TEXT FLY WITHIN
THE BOOK ONLY

UNIVERSAL
LIBRARY

OU_154661

UNIVERSAL
LIBRARY

OUP—552—7-7-66—10,000

OSMANIA UNIVERSITY LIBRARY

Call No. 532
F91E

Accession No. 51790

Author Frisch, Harry L., and

Title Lebowitz, Joel L.
Equilibrium theory of classical fluids. 1964.

This book should be returned on or before the date last marked below.

THE EQUILIBRIUM THEORY OF CLASSICAL FLUIDS

Frontiers in Physics

DAVID PINES, Editor

P. W. Anderson	**CONCEPTS IN SOLIDS:** *Lectures on the Theory of Solids*
N. Bloembergen	**NONLINEAR OPTICS:** *A Lecture Note and Reprint Volume*
N. Bloembergen	**NUCLEAR MAGNETIC RELAXATION:** *A Reprint Volume*
R. Brout	**PHASE TRANSITIONS**
G. F. Chew	**S-MATRIX THEORY OF STRONG INTERACTIONS:** *A Lecture Note and Reprint Volume*
R. P. Feynman	**QUANTUM ELECTRODYNAMICS:** *A Lecture Note and Reprint Volume*
R. P. Feynman	**THE THEORY OF FUNDAMENTAL PROCESSES:** *A Lecture Note Volume*
H. L. Frisch and J. L. Lebowitz	**THE EQUILIBRIUM THEORY OF CLASSICAL FLUIDS:** *A Lecture Note and Reprint Volume*
H. Frauenfelder	**THE MÖSSBAUER EFFECT:** *A Review—With a Collection of Reprints*
S. C. Frautschi	**REGGE POLES AND S-MATRIX THEORY**
M. Gell-Mann and Y. Ne'eman	**THE EIGHTFOLD WAY:** *A Review—With a Collection of Reprints*
R. Hofstadter	**ELECTRON SCATTERING AND NUCLEAR AND NUCLEON STRUCTURE:** *A Collection of Reprints with an Introduction*
M. Jacob and G. F. Chew	**STRONG-INTERACTION PHYSICS:** *A Lecture Note Volume*
L. P. Kadanoff and G. Baym	**QUANTUM STATISTICAL MECHANICS:** *Green's Function Methods in Equilibrium and Nonequilibrium Problems*

A. M. Lane — **NUCLEAR THEORY:** *Pairing Force Correlations to Collective Motion*

P. Nozières — **THORY OF INTERACTING FERMI SYSTEMS**

R. Omnès and M. Froissart — **MANDELSTAM THEORY AND REGGE POLES:** *An Introduction for Experimentalists*

G. E. Pake — **PARAMAGNETIC RESONANCE:** *An Introductory Monograph*

D. Pines — **THE MANY-BODY PROBLEM:** *A Lecture Note and Reprint Volume*

J. R. Schrieffer — **THEORY OF SUPERCONDUCTIVITY**

E. J. Squires — **COMPLEX ANGULAR MOMENTA AND PARTICLE PHYSICS:** *A Lecture Note and Reprint Volume*

L. Van Hove, N. M. Hugenholtz, and L. P. Howland — **PROBLEMS IN QUANTUM THEORY OF MANY-PARTICLE SYSTEMS:** *A Lecture Note and Reprint Volume*

THE EQUILIBRIUM THEORY OF CLASSICAL FLUIDS

A LECTURE NOTE AND REPRINT VOLUME

HARRY L. FRISCH

Bell Telephone Laboratories

JOEL L. LEBOWITZ

Belfer Graduate School of Science

Yeshiva University

W. A. BENJAMIN, INC.
NEW YORK AMSTERDAM
1964

THE EQUILIBRIUM THEORY OF CLASSICAL FLUIDS
A Lecture Note and Reprint Volume

Copyright © 1964 by W. A. Benjamin, Inc.
All rights reserved.

Library of Congress Catalog Card Number 65-13012
Manufactured in the United States of America

The camera copy for this book was furnished by the authors on August 8, 1964; this volume was published on November 27, 1964.

EDITOR'S FOREWORD

The problem of communicating in a coherent fashion the recent developments in the most exciting and active fields of physics seems particularly pressing today. The enormous growth in the number of physicists has tended to make the familiar channels of communication considerably less effective. It has become increasingly difficult for experts in a given field to keep up with the current literature; the novice can only be confused. What is needed is both a consistent account of a field and the presentation of a definite "point of view" concerning it. Formal monographs cannot meet such a need in a rapidly developing field, and, perhaps more important, the review article seems to have fallen into disfavor. Indeed, it would seem that the people most actively engaged in developing a given field are the people least likely to write at length about it.

