a set of phase transitions, though within the disordered phase d and ξ still increase on approach to it. We also note that b is positive only beyond the surface $j + 12m = 0$. In this region the work required to insert an amphiphile molecule becomes sufficiently small for much amphiphilic film to be formed, and the disordered phase then begins to take on some of the appearance of the liquid–crystalline phases.

Indeed, it is possible to divide that region occupied by the disordered phase into two regions. The dividing surface is defined by the values of the parameters j, m, $\mathcal{P}$ for which $d = 0$. On one side of this surface the correlations decay according to (48); on the other the decay is a simple exponential. This dividing surface is independent of $\mathcal{P}$ and is given by

$$4c - b^2 \tag{56}$$

This, then, is the more formal definition of one of the two boundary surfaces introduced in (45) and (46). We emphasize, though, that this is an approximation based on Gaussian fluctuation theory. Higher orders of loop contributions introduce $\mathcal{P}$ dependence. However, the simulation results of Figs. 7 and 8 indicate that the result is valid for moderately small $\mathcal{P}$. Since the Gaussian fluctuation calculations are known to be quantitatively unreliable, we have chosen to calculate the structure factors using Monte Carlo simulation. Some results are given in Figs. 7 and 8 for two different values of the parameters. For orientation we have also included a phase diagram (Fig. 9) obtained from simulation and renormalization group calculation. The first result, Fig. 7, corresponds to $\mathcal{P} = 0, j = 1.3$, $m = -0.4$. Note that this curve corresponds to a point in the (j,m) plane that is well away from the ordered phases and, consequently, exhibits little sign of a peak at finite q. The curve in Fig. 8 ($\mathcal{P} = 0, j = 1.2$, $m = -0.4$), on the other hand, corresponds to a (j,m) point much closer to a short-wavelength ordered phase and possesses a pronounced peak at finite wave number. Results have been calculated for the same pair of (j,m) values but for larger values of $\mathcal{P}$. It transpires that the structure factors are only a little different from those for $\mathcal{P} = 0$, an observation that is consistent with the mean-field susceptibility calculation outlined in (51). Thus the four-body term in the mean-field free energy makes little contribution to the inverse susceptibility of the paramagnetic phase [Eq. (17)].

One can make additional comparisons between the predictions of the lattice model and experiment. Note that the constants in (47) may be

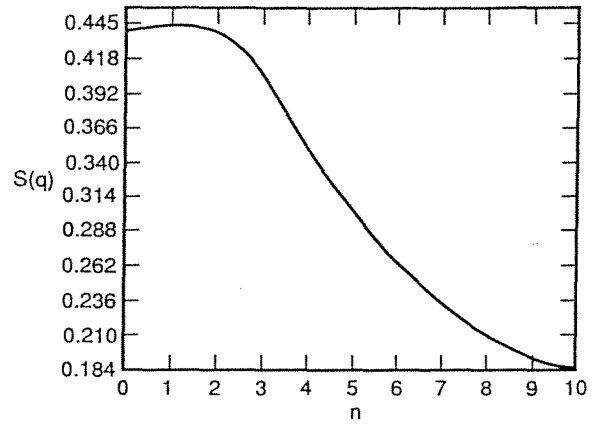


Fig. 7 Structure factor versus wave number calculated by Monte Carlo simulation. The values of the parameters are $j = 1.3$, $m = -0.4$, $\mathcal{P} = 0.0$, and the lattice size, $L = 20$.

determined experimentally. Consequently, their evolution as a function of increasing concentration may also be determined. Teubner and Strey [41] exhibit some such data for the bicontinuous phase of a water–octane–$C_{12}E_5$ mixture (see Table 1 of Ref. 10). From these one may infer that in agreement with the model, b is a positive number that increases with

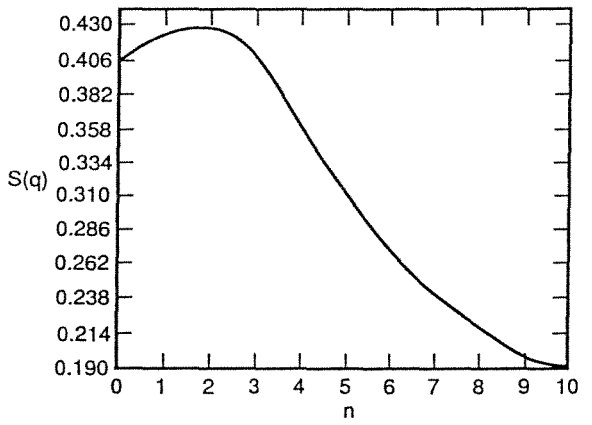


Fig. 8 Structure factor versus wave number calculated by Monte Carlo simulation. The values of the parameters are $j = 1.2$, $m = -0.4$, $\mathcal{P} = 0.0$, and the lattice size, $L = 20$. Note the emergence of a rather pronounced peak at finite q.

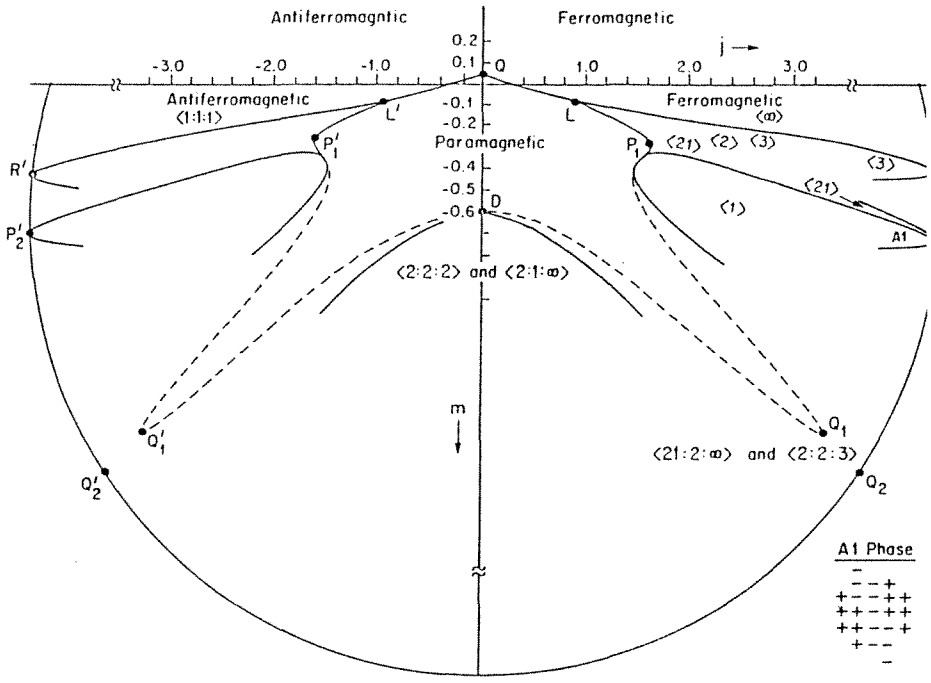


Fig. 9 Phase diagram in j,m space for $\mathcal{P} = 0$ as calculated by simulation, low-temperature expansion, and renormalization group calculations. This diagram emphasizes the observation that although the topology of the mean-field phase diagram is correct, the phase transitions are strongly shifted. For this reason, when correlation functions are being discussed one should make reference to the present phase diagram rather than the mean-field diagrams of Figs. 3 to 5.

decreasing concentration of amphiphile. A quantitative comparison would, however, require the evaluation or estimate of m for this system. Such estimates can be made independently on the basis of, for example, experiments that measure interfacial fluctuations. However, we leave comparisons of this nature for a later paper. Rather, we examine those ratios that are indicated by the lattice model to be constant, or nearly so. For example, the ratio b/c $(= r)$ is, from (51) and (52),

$$r = \frac{j}{m} + 12 \geq 0 = \frac{1}{2M}\left[\mu_{AB} - \tfrac{1}{2}(\mu_{AA} + \mu_{BB})\right] - \tfrac{1}{2} \qquad (57)$$

However, we have earlier noted that a second length scale (d) emerges

only when $j + 12m < 0$, and $-b^2 + 4c < 0$. This, along with the condition of (46), defines a sector within which we expect to find the bicontinuous phase. Thus, for $M_1 = 2M$ and $\mathcal{P} = 0$, we find that

$$j + 10m \geq 0 \tag{58}$$

Consequently, for bicontinuous microemulsion we expect the ratios to lie within the range*

$$0 < r \leq 2 \tag{59}$$

Also, for fixed $I(0)$, d diverges as r becomes large, so large domain-size microemulsions will have smaller values of this ratio. The relation (57) will be reflected in constraints between numerous other properties of the microemulsions. For example, it has already been shown that the ratio d/ξ is roughly constant. However, since such observations stem from (57) it would be convenient if future experimental measurements included the measured value of r. We can study this issue by rearranging (48) and (49). Thus we find that

$$c^{1/2} = \frac{1}{1/\xi^2 + (2\pi/d)^2} \tag{60}$$

and

$$b = 2c^{1/2} - \frac{4c}{\xi^2} \tag{61}$$

determine b and c in terms of the measured ξ and d. In this manner the ratio b/c (in units of angstroms) may be plotted against concentration for the water–octane–$C_{12}E_5$ mixture. Note that for the smallest concentration r is very small, indicating that one is almost on the sheet $j + 12m = 0$. As expected, with increasing concentration r increases (d decreases) and the highest concentration corresponds roughly to the sheet $j + 10m = 0$. Since we assume fixed m, this means that j is varying as a consequence of changes of the amphiphile chemical potential relative to that of oil and water. This corresponds to a horizontal trajectory in the interior of the ellipse in Figs. 3 and 4. Since one is moving away from any order–disorder transition, the overall scattering intensity is expected to decrease, an observation that is consistent with, for example, Fig. 5 of Ref. 10. Indeed,

* Note that one now has a relationship between the activities of oil, water, and amphiphile and the interaction energy m. Experimental tests of this have not yet been made.

one can rewrite (53) as

$$I(0) = \frac{1}{1 - 6(b/c - 7)m} \tag{62}$$

Now, assuming that m is constant for this sequence of experiments, its value may be determined from one experimental value $I(0)$, and all the others are determined uniquely. The results, in the form of plots of $I(0)$ against concentration, are in good qualitative agreement with experiment. Equation (58) may also be rearranged to give

$$m = \frac{1}{6(b/c - 7)[1 - I(0)]} \tag{63}$$

a quantity which, besides being nearly constant in such experiments, also yields another estimate of the microscopic coupling in the lattice model. Experimental determination of this quantity as a function of the concentration of added cosurfactant would also be most useful in the experimental studies. This would enable one to probe the increasing flexibility ($| m |$ becomes smaller) of the amphiphilic film as cosurfactant is added. Finally, note that if we examine (57), (59), and (63), we have a highly nontrivial relation between the chemical potentials and the $I(0)$ from scattering data. Experimental probes of this relation have not yet been undertaken. Such relations would represent one of the most rigorous tests of the theory because they provide relations between experimental quantities that possess no superficial correlations with each other.

The results of the calculations of the structure factors of the model may also be understood more qualitatively as follows. At sufficiently low temperatures, the presence of competing interactions causes lamellae ordering with varying distances between the lamellae, depending on the couplings and temperature. Above the order–disorder transition the hitherto flat lamellae crumple and a disordered phase results. Although there may be no long-range order, there remain correlations that are reminiscent of the layered ordering. There are, consequently, two length scales: one, d, that reflects a tendency of the layers to maintain a fixed interlayer distance, and the other, ξ, a bulk correlation length that determines the length on which this layer-like order falls off. To see this more clearly one might contemplate the spin-spin correlation function equation (48). This function dies off exponentially with a length, ξ, that determines the

* These values of $I(0)$ versus concentration would be useful experimental data for comparison with the model. A broad range of systems should be studied experimentally in this manner before any final conclusions are drawn. However, the preliminary seem very good, and deviations may well be accounted for by the fact that (62) and (63) are not yet sufficiently accurate descriptions of the lattice model.

average number of plus (or minus) spins that are coaligned with the central plus (or minus) spin. The correlation function then passes through zero at some value of the interlattice spacing that determines the most probable distance for a domain wall separating antialigned spins to occur. In the fluid picture we may consider a central oil (water) cube to be correlated with other oil (water) cubes up to the length scale set by the change of sign of the spin–spin correlation function. Thereafter, the fluid around the central oil (water) cube is considered to be essentially water (oil). Any central oil (water) site is, therefore, on average surrounded by concentric shells composed largely of water (oil), and so on. The domain walls in the microemulsion phase are disordered and there is no preferred direction of ordering, so, on average, the width of these concentric shells reflects the characteristic interlamellar ordering (see Figs. 10 and 11). Of course,

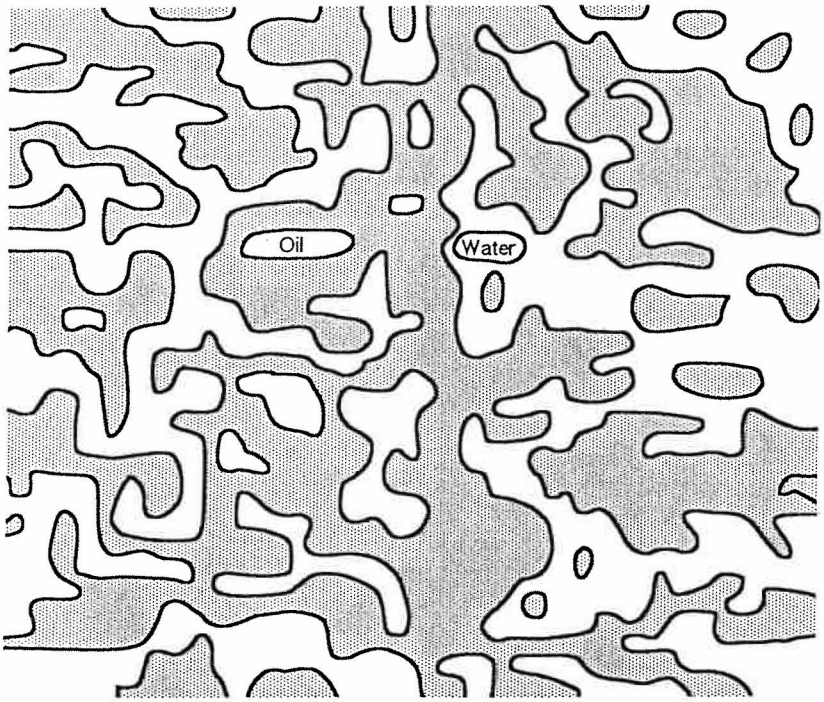


Fig. 10 Instantaneous configuration of bicontinuous microemulsion from Monte Carlo simulation for point $j = 1.3$, $m = -0.4$, $\mathcal{P} = 0.0$, the point to which Fig. 7 also corresponds. The regions occupied by positive spins (oil) have been shaded to give a better visual representation of the structure. The boundary between oil and water has also been smoothed.

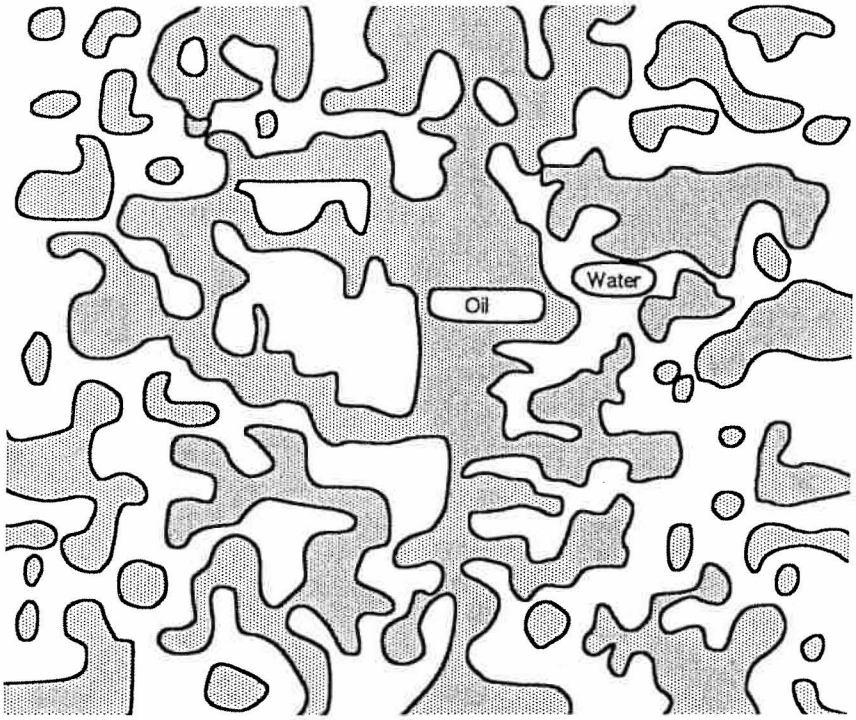


Fig. 11 Instantaneous configuration of bicontinuous microemulsion from Monte Carlo simulation for point $j = 1.2$, $m = -0.4$, $\mathcal{P} = 0.0$, the point to which Fig. 8 corresponds. Note that there are only relatively subtle structural differences from Fig. 10, but Fig. 8 possesses a pronounced finite q peak, whereas Fig. 7 does not.

as one becomes sufficiently far (in units of ξ) removed from the central site, the amplitudes of the oscillation in the spin–spin correlation function become small and the distinction between on-average oil and water domains becomes less distinct because of the fluctuations of the domain walls. However, if ξ is moderately large in comparison to d, the picture that one has of the isotropic phase is that of fairly distinct domains of oil and water separated by amphiphilic film layers that although somewhat crumpled, tend to keep a fixed average distance from neighboring layers. Since the concentration of amphiphile at the bond $n + \frac{1}{2}$ is given by

$$\rho^{AB}_{n+1/2} = \tfrac{1}{2}(1 - \langle \sigma_n \sigma_{n+1/2} \rangle) \tag{64}$$

we see that for low amphiphile volume fractions, one requires $\langle \sigma_n \sigma_{n+1/2} \rangle$

to be close to unity (i.e., spins on nearest-neighbor sites should be well correlated). This is favored by large ξ and d and therefore we expect to find the bicontinuous microemulsion phase with low amphiphile concentration near an order–disorder transition where the period of the lamellar ordering is long. On the basis of simulations this is indeed found to be the case for the lattice model. It would appear that this observation is also consistent with experimental phase behavior.

The discussion above may be illustrated using instantaneous configurations from a simulation of the lattice model. In Figs. 10 and 11 we have shown two sample configurations for those values of the parameters used to compute the structure factors in Figs. 7 and 8. Figure 10 corresponds to the structure factor in Fig. 7, and Fig. 11 corresponds to that of Fig. 8. One striking observation that one might make about the bicontinuous microemulsion structure is that it is difficult to identify distinct well-formed layers separated by oil and water. Rather, there are domains which in some regions appear to be lamellar-like, in others almost globular. The indefiniteness of lamellae should be taken seriously. Indeed, since they reproduce the scattering data so well, one should view the images in Figs. 10 and 11 as truly representative of that phase we call bicontinuous microemulsion.

The fact that the bicontinuous microemulsion is characterized by two such lengths means that one must be careful about any definition of a domain size. Thus in simple systems there is a direct correspondence between the correlation length and the domain. In the present case, however, the average size of oil or water regions is essentially d and fluctuations about it are determined by ξ. Thus the lamellar structures have well-characterized domains of oil or water because the fluctuations of the amphiphilic film are modest. In the microemulsion there remains an average d, but if ξ is small the amphiphile fluctuations are so large that it is inappropriate to regard this as characteristic of an oil or water domain. It is possible to study this issue a little more carefully by calculating the amphiphile–amphiphile correlation function,

$$\langle \rho^{AB}_{n+1/2} \rho^{AB}_{n+1/2} \rangle \tag{65}$$

This function directly probes the issues discussed above. The distance between peaks (say $i = -j$) is precisely the size of an oil region, and the dispersion is also readily calculated. However, to retain direct contact with the neutron-scattering data, we shall, for the remainder of this chapter, rely more on heuristic definitions that are based on the two lengths, ξ and d. We shall use these ideas once more in Section VII, where we discuss the definition of microemulsion.

V. INTERFACIAL TENSIONS AND STRUCTURE OF THE INTERFACES BETWEEN PHASES IN THE LATTICE MODEL

Within the context of mean-field theory one may also construct the interfacial profiles and tensions between the various phases that are in equilibrium in the phase diagrams, Figs. 3 to 6. The interfacial tensions derived from such calculations are expected to be qualitatively correct [31]. One may also heuristically associate the mean-field profiles with the interfacial structure of small regions of the interface. As we shall see, the tensions calculated for the present lattice model are characteristically low, even when there is no proximate critical point. Thus in the region of three-phase oil–water–lamellar coexistence, one finds that the oil–lamellar phase or water–lamellar phase tensions are all ultralow (of order $10^{-4}kT$). One also finds interfacial structure rather reminiscent of liquid–crystalline order at the interface between the oil-rich and water-rich phases [32,33]. We have already alluded to the fact that this implies the existence of a nontrivial length scale, besides the correlation length, at the interface between isotropic phases. The precise evolution of these length scales as a function of the parameters of the model is readily determined numerically, or approximately by a perturbation theory based on the susceptibility of the isotropic phases [33]. It is also worth noting that this phenomenon is intimately related to the peak in the scattering intensity that was discussed in the preceding section. In both cases one finds that the soft modes, as determined by the momentum-dependent susceptibility, occur at nonzero momentum. In the case of the interface they lie in the complex-momentum plane and the real and complex pieces then determine the two length scales present in the interface. In the calculations presented below, in Figs. 12 to 15, we plot the particle density for each layer, z. These results are obtained by averaging the free-energy functional (13) over the x,y directions and then minimizing with respect to the layer densities, subject to the asymptotic conditions for the two bulk phases that are in equilibrium.

In Fig. 12 we plot a typical oil–water interface along the three-phase equilibrium surface between oil, water, and lamellar phase. We note the symmetrical interfacial structure. Figures 13 and 14 correspond, respectively, to an oil–water interface and oil–microemulsion interface that are on the three-phase equilibrium sheet but somewhat away from the lamellar phases. For this reason the tension is still not extremely low, nor is there any discernible interfacial structure. Also, for these values of the parameters the oil–water interface is found to be wet by the third, microemulsion

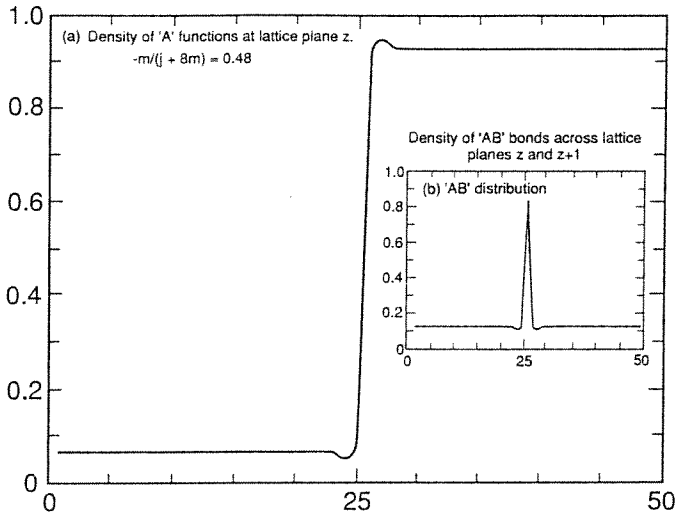


Fig. 12 Typical density plot for oil–water interface for a point on the oil, water, liquid–crystal equilibrium surface. Note the presence of an interfacial structure. On the inset we have presented the amphiphile density distribution across the interface. The interfacial tension corresponding to this profile is ultralow.

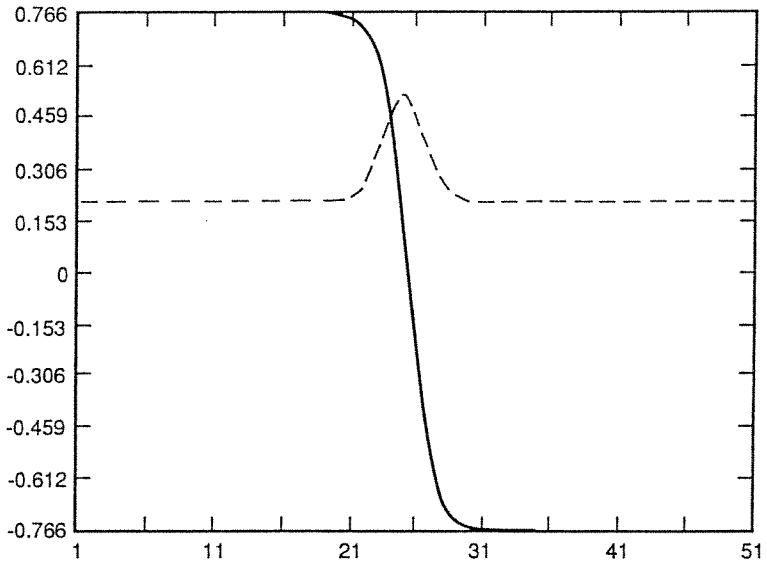


Fig. 13 Typical oil–water interface for a point on the three-phase equilibrium surface between oil, water, and isotropic phase, but far from curve L.

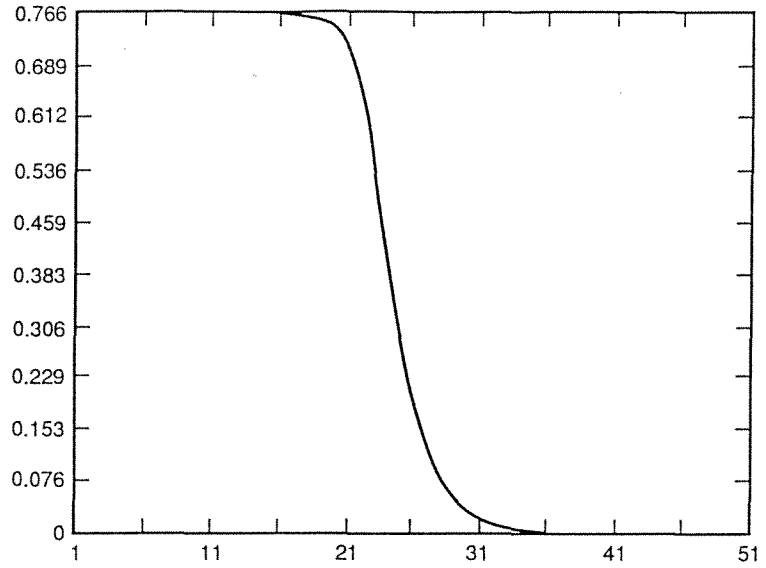


Fig. 14 Typical microemulsion–water interface for a point on the three-phase equilibrium surface between oil, water, and isotropic phase, but far from curve *L*.

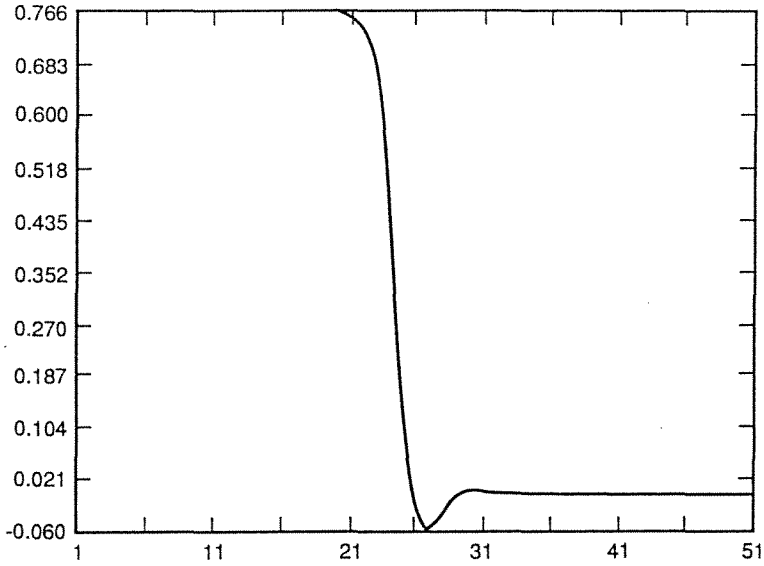


Fig. 15 Typical microemulsion–water interface for a point on the three-phase equilibrium surface between oil, water, and isotropic phase close to curve *L*. At this point the isotropic phase possesses the features of a bicontinuous microemulsion, and this is associated with structure in the microemulsion side of the interface.

phase, a matter to which we shall return in some detail at the end of this section.