"Frontiers in Physics" has been conceived in an effort to improve the situation in several ways. First, to take advantage of the fact that the leading physicists today frequently give a series of lectures, a graduate seminar, or a graduate course in their special fields of interest. Such lectures serve to summarize the present status of a rapidly developing field and may well constitute the only coherent account available at the time. Often, notes on lectures exist (prepared by the lecturer himself, by graduate students, or by postdoctoral fellows) and have been distributed in mimeographed form on a limited basis. One of the principal purposes of the "Frontiers in Physics" series is to make such notes available to a wider audience of physicists.

It should be emphasized that lecture notes are necessarily rough and informal, both in style and content, and those in the series will prove no exception. This is as it should be. The point of the series is to offer new,

rapid, more informal, and, it is hoped, more effective ways for physicists to teach one another. The point is lost if only elegant notes qualify.

A second way to improve communication in very active fields of physics is by the publication of collections of reprints of recent articles. Such collections are themselves useful to people working in the field. The value of the reprints would, however, seem much enhanced if the collection would be accompanied by an introduction of moderate length, which would serve to tie the collection together and, necessarily, constitute a brief survey of the present status of the field. Again, it is appropriate that such an introduction be informal, in keeping with the active character of the field.

A third possibility for the series might be called an informal monograph, to connote the fact that it represents an intermediate step between lecture notes and formal monographs. It would offer the author an opportunity to present his views of a field that has developed to the point at which a summation might prove extraordinarily fruitful, but for which a formal monograph might not be feasible or desirable.

Fourth, there are the contemporary classics—papers or lectures which constitute a particularly valuable approach to the teaching and learning of physics today. Here one thinks of fields that lie at the heart of much of present-day research, but whose essentials are by now well understood, such as quantum electrodynamics or magnetic resonance. In such fields some of the best pedagogical material is not readily available, either because it consists of papers long out of print or lectures that have never been published.

"Frontiers in Physics" is designed to be flexible in editorial format. Authors are encouraged to use as many of the foregoing approaches as seem desirable for the project at hand. The publishing format for the series is in keeping with its intentions. Photo-offset printing is used throughout, and the books are paperbound, in order to speed publication and reduce costs. It is hoped that the books will thereby be within the financial reach of graduate students in this country and abroad.

Finally, suggestions from interested readers as to format, contributors, and contributions will be most welcome.

DAVID PINES

Urbana, Illinois
August 1964

PREFACE

The purpose of this book is twofold: (1) to serve as a convenient reference book for scientists with an active interest in the field of classical fluids or in related fields (such as magnetism, the quantum mechanical many-body problem, and so on, and (2) to serve as a summary and guide to the present status of some active areas of research in this field. Selection implies omission, and the topics covered in this volume reflect by choice and necessity the interests of the editors and some of their friends (as well as the blindspots of the editors, no doubt). In addition to this, however, the inclusion of specific articles has been based on the editors' estimation of their present utility. In many cases, we have not used the original work because we felt that subsequent derivations, some of them published for the first time here, are simpler or more general. This, of course, in no way detracts from the credit due to the originators of the theory. We particularly would like to mention here the name of Yvon, whose very original work has not received sufficient attention in English-language literature.

Space limitations have precluded any treatment of certain topics — among which are a discussion of the structural theories of liquids (Bernal, hole, cell, free volume, significant structures theories) — which, in any case, have been recently reviewed (cf. Bibliography). No surface properties of fluids, like capillarity, are discussed, nor are complex fluids and, unfortunately, we had to omit detailed numerical comparisons between theory and experiment.* To aid the reader in filling in some of these omissions we have appended a bibliography of monographs, textbooks, and review articles that have special bearing on the equilibrium theory of classical fluids.

The papers in this monograph are arranged in three sections. Comments, by either the editors or the indicated author, follow at the ends of some of the papers.

We wish to thank all the authors represented in this book — in particular, those whose contributions were instigated by us: Drs. M. Fisher, E. Helfand,

*A just-published review article dealing with this topic particularly from the viewpoint of the cell theory, is by A. Levelt and E. D. G. Cohen in *Studies in Statistical Mechanics*, Volume II, edited by J. De Boer and G. E. Uhlenbeck (Amsterdam: North-Holland Publishing Company, 1964).

J. Percus, G. Stell, and M. Wertheim. We are indebted to Mrs. F. F. Ziehm of the Bell Telephone Laboratories for her help in the translation from the French of L. van Hove's paper in *Physica*, **16**, 137 (1950), and to Dr. van Hove for authorizing this translation. We also wish to acknowledge our indebtedness to Misses J. Catanzaro, C. Doerr, and R. Loehr for the typing and to Mrs. B. Tenore and Mr. C. J. Jernstedt for the drafting, all of whom are from the Bell Telephone Laboratories. Finally, one of the editors, J. L. L., would like to acknowledge support by the United States Air Force Office of Scientific Research for some of the work reported here.

H. L. FRISCH
Murray Hill, New Jersey

J. L. LEBOWITZ
New York, New York

October, 1964

CONTENTS

Editor's Foreword vii
Preface ix

Section I. General Theory (Exact Results)

I–1. D. Ruelle, "Rigorous Results in Equilibrium Statistical Mechanics" I–1
I–2. C. N. Yang and T. D. Lee, "Statistical Theory of Equations of State and Phase Transitions. I. Theory of Condensation," *Phys. Rev.*, **87**, 404–409 (1952) I–13
Comments I–19
I–3. T. D. Lee and C. N. Yang, "Statistical Theory of Equations of State and Phase Transitions. II. Lattice Gas and Ising Model," *Phys. Rev.*, **87**, 410–419 (1952) I–21
Comment I–31
I–4. J. Groeneveld, "Two Theorems on Classical Many-Particle Systems," *Phys. Lett.*, **3**, 50–51 (1962) I–32
Comment I–34
I–5. O. Penrose, "Convergence of Fugacity Expansions for Fluids and Lattice Gases," *J. Math. Phys.*, **4**, 1312–1320 (1963) I–35
I–6. J. L. Lebowitz and O. Penrose, "Convergence of Virial Expansions," *J. Math. Phys.* **5**, 841–847 (1964) I–44

I–7. L. Van Hove, "Configuration Integral For One-Dimensional System of Particles," *Physica,* **16,** 137–143 (1950) I–46
Comments I–60
I–8. S. F. Edwards and A. Lenard, "Exact Statistical Mechanics of a One-Dimensional System with Coulomb Forces. II. The Method of Functional Integration," *J. Math. Phys.,* **3,** 778–792 (1962) I–61
I–9. A. Lenard, "Exact Statistical Mechanics of a One-Dimensional System with Coulomb Forces. III. Statistics of the Electric Field," *J. Math. Phys.,* **4,** 533–543 (1963)—Abstract Only I–76
Comment I–76
I–10. Comment on the Indirect Problem of Statistical Physics I–77

Section II. Series Expansions and Integral Equations

II–1. J. G. Kirkwood "Statistical Mechanics of Fluid Mixtures." *J. Chem. Phys.,* **3,** 300–313 (1935) II–3
Comments II–17
II–2. J. E. Mayer, "Integral Equations Between Distribution Functions of Molecules," *J. Chem. Phys.,* **15,** 187–201 (1947) II–18
Comment II–32
II–3. J. K. Percus, "The Pair Distribution Function in Classical Statistical Mechanics II–33
II–4. G. Stell, "Cluster Expansions for Classical Systems in Equilibrium" II–171
II–5. M. Wertheim, "Analytic Solution of Percus-Yevivck Equation" *J. Math. Phys.,* **5,** 643–651 (1964) II–268
"Further Comments on the Preceding Paper" II–276
II–6. The Reiss-Frisch-Lebowitz Theory for Mixtures of Hard Spheres II–299
II–7. B. R. A. Nijbor and L. Van Hove, "Radial Distribution of a Gas of Hard Sphere and the Superposition Approximation," *Phys. Rev.* **85,** 777–783 (1952) II–303
Comment II–310
II–8. F. H. Ree and W. G. Hoover, "Fifth and Sixth Virial Coefficients for Hard Spheres and Hard Disks," *J. Chem. Phys.,* **40,** 939–949 (1964) II–311
II–9. Approximate Integral Equations for g(r) and the Equation of State II–323

Section III. Phase Transition and Critical Phenomena

III–1. L. S. Ornstein and F. Zernike, "Accidental Deviations of Density and Opalescence at the Critical Point of a Single Substance," *Proc. Akad. Sci.* (Amsterdam), **17**, 793–806 (1914) III–3

III–2. F. Zernike, "The Clustering-Tendency of the Molecules in the Critical State and the Extinction of Light Caused Thereby," *Proc. Akad. Sci.* (Amsterdam), **18**, 1520–1527 (1916) III–18