The next result, Fig. 15, is derived for a point on the isotropic three-phase equilibrium surface that is much closer to the lamellar phase. Consequently, one might expect the appearance of interfacial structure. However, although the oil–water interface (not shown) is structureless, the oil–microemulsion or water–microemulsion (Fig. 15) interface has structure, confined primarily to the microemulsion side. This is an observation that might be checked experimentally. One can show, in fact, that within the present model this particular type of interfacial structure is commonly associated with Winsor III equilibria and may, in some measure, be used to differentiate between these and conventional three-phase equilibria.

At this point we may pause to review the origins of ultralow interfacial tension, at least within the framework of the present lattice model. Actually, many of the important applications of amphiphilic dispersions stem from the remarkably low interfacial tensions found between the phases in the three-phase triangle, so this question has caused considerable interest and controversy in the literature. More recently, it has become common to attribute the phenomenon to the near-cancellation of the bare surface tension and the transverse component of the pressure [34–36]. There is both experimental [37] and theoretical evidence [36] that such effects are indeed significant. However, the predictions of the lattice model imply that resolution of the question may not be so simple and that such cancellations lead to low, but not ultralow, tension. To see that this is the case we recall that because of the constraint that there are no direct oil–water contacts, all of the amphiphilic film satisfies the Schulman condition [38].

We have, in addition, chosen the zero of energy to be that of flat amphiphilic film, so that at any temperature the tension of a perfectly flat interface would be zero. This choice ensures that the conjectured condition for ultralow tension is automatically satisfied by construction of the lattice Hamiltonian. However, finite temperature effects cause fluctuations that disrupt and bend the interface in a number of ways, and it thereby acquires an effective tension. This tension has been calculated for the three-phase equilibrium surface between oil–water and the isotropic phase of Fig. 3. The tension is low across most of the surface, becoming ultralow only near a tricritical point or near the multistate surface defined by (10) or (46). We had earlier identified this multiphase surface by requiring that the work required to insert an amphiphile into the film vanish. Note, therefore, that one appears to require this further constraint on the microscopic parameters and chemical potentials of the model before the tensions become comparable to those that are conven-

tionally called ultralow. Given the integrity of the predictions of the model so far, one expects that such a constraint must also be satisfied by experimental systems with ultralow tension, an observation that could readily be evaluated experimentally. In fact, as we shall see later, the picture offered by the microemulsion model leads one to suppose that the practical rules of thumb at present used to predict ultralow tensions are actually a reflection, albeit an imprecise one, of the constraints implied by the present model. We note finally that since this explanation of the origin of ultralow tension does not require proximity to any critical point, there is no implication that the compositions of the three phases should be nearly the same. In fact, typically the three-phase Winsor III equilibrium in this model are far from being critical when one considers compositions of the components. This point is important since it rationalizes some of the apparently conflicting observations on these systems that were mentioned in Section I. This point is connected to another puzzling observation; thus the oil–water interface of a Winsor III state is typically nonwet by the microemulsion [7], although there has been at least one observation of a transition to wetting [39].

In the calculations described above one notes that for values of the parameters near the multistate surface, and consequently for a structured interface, the oil–water interface is not wet by the microemulsion phase. We may pursue the issue of wetting across the Winsor progression somewhat more systematically using mean-field calculations of the type described above.* In this case, however, one must study the phase diagram as a function of the coupling constants H and L that break the inversion symmetry. This permits us to locate the critical endpoints that bound the Winsor progression. As mentioned in Section II, these parameters involve, respectively, the difference in chemical potentials of oil and water and the spontaneous curvature of the amphiphilic film. These quantities in turn reflect the changes in the concentration of oil and water and in the amount of brine or cosurfactant present in the system. For ionic surfactants the Debye screening induced by the addition of salt can significantly affect the propensity of the amphiphilic film to bend toward water regions, thereby reducing the magnitude and removing the symmetry in the edge and corner energies of the lattice model. For our purposes it is important to note that most of the freedom present in the parameter space

* Note carefully that the Winsor progression is parameterized by only two of the microscopic variables. This is a consequence of the fact that since we are dealing with uniform phases, j and m appear in the phase-equilibria equalities only as the total coefficient of the two-body term. On the other hand, the actual calculation of interfacial tension is for a fixed choice of the ratio j/m. Thus there exist Winsor III progressions for which the bulk volume fractions of the components are precisely the same, but for which the interfacial tensions are quite different.

of the Hamiltonian is removed if one chooses to study such phase equilibria. The interfacial tensions and contact angles measured for these phase equilibria are dependent on relatively few or none of the parameters of the model. This is a useful consistency check on our understanding of the meaning of the microscopic parameters in the model.

We examine the surface of isotropic three-phase equilibria of the symmetric model. The three interfacial tensions may be computed and a curve of wetting transitions determined (see Figs. 16 and 17). To one side of the curve $w(j,m,\mathcal{P})$ the isotropic surfactant-rich phase wets the oil–water interface. On the other side (shaded) the oil–water interface is nonwet, as might be expected of a microemulsion. Note that as one proceeds to the line of tricritical points, the extent of the region for which the interface

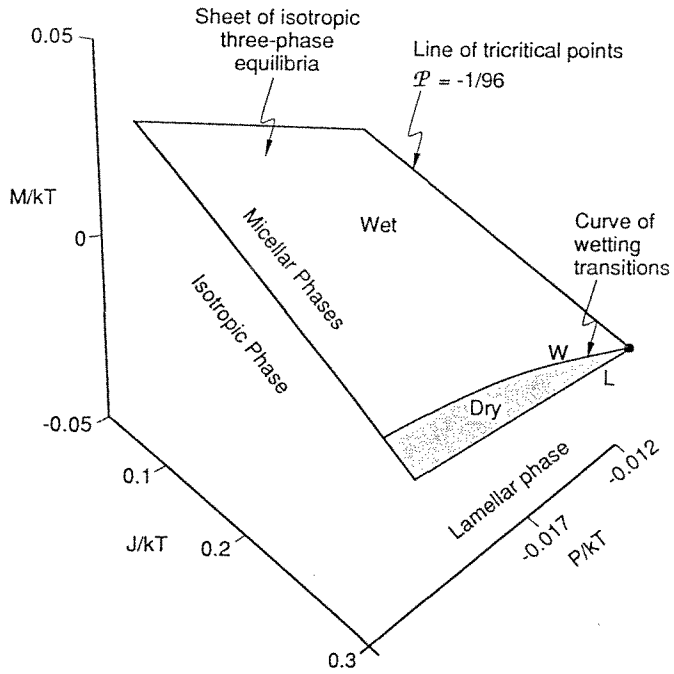


Fig. 16 Sheet of isotropic three-phase equilibria in space of parameters j, m, $\mathcal{P}$. Above the surface one has the oil and water micellar phases and beneath it the isotropic phase. In the vicinity of curve L the isotropic phase possesses the properties of microemulsion. Note that on the shaded portion of the surface the oil–water interface is not wet by the isotropic (microemulsion) phase. Curve $w(j,m,\mathcal{P})$ thus represents a curve of second-order wetting transitions. Curves L and W are fairly close together, so in accordance with most experimental observations, we conclude that microemulsion tends not to wet the oil–water interface.

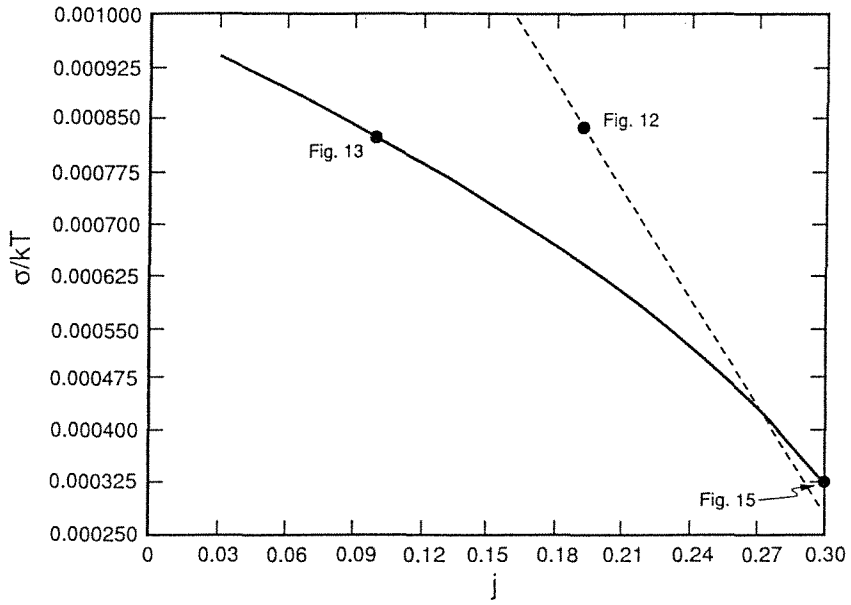


Fig. 17 Location of a wetting transition for the oil–water interface. The dashed line is $2\sigma_{ow}$, while the solid curve is $\sigma_{\mu w}$. At the crossing the microemulsion wets the oil–water interface. The points marked on the curves indicate the figures where density profiles for those values of the parameters have been presented.

is nonwet shrinks, reflecting the normal trend that wetting accompanies the vanishing of interfacial tensions near a tricritical point. However, there is a novel phenomenon present in the lattice model of microemulsions. The interfacial tensions become small as one proceeds in either of two directions on the equilibrium sheet. Thus, as $\mathcal{P}$ becomes small the tensions vanish as one approaches the line of tricritical points. However, for fixed $\mathcal{P}$ the tension also becomes small as one approaches the multistate sheet (or tends to the curve L) because, as we have explained, the work of inserting an amphiphile molecule into the amphiphilic film is vanishing. This second mechanism for lowering the tensions is entirely unrelated to near-criticality and therefore carries with it no implication that the interface should be wet nor that the volume fractions of the components in different phases should be the same. On the contrary, the contact angles actually increase with decreasing oil– and water–microemulsion tensions. One can readily see that the issues of wetting, low interfacial tensions, and near-criticality (as determined by the volume fractions of the components) are subtle and potentially confusing if one

does not understand the global nature of the phase diagram. Presumably, the most common three-phase equilibria that are called Winsor III states are those that lie inside the nonwet region of the equilibrium surface. In some cases they may also be proximate to a tricritical point, reflecting a small corner energy term. Also, in some experiments one is probably close to both of the critical endpoints, a matter to which we shall presently return. In either case the tensions would be lowered as a consequence of being close to a critical region that we have mentioned. However, it is probable that the dominant effect in lowering the tension is proximity to the multistate surface rather than any of these critical points. The interface would then be nonwet because the third phase, whether it be liquid–crystal or microemulsion, always has a higher tension with oil or water than in the case of critical wetting.

Although this phenomenon has been reproduced with numerical calculations, it is also useful to develop a more intuitive understanding. The tensions between all of these interfaces are quite low at low temperatures, and for the case of the oil–water interface is significantly modified only by the in-layer fluctuations of the interface. However, the tensions with the lamellar phases or microemulsion are also affected by the fact that the principal fluctuating layer that defines the interface is hindered by the other amphiphilic films in the middle phase. This observation is closely related to the fact that there is a second, well-characterized length scale (d) present in the lamellar and bicontinuous phases (i.e., the average distance between the layers). As a consequence, these interfacial tensions tend to be higher than those between the simple oil–water interface, where one has only micelles on either side of the interface. This means that if the third phase is sufficiently structured, that is, contains sufficient flat amphiphilic film, it will tend to form a lens rather than spreading out to wet the oil–water interface. Thus even though all the interfacial tensions may be becoming lower on approach to the multistate surface, the amphiphilic film within the bicontinuous phase is becoming more flat and has a greater tendency to dampen the interfacial fluctuations of the principal amphiphilic monolayer that defines the interface. Furthermore, the propensity of the middle phase to wet the oil–water interface is to a large degree determined by the lengths ξ and d since these determine the degree to which the principal monolayer is hindered. Finally, in Fig. 18 we have presented an instantaneous simulation configuration of an interface between water-rich and microemulsion phases. This example is presented to give a qualitative impression of the water/bicontinuous microemulsion interface at the Winsor III state. The configuration was prepared by first locating the Winsor III equilibrium by Monte Carlo simulation of the heat capacity and energy. An initial condition for the interfacial simulation was

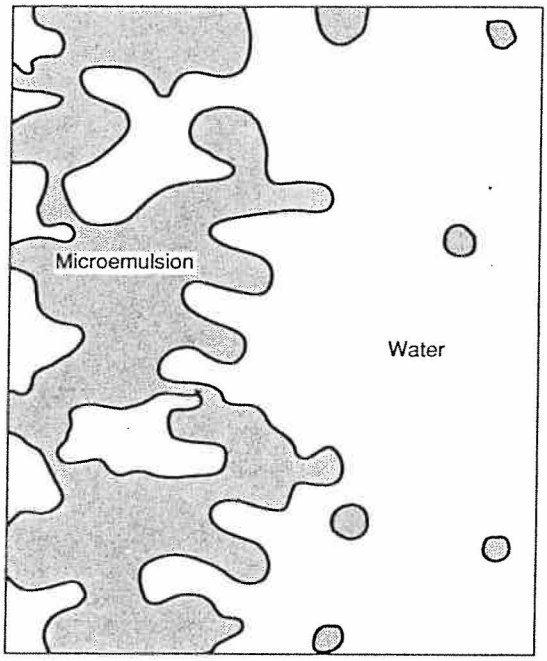


Fig. 18 Instantaneous configuration from a Monte Carlo simulation of the bicontinuous microemulsion–water interface. This calculation was carried out for a point on the three-phase equilibrium surface represented by Fig. 16. Notice the degree of roughness of the interface, a phenomenon that would not be present for a simple liquid–gas interface. The structure in the microemulsion side of the interface that was apparent in the mean-field calculations (Fig. 15) arises from the large domains that aggregate behind the principal amphiphilic film. Note carefully that this particular point was located using Monte Carlo simulation. It thus represents an interface at the Winsor III phase equilibrium.

then constructed by filling half of the cube with equilibrated microemulsion and half the cube with an equilibrated water configuration. Average-structure von Neumann boundary conditions are then applied and the simulation reequilibrated. However, we note that the interface is much rougher than one would expect from liquid–gas interface. In addition, there is a tendency for large micellar aggregate to form on the microemulsion side, this being reflected in the mean-field density profile of Fig. 15.

Note that we can now understand one more of the interrelations between experimental observations. Thus the condition [Eq. (46)] that results in low tensions by stabilizing large amounts of amphiphilic film also

results in a nonwet oil–water interface. Such relationships may mean that the various attempts in the literature to resolve between classical three-phase equilibria and the Winsor III state for oil–, water–, and bicontinuous-microemulsion are quite closely connected, even if they are not in quantitative agreement. This issue of how one should distinguish microemulsion is an interesting matter, and we shall return to it in Section VI.

We now return to the general topic of interfacial tensions in the Winsor III three-phase equilibrium. In particular, we wish to establish the idea that the present lattice model is capable of reproducing the characteristic pattern of tensions to which we referred in Section I. One would certainly expect this to be so since as we commented previously, the extended parameter space contains the oil–microemulsion and water–microemulsion critical endpoints. In addition, we know that the symmetric model ($H = L_1 = L_2 = 0$) that contains the bicontinuous microemulsion phase possesses a region of nonwet three-phase equilibria. The essential ingredients of the interfacial tension plots are therefore already present. In Figs. 19 to 21 we have presented calculations of the oil–water (σ_{ow}), oil–isotropic phase ($\sigma_{o\mu}$), and water–isotropic phase ($\sigma_{w\mu}$) tensions between the critical endpoints. In the absence of constraints beyond those implied by the phase equilibria, it is possible to choose an arbitrary relationship between j and m. This will in turn select a trajectory that for the symmetric three-phase condition ($l_1 = l_2 = h = 0$) corresponds to a point on the three-phase equilibrium sheet of Fig. 16. We therefore present plots for $j + 11m = 0$ (Fig. 19), $j + 12m = 0$ (Fig. 20), and $j + 13m = 0$ (Fig. 21), corresponding to oil–water interfaces that are dry, undergoing a wetting transformation, and wet by isotropic phase, respectively. Recall also that the isotropic phase tends to have the properties of microemulsion only for the first two choices of the j,m relation, these being quite close to curve L of Figs. 3 and 16. The oil–water interface is nonwet for the first case ($j + 11m = 0$), in accordance with the experimental observation that, generally, Winsor III oil–water interfaces are not wet by microemulsion. It would in principal, be possible to make another selection that would produce a wet interface. It should be noted that although the Winsor III interface is typically nonwet, a transition to wetting has been observed in at least one case [39]. Such a situation is possible in the present model, but one would then predict that the peak in the SANS data would move to a smaller wave number. It would be interesting to check these ideas with measurements of tensions and of SANS experiments.

In summary, then, we have observed that Winsor III states tend to have nonwet oil–water interfaces because of the proximity to the mul-

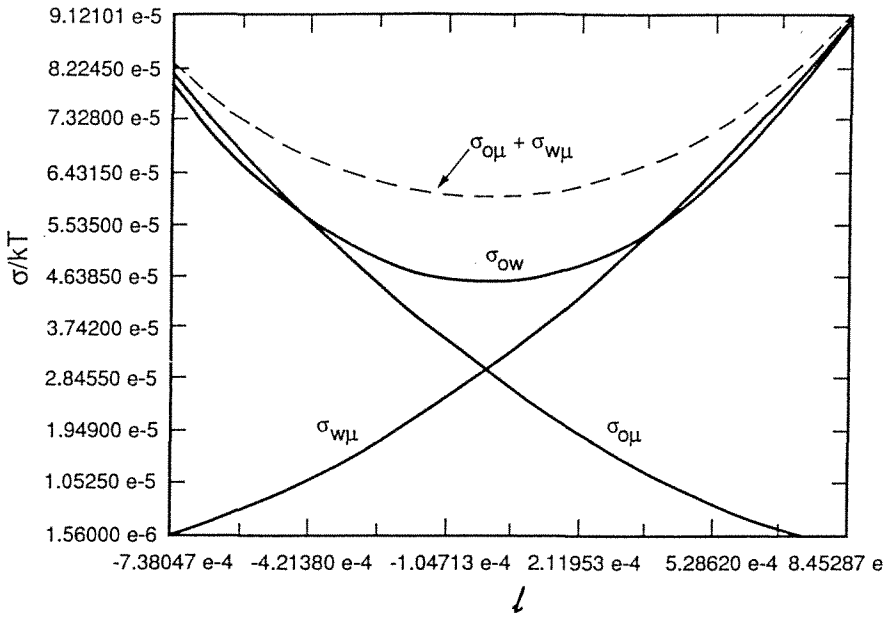


Fig. 19 Interfacial tensions across the Winsor progression. The left- and right-hand sides correspond, respectively, to water–microemulsion and oil–microemulsion critical endpoints. This particular trajectory ($j + 11m = 0$) crosses the three-phase surface (Fig. 16) to the right of the wetting curve, so the oil–water interface is not wet by isotropic phase. Also, since we are close to curve L, the isotropic phase possesses the properties of microemulsion. These predictions are all in accord with experimental observation.

tiphase sheet. Also, for Winsor III states, there are peaks in the experimental SANS data, and these tend to occur at small wave number. Similarly, in the present model one finds that as one proceeds along the three-phase surface toward the multiphase sheet, one begins to see the emergence of a second length scale in bulk and interfacial properties, this reflecting the proximity of an ordered phase on the phase diagram. The two phenomena described above, the general absence of a tendency to wet and a secondary length scale, are related in the present model and, we believe, in the experiments. Thus in this section we have commented that the flattening of the amphiphilic film causes the oil–water interface to be nonwet, whereas in Section IV we have shown that this same aspect is accompanied by a second length scale in the structure factor.

Having established the capacity of the model to describe these phe-

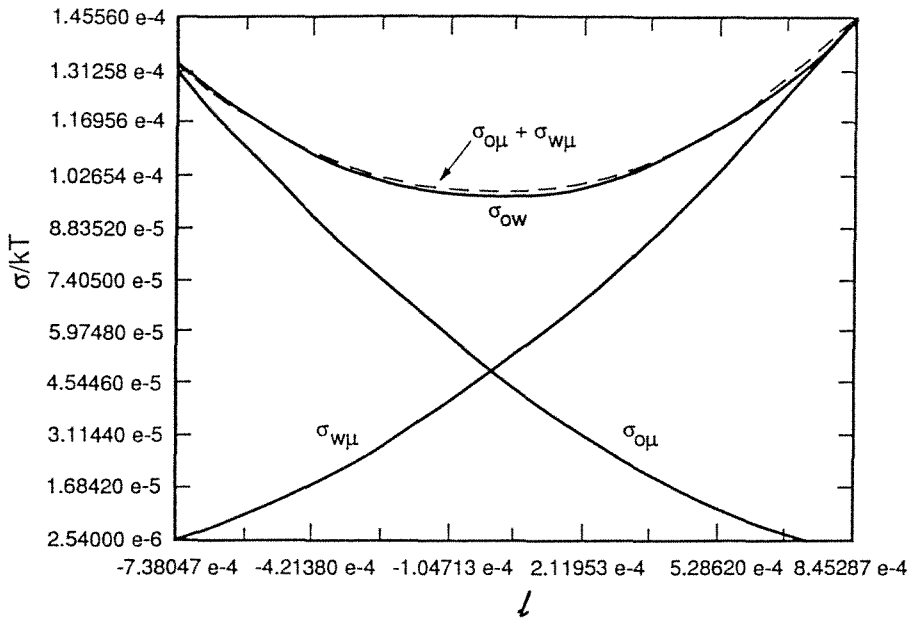


Fig. 20 Interfacial tensions across the Winsor progression. The left- and right-hand sides correspond, respectively, to water–microemulsion and oil–microemulsion critical endpoints. This trajectory ($j + 12m = 0$) crosses the three-phase surface (Fig. 16) very close to the wetting curve, so the oil–water interface is almost wet by isotropic phase. Also, since we are fairly close to curve L, the isotropic phase still possesses the properties of microemulsion. Indeed, this may be viewed as a marginal microemulsion (see Section IV).

nomena, we may now turn to a rather old but practical question about the nature of an "optimal" microemulsion. In early experimental studies it was realized that a number of technologically important features of microemulsion are associated with the minimum in the oil–water interfacial tension that is found in the symmetric bicontinuous portion of the Winsor III state. From the arguments above, one may establish an understanding of the origins of this minimum in the oil–water tension. This permits us to establish the correspondence of microscopic interactions to the classical definition of an optimal microemulsion. However, in the space of parameters of this model there exist other degrees of freedom that might be exploited in establishing a more refined definition. A number of such questions have yet to be studied.

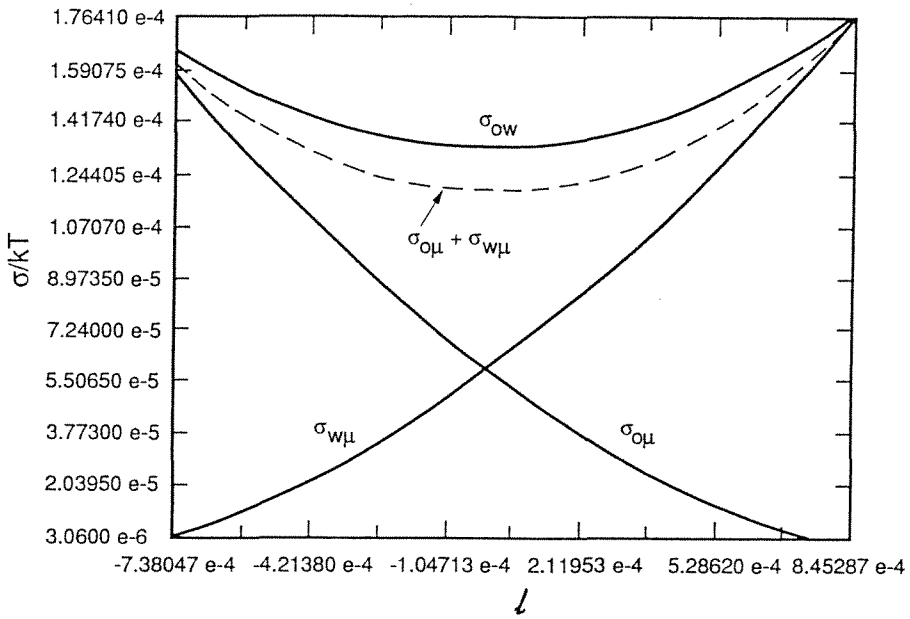


Fig. 21 Interfacial tensions across the Winsor progression. The left- and right-hand sides correspond, respectively, to water–microemulsion and oil–microemulsion critical endpoints. This trajectory ($j + 13m = 0$) crosses the three-phase surface (Fig. 16) to the left of the wetting curve, so the oil–water interface is wet by isotropic phase. Also, since we are far from curve L, the isotropic phase should not be viewed as microemulsion.

VI. CONCLUSIONS AND REFLECTIONS ON VARIOUS ASPECTS OF MODELING COMPLEX FLUIDS

While our understanding of simple fluids is by now fairly satisfactory, the same cannot be said of very complex fluids. In particular, for amphiphilic systems our understanding of the structure of the phases and the transitions between them is comparatively primitive. The large number of degrees of freedom involved in these systems render the conventional methods of calculation and simulation less effective and if one seeks to understand the global phase-diagram one must model the system in a different manner than that conventionally used for simple fluids. Such models should certainly make contact with the microscopic energy scales, though at the minute one cannot hope that the correspondence will be very direct. It is also essential that they be analytically and computa-

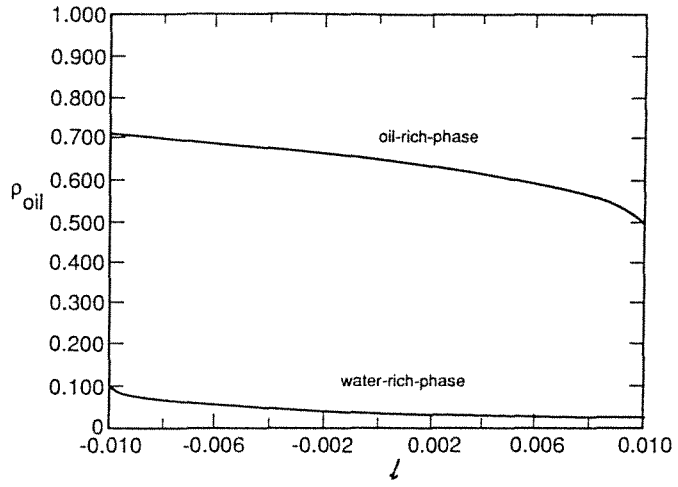


Fig. 22 Volume fraction of oil in the oil-rich phase (upper curve) and water-rich phase (lower curve) across the Winsor progression of Fig. 19. For $l = 0$ one has the Winsor III equilibrium, and here the difference in the fractions of oil in the phases is still substantial. We have thus reproduced the effect that the interface may have ultralow tension and be nonwet (Fig. 19) while on the basis of composition, still noncritical. This was one of the puzzling experimental observations with which we opened this chapter.

tionally tractable since the phenomena we seek to study are frequently quite subtle. That the present mode, based as it is on so few basic hypothesis about the interactions, should be so successful in reproducing and predicting the phenomena is quite gratifying. There is every reason to believe that it encompasses most of the experimental data for these systems. One should, however, also be aware of the limitations of such studies. These are both fundamental and, in some cases, of a more technical nature.