III–3. M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, "On the van der Waals Theory of the Vapor-Liquid Equilibrium. I. Discussion of a One-Dimensional Model," *J. Math. Phys.*, **4**, 216–228 (1963) III–26

III–4. G. E. Uhlenbeck, P. C. Hemmer, and M. Kac, "On the van der Waals Theory of the Vapor-Liquid Equilibrium. II: Discussion of the Distribution Functions," *J. Math. Phys.*, **4**, 229–247 (1963)—Abstract Only III–39

III–5. P. C. Hemmer, M. Kac, and G. E. Uhlenbeck, "On the van der Waals Theory of the Vapor-Liquid Equilibrium. III. Discussion of the Critical Region," *J. Math. Phys.*, **5**, 60–74 (1964)—Abstract Only III–40

III–6. E. Helfand, "Statistical Mechanics of Systems with Long-Range Interactions" III–41

III–7. B. J. Adler and T. E. Wainwright, "Phase Transition in Elastic Disks," *Phys. Rev.*, **127**, 359–361 (1962) III–71
Comment III–74

III–8. M. E. Fisher, "Correlation Function and the Critical Region of Simple Fluids," *J. Math. Phys.*, **5**, 944–962 (1964) III–75

III–9. E. A. Guggenheim, "The Principle of Corresponding States," *J. Chem. Phys.*, **13**, 253–261 (1954) III–94

Bibliography III–103

SECTION I

RIGOROUS RESULTS IN EQUILIBRIUM STATISTICAL MECHANICS*

by

David Ruelle
Institute for Advanced Study
Princeton, New Jersey

1. INTRODUCTION

The purpose of my talk is to present some rigorous results obtained recently in statistical mechanics. These results deal with the problem of taking the limit of an infinite system, the so-called thermodynamic limit.

The problem that equilibrium statistical mechanics aims at solving is basically the following. If particles have given microscopic interactions, what are the thermodynamic properties of a system containing a very large number of such particles? To this end, rules have been constructed which we illustrate by the example of a classical system of particles interacting via a two-body potential.

If we have n particles with positions $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_n$ and momenta $\underset{\sim}{p}_1,\ldots,\underset{\sim}{p}_n$ the Hamiltonian function is

$$H(p,x) = \sum_{i=1}^{n} \frac{\underset{\sim}{p}_i^2}{2m} + \sum_{1\le i<j\le n} \Phi(\underset{\sim}{x}_j - \underset{\sim}{x}_i) \qquad (1)$$

where Φ is the two-body interaction potential. Let Λ be a box of volume V. If we use the so-called canonical ensemble we have to construct the quantity

---

*Lecture given at the Second Eastern Theoretical Physics Conference held in Chapel Hill, N. C., October 1963, whose proceedings are edited by E. Merzbacher.

$$(2\pi\hbar)^{-3n}(n!)^{-1}\int dp_1\ldots dp_n\int_{\Lambda^n} dx_1\ldots dx_n e^{-\beta H(p,x)} = Q(\beta;\ \Lambda,n)$$

$$= \exp[-\beta n f(\beta;\ \Lambda,n)] \qquad (2)$$

Then we have to take the limit of $f(\beta;\ \Lambda,n)$ when the volume V of Λ tends to infinity while V/n tends to a constant v, Λ keeping a reasonable shape. This limit $f(\beta,v)$ should depend on β and v only and has to be identified with the thermodynamic free energy per particle at specific volume v and temperature $T = (k\beta)^{-1}$ where k is the Boltzmann constant. If we use the grand canonical ensemble we have to construct

$$\Xi(\beta,\mu;\ \Lambda) = \sum_{n=0}^{\infty} \lambda^n Q(\beta;\ \Lambda,n) = \exp[\beta V p(\beta,\mu;\ \Lambda)] \qquad (3)$$

where $\lambda = e^{\beta\mu}$, then take the limit of $p(\beta,\mu;\ \Lambda)$ when V tends to infinity, Λ keeping a reasonable shape. This limit, $p(\beta,\mu)$, has to be identified to the pressure at temperature $(k\beta)^{-1}$ and chemical potential μ. $p(\beta,\mu)$ and $f(\beta,v)$ are related by the usual thermodynamic relations, which hold however only after the limit of infinite volume has been taken.

The justification of the rules sketched above is an extremely difficult problem which we shall not touch upon. It is however universally admitted that these rules or generalizations of them give the right answer in the cases of physical

interest. We will here only consider the problems connected with the limit of an infinite system. Their special interest lies in the fact that it is only in this limit that thermodynamic behavior appears and that phase transitions occur.

2. CONDITIONS ON THE POTENTIAL

We will from now on restrict ourselves to systems of particles having only the translation degrees of freedom and interacting through two-body central forces in a ν-dimensional space. In order for the thermodynamic limit to exist the potential Φ has to satisfy rather stringent conditions: it should be sufficiently repulsive at the origin and decrease sufficiently fast at infinity.

If we take a potential which is an attractive square well with radius a (Fig. 1) and if we put all n particles in a region of diameter a the energy will be negative and quadratic in n, producing a huge effect in the exponential defining Q. Although the region of integration is small this effect is catastrophic and no thermodynamic limit exists, as is well known. However, exactly the same effect takes place - the grand partition function Ξ diverges - whenever one has a continuous (or upper semicontinuous) potential Φ for which one can find an integer k and k points $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k$ such that

$$\sum_{i=1}^{k}\sum_{j=1}^{k}\Phi(\underset{\sim}{x}_j-\underset{\sim}{x}_i) < 0 \qquad (4)$$

For instance, by taking the points $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k$ at the vertices of a sufficiently large piece of a face-centered cubic lattice one sees that the potential of Fig. 2 is catastrophic in three dimensions. Let me add that this potential would remain catastrophic in quantum theory, with either Bose or Fermi statistics.

One can insure the convergence of the grand partition function by assuming that there exists an additive lower bound for the potential energy $U(x)_n = U(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_n)$. This means that there should exist a constant $B \geq 0$ such that for any n

$$U(x)_n = \sum_{1\leq i<j\leq n} \Phi(\underset{\sim}{x}_j - \underset{\sim}{x}_i) \geq -nB \qquad (5)$$

Incidentally the existence of an additive lower bound was established for systems of electrically charged particles with hard cores by Onsager [1] in 1938. This is probably the first example of such a calculation. Here however we consider particles of only one kind.

There are several known criteria on Φ under which (5) holds. The most useful one is the following, due to M. Fisher [2]. ν being the number of dimensions of space, if Φ is bounded from below and $\Phi(r) > C_1 r^{-\nu-\varepsilon}$ for sufficiently small r and $|\Phi(r)| < C_2 r^{-\nu-\varepsilon}$ for sufficiently large r, where C_1, C_2, ε are positive constants, then (5) holds. This criterion is satisfied in particular by the Lennard-Jones potential.

We will, in what follows, constantly assume that (5) holds. This is the way, perhaps not the best possible, in which we express that the potential is sufficiently repulsive at the origin. As far as the behavior at infinity is concerned we will use different conditions according to the problem treated. These conditions, of which the strongest is to assume finite range of the potential, are in general too strong, and could be weakened if better proofs were available.

3. THERMODYNAMIC LIMIT

Two proofs of the existence of the thermodynamic limit have been published more than ten years ago. The first by Van Hove [3] for the canonical ensemble, the second by Lee and Yang [4] for the grand canonical ensemble. In these proofs the potential is assumed to have a hard core and to be of finite range. Recently, however, Van Kampen [5] has pointed out a difficulty in Van Hove's proof. On the other hand a slight improvement was brought by Witten [6] to the proof of Lee and Yang: one assumes that the potential diverges at least like $r^{-11.4}$ at the origin (in three dimensions) instead of requiring a hard core.

In order to be able to treat more general potentials, the author [7], and independently M. Fisher [5], have used the following idea. Suppose that the potential energy has an additive lower bound and that $\Phi(r) \leq 0$ if $r \geq R$, i.e., the forces become attractive at large distances (which is the case for Van der Waals forces). Let $Q(\beta; \Lambda, n)$ be the canonical partition

function for a cubic box Λ. Consider now a box Λ' obtained by putting together eight boxes Λ at a distance R (Fig. 3). It is immediately clear the $Q(\beta; \Lambda', 8n) \geq [Q(\beta; \Lambda, n)]^8$ because the interaction between particles in different boxes is attractive and the domain of integration is increased. This implies $f(\beta; \Lambda, n) \geq f(\beta; \Lambda', 8n)$. If we iterate this process we obtain a decreasing sequence of numbers f and it follows from the existence of an additive lower bound for the potential energy that this sequence is bounded from below. On the other hand the density tends to a constant. We have therefore established the existence of the thermodynamical limit for one special sequence of boxes. To prove it for an arbitrary sequence of cubic boxes (or even boxes of more general shape) is a technical problem with which I shall not deal here.