At the fundamental level one must accept that the lattice models, though they exhibit the generic phenomena and elucidate the origins and interrelations between them, cannot be quantitative. Such models are studied as a function of their parameters, so one may make predictions about the effect of changing various experimental parameters, but there is, as yet, no way of a priori calculating their values. On the other hand, the model requires very few parameters, and the unknowns are usually overdetermined with just one physical measurement. For example, SANS data can be used to determine $I(0)$, b, and c, and two of these fix the

microscopic couplings and concentration.* The concentration may be computed independently from the model and one can also use the parameters to calculate the interfacial tensions, contact angles, and other observable quantities and compare them to experiment. Thus in the future it is most desirable that SANS experiments be carried out simultaneously with interfacial, composition, conductivity, heat capacity, and other measurements. Greater emphasis might also be given to specular reflection experiments that will probe the structure of the oil– and water–microemulsion interfaces. This should permit one to determine if there are indeed secondary long length scales present, and to relate this phenomenon to the issue of wetting.

We might also point out that the lattice models prompt the development of a more intuitive level of understanding of complex systems, but such arguments may, in the context of the lattice model, always be reduced to concise well-formulated mathematical questions. This is a powerful combination with which to attack such complex problems.

There are, of course, also some technical problems that have yet to be overcome. Mean-field calculations offer the most complete description of the lattice Hamiltonian, but they are frequently quantitatively unreliable. In the region of the Winsor III states mean-field theory offers no information about the structure of the phases, and although the length scales we study are long compared to the lattice spacing, there are few useful fixed points of a renormalization group flow. Low-order loop corrections have also been shown to be of some limited quantitative value [42]. Many of the important observables, when calculated with simulations, are susceptible to anomalously large finite-size effects, although this problem could certainly be ameliorated by systematic MCRG calculations. Finally, the presence of the underlying lattice, while essentially irrelevant for the bicontinuous microemulsion, affects the liquid–crystalline phases in a number of important ways. First, only order with a symmetry that is appropriate for the lattice is reproduced by the lattice Hamiltonian. Thus tubular phases become arrays of square tubes rather than cylinders, and presumably some classes of phases are entirely excluded. More significantly, even for phases with the correct symmetry, the fluctuations may be disturbed by the presence of a lattice. For example, realistic layers of the fluid lamellar phases undergo long-wavelength transverse fluctuations that control their swelling on the addition of oil or water. At low temperatures the lattice suppresses such fluctuations, although they are restored above the roughening transition in the lattice models. Microemulsion exists only well above the natural rough-

* Heat capacity and composition measurements may be used to fix the value of $\mathcal{P}$.

ening temperatures of the lattice, so these considerations do not affect this phase. One must, however, exercise caution in interpreting anything but long length-scale phenomena when dealing with a lattice model. The phases that lie at high amphiphile concentration have order on relatively short lengths. Nevertheless, we would contend that the liquid–crystalline order predicted by the lattice model is a useful caricature of the true ordered structures.

We now turn to a discussion of some of the results presented in this chapter and their broader interpretation. In Section II we referred to the question of the definition of microemulsion. Recall that it is possible to follow the single-phase corridor from oil-in-water to water-in-oil micelles via the bicontinuous microemulsion phase without passing through any phase transition. One must thus seek a definition that is based on criteria other than singularities of derivatives of the free energy. From the experimentalist perspective, microemulsion is often defined heuristically in terms of its physical properties. Thus the fluid is an isotropic nonviscous dispersion of oil and water that exhibits certain interfacial properties and whose location in the phase diagram has a characteristic topological relationship with the other phases.

More recently, Kahlweit [2] has suggested another interesting description based on, for example, the convergence of certain physical measurements (measured within the water or oil domains) to their bulk values. One type that has been suggested is the dielectric measurement of the water domains. Evidently, such a definition will ensure that the oil or water domain size is quite large. This scheme must ultimately be related to the sizes of the length scales ξ and d mentioned in Section IV since these determine the extent and nature of the dispersed oil or water regions.

Various definitions based on theoretical considerations have also been advanced. Widom has pointed out that since true Winsor III microemulsions tend not to wet the oil–water interface, it may be appropriate to use this as the criterion for their definition. On the other hand, some materials that are commonly considered to be microemulsion do wet the interface, while some fluids that are not, on the basis of other physical properties, do not wet the interface. This classification is evidently not unique, but from the results in Section V we see that it is suggestive.

Other classifications are based on the observation that the microemulsion phase tends to yield a maximum in the SANS scattering intensity. This is also found to be the case for the lattice models and there one can calculate a surface that divides the isotropic phase into two regions, one where the maximum of the structure factors is at $q_{\max} = 0$ and the other, $q_{\max} > 0$. This definition is essentially based on the emergence of a second length scale, d. Although this suggestion is mathematically ap-

pealing, it is unlikely that materials with very small d will possess the other physical properties of microemulsion. It is also unclear that a large d but small ξ would represent an acceptable definition. At the minute the most reasonable choice seems to be based on an expression of Kahlweit's ideas. Thus the second length scale ($d = d_c$) ought to be sufficiently large that the implied water or oil domains take on certain macroscopic properties. However, for this criterion to be workable one must presumably also require that ξ be comparable to d since this will ensure a fairly monodisperse domain size. When such a choice is made, one can again calculate a sheet that divides the isotropic phase into two regions but with $q_{max} < q_c$ or $q_{max} > q_c$ where $q_c = 2\pi/d_c$. It transpires that such a surface lies to the right of the sheet $j + 12m = 0$ in Figs. 3 and 4. Finally, if one is sufficiently close to the order–disorder transition and near this sheet, it is clear that both d and ξ will have the desired properties.

Ultimately, however, one must accept that the progression from conventional three-phase equilibria to the Winsor III state is so gradual that it is difficult to resolve the question in a completely satisfactory manner. All of the definitions mentioned above are based on some important elements of microemulsion, and each has their own merits. It is probably more important to recognize that all are simple reflections of a remarkably rich underlying scheme that is made clearly visible by the lattice models, an observation that is in itself sufficient to justify their development.

In summation we may reflect on those tasks that have yet to be achieved before a truly quantitative theory of microemulsion emerges. Although the essential features of microemulsion are described by the present model there are aspects of the phase-diagram and scattering functions that must be calculated using methods that can accurately describe the long but not diverging length scale. Some of this work has been carried out [35,40]; some has yet to be undertaken. There is, in particular, a need for more accurate but simple analytical expressions, such as (51) to (53), to which experimental data may be fit.

Finally, we have hitherto directed out attention to the equilibrium properties of these systems, and the study of their nonequilibrium behavior is its infancy. It is already clear (J. R. Gunn, C. M. McCallum, K. A. Dawson, unpublished) that there exist a number of important phenomena, their elucidation being important technologically and of substantial intrinsic scientific interest.

ACKNOWLEDGMENTS

During the period that these ideas were developed, one of us, K.A.D., benefited greatly from conversations with Professor B. Widom, who introduced him to the study of microemulsion. Numerous other helpful

insights were gained from conversations with Professor M. E. Fisher, Professor N. Boden, and Dr. M. D. Lipkin. The hospitality of the Institute for Theoretical Physics is also acknowledged, as are interesting conversations with Professor M. Schick that took place during that stay.

The manuscript benefited greatly from the advice and encouragement of Professor S. S. Elberg. Its preparation was assisted by the staff of the Tahoe Marina Lodge, particularly Joanne Fanucchi.

Financial support for the research was provided by the Dreyfus and Sloan Foundations and by a Fellowship to C.J.M. from the Regents of the University of California.

REFERENCES

1. (1988). General patterns of the phase behavior of mixtures of H_2O, nonpolar solvents, amphiphiles, and electrolytes, 1, *Langmuir 4*: 499. (1989). General patterns of the phase behavior of mixtures of H_2O, nonpolar solvents, amphiphiles, and electrolytes, 2, *Langmuir 5*: 805.
2. Kahlweit, M., Strey, R., and Busse, G. Microemulsions: a qualitative thermodynamic approach, *J. Phys. Chem.* (in press).
3. Shinoda, K., and Fruberg, S. (1975). Microemulsions: colloidal aspects, *Adv. Colloid Interface Sci. 4*: 281.
4. Langevin, D. (1987). Microemulsion, *Acc. Chem. Res. 21*: 255.
5. Winsor, P. A. (1954). *Solvent Properties of Amphiphilic Compounds*, Butterworth, London.
6. Reed, R. L., and Healy, R. N. (1977). In *Improved Oil Recovery by Surfactant and Polymer Flooding* (D. O. Shah and R. S. Scheckter, eds.), Academic Press, New York.
7. Kunieda, H., and Shinoda, K. (1982). Correlation between critical solution phenomena and ultralow interfacial tensions in a surfactant/water/oil system, *Bull. Chem. Soc. Jpn. 55*: 1777.
8. Pouchelon, A., Meunier, J., Langevin, D., Chatenay, D., and Cazabat, A. M. (1980). Low interfacial tensions in three phase systems obtained with oil–water surfactant mixtures, *Chem. Phys. Lett. 76*: 277. Cazabat, A. M., Langevin, D., Meurier, J., and Pouchelon, A. (1982). Microemulsion: colloidal aspects, *Adv. Colloid Interface Sci. 16*: 175. Bellocq, A. M., Bourbon, D., Lenamceau, B., and Fourche, G. (1982). *J. Colloid Interface Sci. 89*: 427.
9. Bennett, J. E., Hatfield, J. C., Davis, H. T., Macosko, C. M., and Scriven, L. E. (1982). In *Microemulsions* (I. D. Robb, ed.), Plenum Press, New York, p. 65. Safran, S., Webman, I., and Grest, G. S. (1985). Percolation in interacting colloids, *Phys. Rev.* Lagues, M., Ober, R., and Taupin, C. (1978). Study of structure and electrical conductivity in microemulsions: evidence for percolation and phase inversion, *J. Phys. Lett. 39*: L-487. Cazabat, A. M., Chatenay, D., Langevin, D., and Meunier, J. (1982). Percolation and critical points in microemulsion, *Faraday Discuss. Chem. Soc. 76*: 291.

Clarkson, M. T., Beaglehole, D., and Callaghan, P. T. (1985). Molecular diffusion in microemulsions, *Phys. Rev. Lett. 54*: 1722.

10. Teubner, M., and Strey, R. (1987). Origin of the scattering peak in microemulsions, *J. Chem. Phys. 87*: 3195.
11. Chen, S. H., Chang, S. L., and Strey, R. On the interpretation of scattering peaks from bicontinuous microemulsions, preprint.
12. Chen, S.-H., Chang, S.-L., Strey, R., and Kahlweit, M. Structural evolution within the one-phase region of a three-component microemulsion system: water-*n*-decade-AOT.
13. Chang, N. J., and Kaler, E. W. (1986). Quasi-elastic light scattering study of five-component microemulsion, *Langmuir 3*: 184. Chang, N. J., Billman, J. F., Licklider, R. A., and Kaler, E. W. (1987). In *Statistical Thermodynamics of Micellar and Microemulsion Systems* (S.-H. Chen, ed.), Springer-Verlag, New York. Chang, N. J., and Kaler, E. W. (1987). *J. Chem. Phys.* Huang, J. S. (1969). Ph.D. thesis, Cornell University. Cazabat, A. M., Langevin, D., and Pouchelon, A. (1980). Light scattering study of water–oil microemulsions, *J. Colloid Interface Sci. 73*: 1.
14. Jahn, W., and Strey, R. (1987). In *Physics of Amphiphilic Systems* (J. Meunier, D. Langevin, and N. Soccara, eds.), Springer-Verlag, Heidelberg.
15. Widom, B. (1989). Correlation and scattering functions in a lattice model of microemulsions, *J. Chem. Phys. 90*: 2437.
16. Widom, B. (1986). Lattice model of microemulsions, *J. Chem. Phys. 84*: 6943, and earlier references therein. In particular, Wheeler, J. C., and Widom, B. (1968). Phase transitions and critical points in a model three component system, *J. Am. Chem. Soc. 40*: 3064. See also, Dawson, K. A., Lipkin, M. D., and Widom, B. (1988). Phase diagram of a lattice microemulsion model, *J. Chem. Phys. 88*: 5149.
17. Widom, B., Dawson, K. A., and Lipkin, M. D. (1986). Hamiltonian and phenomenological models of microemulsions, *Physica 140 A*: 26.
18. Schick, M., and Shih, W. H. (1986). Spin-1 model of a microemulsion, *Phys. Rev. B 34*: 1797. Schick, M., and Shih, W. H. (1987). Simple microscopic model of a microemulsion, *Phys. Rev. Lett. 59*: 1205. Garneiro, G. M., and Schick, M. (1988). Simple model for ternary mixtures with non-ionic surfactants, *J. Chem. Phys. 89*: 4638, and preprints.
19. Stockfish, T. P., and Wheeler, J. C. (1988). A microscopic lattice model for microemulsions, *J. Phys. Chem. 92*: 3292, and preprints.
20. Chen, K., Ebner, C., Jayaprakash, C., and Pandid, R. (1987). Microemulsion in oil–water–surfactant mixtures: an Ising lattice gas model, *J. Phys. C 20*: 1361. Chen, K., Ebner, C., Jayaprakash, C., and Pandid, R. (1988). Microemulsion in oil–water–surfactant mixtures: systematics of a lattice gas model, *Phys. Rev. A 38*: 6240.
21. Robledo, A. (1987). Spin Ising transcription of a lattice model of micellar solutions, *Phys. Rev. A 36*: 4067, and preprints.
22. Ciach, A., Hoye, J. S., and Stell, G. (1989). *J. Chem. Press. 90*: 1214. Ciach, A., Hoye, J. S., and Stell, G. (1988). Microscopic models for microemulsions, *J. Phys. A, 21*. Ciach, A., and Hoye, J. S. (1989). Microscopic model

for microemulsion. II. Behavior at low temperature and critical points, *J. Chem. Phys. 90*: 1222, and preprints.
23. Matsen, M. W., and Sullivan, D. E. (1990). Lattice model for surfactant in solution, *Phys. Rev. 441*: 2021.
24. Gunn, J., and Dawson, K. A. (1992). Lattice model description of amphiphilic mixtures, *J. Chem. Phys. 96*: 3152.
25. Dawson, K. A. (1987). Extended Ising model, *Phys. Rev. A 36*: 3383. Balbuena, P., and Dawson, K. A. (1988). Spatially frustrated lattice models, *Phys. Rev. B 38*: 11432.
26. Lipkin, M. D. (1988). Ground states of a frustrated Ising model with one-, two-, and three-body interactions, *Physica A 150*: 18.
27. Berera, A., and Dawson, K. A. (1990). Low-temperature expansion of a spatially frustrated isotropic lattice model, *Phys. Rev. A 41*: 626.
28. Fisher, M. E., and Selke, W. (1981). *Philos. Trans. R. Soc. London 302*: 1.
29. Helfrich, W. (1973). *Z. Naturforsch. C 28*: 693.
30. Levin, Y., and Dawson, K. A. Fluctuation-induced transitions in an isotropic spatially frustrated lattice model, *Phys. Rev. A 42* (4): 1976–1981.
31. Dawson, K. A., Walker, B., and Berera, A. (1990). Accounting for fluctuations in a lattice model of microemulsions, *Physica A 165*: 320.
32. Berera, A., and Dawson, K. A. Low temperature analysis of three-phase coexistence, *Phys. Rev. Lett.* (in press).
33. Dawson, K. A. (1987). Interfaces between phases in a lattice model of microemulsions, *Phys. Rev. A 35*: 1766.
34. Cazabat, A. M., Langevin, D., Meunier, J., and Pouchelon, A. (1982). Critical behavior in microemulsions, *Adv. Colloid Interface Sci. 126*: 175.
35. Aveyard, R., Binks, B. P., Clark, S., and Mead, J. (1986). Interfacial tension minimum in oil–water–surfactant systems, *J. Chem. Soc. Faraday Trans. 82*: 125.
36. Gunn, J. R., and Dawson, K. A. (1989). Microscopic model of amphilic assembly, *J. Chem. Phys. 91*: 6393.
37. Guest, D., Langevin, D., and Meunier, J. (1987). Liquid interfaces: role of fluctuations and analysis of ellipsometry and reflectivity measurements, *J. Phys. (Paris) 48*: 1819.
38. Hoar, T. P., and Schulman, J. H. (1943). *Nature 152*: 102. Schulman, J. H., Stockenius, W., and Prince, L. M. (1959). Mechanism of formation and structure of microemulsion by electron microscopy, *J. Phys. Chem. 63*: 1677. See, for instance, Eicke, H. F., and Rehak, J. (1976). On the formation of water/oil-microemulsion, *Helv. Chim. Acta 59*: 2883.
39. van Nienwkoop, J., and Snoei, G. (1983). Shell Research B.V., report 674, January, unpublished.
40. Levin, Y., and Dawson, K. A. Sine-Gordon renormalization of the orientational roughening transition, *Phys. Rev. A 42* (6): 3507–3511.
41. M. Teubner and R. Strey, Origin of the Scattering in Microemulsion, *J. Chem. Phys. 87* (5), 3195–3200 (1987).
42. Y. Levin, C. J. Mundy, and K. A. Dawson, Renormalization of Landau-Ginzburg-Wilson Theory of Microemulsion, *Phys. Rev. A* (in press).

13

Kinetic Theory of Strongly Inhomogeneous Fluids

H. Ted Davis
University of Minnesota
Minneapolis, Minnesota

I. INTRODUCTION

Fluids confined to spaces a few molecules wide in at least one dimension are strongly inhomogeneous. Examples of these are fluids between clay platelets, in zeolites, in porous vicor glasses, and between coalescing colloidal particles. Also, there are strong density or composition inhomogeneities in interfaces between bulk phases. In recent years a great deal of progress has been made toward understanding the molecular structure of fluid interfaces. Much of the recent advances in molecular theory and molecular simulations is presented in other chapters of this book. As usual, our understanding of equilibrium properties is ahead of our understanding of transport and flow properties. Thus, instead of trying to expose the general theory of flow and transport in strongly inhomogeneous fluids, I have set the more modest goal of describing flow and tracer diffusion in terms of the modified Enskog kinetic theory.

The modified Enskog theory is outlined in Section II. In Section III it is shown that the theory yields the exact equilibrium Yvon–Born–Green (YBG) equations. These equations represent hydrostatic equilibrium for each molecular species. Closure approximations to the YBG equations are discussed and it is demonstrated that the Fischer Methfessel approximation predicts density profiles and salvation forces (normal pressures)

in good agreement with computer simulations of simple fluids confined to planar slit pores. The theory of tracer diffusion is presented in Section IV for the modified Enskog model and for three square-well fluids: fluids whose molecules interact with a central attractive square-well and a smooth, loaded, or rough hard-sphere repulsion. The modified Enskog theory of flow is presented in Section V, and the special problems of confinement-induced solidification are mentioned in Section VI.

II. MODIFIED ENSKOG EQUATION

Enskog's equation has been used for the study of fluids that are homogeneous on the molecular scale. Applications to homogeneous fluids have shown that despite the simplifications it is based on, Enskog's theory predicts the self-diffusion and viscosity coefficients rather well [1]. There is, however, nothing in the approximations underlying the equation that precludes applications to strongly inhomogeneous fluids (i.e., fluids whose component densities vary significantly on a molecular scale). In fact, recent applications of the theory to thin fluid films confined to molecularly thin pores indicate that the theory yields meaningful results for remarkably thin films [2–5]. Enskog's kinetic equation for the singlet distribution function, modified to include attractive interactions through a mean-field approximation, forms the basis of the theory presented herein.

In what follows we assume that the fluid is composed of classical, structureless molecules that interact via pair potentials of the form

$$u(s) = \begin{cases} \infty & s < a \\ u^A(s) & s > a \end{cases} \tag{1}$$

where s is the intermolecular separation, a the hard-core diameter of the molecules, and $u^A(s)$ a continuous potential. Barker and Henderson [6] and Weeks et al. [7] have described systematic ways to approximate continuous pair potentials having strong short-range repulsive interactions by a hard-core cutoff model of the type given by (1).

The singlet distribution function, $f_i(\mathbf{r},\mathbf{v},t)$, denotes the probability density in configuration and velocity space [i.e., $f_i(\mathbf{r},\mathbf{v},t)\, d^3r\, d^3v$ is the probable number of particles of type i lying between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ with velocity between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ at time t]. Similarly, $f_{ij}^{(2)}(\mathbf{r},\mathbf{r}',\mathbf{v}_i,\mathbf{v}_j,t)\, d^3r\, d^3r'\, d^3v_i\, d^3v_j$ denotes the probable number of pairs of particles, one of which lies between $\mathbf{r}$ and $\mathbf{r} + d\mathbf{r}$ with a velocity between $\mathbf{v}_i$ and $\mathbf{v}_i + d\mathbf{v}_i$ and the other of which lies between $\mathbf{r}'$ and $\mathbf{r}' + d\mathbf{r}'$ with a velocity between $\mathbf{v}_j$ and $\mathbf{v}_j + d\mathbf{v}_j$. $f_{ij}^{(2)}$ is the doublet distribution function. In a multicomponent fluid subjected to external forces (e.g., electrical, gravitational and mag-

netic fields or the forces of confinement exerted by solid surfaces), the modified Enskog equation is

$$\frac{\partial f_i}{\partial t} + \mathbf{v}_i\cdot\nabla f_i - \frac{1}{m_i}\nabla u_i^e\cdot\nabla_{v_i} f_i - \sum_j^c \frac{1}{m_i}\int \nabla u_{ij}^A(\mathbf{r}-\mathbf{r}')$$

$$\cdot\nabla_{v_i} f_{ij}^{(2)}(\mathbf{r},\mathbf{r}',\mathbf{v}_i,\mathbf{v}_j,t)\, d^3r'\, d^3v_j = \sum_j^c \int_{\mathbf{v}_{ij}\cdot k>0} [f_{ij}^{(2)}(\mathbf{r},\mathbf{r}+a_{ij}\mathbf{k},\mathbf{v}_i',\mathbf{v}_j',t) \quad (2)$$

$$- f_{ij}^{(2)}(\mathbf{r},\mathbf{r}-a_{ij}\mathbf{k},\mathbf{v}_i,\mathbf{v}_j,t)]a_{ij}^2\mathbf{v}_{ij}\cdot\mathbf{k}a^2k\, d^3v_j$$

where ∇ and ∇_{v_i} are gradient operators with respect to $\mathbf{r}$ and $\mathbf{v}_i$, m_i the molecular mass of species i, u_i^e the potential of the external force on i, u_{ij}^A the continuous part of the pair potential between i and j, and $a_{ij} \equiv (a_i + a_j)/2$, where a_i is the hard-core diameter of i. $\mathbf{k}$ is a unit vector directed from the center of i to the center of j, and $\mathbf{v}_i'$ is the velocity of i after a hard-sphere collision with j [i.e., $\mathbf{v}_i' = \mathbf{v}_i - 2M_j\mathbf{v}_{ij}\cdot\mathbf{kk}$ and $\mathbf{v}_j' = \mathbf{v}_j + 2M_i\mathbf{v}_{ij}\cdot\mathbf{kk}$[8], where $\mathbf{v}_{ij} \equiv \mathbf{v}_i - \mathbf{v}_j$, $M_i = m_i/(m_i + m_j)$ and $M_j = m_j/(m_i + m_j)$]. c denotes the number of molecular species in the fluid.

The assumptions underlying (2) are that the continuous interactions are sufficiently slowly varying that they contribute a mean field in which the molecules move but do not contribute to collisional dissipation and that the hard-core collisions are uncorrelated. As Enskog did, we make the further assumption that the doublet distribution function factors into the product

$$f_{ij}^{(2)}(\mathbf{r},\mathbf{r}',\mathbf{v}_i,\mathbf{v}_j,t) = f_i(\mathbf{r},\mathbf{v}_i,t)f_j(\mathbf{r}',\mathbf{v}_i,t)g_{ij}(\mathbf{r},\mathbf{r}',t) \quad (3)$$

where $g_{ij}(\mathbf{r},\mathbf{r}',t)$ is the pair correlation function. This assumption neglects the velocity dependence of the correlation function. Unlike the usual Enskog theory of weakly inhomogeneous fluids, we do not assume that g_{ij} is the correlation function of homogeneous fluid at some local composition.

The local number density n_i, pair density $n_{ij}^{(2)}$, and average velocity $\mathbf{u}_i$ of species i are obtained from the expressions

$$n_i(\mathbf{r},t) = \int f_i(\mathbf{r},\mathbf{v}_i,t)\, d^3v_i \quad (4)$$

$$n_{ij}^{(2)}(\mathbf{r},\mathbf{r}',t) = \int f_{ij}^{(2)}(\mathbf{r},\mathbf{r}',\mathbf{v}_i,\mathbf{v}_j,t)\, d^3v_i\, d^3v_j \quad (5)$$

$$\mathbf{u}_i(\mathbf{r},t) = \frac{1}{n_i}\int \mathbf{v}_i f_i(\mathbf{r},\mathbf{v}_i,t)\, d^3v_i \quad (6)$$

The spatial pair correlation function is given by $g_{ij}(\mathbf{r},\mathbf{r}',t)$

$$g_{ij}(\mathbf{r},\mathbf{r}',t) = \frac{n_{ij}^{(2)}(\mathbf{r},\mathbf{r}',t)}{n_i(\mathbf{r},t)n_j(\mathbf{r}',t)} \quad (7)$$

Multiplying (2) by d^3v_i, integrating, and using (4) and (5), we find the continuity (mass conservation) equation

$$\frac{\partial n_i}{\partial t} + \nabla\cdot(n_i\mathbf{u}_i) = 0 \tag{8}$$

for species i. This is an exact result even though we derived it from an approximate kinetic equation. In deriving (8) we used the properties $\int\nabla_{\mathbf{v}_i}f_i\,d^3v_i = \int\nabla_{\mathbf{v}_i}f_{ij}^{(2)}\,d^3v_i = 0$, which result from the condition $f_i, f_{ij}\to 0$ as $\mathbf{v}_i\to\infty$.

The momentum conservation equation for species i can be obtained by multiplication of (2) by $\mathbf{v}_i\,d^3v_i$ and integration. The result is

$$\begin{aligned}
&\frac{\partial(m_in_i\mathbf{u}_i)}{\partial t} + \nabla\cdot(m_in_i\langle\mathbf{v}_i\mathbf{v}_i\rangle) + n_i\nabla u_i^e + \sum_j\int\nabla u_{ij}^A(|\mathbf{r}-\mathbf{r}'|)n_{ij}^{(2)}(\mathbf{r},\mathbf{r}',t)\,d^3r' \\
&= \sum_j m_i\int_{\mathbf{v}_{ji}\cdot\mathbf{k}>0}\mathbf{v}_i[f_{ij}^{(2)}(\mathbf{r},\mathbf{r}+a_{ij}\mathbf{k},\mathbf{v}_i',\mathbf{v}_j',t) \\
&\quad - f_{ij}^{(2)}(\mathbf{r},\mathbf{r}-a_{ij}\mathbf{k},\mathbf{v}_i,\mathbf{v}_j,t)]a_{ij}^2(\mathbf{v}_{ji}\cdot\mathbf{k})\,d^2k\,d^3v_i\,dv_j^3 \\
&= \sum_j\frac{2m_im_j}{m_i+m_j}\int_{\mathbf{v}_{21}\cdot\mathbf{k}>0}\mathbf{k}f_{ij}^{(2)}(\mathbf{r},\mathbf{r}-d^2k\,d^3v_1\,d^3v_2
\end{aligned} \tag{9}$$

where

$$n_i\langle\mathbf{v}_i\mathbf{v}_j\rangle \equiv \int a_{ij}\mathbf{k},\mathbf{v}_1,\mathbf{v}_2)a_{ij}^2(\mathbf{v}_{21}\cdot\mathbf{k})^2\,\mathbf{v}_i\mathbf{v}_if_i(\mathbf{r},\mathbf{v}_i,t)\,d^3v_i \tag{10}$$

In integrating the terms involving u_i^e and u_{ij}^A, we used the property $\int\mathbf{v}_i\nabla\mathbf{v}_if_i\,d^3v_i = -\mathbf{U}\int f_i\,d^3v_i$, again resulting from the condition $f_i\to 0$ as $\mathbf{v}_i\to\infty$. $\mathbf{U}$ is the unit tensor, which has the property $\mathbf{U}\cdot\mathbf{a} = \mathbf{a}$ for arbitrary vector $\mathbf{a}$.