Let me instead mention a few more results in this direction. The grand canonical ensemble may be treated in the same way as the canonical ensemble and is in fact easier. If some additional assumptions are made one can generalize the results of Lee and Yang [4] about the zeros of the grand partition function and the thermodynamic connection between the canonical and grand canonical ensemble may be established.

The condition that $\Phi(r) \leq 0$ for r large enough has been lifted by a student of Van Kampen, W. Wils [8], who replaces it by $\Phi(r) < Cr^{-\nu-\varepsilon}$ for some positive C and ε and for

sufficiently large r. The existence of the thermodynamic limit for the canonical ensemble is thus ensured if Φ satisfies the criterion of M. Fisher.

Also, if one forgets about the kinetic energy, the microcanonical ensemble can be treated even if no additive lower bound for the energy is assumed provided the potential has finite range. This result may have interesting developments.

Finally, let me mention that spin systems on a lattice may be treated in a manner analogous to the above. This has been done by Griffiths [9].

4. CORRELATION FUNCTIONS AND VIRIAL EXPANSION

The first paper proving in a case of some generality the convergence of the virial expansion for a gas is a short note by J. Groeneveld [10]. I regret not to be able to go into the details of the proof. The main results are as follows. If Φ is nonnegative and the second virial coefficient is finite, the activity expansions of the density and pressure have a non-vanishing radius of convergence. Furthermore the b-coefficients appearing in these expansions have alternate signs. In what follows we will consider the possibility of extending the first conclusion to potentials which may have an attractive part.

Up to now we have considered only the construction of thermodynamic functions. Statistical mechanics also yields rules for the construction of correlation functions. These functions $\rho(\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_k)$ give the probability density for

finding k different particles at prescribed positions. For the grand canonical ensemble (and at least for the fluid region) one has first to write

$$\rho_\Lambda(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k) = \Xi^{-1}\sum_{n=0}^{\infty}\frac{z^{k+n}}{n!}\int d\underset{\sim}{x}_{k+1}\cdots d\underset{\sim}{x}_{k+n}\, e^{-\beta U(x)_{k+n}} \tag{6}$$

where z is the activity, then to take the limit of an infinite volume in such a way that Λ keeps a reasonable shape, contains $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k$, and the distance of $\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k$ from the boundary of Λ tends to infinity.

To prove the existence of this limit in general may be a very difficult task. The problem can be treated however if we restrict ourselves again to two-body central forces [11], with very unrestrictive conditions at infinity and z sufficiently small. It is known that the functions $\rho_\Lambda(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k)$ satisfy different systems of integral equations. One of them, formed by the Kirkwood-Salsburg equations, will be of particular interest to us. The Kirkwood-Salsburg equations are obtained by rewriting the factors $e^{-\beta\Phi(\underset{\sim}{x}_j-\underset{\sim}{x}_1)}$ in the exponential $e^{-\beta U(x)_{k+n}}$ of (6) in the form $1 + f(\underset{\sim}{x}_j-\underset{\sim}{x}_1)$ where $f(x) = e^{-\beta\Phi(\underset{\sim}{x})} - 1$ is the Mayer function, then performing some algebra.

If we call ρ_Λ the sequence of the functions $\rho_\Lambda(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_k)$, and consider ρ_Λ as an element of a vector space, we can write the Kirkwood-Salsburg system in the form of an inhomogeneous linear equation

$$\rho_\Lambda = z\alpha_\Lambda + z\mathcal{K}_\Lambda\rho_\Lambda \tag{7}$$

in this vector space (some modifications are required into which I shall not go). Equation (7) has been derived under the assumption that the potential energy has an additive lower bound. Our only further assumption is about the behavior of Φ at infinity and is very weak:

$$\int d\underset{\sim}{x}\left|e^{-\beta\Phi(\underset{\sim}{x})} - 1\right| < +\infty \tag{8}$$

It is a remarkable fact that a norm can be introduced on the vector space in which (7) holds, making it a Banach space and such that the norm $\|\mathcal{K}_\Lambda\|$ of $\mathcal{K}_\Lambda$ is finite. For z small enough (7) may thus be solved in the form

$$\rho_\Lambda = z(1-z\mathcal{K}_\Lambda)^{-1}\alpha_\Lambda \tag{9}$$

and this solution is analytic in z in a circle centered at the origin of the complex z-plane.

In the limit of an infinite volume, (7) becomes formally

$$\rho = z\alpha + z\mathcal{K}\rho \tag{10}$$

which can be solved in the same way. It is then a technical problem to show that ρ_Λ tends to ρ. ρ is thus the infinite volume limit of the sequence of the correlation functions when z is real and positive.