In arriving at the final form of the right-hand side of (9) we used binary hard-sphere collision properties [8], namely,

$$\begin{aligned}
&\int_{\mathbf{v}_{ij}\cdot\mathbf{k}>0}\mathbf{v}_if_{ij}^{(2)}(\mathbf{r},\mathbf{r}+a_{ij}\mathbf{k},\mathbf{v}_i'\cdot\mathbf{v}_j',t)a_{ij}^2\mathbf{v}_{ji}\mathbf{k}\,d^2k\,d^3v_i\,d^3v_j \\
&\qquad = \int_{\mathbf{v}_{ji}\cdot\mathbf{k}>0}\mathbf{v}_i'f_{ij}^{(2)}(\mathbf{r},\mathbf{r}-a_{ij}\mathbf{k},\mathbf{v}_i\cdot\mathbf{v}_j,t)a_{ij}^2\mathbf{v}_{ji}\cdot\mathbf{k}\,d^2k\,d^2v_i\,d^3v_j
\end{aligned} \tag{11}$$

and $\mathbf{v}_i' = \mathbf{v}_i + M_i\mathbf{v}_{ji}\cdot\mathbf{kk}$, $M_i = m_i/(m_i+m_j)$. We have also replaced $\mathbf{v}_1$ and $\mathbf{v}_2$ in the final form of the right-hand side of (8) to emphasize that they are simply dummy variables once integration is carried out over both of them.

With the aid of several manipulations, (8) can be put into a form more like what is expected in continuum mechanics. First, let us introduce the

coordinate transformation $\mathbf{r},\mathbf{r}' \rightarrow \mathbf{r}, \mathbf{s} \equiv \mathbf{r}' - \mathbf{r}$ to obtain

$$\begin{aligned}\sum_j \int \nabla u_{ij}^A(|\mathbf{r} - \mathbf{r}'|) n_{ij}^{(2)}(\mathbf{r},\mathbf{r}',t)\, d^3r' \\ = -\sum_j \int \frac{\mathbf{s}}{s}\frac{du_{ij}^A(s)}{ds} n_{ij}^{(2)}(\mathbf{r},\mathbf{r}+\mathbf{s},t)\, d^3s \\ = -\sum_j \frac{1}{2}\int \frac{\mathbf{s}}{s}\frac{u_{ij}^A(s)}{ds}[n_{ij}^{(2)}(\mathbf{r},\mathbf{r}+s) - n_{ij}^{(2)}(\mathbf{r},\mathbf{r}-\mathbf{s})]\, d^3s\end{aligned} \tag{12}$$

The final form of the right-hand side of (12) is obtained by using the transformation $\mathbf{s} \rightarrow -\mathbf{s}$ and taking the mean of the original and the transformed expressions.

Next we use the identity

$$\frac{d}{d\alpha} n_{ij}^{(2)}(\mathbf{r}-\alpha\mathbf{s},\mathbf{r}+(1-\alpha)\mathbf{s},t) = -\mathbf{s}\cdot\nabla n_i^{(2)}(\mathbf{r}-\alpha\mathbf{s},\mathbf{r}+(1-\alpha)\mathbf{s},t) \tag{13}$$

from which one obtains another identity,

$$\begin{aligned}n_{ij}^{(2)}(\mathbf{r},\mathbf{r}+\mathbf{s},t) = n_{ij}^{(2)}(\mathbf{r}-\mathbf{s},\mathbf{r},t) - \mathbf{s}\cdot\nabla\int_0^1 d\alpha\, n_{ij}^{(2)} \\ \times(\mathbf{r}-\alpha\mathbf{s},\mathbf{r}+(1-\alpha)\mathbf{s},t)\end{aligned} \tag{14}$$

Combining (12) and (14), we find that

$$\sum_j \int \nabla u_{ij}^A(|\mathbf{r}-\mathbf{r}'|) n_{ij}^{(2)}(\mathbf{r},\mathbf{r}',t)\, d^3r' = -\nabla\cdot\mathbf{T}_i^A - \mathbf{B}_i^A \tag{15}$$

where

$$\mathbf{T}_i^A = \sum_j \frac{1}{2}\int_0^1 \int \frac{\mathbf{ss}}{s}\frac{du_{ij}^A(s)}{ds} n_{ij}^{(2)}(\mathbf{r}-\alpha\mathbf{s},\mathbf{r}+(1+\alpha)\mathbf{s},t)\, d^3s\, d\alpha \tag{16}$$

and

$$\mathbf{B}_i^A = \sum_j \frac{1}{2}\int \frac{\mathbf{s}}{s}\frac{du_{ij}^A(s)}{ds}[n_{ji}^{(2)}(\mathbf{r},\mathbf{r}-\mathbf{s}) - n_{ij}^{(2)}(\mathbf{r},\mathbf{r}-\mathbf{s})]\, d^3s \tag{17}$$

$\mathbf{T}_i^A$ is the contribution of the ith species to the part of the stress tensor arising from the continuous intermolecular interactions. $\mathbf{B}_i^A$ is the asymmetric force on species i contributed by the continuous intermolecular interactions [9]. The asymmetric force arises from the density inhomogeneities of different species. Even if the pair correlation function depends

only on s, $\mathbf{B}_i^A$ is not zero in a multicomponent solution since the quantity

$$n_{ji}^{(2)}(\mathbf{r},\mathbf{r} - \mathbf{s}) - n_{ij}^{(2)}(\mathbf{r},\mathbf{r} - \mathbf{s}) = g_{ij}(s,t)[n_i(\mathbf{r} - \mathbf{s})n_j(\mathbf{r}) - n_i(\mathbf{r})n_j(\mathbf{r} - \mathbf{s})] \tag{18}$$

is in general not zero if $i \neq j$. We remark also that $\mathbf{B}_i^A$ cannot be cast into the form of a divergence of a function of $\mathbf{r}$. Thus $\mathbf{B}_i^A$ differs in a fundamental way from the local stress tensor contribution to the momentum balance.

Manipulations similar to those leading to (14) can be used to obtain

$$\text{right-hand side of (9)} = \nabla\cdot\mathbf{T}_i^R + \mathbf{B}_i^R \tag{19}$$

where according to the Enskog hard-sphere theory the contribution of the repulsive forces to the stress tensor and asymmetric force are

$$\mathbf{T}_i^R = \sum_j \frac{m_j m_j}{m_i + m_j} \int_0^1 \int \int_{\mathbf{v}_{21}\cdot\mathbf{k}>0} \mathbf{kk} f_{ij}^{(2)}(\mathbf{r} + \alpha a_{ij}\mathbf{k},\mathbf{r} - (1 - \alpha)a_{ij}\mathbf{k},\mathbf{v}_1,\mathbf{v}_2,t) \times a_{ij}^3(\mathbf{v}_{21}\cdot\mathbf{k})^2\, d^2k\, d^3v_1\, d^3v_2\, d\alpha \tag{20}$$

$$\mathbf{B}_i^R = \sum_j \frac{m_i m_j}{m_i + m_j} \int_{\mathbf{v}_{21}\cdot\mathbf{k}>0} \mathbf{k}[f_{ij}^{(2)}(\mathbf{r} + a_{ij}\mathbf{k},\mathbf{r},\mathbf{v}_1,\mathbf{v}_2,t) - f_{ji}^{(2)}(\mathbf{r} - a_{ij}\mathbf{k},\mathbf{r},\mathbf{v}_1,\mathbf{v}_2,t)]a_{ij}^2(\mathbf{v}_{21}\cdot\mathbf{k})^2\, d^2k\, d^3v_1\, d^3v_2 \tag{21}$$

In deriving (20) and (21) we used the transformations $\mathbf{k} \rightarrow -\mathbf{k}$ and $\mathbf{v}_1\mathbf{v}_2 \rightarrow \mathbf{v}_2,\mathbf{v}_1$.

Finally, the local kinetic stress tensor of component i is defined as

$$\mathbf{T}_i^K = -m_i n_i \langle(\mathbf{v}_i - \mathbf{u})(\mathbf{v}_i - \mathbf{u})\rangle \equiv -m_i \int (\mathbf{v}_i - \mathbf{u})(\mathbf{v}_i - \mathbf{u}) f_i(\mathbf{r},\mathbf{v}_i,t)\, d^3v_i \tag{22}$$

where $\mathbf{u}((\mathbf{r},t)$ is the local mass average velocity,

$$\mathbf{u}(\mathbf{r},t) = \frac{\sum_i m_i n_i(\mathbf{r},t)\mathbf{u}_i(\mathbf{r},t)}{\sum_i m_i n_i(\mathbf{r},t)} \tag{23}$$

The momentum balance equation of species i can now be expressed as [9]

$$\frac{\partial(m_i n_i \mathbf{u}_i)}{\partial t} + \nabla\cdot[m_i n_i(\mathbf{u}\mathbf{u}_i - \mathbf{u}_i\mathbf{u} - \mathbf{u}\mathbf{u})] = -m_i n_i \nabla u_i^e + \mathbf{B}_i + \nabla\cdot\mathbf{T}_i \tag{24}$$

$$\mathbf{T}_i = \mathbf{T}_i^K + \mathbf{T}_i^A + \mathbf{T}_i^R \tag{25}$$

and

$$\mathbf{B}_i = \mathbf{B}_i^A + \mathbf{B}_i^R \tag{26}$$

Summing (24) over all species and noting that the asymmetric forces cancel,

$$\sum_i \mathbf{B}_i = 0 \tag{27}$$

we obtain the usual momentum balance equation of fluid mechanics:

$$\frac{\partial(\rho \mathbf{u})}{\partial t} + \nabla\cdot(\rho \mathbf{u}\mathbf{u}) = \rho \hat{\mathbf{F}}^e - \nabla\cdot\mathbf{T} \tag{28}$$

where the local mass density is

$$\rho(\mathbf{r},t) = \sum_i m_i n_i(\mathbf{r},t) \tag{29}$$

the local body force density is

$$\rho \hat{\mathbf{F}}^e(\mathbf{r},t) = -\sum_i m_i n_i(\mathbf{r},t)\nabla u^e \tag{30}$$

and the total local stress tensor is

$$\mathbf{T} = \sum_i \left[\mathbf{T}_i^K + \mathbf{T}_i^A + \mathbf{T}_i^R \right] \tag{31}$$

The continuity equation for ρ is obtained by multiplying (5) by m_i and summing over species. The well-known result is

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho \mathbf{u}) = 0 \tag{32}$$

Although continuum mechanics can be used to derive (29), the total momentum balance equation, it cannot be used to decompose the species momentum balance into local species stress and asymmetric force contributions as given by (24). At equilibrium (24) provides hydrostatic equations governing the density distributions of the species in the presence of external fields or in fluid interfaces. We address equilibrium properties of strongly inhomogeneous fluids in the next section.

III. EQUILIBRIUM THEORY OF INHOMOGENEOUS FLUIDS

A. Equations of Hydrostatics and the Yvon–Born–Green Equations

At equilibrium the singlet and doublet distribution functions are given by

$$f_i = \varphi(\mathbf{v}_i) n_i(\mathbf{r}) \quad \text{and} \quad f_{ij}^{(2)} = \varphi(\mathbf{v}_i)\varphi(\mathbf{v}_j) n_{ij}^{(2)}(\mathbf{r},\mathbf{r}') \tag{33}$$

where φ is the Maxwellian velocity distribution function

$$\varphi(\mathbf{v}) = \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-mv^2/2kT} \tag{34}$$

k is Boltzmann's constant and T is the absolute temperature. There is mean flow at equilibrium. Thus $\mathbf{u}_i = \mathbf{u} = 0$, so the momentum balance equation of species i reduces to

$$0 = -m_i n_i \nabla u_i^e + \mathbf{B}_i + \nabla \cdot \mathbf{T}_i \tag{35}$$

the equation of hydrostatic for species i. The set of these equations governs the component density distributions $n_i(\mathbf{r})$ for a fluid at equilibrium.

Although (35) identifies hydrostatic equilibrium of each of the species as the determinant of the density distributions of inhomogeneous fluid at equilibrium, it is more convenient for computational purposes to use the equivalent set of equations obtained directly from (9) in the equilibrium limit. At equilibrium the velocity integrals in (9) can be performed. The results are

$$n_i m_i \langle \mathbf{v}_i \mathbf{v}_i \rangle = n_i kT\mathbf{U} \tag{36}$$

and

$$\sum_j \frac{2m_i m_j}{m_i + m_j} \int_{\mathbf{v}_{ij}\cdot\mathbf{k}>0} \mathbf{k} f_{ij}^{(2)}(\mathbf{r}, \mathbf{r} - d_{ij}\mathbf{k}, \mathbf{v}_i, \mathbf{v}_j) a_{ij}^2 (\mathbf{v}_{ij}\mathbf{k})^2 \, d^2k \, d^3v_i \, d^3v_j$$
$$= \sum_j kT \int n_{ij}^{(2)}(\mathbf{r}, \mathbf{r} - a_{ij}\mathbf{k}) a_{ij}^2 \mathbf{k} \, d^2k \tag{37}$$

With these results the equilibrium version of (9) becomes

$$kT \nabla n_i + n_i \nabla u_i^e + \sum_j \int \nabla u_{ij}^A n_{ij}^{(2)}(\mathbf{r}, \mathbf{r}') \, d^3r'$$
$$+ \sum_j kT \int n_{ij}^{(2)}(\mathbf{r}, \mathbf{r} + a_{ij}k) a_{ij}^2 \mathbf{k} \, d^2k = 0 \tag{38}$$

$i = 1,2, \ldots, c$. Equation (38) is known as the Yvon–Born–Green (YBG) equation for the density distribution n_i. The equation is rigorously valid for multicomponent fluids whose particles obey the pair potential given by (1). Thus, even though the modified Enskog equation is an approximation for a nonequilibrium system, it yields the exact equations of hydrostatics.

At equilibrium the stress tensor and asymmetric force of species i are of the forms

$$\mathbf{T}_i = -n_i kT\mathbf{U} + \sum_j \frac{1}{2} \int_0^1 \int \frac{\mathbf{ss}}{s} \frac{du_i^A}{ds} n_{ij}(\mathbf{r} - \alpha\mathbf{s}, \mathbf{r} + (1 - \alpha)\mathbf{s}) \, d^3 s d\alpha$$

$$- \sum_j \frac{1}{2} kT \int_0^1 \int \mathbf{k}\mathbf{k} n_{ij}^{(2)}(\mathbf{r} - \alpha a_{ji}\mathbf{k}, \mathbf{r} + (1 - \alpha) a_{ij}\mathbf{k}) a_{ij}^3 \, d^2k \, d\alpha \quad (39)$$

and

$$\mathbf{B}_i = \text{Eq. (17)} + \sum_j \frac{1}{2} \int \mathbf{k}[n_{ji}^{(2)}(\mathbf{r}, \mathbf{r} + d_{ij}\mathbf{k}) - n_{ij}^{(2)}(\mathbf{r}, \mathbf{r} - d_{ij}\mathbf{k})] \, d_{ij}^2 \, d^2k + \sum_j \frac{1}{2} \int \frac{\mathbf{s}}{s} \frac{du_{ij}^A}{ds} [n_{ji}^{(2)}(\mathbf{r}, \mathbf{r} - \mathbf{s}) - n_{ij}^{(2)}(\mathbf{r}, \mathbf{r} - \mathbf{s}) \, d^3s \quad (40)$$

These results are again exact for fluids whose molecules obey pair potentials of the form of (1).

The total asymmetric force is zero, as noted before, and the total stress tensor is

$$\mathbf{T} = \sum_i n_i kT \mathbf{I} + \sum_{i,j} \frac{1}{2} \int_0^1 \left[\int \frac{\mathbf{ss}}{s} \frac{du_{ij}^A}{ds} n_{ij}^{(2)}(\mathbf{r} - \alpha \mathbf{s}, \mathbf{r} + (1 - \alpha)\mathbf{s}) \, d^3s + kT \int_0^1 \int \mathbf{k}\mathbf{k} n_{ij}^{(2)}(\mathbf{r} - \alpha a_{ij}\mathbf{k}, \mathbf{r} + (1 - \alpha) a_{ij} k) a_{ij}^3 \, d^2k \right] d\alpha \quad (41)$$

B. Local Average Density Approximation

The YBG equations for the density distributions are not a closet set since they contain contributions from the doublet density distribution function. Thus come closure approximation must be found if we wish to solve the YBG equations. One way to accomplish closure is to use the mean-spherical, hypernetted chain or the Percus–Yevick approximation. All of these begin with the exact Ornstein–Zernicke equation

$$g_{ij}(\mathbf{r}, \mathbf{r}') = 1 + c_{ij}(\mathbf{r}, \mathbf{r}') + \sum_k \int [g_{ij}(\mathbf{r}, \mathbf{r}'') - 1] n_k(\mathbf{r}'') c_{ki}(\mathbf{r}'' \mathbf{r}') \, d^3r'' \quad (42)$$

The Percus–Yevick approximation is, for example,

$$c_{ij}(\mathbf{r}, \mathbf{r}') = \begin{cases} 0 & | \mathbf{r} - \mathbf{r}' | > a_{ij} \\ g_{ij}(\mathbf{r}, \mathbf{r}') - e^{\beta u_{ij}(r - \mathbf{r}')} g_{ij}(\mathbf{r}, \mathbf{r}') & | \mathbf{r} - \mathbf{r}' | < a_{ij} \end{cases} \quad (43)$$

Equations (42) and (43) combine with the YBG equation [Eq. (38)] to yield a closed set of equations for n_i and g_{ij}. The problem with solving these equations is that they are computationally very costly even for a one-component fluid in planar geometry (i.e., n varying only in one direction).

A much simpler approximation that has been introduced in the study of the liquid–vapor interface is [9]

$$n_{ij}^{(2)}(\mathbf{r}, \mathbf{r}') \approx n_i(\mathbf{r}) n_j(\mathbf{r}') g_{ij} \left(| \mathbf{r} - \mathbf{r}' |, \mathbf{n}\left(\frac{\mathbf{r} + \mathbf{r}'}{2} \right) \right) \quad (44)$$

where

$$g_{ij}\left(\left|\mathbf{r}-\mathbf{r}'\right|,\mathbf{n}\left(\frac{\mathbf{r}+\mathbf{r}'}{2}\right)\right)$$

is the pair correlation function of homogeneous fluid evaluated at the component densities $\mathbf{n}$ ($\equiv \{n_1, \cdots, n_c\}$) at position $(\mathbf{r}+\mathbf{r}')/2$. Often, the problem has been simplified even further by assuming that in evaluating the mean-field contribution involving u_{ij}^A, the structureless fluid approximation (i.e., $g_{ij} = 0, \left|\mathbf{r}-\mathbf{r}'\right| < a_{ij}$ and $g_{ij} = 1, \left|\mathbf{r}-\mathbf{r}'\right| > a_{ij}$) can be used and that in the hard-sphere term the contact value of the hard-sphere correlation function can be used. With these "van der Waals" approximations the YBG equation becomes

$$\nabla[kT \ln n_i + u_i^e + \sum_j \int u_{ij}^A(s) n_j(\mathbf{r}+\mathbf{s})\, d^3s] + \sum_j kT \int n_j(\mathbf{r}-a_{ij}\mathbf{k}) g_{ij}^{\mathrm{HS}}(a_{ij},\mathbf{n}(\mathbf{r}+\tfrac{1}{2}a_{ij}\mathbf{k})) a_{ij}^2 \mathbf{k}\, d^2k = 0 \tag{45}$$

Although "local density functional approximations" such as (45) have been shown to give qualitatively correct predictions of density and pressure profiles and surface tensions for one-component liquid–vapor interfaces [9], they are known to fail for fluid–solid interfaces where the density distribution tends to be highly oscillatory rather than monotonically varying in space [10]. Consequently, "nonlocal density functional approximations" have been explored. Typical of this approach is the Fischer–Methfessel approximation, namely [11],

$$n_{ij}^{(2)}(\mathbf{r},\mathbf{r}') \approx n_i(\mathbf{r}) n_j(\mathbf{r}') g_{ij}\left(\left|\mathbf{r}-\mathbf{r}'\right|,\bar{\mathbf{n}}\left(\left(\frac{\mathbf{r}+\mathbf{r}'}{2}\right)\right)\right) \tag{46}$$

where it is assumed that the pair correlation function is that of homogeneous fluid, but, instead of being evaluated at some local composition, it is evaluated at a locally averaged composition $\bar{\mathbf{n}}(\mathbf{R})$. In its most general form, the local average can be defined as

$$\bar{n}_i(\mathbf{r}) = \int w_i(\mathbf{r},\mathbf{r}',\{\mathbf{n}\}) n_i(\mathbf{r}')\, d^3r' \tag{47}$$

where the local weighting function $w_i\mathbf{r},\mathbf{r}'\{\mathbf{n}\})$ is a functional of the density distributions $\mathbf{n}$. To ensure that $\bar{n}_i = n_i$ in a homogeneous fluid, it is required that

$$\int w_i(\mathbf{r},\mathbf{r}',\{\mathbf{n}\})\, d^3r' = 1 \tag{48}$$

In the context of either the YBG equations or of density functional free-energy theory, the nonlocal density functional theory has been in

essence an attempt to pick the "best" choice of weighting functions w_i. The approaches of Nordholm et al. [12], Tarazona [13], Percus [14], Vanderlick et al. [15], Kroll [16], and Fischer and Methfessel's [11] works have proceeded along these lines. In particular, for a one-component fluid, Fischer and Methfessel chose

$$\bar{n}(\mathbf{r}) = \left[\frac{4}{3}\pi\left(\frac{a}{2}\right)^3\right]^{-1} \int_{s<a/2} n(\mathbf{r} + \mathbf{s})\, d^3s \tag{49}$$

or

$$w(\mathbf{r},\mathbf{r}') = 1 - \eta\left(|\mathbf{r} - \mathbf{r}'| - \frac{a}{2}\right) \tag{50}$$

where η is the Heaviside function, that is, $\eta(x) = 0$, $x < 0$ and $= 1$, $x > 0$. A generalization of the Fischer–Methfessel formula to a multicomponent fluid is

$$g_{ij}(\mathbf{r},\mathbf{r}') = g_{ij}\left(|\mathbf{r} - \mathbf{r}'|, \bar{n}\left(\frac{a_j\mathbf{r} + a_i\mathbf{r}'}{2a_{ij}}\right)\right) \tag{51}$$

where

$$\bar{n}_i(\mathbf{r}) = \left[\frac{4}{3}\pi\left(\frac{a_i}{2}\right)^3\right]^{-1} \int_{s<a_i/2} n_i(\mathbf{r} + \mathbf{s})\, d^3s \tag{52}$$

a_i is the diameter of a particle of species i and $a_{ij} = (a_i + a_j)/2$. This form of $\bar{n}_i$ is suggested by the exact theory of an inhomogeneous fluid of hard rods [14,15]. In fact, one can show [17] that the YBG equation for a one-component fluid of hard rods yields the exact density distribution function when the Fischer–Methfessel approximation is used. The work of Nordholm and Johnson suggests averaging n_i over the molecular diameter in (54) instead of over the molecular radius. Tarazona's formula for $\bar{n}_i$ is considerably more complicated since he requires the free-energy functional derivatives with regret to density to yield the Percus–Yevick direct correlation function for hard spheres.

As in the van der Waals theory, Fischer and Methfessel assume (1) that the pair correlation function can be replaced by a step function in the term involving the continuous interaction u_{ij}^A (which is presumably long range and therefore gives a contribution less sensitive to short-range pair correlations) and (2) that the hard-sphere contact value of the pair correlation can be used in the hard-sphere part of the YBG equation. Thus the Fischer–Methfessel model is

$$\nabla[kT \ln n_i + u_i^e + \sum_j \int u_{ij}^A(s) n_j(\mathbf{r} + \mathbf{s})\, d^3s] + \sum_j kT \int n_j(\mathbf{r} + a_{ij}\mathbf{k}) g_{ij}^{\mathrm{HS}}(a_{ij}, \bar{\mathbf{n}}(\mathbf{r} + \tfrac{1}{2}a_i\mathbf{k}) a_{ij}^2 \mathbf{k}\, d^2k = 0 \tag{53}$$

The equations looks very similar to the local density functional equation (45) but in fact predicts dramatically different and qualitatively correct density distribution for fluids near solid surfaces.

For a one-component fluid, an improvement over the Fischer–Methfessel model can be obtained by approximating the pair correlation function in the mean-field term by

$$g(|\mathbf{r} - \mathbf{r}'|,\bar{\mathbf{n}}) \simeq 1 - g^{\mathrm{HS}}(|\mathbf{r} - \mathbf{r}'|,\bar{\mathbf{n}}) \left[\frac{\alpha(\bar{n})u^A(|\mathbf{r} - \mathbf{r}|)}{kT}\right] \tag{54}$$

using the Barker–Henderson method [6] to determine the hard-sphere diameter a and attractive potential u^A from the continuous pair potentials representing the actual fluid particles. The quantity $\alpha(n)$ is given by $\alpha(n) = kT[\partial P^{\mathrm{HS}}(n)/\partial n]$, where $P^{\mathrm{HS}}(n)$ is the pressure of a homogeneous hard-sphere fluid. According to the Barker–Henderson method, one computes the effective (temperature dependent) hard-sphere diameter from the formula

$$a = \int_0^{\sigma} [1 - e^{u(s)/kT}]\, ds \tag{55}$$

where u is the pair potential between particles and σ is the particle separation at which $u(s) = 0$. Then one sets

$$u^A(s) = \begin{cases} 0 & s \leq \sigma \\ u(s) & s \geq \sigma \end{cases} \tag{56}$$

and equates $g(s,\bar{\mathbf{n}})$ to the correlation function $g^{\mathrm{HS}}(s,\bar{\mathbf{n}})$ of hard spheres whose collision diameters are a. With (54) the YBG equation becomes

$$\begin{aligned} \nabla [kT \ln n + u^e] + \int \nabla u^A(s) n(\mathbf{r} + \mathbf{s}) g^{\mathrm{HS}}(s,\bar{\mathbf{n}}(\mathbf{r} + \tfrac{1}{2}\mathbf{s}) \\ \times \left[1 - \frac{\alpha(\bar{n}(\mathbf{r} + \frac{1}{2}\mathbf{s}))u^A(s)}{kT}\right] d^3s \\ + kT \int n(\mathbf{r} + a\mathbf{k}) g^{\mathrm{HS}}(a,\bar{\mathbf{n}}(\mathbf{r} + \tfrac{1}{2}a\mathbf{k})) a^2 \mathbf{k}\, d^2k = 0 \end{aligned} \tag{57}$$

For a homogeneous fluid, (57) yields a more accurate equation of state than does (53). Also, the free-energy analog of (57) has been shown to be more accurate than the free-energy analog of (53) in predicting the density profile of Lennard–Jones and dipolar fluids near a hard, flat wall [18]. Thus it is our expectation that (57) will generally be more accurate than (53).

C. Application of the Fischer–Methfessel Model

The Fischer–Methfessel model has been compared [19] with molecular dynamics studies of a Lennard–Jones fluid confined by planar Lennard–Jones walls [20]. The wall–particle potential used in the studies is that of 6–12 LJ-fluid particle interacting with semi-infinite slabs of solids composed of 6–12 LJ particles. In particular, the wall potential used was [21]

$$u^{w}(x) = \theta^{w}(x) + \theta^{w}(L = x) \tag{58}$$

where

$$\theta^{w}(x) = 2\pi\epsilon_{w}\left[\frac{2}{5}\left(\frac{\sigma_{w}}{x}\right)^{10} - \left(\frac{\sigma}{x}\right)^{4} - \frac{\sqrt{2}\sigma^{3}}{3[x + (0.61/\sqrt{2})\sigma_{w}]^{3}}\right] \tag{59}$$

x is the distance away from the left-hand wall and ϵ_{w} and σ_{w} are energy and length parameters characterizing the wall–particle interaction. $\theta^{w}(L - x)$ is the potential energy between the right-hand wall and a particle at x. L denotes the separation of the planar walls.

The molecular dynamics studies have been performed for fluid molecules whose pair potential is

$$u(r_{ij}) = 4\epsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{60}$$

In the theoretical work to be presented here, it is assumed that the hard-sphere interaction replaces the repulsive part of (60) and the attractive part is equated to u^{A}. Thus, in the theoretical work,

$$u^{A}(r_{ij}) = \begin{cases} 0 & r_{ij} < a \\ -4\epsilon\left(\dfrac{a}{r_{ij}}\right)^{6} & r_{ij} > a \end{cases} \tag{61}$$

In the following comparison of theory and simulation, the hard-cord diameter d is equated to the Lennard–Jones parameter (i.e., we set $a = \sigma$). We expect improved results if (57), a model based on the Barker–Henderson translation of a continuous potential model into a hard-sphere cutoff model, were solved instead of the Fischer–Methfessel model. However, (57) is numerically more complex than the Fischer–Methfessel model and has not been investigated yet.

Between planar walls $n = n(x)$, so the Fischer–Methfessel equation simplifies considerably to

$$\frac{d}{dx}\left[kT \ln n(x) + u^e(x) + \int \bar{u}^A(|x - x'|) n(x')\, dx' + 2\pi kT \int_0^x dx'' \int_{-a}^{a} dx'\, x' n(x' + x'') g^{HS}(d, \bar{n}(x'' + \tfrac{1}{2}x'))\right] = 0 \quad (62)$$

where

$$\begin{aligned} \bar{u}^A(|x - x'|) &\equiv \int_{-\infty}^{\infty} \int \bar{u}(|\mathbf{r} - \mathbf{r}'|)\, dy'\, dz' \\ &= -2\pi\epsilon a^2 \qquad |x - x'| < a \\ &= -\frac{2\pi\epsilon a^6}{|x - x'|^4} \qquad |x - x'| > a \end{aligned} \quad (63)$$

Equation (62) can be integrated to yield

$$\kappa = u^e(x) + kT \ln n(x) + \int n(x')\bar{u}^A(|x - x'|)\, dx' + 2\pi kT \int_0^x dx'' \int_{-a}^{a} dx'\, x' n(x' + x'') g^{HS}(a, \bar{n}(x'' + \tfrac{1}{2}x')) \quad (64)$$

The quantity κ is a constant of integration and constitutes a thermodynamic field variable for the system.

It remains to choose the contact pair correlation function for hard spheres. The best available choice is the Carnahan–Starling formula,

$$g^{HS}(a, n) = \frac{1 - (\pi/12)na^3}{[1 - (\pi/6)na^3]^3} \quad (65)$$

The Fischer–Methfessel expression for $\bar{n}$ in the planar case is

$$\bar{n}(x) = \frac{6}{a^3} \int_{|x' - x| \le a/2} \left[\left(\frac{a}{2}\right)^2 - (x' - x)^2\right] n(x')\, dx' \quad (66)$$

With (65) and (66), (64) becomes a solvable nonlinear equation for the density profile $n(x)$ once the constant κ is set. One way to set κ is to require that the number of particles in the simulated system be the same as the number in the theoretical system [i.e., to impose the constant that $\int_0^L n(x)\, dx$ = number of particles in the molecular simulation]. In this way, κ enters (64) as an unknown, but the density constraint adds an equation to compensate for the new unknown.

Density profiles determined by molecular dynamics by Magda et al. [20] were compared to those determined by Vanderlick et al. [19] by solving (64). In the studies the parameters ϵ_w and σ_w were set equal to ϵ

and σ, and as mentioned above, a was taken to be the same as σ. The temperature of the confined fluid film was taken to be $T = 1.2\epsilon/k$. Equation (64) was solved by discretizing the integrals with the trapezoidal rule and solving the resulting algebraic system with the Newton–Raphson method. Comparisons for pores of width $L = 3a$ and $4a$ are shown in Fig. 1. In agreement with the computer simulations, the predicted density profiles show strong layering as a result of fluid–wall interactions.

In another paper, Davis et al. [5] have examined a family of density profiles, pore average densities, and normal pressures or solvation forces. The family is generated at constant κ by varying pore width L. It has been shown that for a one-dimensional van der Waals fluid, the quantity κ defined by (64) is, in fact, the chemical potential, so a family at constant κ would be a family at constant chemical potential. Davis et al. assumed that (64) is also in three dimensions the chemical potential corresponding to the Fischer–Methfessel theory. With this assumption, the results predicted for constant κ can be compared to the molecular dynamics studies of Magda et al., since these studies were carried out at approximately constant chemical potential.

The density of the bulk fluid in equilibrium with the confined fluids in the molecular dynamics simulations was $n_b\sigma^3 = 0.5925$ and the temperature was $kT/\epsilon = 1.2$. Thus the value of κ was calculated by setting $n_b a^3 = 0.5925$ and $kT/\epsilon = 1.2$ in the bulk fluid formula. Also, $\kappa = \mu_0^{HS}(n_b) + n_b \int \bar{u}^A(|s_x|)\, ds_x$ [Eq. (67)], where μ_0^{HS} is the chemical po-

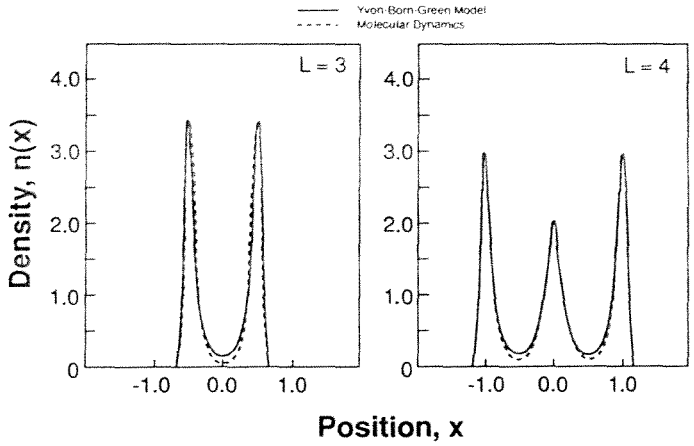


Fig. 1 Comparison of the density profiles predicted by the Fischer–Methfessel approximation with molecular dynamics simulations. (Redrawn from Ref. 19.)

tential of homogeneous hard sphere fluid at density n_b,

$$\mu_o^{HS}(n_b) = kT \ln n_b + \int_0^{n_b} \frac{1}{n} \frac{\partial}{\partial n} \left[\frac{2\pi}{3} n^2 a^3 g^{HS}(n) \right] dn \qquad (67)$$

One can easily show that the equations for the chemical potential of homogeneous fluid implied by (64) is (67). We can alternatively show that the pressure $P_0(n)$ of homogeneous fluid implied by the Fischer–Methfessel approximation yields, with the aid of the Gibbs–Duhem equation $dP_0(n) = nd\mu_0(n)$, (67) for the chemical potential of homogeneous fluid.

A series of density profiles calculated for various pore widths L for the fixed values of κ and T are shown in Fig. 2. The number of fluid layers (i.e., the number of peaks in the density profile) is seen to be a sensitive function of pore width L. Between $L = 2a$ and $L = 2.6a$, fluid film undergoes a transition from a single layer to a double layer. At $L = 2.95a$ the double layer is more favored than at $L = 2.6a$, as indicated by the fact that for $L = 2.95a$. At $L = 3.5a$, two more layers appear to be trying to form, but when the pore width is increased to $L = 4D$, it is a three-layered structure that forms. At $L = 5a$, a well-developed four-layered structure is formed.

The normal pressure of the confined fluid can be computed from

$$P_N = -\int_0^L n(x) \frac{\partial \theta^w}{\partial x}(x)\, dx \qquad (68)$$

where $n(x)$ is the predicted density profile and $\theta^w(x)$ is the particle wall potential given by Eq. (59). The solvation pressure is defined as

$$P_s = P_N - P_0(n_b) \qquad (69)$$

where $P_o(n_b)$ is the pressure of bulk phase ($L = \infty$). For $n_b a^3 = 0.5925$ and $kT/\epsilon = 1.2$, the predicted solvation pressure is shown in Fig. 3 as a function of pore width L. The solvation pressure oscillates strongly as a function of L. The maxima in the solvation pressure coincide with pore widths most favorable to the layered structures. These widths are approximately an integral multiple of a. At these favorable widths the layering is strong, the density peaks near the wall are large, and therefore the normal pressure is large. For L between the favorable widths, the layering is somewhat frustrated and the density peaks near the walls are smaller, with the consequence that the pressure is smaller [e.g., in Fig. 2 we see that the density peaks near the wall for the $L = 2.6a$ pore are smaller than for the $L = 2.95a$ pore, and as shown in Figs. 2 and 3 (e.g., for pore width $L = 2.6a$) the density peaks and pressure are smaller than the density peaks and pressure for $L = 2.95a$, and for pore width $L =$

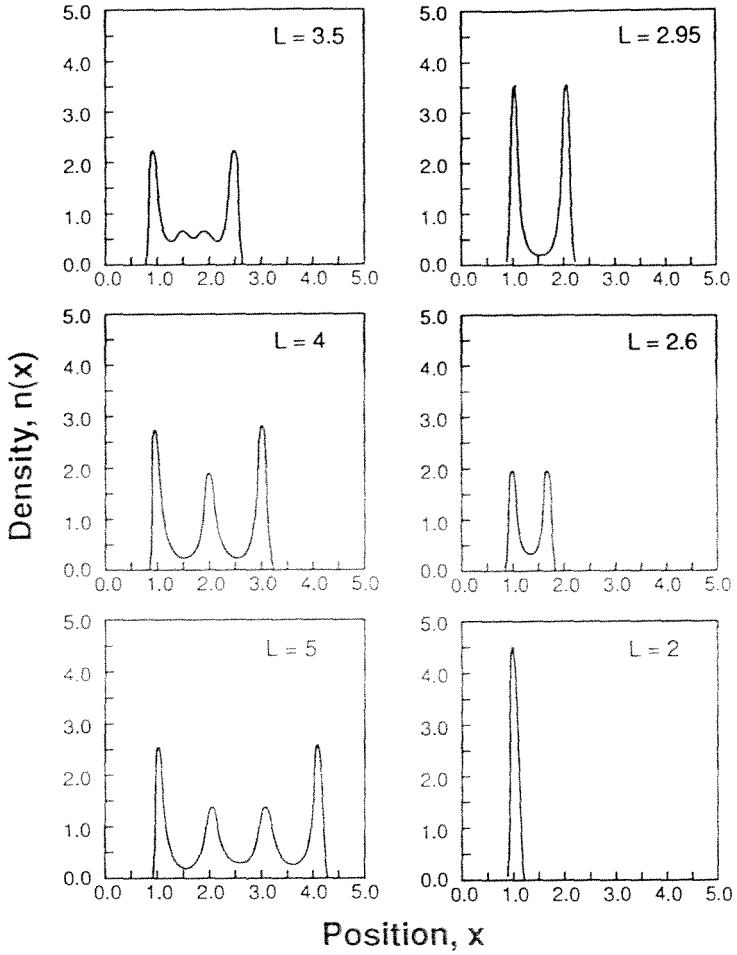


Fig. 2 Density profiles illustrating the effect of pore width L on layering structure. Predicted for the 6–∞ LJ fluid confined by planar 3–4–10 LJ walls. (Redrawn from Ref. 5.)

$3.5a$ the density peaks and pressure are smaller than those for $L = 4a$].

Also shown in Fig. 3 are the molecular dynamics results of Magda et al., for the confined 6–12 Lennard–Jones fluid. The predicted trends in the solvation pressure agree quite well with those calculated in the computer simulations. This oscillatory behavior of the solvation pressure has been observed by the surface forces apparatus, with which one measures the forces required to confine thin films of fluids between opposed mica

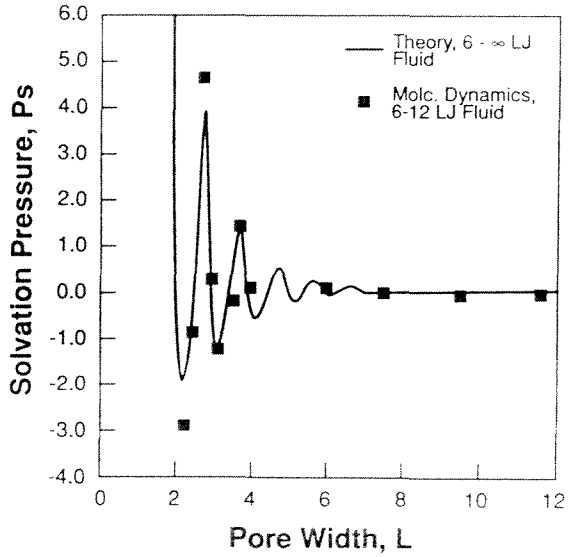


Fig. 3 Solvation pressure, $P_s = P_N - P_0(n_b)$, versus pore width. Comparison of Fischer–Methfessell approximation and molecular dynamics for fluid confined by planar 3–4–10 LJ walls. Pressure in units of ϵ/σ^3 and pore widths in units of σ. (Redrawn from Ref. 5.)

surfaces [22,23]. The theory thus explains these observed trends in terms of fluid layering induced by interaction between the fluid and the confining walls. In a later section we examine the implications of this layering for the transport behavior of confined fluids.

IV. THEORY OF TRACER DIFFUSION

A. Some General Properties

In tracer diffusion the diffusing species is present in such small concentrations that interactions among the diffusing species are negligible. The tracer diffusion coefficient in a strongly inhomogeneous fluid is a second rank tensor since diffusion depends on direction in this case. It follows from the theory [24] of Brownian motion that the local diffusivity tensor for tracer diffusion can be computed from

$$\mathbf{D}(\mathbf{r}) = \lim_{t\to\infty} \frac{\langle[\mathbf{r}_1(t) - \mathbf{r}_1(0)][\mathbf{r}_1(t) - \mathbf{r}_1(0)]\rangle^{\mathbf{r}}}{2t} \tag{70}$$

where $\langle\alpha\rangle^{\mathbf{r}}$ denotes the equilibrium ensemble average of the quantity α

for a tracer particle initially placed at $\mathbf{r}$. This ensemble average can be expressed as

$$\langle \alpha \rangle^{\mathbf{r}} = \int \cdots \int \alpha \left[\frac{\mathcal{P}_N^{eq}(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_N)}{p^{eq}(\mathbf{r}_1)} \right] \times \delta(\mathbf{r} - \mathbf{r}_1)\, d^3r_1 \cdots d^3r_N\, d^3v_1 \cdots d^3v_N \tag{71}$$

where $\mathcal{P}_N{}^{eq} d^3r_1 \cdots d^3v_N$ denotes the probability that in an equilibrium ensemble the tracer particle lies between $\mathbf{r}$ and $\mathbf{r}_1 + d\mathbf{r}_1$ with a velocity between $\mathbf{v}_1$ and $\mathbf{v}_1 + d\mathbf{v}_1$, while the other $i = 1, 2, \ldots, N$ particles lie between $\mathbf{r}_i$ and $\mathbf{r}_i + d\mathbf{r}_i$ with velocities between $\mathbf{v}_i$ and $\mathbf{v}_i + d\mathbf{v}_i$. The quantity $p^{eq}(\mathbf{r}_1)\, d^3r_i$ is the singlet probability that the tracer particle lies between $\mathbf{r}_1$ and $\mathbf{r}_1 + d\mathbf{r}_1$.

In an isotropic fluid $p^{eq}(\mathbf{r}_1) =$ constant and the diffusivity is independent of $\mathbf{r}$ and is equal to

$$\begin{aligned} D &= \lim_{t \to \infty} \frac{\langle [\mathbf{r}_1(t) - \mathbf{r}_1(0)]^2 \rangle}{6t} \\ &= \lim_{t \to \infty} \frac{1}{6t} \int \cdots \int [\mathbf{r}_1(t) - \mathbf{r}_1(0)]^2 \mathcal{P}_N{}^{eq}\, d^3r_1 \cdots d^3v_N \end{aligned} \tag{72}$$

The right-hand side of (70) can be rewritten in terms of the velocity autocorrelation function. By integration along the dynamical path of the diffusing particle, we obtain

$$\mathbf{r}_1(t) - \mathbf{r}_1(0) = \int_0^t \mathbf{v}_1(\tau)\, d\tau$$

so

$$\mathbf{D}(\mathbf{r}) = \lim_{t \to \infty} \frac{1}{2} \int_0^t \int_0^t \langle \mathbf{v}(\tau)\mathbf{v}(\tau') \rangle^{\mathbf{r}}\, d\tau\, d\tau' \tag{73}$$

With the aid of coordinate transformation $\tau, \tau' \to \tau, \tau'' = \tau' - \tau$, the condition of stationarity $\langle \mathbf{v}(\tau\mathbf{v}(\tau') \rangle = \langle \mathbf{v}(0)\mathbf{v}(\tau' - \tau) \rangle$, and the property $\int_0^t d\tau \int_{-\tau}^{t-\tau} d\tau'' \, \alpha(\tau'') = 2 \int_0^t d\tau''(t - \tau'')\alpha(\tau'')$ resulting from interchange of the order of integration, (73) becomes

$$\begin{aligned} \mathbf{D}(\mathbf{r}) &= \lim_{t \to \infty} \int_0^t \langle \mathbf{v}(0)\mathbf{v}(\tau'') \rangle^{\mathbf{r}} \left(1 - \frac{\tau''}{t} \right) d\tau'' \\ &= \int_0^\infty \langle \mathbf{v}(0)\mathbf{v}(\tau'') \rangle^{\mathbf{r}}\, d\tau'' \end{aligned} \tag{74}$$

This is the Green–Kubo formula generalized to tracer diffusion in an anisotropic fluid.

The velocity autocorrelation function is also related to the friction tensor ζ or, more precisely, the mobility tensor ζ^{-1}. If a tracer particle of mass m_1 is subjected to a small constant external field $\mathbf{F}_1{}^P$ that does not affect the other particles of the medium, it will experience at steady state a drift velocity $\mathbf{u}_1$ given by

$$\mathbf{F}_1{}^P = \zeta \cdot \mathbf{u}_1 \tag{75}$$

To relate ζ to the velocity autocorrelation function, we solve the Liouville equation for a small force $\mathbf{F}_1{}^e$ applied at $t = 0$ to an equilibrium system. The Liouville equation for the n-particle probability distribution function is

$$\frac{\partial \mathcal{P}_N}{\partial t} = i\mathcal{L}\mathcal{P}_N \tag{76}$$

where $i \equiv \sqrt{-1}$ and $\mathcal{P}_N(\mathbf{r}_1, \ldots, \mathbf{r}_N, \mathbf{v}_1, \ldots, \mathbf{v}_j, t)\, d^3r_1, \ldots, d^3v_N$ is the probability that particles 1,2, . . . lie between $\mathbf{r}_1$ and $\mathbf{r}_1 + d\mathbf{r}_1$, $\mathbf{r}_2$ and $\mathbf{r} + d\mathbf{r}_2$, . . . and have velocities between $\mathbf{v}_1$ and $\mathbf{v}_1 + d\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_2 + d\mathbf{v}_2$, . . . at time t. The Liouville operator $\mathcal{L}$ is defined by

$$\mathcal{L} \equiv -i \sum_{j=1}^{N} \left[\mathbf{v}_j \cdot \nabla_j + \frac{\mathbf{F}_j}{m_j} \cdot \nabla \mathbf{v}_j \right] \tag{77}$$

From the definition of $\mathcal{L}$ and Newton's law $d\mathbf{v}_j/dt = \mathbf{F}_j/m_j$, with $d\mathbf{r}_j/dt = \mathbf{v}_j$, it follows that

$$\frac{d\mathbf{v}_j}{dt} = i\mathcal{L}\mathbf{v}_j \quad \text{and} \quad \frac{d\mathbf{r}_j}{dt} = i\mathcal{L}\mathbf{r}_j \tag{78}$$

or

$$\mathbf{v}_j(t) = e^{i\mathcal{L}t}\mathbf{v}_j(0) \quad \text{and} \quad \mathbf{r}_j(t) = e^{i\mathcal{L}t}\mathbf{r}_j(0) \tag{79}$$

from which it follows that

$$h(\mathbf{r}_1(\pm t), \ldots, \mathbf{v}_N(\pm t)) = e^{\pm i\mathcal{L}t} h(\mathbf{r}_1(0), \ldots, \mathbf{v}_N(0)) \tag{80}$$

where h is any function of the dynamical variables $\mathbf{r}_1, \ldots, \mathbf{v}_N$.

In the absence of the perturbing force $\mathbf{F}_1{}^P$ and at equilibrium, the solution to (76) is

$$\mathcal{P}_N{}^{eq} = \frac{e^{-H/kT}}{\int e^{-H/kT}\, d^3r_1 \cdots d^3v_N} \tag{81}$$

where $H(\mathbf{r}_1, \ldots, \mathbf{v}_N)$ is the unperturbed Hamiltonian of the fluid. Suppose that a uniform constant, perturbing force $\mathbf{F}_1^P$ is turned on at time $t = 0$. The distribution function can be expressed in the form

$$\mathcal{P}_N = \mathcal{P}_N^{eq} + \Delta\mathcal{P}_N \tag{82}$$

where for sufficiently small perturbation $\Delta\mathcal{P}_N$ is linear in the perturbing force $\mathbf{F}_1^P$. Denoting the Liouville operator by $\mathcal{L}' = \mathcal{L} + \Delta\mathcal{L}$, where $\mathcal{L}$ is the unperturbed operator and

$$\Delta\mathcal{L} = -i\frac{\mathbf{F}_1^P}{m_1}\cdot\nabla\mathbf{v}_1 \tag{83}$$

we can linearize the Liouville equation to obtain

$$\begin{aligned}\frac{\partial\Delta\mathcal{P}_N}{\partial t} &= -i\mathcal{L}\,\Delta\mathcal{P}_N - \frac{\mathbf{F}_1^P}{m_1}\cdot\nabla_{\mathbf{v}_1}\mathcal{P}_N^{eq} \\ &= -i\mathcal{L}\,\Delta\mathcal{P}_N + \frac{1}{kT}\mathbf{F}_1^P\cdot\mathbf{v}_1\mathcal{P}_N^{eq}\end{aligned} \tag{84}$$

The formal solution to this equation is

$$\Delta\mathcal{P}_N(\mathbf{r}_1, \ldots, \mathbf{v}_N, t) = \frac{\mathbf{F}_1^P}{kT}\int_0^t [e^{-i\mathcal{L}(t-\tau)}\mathbf{v}_1\, d\tau]\mathcal{P}_N^{eq}(\mathbf{r}_1, \ldots, \mathbf{v}_N) \tag{85}$$

Since $\mathcal{L}H = 0$, $e^{-i\mathcal{L}(t-\tau)}\mathcal{P}_N^{eq}(\mathbf{r}_1, \ldots, \mathbf{v}_N) = \mathcal{P}_N^{eq}(\mathbf{r}_1, \ldots, \mathbf{v}_N)$ so the operator $e^{-i\mathcal{L}(t-\tau)}$ is applied only to $\mathbf{v}_1$ in (85).

The perturbation Δf to the singlet velocity distribution function can be obtained by multiplying (85) by $N_1\, d^3r_2\cdots d^3r_N\, d^3v_2\cdots d^3v_N \equiv N_1(d^3r\, d^3v)^{N-1}$ and integrating the result,

$$\Delta f = \frac{N_1}{kT}\mathbf{F}_1^P\cdot\int_0^t\int\cdots\int\mathcal{P}_N^{eq}(d^3r\, d^3v)^{N-1}\, dt \tag{86}$$

where we have used the property that $\mathcal{L}\to -\mathcal{L}$ under the transformation $\mathbf{v}_i\to -\mathbf{v}_i$ and have noted that $\exp[iL(t-\tau)]\mathbf{v}_1 = \mathbf{v}_1(t-\tau)$. The local drift velocity $\mathbf{u}_1(\mathbf{r})$ is then given by

$$\begin{aligned}n_1(\mathbf{r})\mathbf{u}_1(\mathbf{r}) &= \int\mathbf{v}_1\,\Delta f\, d^3v_1 \\ &= \frac{N_1}{kT}\mathbf{F}^P\cdot\int_0^t\int\cdots\int[\mathbf{v}_1 e^{-i\mathcal{L}(t-\tau)}\mathbf{v}_1]P_N^{eq}(d^3r)^{N-1}(d^3v)^N\, d\tau \\ &= n_1(\mathbf{r})\zeta^{-1}(\mathbf{r})\cdot\mathbf{F}_1^P\end{aligned} \tag{87}$$

where the local friction tensor $\zeta(\mathbf{r})$ is defined (in the long time limit) by

$$\begin{aligned}\zeta(\mathbf{r}) &= \lim_{t\to\infty} \frac{N_1}{n_1(\mathbf{r})kT}\int_0^t \int \cdots \int [\mathbf{v}_1 e^{-i\mathscr{L}\tau'}\mathbf{v}_1]\mathscr{P}_N{}^{eq}(d^3r)^{N-1}(d^3v)^N\, d\tau' \\ &= \lim_{t\to\infty} \frac{1}{kT}\int_0^t \int \cdots \int \mathbf{v}_1(0)\mathbf{v}_1(\tau')\frac{\mathscr{P}_N{}^{eq}}{p_{r_1}^{eq}}\delta(\mathbf{r}-\mathbf{r}_1)(d^3r\, d^3v)^N\, d\tau' \qquad (88) \\ &= \frac{1}{kT}\int_0^\infty \langle \mathbf{v}_1(0)\mathbf{v}_1(\tau')\rangle^r\, d\tau\end{aligned}$$

Comparison of (74) and (88) reveals the relation

$$\mathbf{D}(\mathbf{r}) = kT\zeta^{-1}(\mathbf{r}) \qquad (89)$$

which is the Einstein relation generalized to anisotropic fluids.

B. Self-Diffusion in Several Model Fluids

In Section IV.C we solve the modified Enskog equation for tracer diffusion. However, under the assumption that diffusive transport involves only uncorrelated binary collisions, we can find the self-diffusion coefficients for a whole class of molecular models in which the forces are impulsive (i.e., models for which the changes in the intermolecular potentials occur exclusively at discontinuities). In particular, we consider fluids whose molecular pairs interact at a separation $r_{ij} = d$ as smooth spheres, rough spheres, or loaded spheres and at a separation $r_{ij} = Ra$, $R > 1$, via a square-well potential. Thus we assume that

$$u^A(r_{ij})\begin{cases} -\epsilon & a < r_{ij} < Ra \\ 0 & r_{ij} > Ra \end{cases} \qquad (90)$$

By "smooth sphere" we mean a hard-sphere particle whose center of mass coincides with its center of geometry. Only the normal components of the translational velocities of a pair of smooth spheres change upon collision. Smooth spheres do not exchange rotational energy. A loaded sphere is also smooth, but its center of mass is offset from its geometrical center, so loaded spheres can exchange rotational energy through collisions. The center of mass and geometrical center of a rough sphere coincide, but the surface of a rough sphere is "perfectly rough" in the sense that the relative velocities of the points of contact reverse in a collision between a pair of rough spheres.