A consequence of these results is that the density and the pressure are analytic functions of the activity z in some circle centered at the origin: the z-expansions of ρ and p are thus convergent, which also implies the convergence of the virial expansion in some part of the gas phase.

O. Penrose [12] has attacked the problem in a slightly different way by writing power series expansions for ρ_Λ and ρ and obtaining by iteration from (7) and (10) excellent upper bounds for the coefficients. These bounds have then been used in a very elegant paper by J. L. Lebowitz and O. Penrose [13] to obtain bounds for the virial coefficients and a lower bound for the radius of convergence of the virial expansion.

Let me finally mention that by an extension of the above methods one can prove [14] a cluster property of the correlation functions (in the activity region considered). For the two-body correlation function this cluster property simply is that $\rho(x_1,x_2) - \rho(x_1)\rho(x_2)$ is an absolutely integrable function of $x_2 - x_1$.

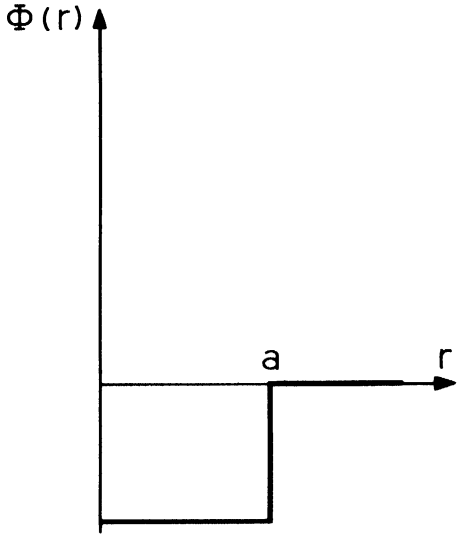


FIG. 1

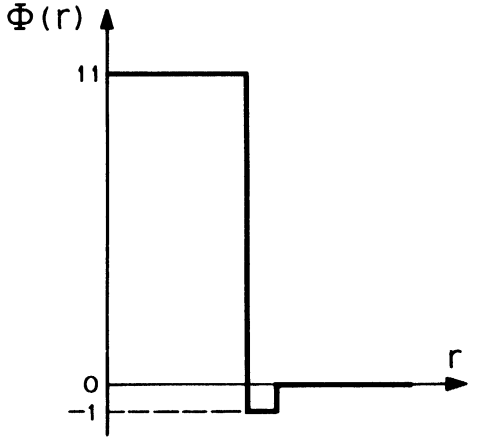


FIG. 2

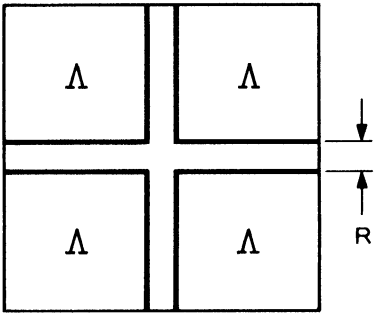


FIG. 3

BIBLIOGRAPHY

[1] L. Onsager. Journal of Physical Chemistry 43, 189 (1939).

[2] M. Fisher. Private communication. For a proof, see D. Ruelle, Lecture notes of the Theoretical Physics Institute (1963), University of Colorado, Boulder.

[3] L. Van Hove. Physica 15, 951 (1949).

[4] C. N. Yang and T. D. Lee. Phys. Rev. 87, 404 (1952).

[5] Private communication of M. Fisher.

[6] L. Witten. Phys. Rev. 93, 1131 (1954).

[7] D. Ruelle. Helv. Phys. Acta. 36, 183 (1963).

[8] W. Wils. Unpublished.

[9] R. Griffiths. To be published.

[10] J. Groeneveld. Physics Letters 3, 50 (1962).

[11] D. Ruelle. Annals of Physics. 25, 103 (1963).

[12] O. Penrose. J. of Math. Phys. 4, 1312 (1963).

[13] J. L. Lebowitz and O. Penrose. J. Math. Phys., July, (1964)

[14] D. Ruelle. Rev. Mod. Phys. 36, 580 (1964).

PHYSICAL REVIEW VOLUME 87, NUMBER 3 AUGUST 1, 1952

Statistical Theory of Equations of State and Phase Transitions. I. Theory of Condensation

C. N. YANG AND T. D. LEE
Institute for Advanced Study, Princeton, New Jersey

(Received March 31, 1952)

A theory of equations of state and phase transitions is developed that describes the condensed as well as the gas phases and the transition regions. The thermodynamic properties of an infinite sample are studied rigorously and Mayer's theory is re-examined.