Isotropic fluid transport coefficients (i.e., transport coefficients in weakly homogeneous fluids) have been studied extensively with smooth, rough, and loaded sphere models [1,25,26]. For tracer diffusion, the rough-

sphere theoretical result has led to a useful approximation formula relating the diffusivity of a polyatomic solute to the product of a molecular shape factor and the diffusivity of hard spheres [27–29]. Brown and Davis have obtained formulas for the self-diffusion coefficients of fluids of square-well particles interacting with smooth, rough, or loaded sphere repulsions [30]. The results presented in this section are extensions of the Brown–Davis theory to anisotropic fluids.

As in isotropic fluids, we assume that the velocity autocorrelation function decays exponentially. Since the relaxation is in an anisotropic system, the formula for exponential decay is

$$\langle \mathbf{v}_1(0)\mathbf{v}_1(t)\rangle^{\mathbf{r}} = \langle v_1(0)^2\rangle \tfrac{1}{3}\mathbf{U}e^{-t\boldsymbol{\tau}^{-1}(\mathbf{r})} \tag{91}$$

where $\mathbf{U}$ is the unit tensor ($\mathbf{U}\cdot\mathbf{a} = \mathbf{a}$) and $\boldsymbol{\tau}(\mathbf{r})$ is the local relaxation tensor. The diffusion coefficient for this approximation is

$$\mathbf{D}(\mathbf{r}) = \tfrac{1}{3}\langle v_1(0)^2\rangle \boldsymbol{\tau}(\mathbf{r}) = \frac{kT}{m}\boldsymbol{\tau}(\mathbf{r}) \tag{92}$$

From (89) and (92) it follows that $\boldsymbol{\tau} = m\boldsymbol{\zeta}^{-1}$ for the exponential approximation. In addition to (91), we assume that uncorrelated binary collisions are responsible for velocity relaxation—these assumptions are equivalent to those on which the Enskog equation is based. Expansion of the autocorrelation function about $t = 0$ yields the relationship

$$\boldsymbol{\tau}^{-1}(\mathbf{r}) = \frac{m}{kT}\lim_{\Delta t\to 0}\frac{\langle \mathbf{v}_1\Delta\mathbf{v}_1\rangle^{\mathbf{r}}}{\Delta t} \tag{93}$$

where $\Delta\mathbf{v}_1$ is the velocity change during binary collisions.

It should be pointed out that for systems of finite width W in any direction x, the quantity $[x(t) - x(0)]^2 \le [W - x(0)]^2$. Thus $\langle[x(t) - x(0)]^2\rangle/2t \to 0$ as $t \to \infty$. This means that the self-diffusion coefficient in the confined direction would appear to be zero if the mean-square displacement is measured for times large compared to the time it takes for the diffusing particle to traverse the distance W. However, if the time to traverse the distance W is long compared to the diffusive relaxation time the quantity $\langle[x(t) - x(0)]^2\rangle/2t$ will first converge to the self-diffusion coefficient and remain constant until t approaches the time to traverse W. Thus when we apply the exponential approximation and compute $\mathbf{D}(\mathbf{r})$ from (92), we have assumed that the time for the diffusing particle to traverse the confined distance is large compared to the relaxation time $\tau_{xx}(\mathbf{r})$.

Rigid particles interacting via a square-well potential undergo four types of binary collisions, so the ensemble average indicated at (88) has

to include a properly weighted average of these. After Davis et al. [31,32], we identify the following four types of binary collisions:

1. Two smooth or loaded spheres initially approaching each other collide at a distance of separation a. The relative velocities of their centers of geometry are reversed in such a collision. A pair of rough spheres will experience a reversal of their surface velocities at the point of contact upon collision at a separation a.
2. A pair of approaching particles initially separated by a distance greater than Ra will, upon reaching the separation $r_{ij} = Ra$, convert an amount ϵ of potential energy into translational energy in the case of smooth and rough spheres and into translational and rotational energy in the case of loaded spheres.
3. A pair of particles initially separated by a distance less than Ra and moving apart can with enough kinetic and rotational energy escape the square well. In this case, at the separation $r_{ij} = Ra$, the smooth or rough spheres will give up translational energy of amount ϵ and the loaded spheres will give up translational and rotational energy totaling ϵ.
4. A pair of spheres initially separated by a distance less than Ra and moving apart cannot undergo an escape collision at $r_{ij} = Ra$ if they do not possess sufficient energy. In this case, the smooth or rough spheres experience at $r_{ij} = Ra$ a reversal of the relative velocity of their centers and the loaded spheres will experience a reversal of the relative points on their surfaces along the line between their centers (geometric centers).

Consider a collision between particles 1 and 2. The relative positions of colliding particles will be identified by a unit vector $\mathbf{k}$ directed from the center (geometrical) of spheres 2 toward the center of sphere 1. The angular velocities $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ must be included in the collision dynamics for rough and loaded spheres. For loaded spheres we must also specify their orientations $\mathbf{e}_1$ and $\mathbf{e}_2$ (unit vectors lying between the center of geometry and the center of mass and the distance v that the center of mass lies from the center of the spheres. The force $\mathbf{F}_{12}$ that particle 2 exerts on particle 1 is impulsive for the models considered, so we can integrate the equation of motion over the infinitesimal times δt to obtain the collisional changes of $\mathbf{v}_1$ and $\boldsymbol{\omega}_1$:

$$\Delta \mathbf{v}_1 = -\frac{1}{m}\int_t^{t+\delta t} \mathbf{F}_{12}\, dt = -\frac{\mathbf{J}}{m} \tag{94}$$

$$\Delta \boldsymbol{\omega}_1 = -\mathbf{I}^{-1} e \mathbf{e}_1 \times \mathbf{J} \qquad \text{for loaded spheres} \tag{95}$$

and

$$\Delta\boldsymbol{\omega}_1 = \frac{a}{2I}\,\mathbf{k} \times \mathbf{J} \qquad \text{for rough spheres} \tag{96}$$

where $\mathbf{J}$ is the linear momentum exchanged in a binary collision and $\mathbf{a} \times \mathbf{b}$ denote the cross product of vectors $\mathbf{a}$ and $\mathbf{b}$. The moment of inertia is a tensor $\mathbf{I}$ for loaded spheres and is a scalar I for rough spheres. The quantity e denotes the distance of the center of mass from the center of geometry of a loaded sphere and $\mathbf{e}_i$ is the orientation of the center of mass of loaded sphere i relative to its center of geometry.

The quantity $\mathbf{J}$ can be determined from the laws of conservation of energy and momentum. For example, for a pair of smooth spheres colliding at a, we have $\mathbf{J} = J\mathbf{k}$, $m\mathbf{v}_1 + m\mathbf{v}_2 = m\mathbf{v}_1' + m\mathbf{v}_2'$, and $\frac{1}{2}mv_1^2 + mv_2^2 = \frac{1}{2}mv_1'^2 + \frac{1}{2}mv_2'^2$. Since $\Delta\mathbf{v}_1 = \mathbf{v}_1' - \mathbf{v}_1 = -\Delta\mathbf{v}_2 = -(\mathbf{v}_2' - \mathbf{v}_2) = -\mathbf{J}/m$, it follows from energy conservation that $\frac{1}{2}mv_1^2 + \frac{1}{2}mv_2^2 = \frac{1}{2}m(v_1^2 - 2\mathbf{J}\cdot\mathbf{v}_1/m + J^2/m^2) + \frac{1}{2}m(v_2^2 + 2\mathbf{J}\cdot\mathbf{v}_2/m + J^2/m^2)$, or $J^2 = -m\mathbf{J}\cdot\mathbf{v}_{21}$, where $\mathbf{v}_{21} = \mathbf{v}_2 - \mathbf{v}_1$. Thus $J^2 = -Jm\mathbf{v}_{21}\cdot\mathbf{k}$ or

$$\mathbf{J} = -m\mathbf{v}_{21}\cdot\mathbf{k}\mathbf{k} \tag{97}$$

Results for all four types of binary collisions for the three models considered here are tabulated in Table 1. The following notation is used in the table:

$$\mathbf{v}_{21} = \mathbf{v}_2 + \mathbf{v}_1 \tag{98}$$

$$\mathbf{g}_{21} = \mathbf{v}_{21} - \frac{a}{2}(\boldsymbol{\omega}_1 + \boldsymbol{\omega}_2) \times \mathbf{k} \tag{99}$$

$$\tilde{\mathbf{g}}_{21} = \mathbf{v}_{21} + \boldsymbol{\omega}_2 \times \mathbf{e}_2 e - \boldsymbol{\omega}_1 \times \mathbf{e}_1 e \tag{100}$$

$$\mu_1 = \mu_2 = \frac{2}{m} \tag{101}$$

$$\mu_3 = \frac{2}{m} + \mathbf{I}^{-1}\colon [e(\mathbf{e}_1 \times \mathbf{k})(\mathbf{e}_1 \times \mathbf{k}) + e(\mathbf{e}_2 \times \mathbf{k})(\mathbf{e}_2 \times \mathbf{k})] \tag{102}$$

$$\kappa = \frac{4I}{m\sigma_1^2} \tag{103}$$

Let us now turn to the evaluation of (93). The probability that a pair of particles in a configuration suitable for a binary collision of type i during the time at Δt is

$$\phi(\mathbf{m}_1)\phi(\mathbf{m}_2)p^{eq}(\mathbf{r}_1)p^{eq}(\mathbf{r}_1 - a^{(i)}\mathbf{k})g(\mathbf{r}_1,\mathbf{r}_1 - a^{(i)}\mathbf{k})\,d\mathbf{m}_1\,d\mathbf{m}_2\,d^3r_1\,dV_{\Delta t}^{(2)} \tag{104}$$

where $a^{(i)} = \lim_{\epsilon\to 0}(a + \epsilon) \equiv a+$ for a type 1 collision, $A^{(i)} = \lim_{\epsilon\to 0}(Ra + \epsilon) \equiv Ra+$ for a type 2 collision, and $a^{(i)} = \lim_{\epsilon\to 0}(Ra - \epsilon) \equiv$

Table 1 Momentum Exchange in the Four Binary Collisions for Spherical Molecules Interacting via a Square-Well Potential and Having Smooth, Rough, and Loaded Hard Cores, Respectively

Type of collision	Velocity	J
	Smooth hard core	
1	$\mathbf{v}_{21}\cdot\mathbf{k} > 0$	$\left[-\frac{2}{\mu_1}\cdot\mathbf{v}_{21}\mathbf{kk}\right]$
2	$\mathbf{v}_{21}\cdot\mathbf{k} > 0$	$\frac{1}{\mu_1}\left[-\mathbf{v}_{21}\cdot\mathbf{k} - \sqrt{(\mathbf{v}_{21}\cdot\mathbf{k})^2 + 2\epsilon\mu_1}\right]\mathbf{k}$
3	$\mathbf{v}_{21}\cdot\mathbf{k} < \sqrt{2\epsilon\mu_1}$	$\frac{1}{\mu_1}\left[-\mathbf{v}_{21}\cdot\mathbf{k} + \sqrt{(\mathbf{v}_{21}\cdot\mathbf{k})^2 - 2\epsilon\mu_1}\right]\mathbf{k}$
4	$-\sqrt{2\epsilon\mu_1} < \mathbf{v}_{21}\cdot\mathbf{k} < 0$	$\left[-\frac{2}{\mu_1}\mathbf{v}_{21}\cdot\mathbf{kk}\right]$
	Rough hard core	
1	$\mathbf{v}_{21}\cdot\mathbf{k} > 0$	$-\frac{2}{\mu_2}\frac{\kappa}{\kappa+1}\left(\mathbf{g}_{21} + \frac{1}{\kappa}\mathbf{v}_{21}\cdot\mathbf{kk}\right)$
2	$\mathbf{v}_{21}\cdot\mathbf{k} > 0$	$\frac{1}{\mu_2}\left[-\mathbf{v}_{21}\cdot\mathbf{k} - \sqrt{(\mathbf{v}_{21}\cdot\mathbf{k})^2 + 2\epsilon\mu_1}\right]\mathbf{k}$
3	$\mathbf{v}_{21}\cdot\mathbf{k} < \sqrt{2\epsilon\mu_2}$	$\frac{1}{\mu_2}\left[-\mathbf{v}_{21}\cdot\mathbf{k} + \sqrt{(\mathbf{v}_{21}\cdot\mathbf{k})^2 - 2\epsilon\mu_1}\right]\mathbf{k}$
4	$-\sqrt{2\epsilon\mu_1} < \mathbf{v}_{21}\cdot\mathbf{k} < 0$	$\left[-\frac{2}{\mu_2}\mathbf{v}_{21}\cdot\mathbf{kk}\right]$
	Loaded hard core	
1	$\tilde{\mathbf{g}}_{21}\cdot\mathbf{k} > 0$	$\left[-\frac{2}{\mu_3}\cdot\tilde{\mathbf{g}}_{21}\mathbf{kk}\right]$
2	$\tilde{\mathbf{g}}_{21}\cdot\mathbf{k} > 0$	$\frac{1}{\mu_3}\left[-\tilde{\mathbf{g}}_{21}\cdot\mathbf{k} - \sqrt{(\tilde{\mathbf{g}}_{21}\cdot\mathbf{k})^2 + 2\epsilon\mu_3}\right]\mathbf{k}$
3	$\tilde{\mathbf{g}}_{21}\cdot\mathbf{k} < \sqrt{2\epsilon\mu_3}$	$\frac{1}{\mu_3}\left[-\tilde{\mathbf{g}}_{21}\cdot\mathbf{k} + \sqrt{(\tilde{\mathbf{g}}_{21}\cdot\mathbf{k})^2 - 2\epsilon\mu_3}\right]\mathbf{k}$
4	$-\sqrt{2\epsilon\mu_3} < \tilde{\mathbf{g}}_{21}\cdot\mathbf{k} < 0$	$\left[-\frac{2}{\mu_3}\tilde{\mathbf{g}}_{21}\cdot\mathbf{kk}\right]$

Source: Ref. 30.

$Ra-$ for types 3 and 4 collisions. $\mathbf{m}_i$ and $d\mathbf{m}_i$ are the velocity and the velocity volume element for smooth spheres. For loaded and rough spheres, $\mathbf{m}_i$ denotes the linear and the angular velocities and $d\mathbf{m}_i$ is the corresponding volume element. $\phi(\mathbf{m}_i)$ is the equilibrium distribution of the relevant velocities (Gaussian in $\mathbf{v}_i$ and $\boldsymbol{\omega}_i$).

The pair correlation function is discontinuous at d and Rd. In particular,

$$g(\mathbf{r}_1,\mathbf{r}_1 - Ra\mathbf{k}-) = e^{\epsilon/kT} g(\mathbf{r}_1,\mathbf{r}_1 - Ra\mathbf{k}+) \tag{105}$$

The volume $dV_{\Delta t}^{(2)}$ in which particle 2 must reside to collide with particle 1 during the time Δt is $|\mathbf{v}_{21}\cdot\mathbf{k}|\, a^{(i)2}\, d^2k\, \Delta t$ for smooth and rough spheres and $|\mathbf{g}_{21}\cdot\mathbf{k}|\, a^{(i)2}\, d^2k\, \Delta t$ for loaded spheres. Thus the contribution to $\langle \mathbf{v}_1\, \Delta\mathbf{v}_1\rangle$ from a type i collision for smooth spheres is

$$\lim_{\Delta t\to 0} \frac{\langle \mathbf{v}_1\, \Delta\mathbf{v}_1^{(i)}\rangle^{\mathrm{r}}}{\Delta t} = (N-1)\int\int_{\mathbf{v}_{21}\cdot\mathbf{k}>0} \mathbf{v}_1 \left(\frac{-\mathbf{J}^{(i)}}{m}\right) \mathbf{k}\phi(\mathbf{v}_1)\phi(\mathbf{v}_2) \times p(\mathbf{r} - a^{(i)}\mathbf{k})g(\mathbf{r},\mathbf{r} - a^{(i)}\mathbf{k})a^{(i)2}(\mathbf{v}_{21}\cdot\mathbf{k})\, d^2k\, d^3v_1\, d^3v_2 \tag{106}$$

The factor of $N-1$ in this expressing accounts for the fact that there are $N-1$ particles with which particle 1 can collide. Expressing $\mathbf{v}_1$ in the Cartesian coordinates $\mathbf{v}_1 = v_{1x}\mathbf{i} + v_{1y}\mathbf{j} + v_{1z}\mathbf{k}$, we see that the velocity integrals in (107) are the same as those that occur in the isotropic fluid theory. This enables us to use the results already obtained by Brown and Davis [30]. The same thing is true for the rough and loaded spheres, so in all cases it follows that

$$\tau^{-1}(\mathbf{r}) = \frac{m}{kT}\lim_{\Delta t\to 0}\frac{\langle \mathbf{v}_1\, \Delta\mathbf{v}^{(i)}\rangle^{\mathrm{r}}}{\Delta t} = \left\{\begin{matrix}\text{homogeneous}\\ \text{fluid result}\end{matrix}\right\} \int\int p^{\mathrm{eq}}(\mathbf{r} - a^{(i)}\mathbf{k})g(\mathbf{r},\mathbf{r} - a^{(i)}\mathbf{k})\, d^{(i)2}\mathbf{k}\mathbf{k}\, d^2k \tag{107}$$

Comparison of (107) and the homogeneous results of Brown and Davis yields the following results for a fluid of square-well particles having a smooth, rough, or loaded hard-core repulsion:

Smooth hard cores:

$$\tau^{-1} = \frac{8a^2(\pi mkT)^{1/2}}{3m}\left[\Lambda_{\mathrm{HC}} + \Lambda_{\mathrm{SW}}R^2\Xi\left(\frac{\epsilon}{kT}\right)\right] \tag{108}$$

Rough hard cores:

$$\tau^{-1} = \frac{8a^2(\pi mkT)^{1/2}}{3m}\left[\frac{2\kappa+1}{\kappa+1}\Lambda_{\mathrm{HC}} + \Lambda_{\mathrm{SW}}R^2\,\Xi\left(\frac{\epsilon}{kT}\right)\right] \tag{109}$$

Loaded hard cores:

$$\tau^{-1} = \frac{2U(\alpha)a^2(\pi mkT)^{1/2}}{3m}\left[\Lambda_{HC} + \Lambda_{SW}R^2\Xi\left(\frac{\epsilon}{kT}\right)\right] \tag{110}$$

where

$$\Lambda_{HC} = \frac{3}{4\pi}\int n(\mathbf{r} + \mathbf{k}d)g(\mathbf{r},\mathbf{r} + a\mathbf{k}+)\mathbf{kk}\, d^2k \tag{111}$$

and

$$\Lambda_{SW} = \frac{3}{4\pi}\int n(\mathbf{r}_1 + Ra\mathbf{k})g(\mathbf{r},\mathbf{r} + Ra\mathbf{k}+)\mathbf{kk}\, d^2k \tag{112}$$

$$\Xi\left(\frac{\epsilon}{kT}\right) = e^{\epsilon/kT} - \frac{\epsilon}{2kT} - 2\int_0^\infty x^2\left(x^2 + \frac{\epsilon}{kT}\right)^{1/2} e^{-x^2}\, dx \tag{113}$$

and

$$U(\alpha) = \int_0^\pi \int_0^\pi [1 + \alpha(\sin^2\theta_1 + \sin^2\theta_2)]^{-1/2} \sin\theta_1 \sin\theta_2\, d\theta_1\, d\theta_2 \tag{114}$$

where

$$\alpha = \frac{me^2}{2\Gamma} \tag{115}$$

Γ is the principal moment of inertia of the loaded spheres ($\mathbf{I} = \Gamma\mathbf{e}_1\mathbf{e}_1 + \Gamma\mathbf{e}_2\mathbf{e}_2$). The functions $\Xi(\epsilon/kT)$ and $U(\alpha)$ have been tabulated by Brown and Davis. In obtaining the given forms of Λ_{HC} and Λ_{SW}, we have used the property $n(\mathbf{r}) = Np(\mathbf{r})$.

For a given equilibrium density distribution $n(\mathbf{r})$, $\tau(\mathbf{r})$ can be computed from (108), (109), or (110). The equilibrium YBG equation is the same for all three of these models. The equation is

$$kT\,\nabla n(\mathbf{r}) + n\,\nabla u^e(\mathbf{r}) + kT(1 - e^{\beta\epsilon})(Ra)^2\int n(\mathbf{r})n(\mathbf{r} + Ra\mathbf{k}) \times g(\mathbf{r},\mathbf{r} + Ra\mathbf{k}+)\mathbf{k}\, d^2k + kTa^2\int n(\mathbf{r})n(\mathbf{r} + a\mathbf{k})g(\mathbf{r},\mathbf{r} + a\mathbf{k}+)\mathbf{k}\, d^2k = 0 \tag{116}$$

As pointed out previously, a closure assumption has to be made to render (116) solvable. Several closure approximations that have been used are discussed in Section III.

To apply the theory presented above, once a closure approximation is introduced, (116) can then be solved for $n(\mathbf{r})$, $\tau^{-1}(\mathbf{r})$ can be computed

from Eqs. (108) to (110), and its inverse τ can be used in (92) to predict the local self-diffusion coefficient tensor $\mathbf{D}(\mathbf{r})$. Such calculations have not yet been done for the models introduced in this section. Hopefully, such applications will be made in the future for fluids in slit pores, cylindrical pores, and periodic pores (zeolites)—comparison of the models for various parameters will provide insight into the relative roles of repulsive and attractive forces and of rotational collision dynamics in the diffusion behavior of strongly inhomogeneous fluids. In the next section we solve the modified Enskog equation for tracer diffusion and apply the results to self-diffusion in fluids confined to slit pores as a function of separation of the confining pore walls.

C. Modified Enskog Theory of Tracer Diffusion

Let us return now to the modified Enskog equation for a multicomponent mixture. In a transport situation the vector $\mathbf{H}_i$ acts as a driving force for diffusion, where, by definition,

$$\mathbf{H}_i = kT\nabla \ln n_i - \nabla u_i^e - \sum_j \int n_j(\mathbf{r}+\mathbf{s})g_{ij}(\mathbf{r}+\mathbf{s})\frac{\mathbf{s}}{s}\frac{du_{ij}^A}{ds}d^3s + \sum_j kT\int n_j(\mathbf{r}+a_{ij}\mathbf{k})g_{ij}(\mathbf{r},\mathbf{r}+a_{ij}\mathbf{k})a_{ij}^2\mathbf{k}\,d^2k \tag{117}$$

At equilibrium $\mathbf{H}_i = 0$ and in fact is equivalent to the equilibrium YBG equation.

In tracer diffusion, the tracer component is added in extremely small amounts, $n_1 \ll n_j$, so the other components are not perturbed appreciably from equilibrium, that is,

$$f_j = n_j(\mathbf{r})\varphi_j, \qquad j = 2, \ldots \tag{118}$$

$$\varphi_j = \left(\frac{m_j}{2\pi kT}\right)^{3/2}\exp\left(\frac{-m_j v_j^2}{2kT}\right) \tag{119}$$

where the n_j satisfy the equilibrium YBG equations $\mathbf{H}_j = 0$. We consider steady state, so that $\partial f_1/\partial t = 0$, and write the tracer singlet distribution function as

$$f_1 = n_1(\mathbf{r})\varphi_1(v_1)[1 + \psi(\mathbf{r},\mathbf{v}_1)] \tag{120}$$

where $\psi(\mathbf{r},\mathbf{v}_1)$ is the perturbation of f_1 from equilibrium.

Substitution of (118) and (120) into the modified Enskog equation [Eq. (2)], neglect of terms of order n_1^2, and some rearrangement lead to the

result [2]

$$(1 + \psi)\varphi_1 \mathbf{v}_1 \cdot \mathbf{H}_1 + \varphi_1 \left[kT \mathbf{v}_1 \cdot \nabla\psi - \frac{1}{m_1} \nabla u^e \cdot \nabla_{\mathbf{v}_1}\psi \right.$$
$$\left. + \frac{1}{m^1} \nabla_{\mathbf{v}_1}\psi \cdot \sum_{j \neq 1} \int n_j(\mathbf{r} + \mathbf{s}) g_{ij}(\mathbf{r}, \mathbf{r} + \mathbf{s}) \frac{\mathbf{s}}{s} \frac{du_{ij}^A}{ds} d^3s \right]$$
$$= \sum_{j \neq 1} kT \int_{\mathbf{v}_{ji} \cdot \mathbf{k} > 0} [n_j(\mathbf{r} + a_{ij}\mathbf{k}) g_{ij}(\mathbf{r}, \mathbf{r} + a_{ij}\mathbf{k}) \psi(\mathbf{r}, \mathbf{v}_1')$$
$$- n_j(\mathbf{r} - a_{ij}\mathbf{k}) g_{ij}(\mathbf{r}, \mathbf{r} - a_{ij}(\mathbf{k}) \psi(\mathbf{r}, \mathbf{v}_1)] \varphi(v_1) \varphi(v_j) a_{ij}^2 \mathbf{v}_{ji} \cdot \mathbf{k} \, d^2k \, d^3v_1 \quad (121)$$

To obtain the form of the right-hand side of (121) we used the relationship, $\varphi(v_1')\varphi(v_j') = \varphi(v_1)\varphi(v_j)$, arising from the conservation of kinetic energy in a hard-sphere collision.

We are interested in a linear first-order transport theory. The terms having the factors $\psi\mathbf{H}_1$ or $\nabla\psi$ are nonlinear or second order in the deviation from equilibrium and so are dropped from (121). The resulting equation can be solved by the Chapman–Enskog procedure, which is to expand ψ in a series of Sonine polynomials. The lowest-order approximation has been shown to be excellent for isotropic fluids and will be assumed sufficient here as well. The approximation

$$\psi = \mathbf{a}(\mathbf{r}) \cdot \mathbf{v}_1 \quad (122)$$

preserves the property $n_1 = \int f_1 \, d^3v_1$. The unknown $\mathbf{a}(\mathbf{r})$ is determined by the Galerkin procedure of multiplying (121) by $\mathbf{v}_1 \, d^3v_1$, integrating, and solving the resulting algebraic equation for $\mathbf{a}(\mathbf{r})$. The result is

$$\mathbf{a} = \frac{m_1}{kT} \zeta^{-1} \cdot \mathbf{H}_1 \quad (123)$$

where the local Enskog friction tensor is given by

$$\zeta(\mathbf{r}) = \sum_{j \neq 1} \frac{m_1}{\pi} \left(\frac{2\pi kT}{m_{1j}} \right)^{1/2} \int g_{1j}(\mathbf{r}, \mathbf{r} + a_{1j}\mathbf{k}) n_j(\mathbf{r} + a_{ij}\mathbf{k}) a_{1j}^2 \mathbf{k}\mathbf{k} \, d^2k \quad (124)$$

with $m_{1j} \equiv m_1 m_j/(m_1 + m_j)$. For the special case of self-diffusion, (124) reduces to the result obtained for hard spheres ($\epsilon = 0$) in Section IV.B. The continuous potential u^A in the modified Enskog theory affects the diffusion coefficient through its control of the density distribution $n_i(\mathbf{r})$.

The local drift velocity of the tracer particle is

$$\mathbf{u}_1(\mathbf{r}) = \frac{1}{n_1} \int f_1 \mathbf{v}_1 \, d^3v_1 = -\zeta^{-1} \cdot \mathbf{H}_1 \quad (125)$$

As we asserted earlier, nonzero $\mathbf{H}_1$ provides the driving force for the nonzero drift velocity of tracer particles.

Let $n_1^\circ(\mathbf{r}), n_2^\circ(\mathbf{r}), \cdots$ denote the component densities of system at equilibrium. Suppose the density of tracer particles 1 is caused to deviate slightly from equilibrium (i.e., $n_1 - n_1^\circ \ll n_1^\circ$) and small uniform perturbing force $\mathbf{F}_1{}^P$ is applied to the tracer particles. Then, since $\mathbf{H}_1(\{\mathbf{n}_1^\circ\}) = \mathbf{0}$, one can show that to lowest order

$$\mathbf{H}_1 = kT\,\nabla\left(\frac{n_1}{n_1^\circ}\right) + \mathbf{F}_1{}^P \tag{126}$$

so to lowest order the local flux of tracer particles obeys

$$\mathbf{J}_1(\mathbf{r}) = -n_1^\circ(\mathbf{r})kT[\zeta(\mathbf{r})]^{-1}\cdot\nabla\left(\frac{n_1}{n_1^\circ}\right) - n_1^\circ(\mathbf{r})[\zeta(\mathbf{r})]^{-1}\cdot\mathbf{F}_1{}^P \tag{127}$$

Equation (127) has the form

$$\mathbf{J}_1(\mathbf{r}) = -n_1^\circ(\mathbf{r})\mathbf{D}(\mathbf{r})\cdot\nabla\left(\frac{n_1}{n_1^\circ}\right) - n_1^\circ(\mathbf{r})[\zeta(\mathbf{r})]^{-1}\mathbf{F}_1{}^P \tag{128}$$

where $\mathbf{D}(\mathbf{r})$ is the local diffusion tensor and is, as shown in Section IV.A, related to the friction tensor by the Einstein relation

$$\mathbf{D}(\mathbf{r}) = kT\zeta(\mathbf{r})^{-1} \tag{129}$$

D. Modified Enskog Theory of Self-Diffusion in Slit Pores

By self-diffusion we mean that the molecular properties of the tracer diffusion species are the same as those of the other fluid particles. Thus we consider in detail the case of self-diffusion in a one-component fluid confined by two flat, impermeable, solid walls (i.e., we consider self-diffusion in a slit pore). At the end of this section we examine self-diffusion in a cylindrical pore. The particle–wall potential u^e is assumed to be a function only of the distance x of the particle from the wall. The equilibrium densities, n_1° and n°, of tracer and other fluid particles also depend only on distance x from the wall.

Since the tracer and other fluid particles have the same molecular properties and since we are considering a situation such as $n_1^\circ \ll n^\circ$, the equilibrium YBG equations become, to lowest order of n_1°/n°,

$$kT\frac{dn_1^\circ}{dx} + n_1^\circ\frac{du^e}{dx} - n_1^\circ\int n^\circ(x + s_x)g(x,x + s_x)\frac{s_x}{s}\frac{du^A}{ds}\,d^3s + 2\pi a^2 n_1^\circ kT\int_{-1}^{1} g(x,x + a\xi)n^\circ(x + a\xi)\xi\,d\xi \tag{130}$$

and

$$kT\frac{dn^\circ}{dx} + n^\circ\frac{du^e}{dx} - n^\circ\int n^\circ(x + s_x)g(x,x + s_x)\frac{s_x}{s}\frac{du^A}{ds}\,d^3s$$
$$+ 2\pi a^2 n^\circ kT\int_{-1}^{1} g(x,x + a\xi)n^\circ(x + a\xi)\xi\,d\xi \qquad (131)$$

If (130) is divided by n_1° and (131) by n°, and if the results are subtracted, it follows that

$$\frac{d}{dx}\ln\frac{n_1^\circ}{n^\circ} = 0 \qquad (132)$$

or n_1°/n° = constant. Thus only (131) has to be solved for $n^\circ(x)$. $n_1^\circ(x)$ can then be computed from $(\overline{X}_1^\circ/\overline{X}^\circ)n^\circ(x)$, where $\overline{X}_1^\circ$ and $\overline{X}^\circ$ are the overall mole fractions of tracer and other fluid particles.

To the slit pore the Enskog friction tensor is of the form

$$\zeta = \zeta_N\hat{\imath}\hat{\imath} + \zeta_T(\hat{\jmath}\hat{\jmath} + \hat{k}\hat{k}) \qquad (133)$$

where $\hat{\imath},\hat{\jmath}$ and $\hat{k}$ are unit vectors in the x, y, and z directional, where ζ_N and ζ_T are the normal and transverse components of the friction tensor. The components are given by

$$\zeta_N(x) = 4a^2(\pi mkT)^{1/2}\int_{-1}^{1} g(x,x + a\xi+)n^\circ(x + a\xi)\xi^2\,d\xi \qquad (134)$$

and

$$\zeta_T(x) = 4a^2(\pi mkT)^{1/2}\int_{-1}^{1} g(x,x + a\zeta+)n^\circ(x + a\xi)(1 - \xi^2)\,d\xi \qquad (135)$$

Thus, in the absence of an external perturbing field, the diffusion flux in the slit pore is given by

$$\mathbf{J}_1 = -n_1^\circ D_N\hat{\imath}\frac{\partial}{\partial x}\ln\left(\frac{n_1}{n_1^\circ}\right) - D_T\left[\hat{\jmath}\frac{\partial n_1}{\partial y} + \hat{k}\frac{\partial n_1}{\partial z}\right] \qquad (136)$$

where $D_N \equiv kT/\zeta_N$ and $D_T \equiv kT/\zeta_T$.

Pore wall impermeability yields the condition $J_{1x} = 0$. We can choose y to be the diffusion direction so that $J_{1z} = 0$. These conditions imply that n_1 is independent of z and that $n_1(x,y) = n_1^\circ(x)\chi(y)$. Thus the local diffusion flux is given by

$$J_{1y} = -D_T(x)n_1^\circ(x)\frac{d\chi(y)}{dy} \qquad (137)$$

Experimentally, one cannot usually measure the local diffusion flux. What one can measure is the pore average flux J_{pore} and pore average density

n_{pore} defined by

$$J_{\text{pore}} = \frac{1}{L}\int_0^L J_{1y}\,dx \quad \text{and} \quad \frac{1}{L}\int_0^L n_1\,dx \tag{138}$$

where L is the distance of separation of the solid walls. Multiplying (137) by dx and integrating, we find Fick's law for self-diffusion in a slit pore, namely,

$$J_{\text{pore}} = -D_{\text{pore}}\frac{dn_{\text{pore}}}{dy} \tag{139}$$

where

$$D_{\text{pore}} = \frac{\int_0^L D_T(x)n^\circ(x)\,dx}{\int_0^L n^\circ(x)\,dx} \tag{140}$$

Using molecular dynamics, Magda et al. determined the pore diffusivity for a 6–12 Lennard–Jones fluid confined between flat solid walls interacting with fluid particles according to the 3–4–10 Lennard–Jones potential at (59). Vanderlick and Davis [3] compared the results of the modified Enskog theory using the Fischer–Methfessel theory of the density profile and the Carnahan–Starling formula [Eq. (65)] for $g^{\text{HS}}a,\bar{n}$. They used the 6–∞ fluid–fluid pair potential and chose the effective hard-sphere diameter to be $a = 0.972\sigma$ to ensure that the Enskog theory gives the correct bulk phase diffusivity for the conditions of the simulations. The conditions were $\epsilon = \epsilon_w$, $\sigma = \sigma_w$, $kT/\epsilon = 1.2$, $n_{\text{bulk}}a^3 = 0.5925$, and $\kappa/\epsilon = -3.6254$. The comparison is shown in Fig. 4, which gives the pore diffusion coefficient versus pore width L. The theory predicts an oscillatory approach to bulk diffusivity as the wall separation increases. The simulations also exhibit the oscillatory behavior, although the data are too sparse to resolve more than two periods. The maxima and minima in the diffusivity reflect the minima and maxima in the average pore density Fig. 5. The density maxima occur at pore widths favoring an integral number of fluid layers, and the minima occur as transitions between these widths.

Magda et al. [20] observed that although the density profile is strongly peaked in fluid confined to slit pores, the local diffusivity measured in bins parallel to the pore walls was quite insensitive to distance from the pore wall. Their results are shown in Fig. 6. According to the modified Enskog theory, this insensitivity to the strong oscillations in the density profile occurs because of smoothing in two ways. First, as is apparent in the integration over ξ in (135), the fact that the tracer particle experiences

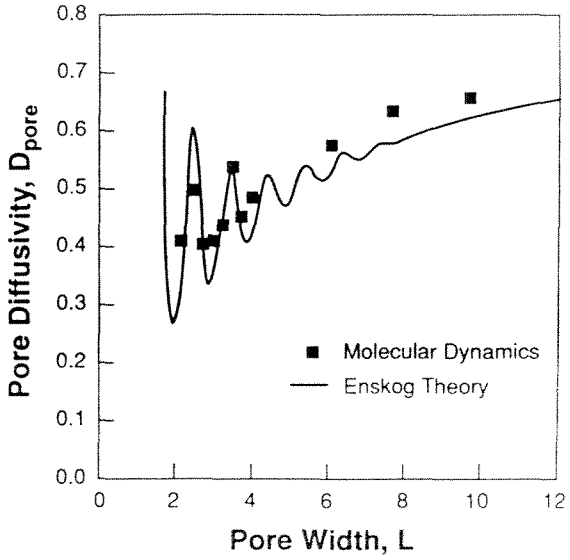


Fig. 4 Pore diffusivity versus pore width. Comparison of theory and molecular dynamics. Diffusivity in units of $(3\sigma/8)(kT/\pi m)^{1/2}$, and pore width L in units of σ. (Redrawn from Ref. 3.).

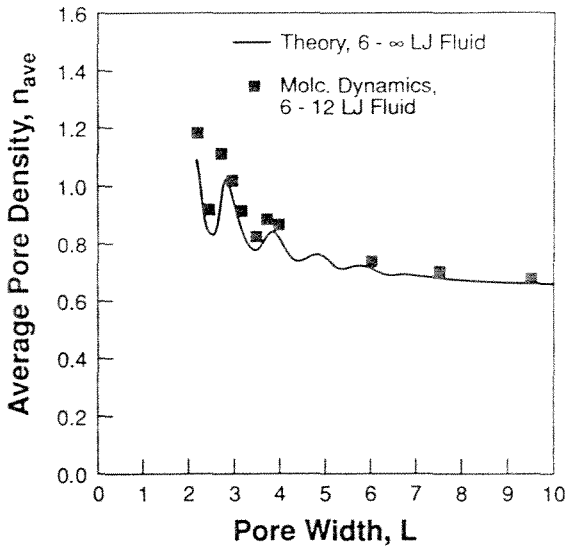


Fig. 5 Average pore density versus pore width. Comparison of theory (Ref. 3) and molecular dynamics (Ref. 20). Density in units of σ^{-3} and pore width in units of σ. (Redrawn from Ref. 3.)

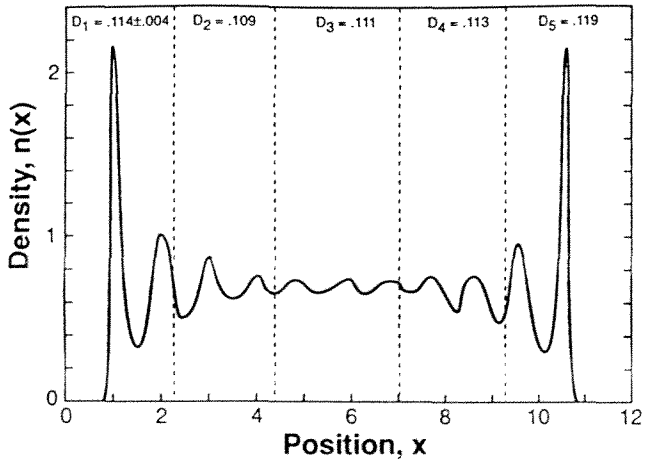


Fig. 6 Molecular dynamics results for the density profile and the local transverse diffusivity of a 6–12 LJ fluid confined in a planar slit pore with 3–4–10 LJ walls. D is the diffusion coefficient parallel to the pore walls. Averaged over the ith slice parallel to the walls. The units of diffusivity, density, and distance from the wall are $\sqrt{\epsilon\sigma^2/m}\ \sigma^{-3}$, and σ. (Redrawn from Ref. 20.)

binary collisions with particles arriving from every direction results in an average of local density over the surface of a sphere of diameter a. The other averaging occurs, for similar reasons, in the density dependence of the pair correlation function. In the context of the Fischer–Methfessel model, the pair correlations are governed by the local density averaged over a sphere of radius $d/2$.

In Fig. 7 are presented the density profiles $n(x)$ and the local transverse diffusivity profiles $D_T(x)$ predicted for a 6–∞ Lennard–Jones fluid and for a hard-sphere fluid confined by 3–4–10 Lennard–Jones walls. The density profiles were computed from the Fischer–Methfessel model and the diffusivity from the modified Enskog theory with the Carnahan–Starling approximation for $g^{HS}(a,\bar{n})$. For both fluids, the transverse diffusivity is only a weak function of pore transition.

V. FLOW IN A STRONGLY INHOMOGENEOUS ONE-COMPONENT FLUID

A. Modified Enskog Theory

Davis has solved the modified Enskog equation for slow flow in a strongly inhomogeneous one-component fluid. We outline his results here. The momentum balance equation is given at (29). According to the modified

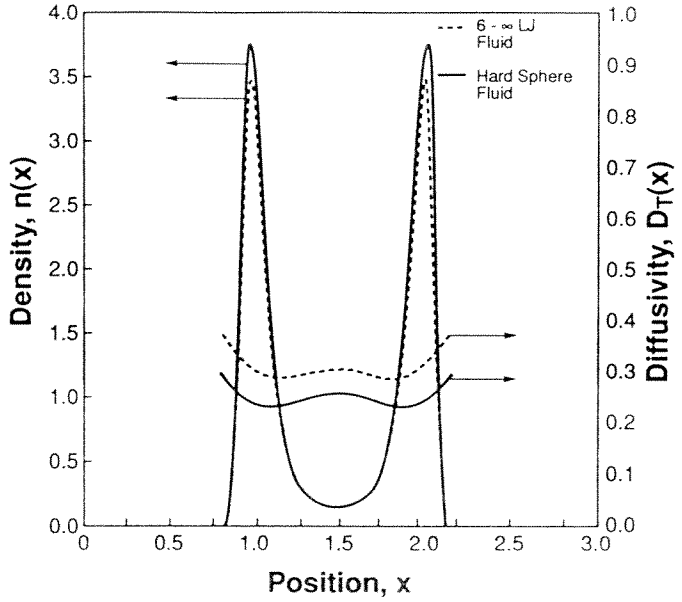


Fig. 7 Fluid density and local transverse diffusivity versus distance from pore wall predicted from the modified Enskog theory using the Fischer–Methfessel closure approximation. Density in units of σ^{-3}, diffusivity in units of $(3\sigma/8)(kT/\pi m)^{1/2}$, and distance in units σ. (Redrawn from Ref. 3.)

Enskog theory for a pure fluid. the stress tensor is of the form

$$\mathbf{T} = -mn\langle(\mathbf{v}_1 - \mathbf{u})(\mathbf{v}_1 - \mathbf{u})\rangle + \frac{1}{2}\int_0^1 \int \frac{\mathrm{ss}}{s}\frac{du^A}{ds} n^{(2)}(\mathbf{r} - \alpha) \times \mathbf{s},\mathbf{r} + (1 - \alpha)\mathbf{s},t)\, d^3s\, d\alpha + \frac{1}{2}a^3 \int_0^1 \iint_{\mathbf{v}_{21}\cdot\mathbf{k}>0} \mathbf{kk} f^{(2)} \times (\mathbf{r} - \alpha a\mathbf{k},\mathbf{r} + (1 - \alpha)a\mathbf{k},\mathbf{v}_1,\mathbf{v}_2,t)(\mathbf{v}_1\cdot\mathbf{k})^2 d^2k\, d^3v_1\, d^3v_2\, d\alpha \tag{141}$$

and the modified Enskog equation is

$$\frac{\partial f}{\partial t} + \mathbf{v}_1\cdot\nabla f - \frac{1}{m}\nabla u^e\cdot\nabla_{\mathbf{v}_1} f + \frac{1}{m}\int \frac{\mathrm{s}}{s}\frac{du^A}{ds}(s)\cdot\nabla_{\mathbf{v}_1} f(\mathbf{r},\mathbf{v},t) f \times (\mathbf{r} + \mathbf{s},\mathbf{v}_2,t)\, d^3s\, d^3v_2 = a^2 \iint_{\mathbf{v}_{21}\cdot\mathbf{k}>0} [g(\mathbf{r},\mathbf{r}, + a\mathbf{k}+,t) f(\mathbf{r},\mathbf{v}_1',t) f \times (\mathbf{r} + a\mathbf{k}+,\mathbf{v}_2',t) - g(\mathbf{r},\mathbf{r} - a\mathbf{k}+,t) f \times (\mathbf{r},\mathbf{v}_1,t) f(\mathbf{r} - a\mathbf{k},\mathbf{v}_2,t)]\mathbf{v}_{21}\cdot\mathbf{k}\, d^2k\, d^3v_2 \tag{142}$$

As in the case for tracer diffusion, the Chapman–Enskog technique

can be used to solve the kinetic equation. The singlet velocity distribution function can be expressed in the form

$$f(\mathbf{r},\mathbf{v},t) = n(\mathbf{r},t)\varphi(\mathbf{V}(\mathbf{r},t))[1 + \psi(\mathbf{r},\mathbf{v},t)] \tag{143}$$

where $\varphi(\mathbf{V})$ is the Maxwell velocity distribution and

$$\mathbf{V} \equiv \mathbf{v} - \mathbf{u}(\mathbf{r},t) \tag{144}$$

is the velocity relative to the local mass average velocity. Since

$$n = \int f\, d^3v, \quad \frac{3nkT}{2} = \int \frac{1}{2} mV^2 f\, d^3v \quad \text{and} \quad n\mathbf{u} = \int \mathbf{v} f\, d^3v \tag{145}$$

it follows that the perturbation function obeys the conditions

$$0 = \int \varphi(\mathbf{V})\psi\, d^3v = \int v^2\varphi(\mathbf{V})\psi\, d^3v = \int \mathbf{V}\varphi(v)\psi\, d^3v \tag{146}$$

According to the Chapman–Enskog method, the perturbation function ψ is approximated by the first few terms of a set of Gaussian-weighted orthogonal polynomials in velocity. The Sonine or Hermite polynomials serve the purpose. Through the second-order polynomials, the approximating function for the singlet velocity distribution function is

$$f(\mathbf{r},\mathbf{v},t) = n(\mathbf{r},t)\varphi(\mathbf{V}(\mathbf{r},t)\left[1 + \mathbf{a}\cdot\mathbf{V} + \mathbf{b}:\left(\mathbf{VV} - \frac{1}{3}V^2\mathbf{U}\right) + c\left(\frac{1}{2}V^2 - \frac{3kT}{m}\right)\right] \tag{147}$$

where $\mathbf{b}$ is a traceless symmetric tensor (i.e., $\mathbf{b} : \mathbf{U} = 0$ and $\mathbf{b} = \mathbf{b}^T$). It follows from the conditions at (146) that $\mathbf{a} = c = 0$, so

$$\psi = \mathbf{b} : (\mathbf{VV} - \tfrac{1}{3}V^3\mathbf{U}) \tag{148}$$

To determine $\mathbf{b}$ by the Galerkin method, we multiply (142) by $(\mathbf{V}_1\mathbf{V}_1 - \frac{1}{3}V_1\mathbf{U})\, d^3v_1$, integrate over $\mathbf{v}_1$, and set the residual equal to zero. If the result is linearized in ψ and if higher-order terms (i.e., those proportional to $\partial\nabla\mathbf{u}/\partial t$ and $\mathbf{u}\cdot\nabla\mathbf{u}$ and those nonlinear in $\nabla\mathbf{u}$) are dropped, the following integral equation is obtained for $\mathbf{b}(r,t)$:

$$\begin{aligned} a^2 \int g(\mathbf{r},\mathbf{r} + a\mathbf{k}+,t)n(\mathbf{r} + a\mathbf{k},t)\{2\mathbf{k}^\circ\mathbf{kkk} : [\mathbf{b},(\mathbf{r},t) + \mathbf{b}(\mathbf{r} + a\mathbf{k},t)] \\ - 2[\mathbf{kk}\cdot\mathbf{b}(\mathbf{r},t) + \mathbf{b}(\mathbf{r},t)\cdot\mathbf{kk} - \frac{2}{3}\mathbf{b}(\mathbf{r},t) : \mathbf{kkU}]\}\, d^2k \\ = \sqrt{\pi}\left(\frac{m}{kT}\right)^{3/2}\left[\boldsymbol{\epsilon}(\mathbf{r},t) + \frac{a^3}{2}\int g(\mathbf{r},\mathbf{r} + a\mathbf{k}-,t) \right. \\ \left. \times\, n(\mathbf{r} + a\mathbf{k},t)\mathbf{k}^\circ\mathbf{kkk} : \boldsymbol{\epsilon}(\mathbf{r},t)\right] \end{aligned} \tag{149}$$

where $\boldsymbol{\epsilon}$ is the traceless symmetric part of $\nabla\mathbf{u}$, that is,

$$\boldsymbol{\epsilon} \equiv \tfrac{1}{2}(\nabla\mathbf{u} + \nabla\mathbf{u}^T) - \tfrac{1}{3}(\nabla\cdot\mathbf{u})\mathbf{U} \tag{150}$$

and $\mathbf{k}^\circ\mathbf{k}$ is the traceless diadic

$$\mathbf{k}^\circ\mathbf{k} \equiv \mathbf{kk} - \tfrac{1}{3}\mathbf{U} \tag{151}$$

[*Note*: The sign of the second term in square brackets on the left-hand side of (149) is negative. It was typeset incorrectly as positive in (6.9) of Ref. 4. Also, the right-hand side of (6.9) of Ref. 4 has an erroneous factor of 2.]

Because of the term $\mathbf{b}(\mathbf{r} - a\mathbf{k},t)$ in (149) the theory results in a difficult to solve integral equation for $\mathbf{b}$. To simplify the theory, we approximate $\mathbf{b}(\mathbf{r} - a\mathbf{k},t)$ by $\mathbf{b}(\mathbf{r},t)$ and obtain an algebraic equation for $\mathbf{b}(\mathbf{r},t)$:

$$\mathbf{H} : \mathbf{b} + \mathbf{K}\cdot\mathbf{b} + \mathbf{b}\cdot\mathbf{K} = \mathbf{L} : \boldsymbol{\epsilon} \tag{152}$$

where

$$\mathbf{H} = a^2 \int g(\mathbf{r},\mathbf{r} - a\mathbf{k}+,t)n(\mathbf{r} - a\mathbf{k},t)[2\mathbf{k}^\circ\mathbf{kkk} - \tfrac{2}{3}\mathbf{Ukk}]\, d^2k \tag{153}$$

$$\mathbf{K} = -a^2 \int g(\mathbf{r},\mathbf{r} - a\mathbf{k}+,t)n(\mathbf{r} - a\mathbf{k},t)\mathbf{kk}\, d^2k \tag{154}$$

and

$$\mathbf{L}(\mathbf{r},t) = \sqrt{\pi}\left(\frac{m}{kT}\right)^{3/2}\left[\mathbb{U} + \frac{a^3}{2}\int g(\mathbf{r},\mathbf{r} - a\mathbf{k}+,t)n(\mathbf{r} - a\mathbf{k},t) \times \mathbf{k}^\circ\mathbf{kkk}\, d^2k\right] \tag{155}$$

The quantity $\mathbb{U}$ is a fourth-rank tensor with the property

$$\mathbb{U} : \mathbf{a} = \mathbf{a} = \mathbf{a} : \mathbb{U} \tag{156}$$

where $\mathbf{a}$ is any second-rank tensor. The general form of $\mathbf{b}$ will be

$$\mathbf{b}(\mathbf{r},t) = \mathbf{B}(\mathbf{r},t) : \boldsymbol{\epsilon} \tag{157}$$

$\mathbf{B}$ is a fourth-rank tensor and is symmetric and traceless with respect to its first two components and its second two components; that is, $B_{\alpha\beta\nu\mu} = B_{\beta\alpha\nu\mu}$, $B_{\alpha\beta\nu\mu} = B_{\alpha\beta\mu\nu}$, $\sum_\alpha B_{\alpha\alpha\nu\mu} = \sum_\nu B_{\alpha\beta\nu\nu} = 0$, where $B_{\alpha\beta\nu\mu}$ is the $\alpha\beta\nu\mu$ Cartesian component of $\mathbf{B}$.

In the case of an isotropic fluid, we can easily solve (149) to obtain $\mathbf{b} = b_0\boldsymbol{\epsilon}$, where

$$b_0 = -\frac{5}{8\sqrt{\pi}}\left(\frac{m}{kT}\right)^{3/2}\left[\frac{1 + \dfrac{4\pi}{15}na^3g(a,n)}{na^2g(a,n)}\right] \tag{158}$$

Once $\mathbf{b}$ has been determined by solving (149), the stress tensor can be evaluated by insertion of the result into (141). We will not carry out this evaluation here since our aim is to obtain the momentum conservation equation. This equation contains $\nabla\cdot\mathbf{T}$, which is simpler than $\mathbf{T}$ since the ξ integration is eliminated in taking the divergence. For a one-component fluid, (9) can be combined with (8) to yield

$$mn\frac{\partial \mathbf{u}}{\partial t} + mn\mathbf{u}\cdot\nabla\mathbf{u} + n\nabla u^e + \nabla\cdot\mathbf{P} = \nabla\cdot\boldsymbol{\tau} \tag{159}$$

where we have equated $\mathbf{T}$ to $\boldsymbol{\tau} - \mathbf{P}$, $\mathbf{P}$ being the part of the pressure tensor that remains in the limit of zero flow and $\boldsymbol{\tau}$ being the stress due exclusively to flow. In particular,

$$\begin{aligned}\nabla\cdot\mathbf{P} = &-\int \frac{\mathbf{s}}{s}\frac{du^A}{ds}(s)n(\mathbf{r},t)n(\mathbf{r}+\mathbf{s},t)g(\mathbf{r},\mathbf{r}+\mathbf{s},t)\,d^3s \\ &+ a^2kT\int \mathbf{k}n(\mathbf{r},t)n(\mathbf{r}+a\mathbf{k},t)g(\mathbf{r},\mathbf{r}+a\mathbf{k}+,t)\,d^2k\end{aligned} \tag{160}$$

and

$$\begin{aligned}\nabla\cdot\boldsymbol{\tau} = &-\nabla\cdot[mn\int \mathbf{V}_1\mathbf{V}_1\psi(\mathbf{r},\mathbf{V}_1)\psi(\mathbf{V}_1)\,d^2V_1] \\ &+ a^2m\iint_{\mathbf{v}_{21}\cdot\mathbf{k}>0} \mathbf{k}n(\mathbf{r},t)n(\mathbf{r}-a\mathbf{k},t)g(\mathbf{r},\mathbf{r}-a\mathbf{k}+,t) \\ &\times\varphi(\mathbf{V}_1(\mathbf{r},t))\{\varphi(\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t)) \\ &-\varphi(\mathbf{V}_2(\mathbf{r},t)) + \varphi(\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t))[\varphi(\mathbf{r},\mathbf{V}_1(\mathbf{r},t),t) \\ &+\psi(\mathbf{r}-\mathbf{k}d,\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t),t) + \mathcal{O}(\psi^2)]\}(\mathbf{v}_{21}\cdot\mathbf{k})^2\,d^2k\,d^3v_1\,d^3v_2\end{aligned} \tag{161}$$

We are interested in the low flow limit in which terms of order $(\nabla\mathbf{u})^2$ and $\nabla^3\mathbf{u}$ are negligible. Thus the term of order ψ^2 in (161) is negligible and the following approximations can be inserted:

$$\varphi(\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t) \approx \varphi(\mathbf{V}_2) + \frac{m}{kT}\varphi(\mathbf{V}_2)[-\mathbf{V}_2\mathbf{k}a : \nabla\mathbf{u} + \mathbf{V}_2\mathbf{k}\mathbf{k}a^2 : \nabla\nabla\mathbf{u}] \tag{162}$$

and

$$\begin{aligned}&\varphi(\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t))[\psi(\mathbf{r},\mathbf{V}_1(\mathbf{r},t),t) + \psi(\mathbf{r}-a\mathbf{k},\mathbf{V}_2(\mathbf{r}-a\mathbf{k},t),t)] \\ &\quad\approx \varphi(\mathbf{V}_2)\{(\mathbf{V}_1\mathbf{V}_1 - \tfrac{1}{3}V_1{}^2\mathbf{U}) : \mathbf{B} : \boldsymbol{\epsilon} + (\mathbf{V}_2\mathbf{V}_2 \\ &\qquad - \tfrac{1}{3}V_1{}^2\mathbf{U}) : [\mathbf{B} : \boldsymbol{\epsilon} - a\mathbf{k}\cdot\nabla(\mathbf{B} : \boldsymbol{\epsilon}]\}\end{aligned} \tag{163}$$

where the quantities $\mathbf{V}_1$, $\mathbf{V}_2$, $\mathbf{u}$, $\mathbf{B}$, and $\boldsymbol{\epsilon}$ on the right-hand side of (162)

and (163) are evaluated at $\mathbf{r}$ and t As in simplifying (149) to solve for $\mathbf{b}$, the approximation $\mathbf{B}(\mathbf{r} - a\mathbf{k},t) : \boldsymbol{\epsilon}(\mathbf{r} - a\mathbf{k},t) = \mathbf{B} : \boldsymbol{\epsilon} - a\mathbf{k} \cdot \nabla(\mathbf{B} : \boldsymbol{\epsilon})$ is not an entirely consistent approximation, but hopefully it will not cause too much error in the final result (it is partly justified by noting that at high densities the contributions from $\mathbf{b}$ are not large compared to the contributions from the terms in (161) and at low density $n\mathbf{b}$ is almost independent of density.

Carrying out the integrations over velocities in (161) with the approximations just discussed, we obtain

$$\nabla\cdot\boldsymbol{\tau} = \nabla\left[-mn\left(\frac{kT}{m}\right)^2\mathbf{b}\right] + \mathbf{M}_1 : \nabla\mathbf{u} + \mathbf{M}_2 : \nabla\nabla\mathbf{u} - \frac{2}{a}\left(\frac{\pi kT}{m}\right)^{1/2}\frac{kT}{m}[2\mathbf{M}_1 : \mathbf{b} + \mathbf{M}_2 : \nabla\mathbf{b}] \tag{164}$$

where

$$\mathbf{M}_1 = 2\left(\frac{mkT}{\pi}\right)^{1/2} a^3 n(\mathbf{r},t)\int n(\mathbf{r} + a\mathbf{k},t)g(\mathbf{r},\mathbf{r} + a\mathbf{k} + ,t)\mathbf{kkk}\, d^2k \tag{165}$$

and

$$\mathbf{M}_2 = \left(\frac{mkT}{\pi}\right)^{1/2} a^4 n(\mathbf{r},t)\int n(\mathbf{r} + a\mathbf{k},t)g(\mathbf{r},\mathbf{r} + a\mathbf{k} + ,t)\mathbf{kkkk}\, d^2k \tag{166}$$

In reducing (163) to its final form, we have used the fact that $\mathbf{b}$ is a symmetric traceless tensor.

The results in this section can be summarized by noting that in the slow-flow limit, the momentum balance equation in a strongly inhomogeneous fluid is of the form

$$mn\frac{\partial\mathbf{u}}{\partial t} + n\nabla u^e + \nabla\cdot\mathbf{P} = \hat{\mathbf{M}}_1 : \nabla\mathbf{u} + \hat{\mathbf{M}}_2 : \nabla\nabla\mathbf{u} \tag{167}$$

where $\hat{\mathbf{M}}_1(\mathbf{r}),t)$ and $\hat{\mathbf{M}}_2(\mathbf{r},t)$ are third and fourth rank tensors that depend on the fluid density distribution $n(\mathbf{r})$. The contribution of $\hat{\mathbf{M}}_1$ to (166) comes solely from the fluid inhomogeneity since $\hat{\mathbf{M}}_1$ vanishes when density is constant.

For weakly inhomogeneous fluids, (166) reduces to

$$mn\frac{\partial\mathbf{u}}{\partial t} + n\,\nabla u^e + \nabla P = \eta_b\nabla(\nabla\cdot\mathbf{u}) + \eta_s\nabla^2\mathbf{u} \tag{168}$$

where P, η_b, and η_s—the pressure, bulk viscosity, and shear viscosity

of isotropic fluid—are for the modified Enskog theory given by

$$P = nkT\left[1 + \frac{2\pi}{3} na^3 g(a,n)\right] + \tfrac{1}{6}n^2 \int s \frac{du^A}{ds} g(s,n)\, d^3s \quad (169)$$

$$\eta_b = \frac{16}{5\pi}\eta^* \left(\frac{2\pi}{3} na^3\right)^2 g(a,n) \quad (170)$$

$$\eta = \eta^* \left\{\frac{[1 + (4\pi/15)\, na^3 g(a,n)]^2}{g(a,n)} + \frac{48}{25\pi}\left(\frac{2\pi}{3n} a^3\right)^2 g(a,n)\right\} \quad (171)$$

where η^* is the ideal gas shear viscosity predicted for hard spheres, namely,

$$\eta^* = \frac{5}{16a^2}\left(\frac{mkT}{\pi}\right)^{1/2}$$

B. Steady Planar Flow in a Slit Pore

Consider a simple fluid confined between identical, planar flat structureless solid walls. The wall–fluid interaction potential u^e depends only on the distance x from a wall. As we saw in Section III, the density profile $n(x)$ of such a confined fluid at equilibrium varies rapidly with distance from the wall.

For the purpose of the analysis to be presented in this section, Bitsanis et al. [33] made a very important observation in molecular dynamics studies of the 6–12 Lennard–Jones fluid undergoing Couette flow between 3–4–10 Lennard–Jones walls. They found that shear does not affect the density profile $n(x)$ appreciably, even at shear rates as high as $10^{10}s^{-1}$. This observation means that in Couette flow the density profile $n(x)$ can be computed from the equation of hydrostatics equilibrium. Thus the pressure tensor $\mathbf{P}$ has the symmetry of the equilibrium fluid,

$$\mathbf{P} = P_N \hat{e}_3\hat{e}_3 + P_T(\hat{e}_1\hat{e}_1 + \hat{e}_3\hat{e}_3) \quad (172)$$

where P_N is the component of pressure normal to the wall and P_T is the component transverse to the wall. The equilibrium density profile is determined by

$$n\frac{du^e}{dx} + \frac{dP_N}{dx} = 0 \quad (173)$$

where dP_N/dx is the x-component of (159).

If the direction of flow is $\hat{e}_1$, the steady ($\partial\mathbf{u}/\partial t = 0$) low-flow ($\mathbf{u}\cdot\nabla\mathbf{u}$

negligible) limit of (158) in the direction of flow becomes

$$\frac{\partial P_T}{\partial y} = \nabla\cdot\boldsymbol{\tau}\cdot\hat{e}_1 \tag{174}$$

The observation that the density profile depends only on x considerably simplifies the structure of $\nabla\cdot\boldsymbol{\tau}$. Assuming that for planar flow $\mathbf{u} = u_y(x)\hat{e}_1$, $n = n(x)$, and $g(\mathbf{r},\mathbf{r} + a\mathbf{k}+,) = g^{\mathrm{HS}}(a;\bar{n}(x + (a/2)\,k_x))$, we can carry out part of the integration over d^2k to evaluate $\mathbf{M}_1$, $\mathbf{M}_2$, and $\mathbf{b}$. For $\mathbf{M}_1$ and $\mathbf{M}_2$ the results are

$$\begin{aligned}\mathbf{M}_1 = {} & 2\pi\left(\frac{mkT}{\pi}\right)^{1/2} a^3 n(x)\int_{-1}^{1} n(x + a\xi)g(x,x + a\xi) \\ & \times [(1 - \xi^2)\xi(\hat{e}_3\mathbf{U}^{(2)} + \mathbf{U}^{(2)}\hat{e}_3) + 2\xi^3\hat{e}_3\hat{e}_3\hat{e}_3\hat{e}_3]\,d\xi\end{aligned} \tag{175}$$

and

$$\begin{aligned}\mathbf{M}_2 = {} & \pi\left(\frac{mkT}{\pi}\right)^{1/2} a^4 n(x)\int_{-1}^{1} n(x + a\xi)g(x,x + a\xi+) \\ & \times \{(1 - \xi^2)\xi^2(\hat{e}_3\hat{e}_3\mathbf{U}^{(2)} + \hat{e}_3\mathbf{U}^{(2)}\hat{e}_3 + \mathbf{U}^{(2)}\hat{e}_3\hat{e}_3) + 2\xi^4\hat{e}_3\hat{e}_3\hat{e}_3\hat{e}_3 \\ & + \sum_{i=1}^{2}[(1 - \xi^2)\xi^2(\hat{e}_3\hat{e}_i\hat{e}_3\hat{e}_i + \hat{e}_i\hat{e}_3\hat{e}_3\hat{e}_i + \hat{e}_i\hat{e}_3\hat{e}_i\hat{e}_3) \\ & + \tfrac{3}{4}(1 - \xi^2)^2\hat{e}_i\hat{e}_i\hat{e}_i\hat{e}_i]\}\,d\xi\end{aligned} \tag{176}$$

where $\hat{e}_1$, $\hat{e}_2$, and $\hat{e}_3$ are Cartesian coordinates, $\hat{e}_3$ lying in the x-direction and $\hat{e}_1$ and $\hat{e}_2$ lying in the y- and z-directions. The quantity $\mathbf{U}^{(2)} \equiv \sum_{i=1}^{2}\hat{e}_i\hat{e}_i$ is the unit dyadic in the yz-plane.

To determine $\mathbf{b}$ we have to take advantage of the symmetry of the problem and solve (149) for the various components of $\mathbf{b}$. The results are

$$b_{xx} = \delta\,\frac{1 + (3\pi a^3/2)\int_{-1}^{1} G(x,\xi)(\xi^2 - \frac{1}{3})^2\,d\xi}{12a^2\int_{-1}^{1} G(x,\xi)\xi^2(\xi^2 - 1)\,d\xi}\,\epsilon_{xx} \tag{177}$$

$$b_{xy} = \delta\,\frac{1 + \pi a^3\int_{-1}^{1} G(x,\xi)\xi^2(1 - \xi^2)\,d\xi}{a^2\int_{-1}^{1} G(x,\xi)(6\xi^2 - 8\xi^4 - 2)\,d\xi}\,\epsilon_{xy} = b_{yz} \tag{178}$$

$$b_{xz} = \frac{b_{xy}\epsilon_{xz}}{\epsilon_{xy}} = b_{zx} \tag{179}$$

$$b_{yz} = -\delta\,\frac{1 + (\pi a^3/4)\int_{-1}^{1} G(x,\xi)(1 - \xi^2)^2\,d\xi}{2a^2\int_{-1}^{1} G(x,\xi)(1 - \xi^4)\,d\xi}\,\epsilon_{yz} = b_{zy} \tag{180}$$

$$b_{yy} = -\frac{\pi\delta}{\alpha_1^2 - \alpha_2^2}\{[\alpha_1(1 + \beta_1) - \alpha_2\beta_2]\epsilon_{yy} + [\alpha_1\beta_2 - \alpha_2(1 + \beta_1)]\epsilon_{zz}\} \tag{181}$$

$$b_{zz} = -\frac{\pi\delta}{\alpha_1^2 - \alpha_2^2}\{[\alpha_1\beta_2 - \alpha_2(1 + \beta_1)]\epsilon_{yy} + [\alpha_1(1 + \beta_1) - \alpha_2\beta_2]\epsilon_{zz}\} \tag{182}$$

where the following abbreviations have been used:

$$\delta = \frac{2}{\sqrt{\pi}}\left(\frac{m}{kT}\right)^{3/2} \tag{183}$$

$$G(x,\xi) = n(x + a\xi)g(x,x + a\xi+) \tag{184}$$

$$\alpha_1 = \pi a^2 \int_{-1}^{1} G(x,\xi)[3(1 - \xi^2)^2 - 4(1 - \xi^4)]\, d\xi \tag{185}$$

$$\alpha_2 = \pi a^2 \int_{-1}^{1} G(x,\xi)[(1 - \xi^2)^2 - 4\xi^2(1 - \xi^2)]\, d\xi \tag{186}$$

$$\beta_1 = \frac{\pi a}{2} \int_{-1}^{1} G(x,\xi)\left[\tfrac{7}{4}(1 - \xi^2)^2 + 2\xi^2 - \tfrac{4}{3}\right] d\xi \tag{187}$$

$$\beta_2 = \frac{\pi a}{2} \int_{-1}^{1} G(x,\xi)\left[\tfrac{5}{4}(1 - \xi^2)^2 + 2\xi^2 - \tfrac{4}{3}\right] d\xi \tag{188}$$

Equation (168) can now be expressed in the form

$$\frac{\partial P_T}{\partial y} = \hat{M}_1 \frac{\partial u_y}{\partial x} + \hat{M}_2 \frac{\partial^2 u_y}{\partial x^2} \tag{189}$$

where

$$\begin{aligned}\hat{M}_1 = {} & 2\pi\left(\frac{mkT}{\pi}\right)^{1/2} a^3 n(x) \int_{-1}^{1} n(x + a\xi)g(x,x + a\xi+)(1 - \xi^2)\xi\, d\xi \\ & - \frac{1}{2}\frac{\partial}{\partial x}\left[mn(x)\left(\frac{kT}{m}\right)^2 B(x)\right] - 4\pi mn(x)a^2\left(\frac{kT}{m}\right)^2 B(x)\int_{-1}^{1} n(x + a\xi) \\ & \times g(x,x + a\xi+)(1 - \xi^2)\xi\, d\xi \\ & - \pi mn(x)a^3\left(\frac{kT}{m}\right)^2 \frac{\partial B(x)}{\partial x}\int_{-1}^{1} n(x + a\xi)g(x,x + a\xi+)(1 - \xi^2)\xi^2\, d\xi\end{aligned} \tag{190}$$

$$\begin{aligned}\hat{M}_2 = {} & \pi\left(\frac{mkT}{\pi}\right)^{1/2} a^4 n(x) \int_{-1}^{1} n(x + a\xi)g(x,x + a\xi+)(1 - \xi^2)\xi^2\, d\xi \\ & - \tfrac{1}{2}mn(x)\left(\frac{kT}{m}\right)^2 B(x) - \pi mn(x)a^3\left(\frac{kT}{m}\right)^2 B(x)\int_{-1}^{1} n(x + \xi a) \\ & \times g(x,x + \xi a+)(1 - \xi^2)\xi^2\, d\xi\end{aligned} \tag{191}$$

and

$$B(x) = \left(\frac{m}{\pi kT}\right)^{1/2} \frac{1 + \pi a^3 \int_{-1}^{1} n(x + a\xi) g(x, x + a\xi +) \xi^2 (1 - \xi^2)\, d\xi}{a^2 \int_{-1}^{1} n(x + a\xi) g(x, x + a\xi +)(6\xi^2 - 8\xi^4 - 2)\, d\xi} \tag{192}$$

Let us now specialize planar Couette flow, in which case $u_y(0) = -u_0, u_y(L) = u_0$, and $\partial P_T/\partial y = 0$. In this case the solution to (184) is

$$u_y(x) + u_0 \frac{\int_0^x dx'' \exp\{-\int_0^{x''} dx' [\hat{M}_1(x')/\hat{M}_2(x')]\}}{\int_0^L dx'' \exp\{-\int_0^{x''} dx' [\hat{M}_1(x')/\hat{M}_2(x')]\}} \tag{193}$$

At the time of this writing (188) has not been applied in its entirety to Couette flow in confined thin films. However, Davis et al. [5] tested (193) for the case in which $B(x)$ is neglected (for homogeneous fluids whose density is greater than about $0.5\sigma^{-3}$, this is known to give a good approximation to the viscosity). Bitsanis et al. [33] carried out molecular

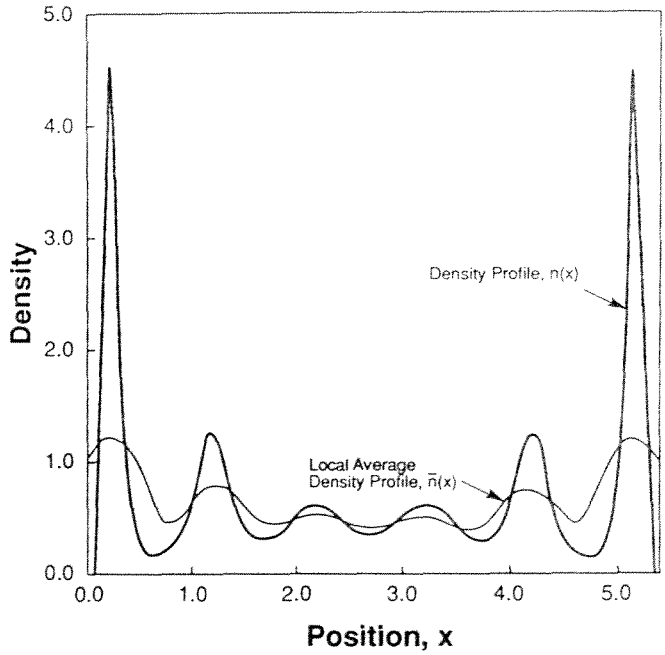


Fig. 8 Density and local average density profiles determined for a 6–12 LJ fluid confined by a planar 3–4–10 LJ walls. Density in units of σ^{-3} and distance from pore wall in units of σ. (Redrawn from Ref. 33.)

dynamics studies of Couette flow of a 6–12 Lennard–Jones fluid confined between structureless 3–4–10 Lennard–Jones walls. They found that the density profiles $n(x)$ and pore diffusivity are independent of flow, even up to shear rates of $10^{10}s^{-1}$. In Fig. 8 is presented their density profile for the case $\epsilon_w = \epsilon$, $\sigma_w = \sigma$, $\epsilon/kT = 1$, wall separation $L = 7.5\sigma$, and pore average density $n_p = 0.692\sigma^{-3}$. Also shown in Fig. 8 is the local average density profile computed from the molecular dynamics density profile according to the Fischer–Methfessel formula (71).

The velocity profile found in the molecular dynamics study is compared in Fig. 9 with the profile predicted by (193) with the terms in $B(x)$ neglected and using the Fischer–Methfessel approximation, $g(x,x + a\xi+) = g^{HS}(a,\bar{n}(x + \frac{1}{2}a\xi))$, for the contact value of the pair correlation function. For the predictions, $\bar{n}(x)$ is computed from (71) using the molecular dynamics results for the density profile $n(x)$. The predicted velocity profile agrees fairly well with the observed profile. The velocity profile is sur-

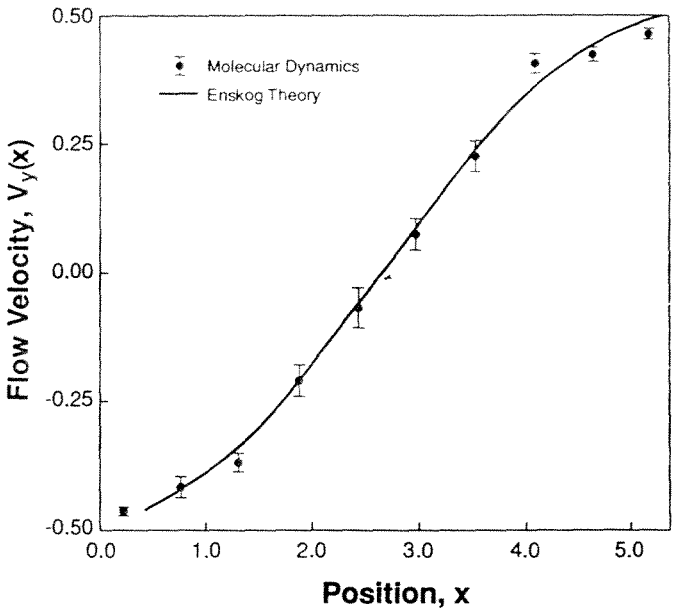


Fig. 9 Comparison of the velocity profile of a confined fluid undergoing planar Couette flow. Molecular dynamics for a 6–12 LJ fluid between planar 3–4–10 LJ walls. Modified Enskog theory for a 6–∞ fluid. Velocity in units of $(kT/m)^{1/2}$ and distance in units of σ. (Redrawn from Ref. 5.)

prisingly insensitive to the sharp oscillations in the density profile. According to the theory, the velocity profile depends on density only through the ratio $\hat{M}_1(x)/\hat{M}_2(x)$. From (190) and (191) it follows that the local density $n(x)$ cancels out of the ratio except in the term $-\partial[mn(kT/m)^2B(x)/2]/\partial x$, which is expected to be small except at very low densities. Otherwise, the density profile is smoothed once by an average over the area of a sphere of a molecular diameter a because the binary hard-sphere collisions occur on contact of a molecular centered at x with other molecules approaching it from all directions. A second and a third smoothing occur through the integrations over x' and x'' in (193). This smoothing is model independent since we expect the planar Couette flow approximation to the momentum balance equation to be of the form given by (189) from symmetry arguments alone.

Hopefully, in the future the effect of the term $B(x)$ (i.e., of **b**) $\hat{M}_1$ and $\hat{M}_2$ will be explored. The quantity $B(x)$ arises from the deviation of the singlet velocity distribution function from local equilibrium. The quantities $\mathbf{M}_1$ and $\mathbf{M}_2$ arise from momentum transport between particles whose velocities are at equilibrium relative to the local mass average velocity.

VI. CLOSING REMARKS

In the modified Enskog theory it is implicitly assumed that the confined film remains fluid. It has been found [34,35], however, by shear experiments with the surface forces apparatus that when the separation of confining solids is only one or two molecular diameters wide, the confined film may not flow under an infinitesimal stress. Instead, there can be solidlike structure, a finite yield stress [36], and the film can shear in a stick-slip mode [35]. Molecular dynamics indicates that the confining solids can induce solidlike structure in films of a few molecules thickness and the shear stress must "melt" the solidlike structure to induce flow [37]. In cylindrical pores of sufficiently small diameter that the molecules have to move in single file, Suh and MacElroy [38] found that their molecular dynamics results for the density dependence of the self-diffusion coefficient agree well with the exact theory for one-dimensional hard rods. In zeolite pores molecular dynamics indicates [39] that molecular motion is not that of a fluid, but rather is an activated hopping motion. Thus a kinetic theory of the Enskog type is not expected to be valid in pores having at least one dimension on the order of one or two molecular diameters. A totally different theoretical approach is needed for such systems.

ACKNOWLEDGMENTS

The author is grateful to the University of Minnesota National Science Foundation and the Minnesota Supercomputer Institute for partial support of this work.

REFERENCES

1. Dymond, J. H. (1982). *J. Phys. Chem. 85*: 3291. Dymond, J. H., and Young, K. Y. (1980). *Int. J. Thermodyn. 4*: 345. Dymond, J. H., and Woolf, A. L. (1982). *J. Chem. Soc. Faraday Trans. 1 78*: 991.
2. Davis, H. T. (1987). *J. Chem. Phys. 88*: 1474.
3. Vanderlick, T. K., and Davis, H. T. (1987). *J. Chem. Phys. 87*: 1791.
4. Davis, H. T. (1987). *Chem. Eng. Commun. 58*: 413.
5. Davis, H. T., Bitsanis, I., Vanderlick, T. K., and Tirrell, M. (1987). In *Supercomputer Research in Chemistry and Chemical Engineering* (K. S. Jensen and D. G. Truhlar, eds.), ACS Symposium Series 352, American Chemical Society, Washington, D.C.
6. Barker, J. A., and Henderson, D. (1967). *J. Chem. Phys. 47*: 4714. (1976). *Rev. Mod. Phys. 48*: 587.
7. Weeks, J. D., Chandler, D., and Anderson, H. C. (1971). *J. Chem. Phys. 54*: 5237; *55*: 5422.
8. Chapman, S., and Cowling, T. G. (1954). *The Mathematical Theory of Non-uniform Gases and Liquids*, Cambridge University Press, New York.
9. Davis, H. T., and Scriven, L. E. (1982). *Adv. Chem. Phys. 49*: 357.
10. Freaser, B. C., and Nordholm, S. (1983). *J. Chem. Phys. 79*: 4431.
11. Fischer, J., and Methfessel, M. (1980). *Phys. Rev. 22*: 2836.
12. Nordholm, S., Johnson, M., and Freasier, B. C. (1980). *Aust. J. Chem. 33*: 2139. Johnson, M., and Nordholm, S. (1981). *J. Chem. Phys. 75*: 1953.
13. Tarazona, P. (1985). *Phys. Rev. A 31*: 2672.
14. Percus, J. K. (1976). *J. Stat. Phys. 15*: 505.
15. Vanderlick, T. K., Davis, H. T., and Percus, J. K. (1989). *J. Chem. Phys. 91*: 7136.
16. Meister, T. F., and Kroll, D. M. (1985). *Phys. Rev. A 31*: 4055.
17. Davis, H. T. (1986). *J. Chem. Phys. 85*: 6808.
18. Tang, Z., Scriven, L. E., and Davis, H. T., *J. Chem. Phys.* (submitted).
19. Vanderlick, T. K., Scriven, L. E., and Davis, H. T. (1989). *J. Chem. Phys. 90*: 2422.
20. Magda, J. J., Tirrell, M., and Davis, H. T. (1985). *J. Chem. Phys. 83*: 1888.
21. Steele, W. A. (1974). *The Interaction of Gases with Solid Surfaces*, Pergamon Press, Elmsford, NY, Chapter 2.
22. Horn, R. G., and Israelachvili, J. N. (1980). *Chem. Phys. Lett. 71*: 1922.
23. Horn, R. G., and Israelachvili, J. N. (1981). *J. Chem. Phys. 75*: 1400.

24. Chandrasekhar, S. (1945). *Revs. Mod. Phys. 15*: 1.
25. Hirschfelder, J. O., Curtiss, C. F., and Bird, R. B. (1954). *Molecular Theory of Gases and Liquids*, Wiley, New York, pp. 644–647.
26. Dahler, J. S., and Sather, N. F. (1963). *J. Chem. Phys. 38*: 2363. Condiff, D. W., Lu, W. K., and Dahler, J. S. (1965). *J. Chem. Phys. 42*: 3445.
27. Chandler, D. (1974). *Chem. Res. 7*: 246.
28. Baleiko, M. O., and Davis, H. T. (1974). *J. Phys. Chem. 78*: 1564.
29. Evans, D. F., Tominaga, T., and Davis, H. T. (1981). *J. Chem. Phys. 74*: 1298.
30. Brown, R. G., and Davis, H. T. (1971). *J. Phys. Chem. 75*: 1970.
31. Davis, H. T., Rice, S. A., and Sengers, J. V. (1961). *J. Chem. Phys. 35*: 2210.
32. Davis, H. T., and Luks, K. D. (1965). *J. Phys. Chem. 69*: 869.
33. Bitsanis, I., Magda, J. J., Tirrell, M., and Davis, H. T. (1987). *J. Chem. Phys. 87*: 1733.
34. Van Alsten, J., and Granick, S. (1988). *Phys. Rev. Lett. 61*: 2570.
35. Israelachivili, J. N., McGuiggan, P. M., and Homola, A. J. (1988). *Science 240*: 189. Gee, M. L., McGuiggen, P. M., Israelachivili, J. N., and Homola, A. *J. Chem. Phys.* (to appear).
36. Schoen, M., Cushman, J. H., Dristler, D. J., and Rhykerd, C. L., Jr. (1988). *J. Chem. Phys. 88*: 1398. Schoen, M., Rhykerd, C. L., Jr., Dristler, D. J., and Cushman, J. H. (1989). *Science 245*: 1223.
37. Thompson, P. A., and Robbins, M. O. (1990). *Science 250*.
38. Suh, S. H., and MacElroy, J. M. D. (1986). *Mol. Phys. 58*: 445. MacElroy, J. M. D., and Suh, S.-H. (1987). *Mol. Phys. 60*: 475.
39. van Tassel, P., McCormick, A., and Davis, H. T. *J. Chem. Phys.* (submitted).

Index