I. INTRODUCTION

THIS and a subsequent paper will be concerned with the problem of a statistical theory of equations of state and phase transitions. This problem has always interested physicists both from the practical viewpoint of seeking for a workable theory of properties of matter (such as a theory of liquids) and also from the more academic viewpoint of understanding the occurrence of the discontinuities associated with phase transitions in the thermodynamic functions.

The work reported in this paper is quite general and fairly abstract. We are returning in a subsequent paper to the illustration and application of the methods here outlined. In order to present the work of this present paper in its proper perspective, it may be helpful if we outline briefly the history of our own thinking on the subject.

About a year ago one of us was able to make progress[1] with the problem of the spontaneous magnetization of the Ising model, taking advantage of some special properties of this problem when treated by the Onsager-Kaufman method.[2] We then noted that the solution there obtained was also the solution of another, physically quite different, but formally identical, problem. This is the problem of a lattice gas with attractive interaction between nearest neighbors. We were thus able to follow in detail the behavior of such a lattice gas, which in many ways should reveal the features of an actual gas. In particular, we were able to study and characterize the condensation phenonenon, and to identify the liquid, gas, and transition regions in the $p-v$ diagram. The isotherms thus obtained are flat in the transition region and rise very rapidly with increasing density in the liquid phase. At this point, we were led to compare the specific solution with the well-known work[3] of Mayer on the theory of condensation of gases. In particular we were led to inquire as to why, in Mayer's theory, the isotherms stay flat beyond the condensation point and do not give the equation of state for the liquid phase. It soon became apparent that this difference lay, not in the difference of the models, but in the inadequacy of Mayer's method for dealing with a condensed phase. This led to a study of the analytical behavior of the grand partition function of an assembly of interacting atoms, and we were able, as in the special case mentioned above, to identify and characterize quite generally the condensation phenomena. These general conclusions will be presented in the present paper.

The problem is approached by allowing the fugacity to take on complex values. Although only real values of the fugacity are of any physical interest, the analytical behavior of the thermodynamic functions can only be completely revealed by going into the complex plane, whereby one is able to obtain a description of the condensed phases as well as the gas phase and the transition regions. This approach is of a very general nature and can be applied to other problems of phase transitions such as ferromagnetism, order disorder transition, etc. It will be emphasized that also this approach can lead to practical approximation methods for the description of systems undergoing transitions. These points will be discussed in paper II.

The physical conclusions of this paper derive from some mathematical results which we shall state in the form of two theorems. Due to the nature of the problem (which involves a double limiting process) it is imperative to have mathematical rigor preserved throughout. The proofs are necessarily of a mathematical nature and will be given in the appendix.

II. INTERACTION

We consider a monatomic gas with the interaction

$$U=\sum u(r_{ij}), \tag{1}$$

where r_{ij} is the distance between the ith and jth atoms. The following assumptions are made about the nature of these interactions:

(1) The atoms have a finite impenetrable core of diameter a, so that $u(r)=+\infty$ for $r\leqq a$.

(2) The interaction has a finite range b so that

$$u(r)=0 \text{ for } r\geqq b.$$

(3) $u(r)$ is nowhere minus infinity.

The theory can be easily generalized to include many body forces and forces with a weak long tail such as

[1] C. N. Yang, Phys. Rev. 85, 808 (1952).
[2] L. Onsager, Phys. Rev. 65, 117 (1944); B. Kaufman, Phys. Rev. 76, 1232 (1949).
[3] J. E. Mayer, J. Chem. Phys. 5, 67 (1937); J. E. Mayer and Ph. G. Ackermann, J. Chem. Phys. 5, 74 (1937); J. E. Mayer and S. F. Harrison, J. Chem. Phys. 6, 87, 101 (1938); B. Kahn and G. E. Uhlenbeck, Physica 5, 399 (1938). M. Born and K. Fuchs, Proc. Roy. Soc. (London) A166, 391 (1938).

van der Waals' force. But for clarity we shall first treat only interactions with the properties enumerated above.

Consider a box of volume V kept at a constant temperature T. If it is allowed to exchange atoms with a reservoir at a given chemical potential μ per atom, the relative probability of having N atoms in the box is

$$Q_N y^N/N!,$$

where

$$Q_N=\int\cdots\int_V d\tau_1\cdots d\tau_N \exp(-U/kT) \qquad (2)$$

is the configurational part of the partition function for N atoms and

$$y=(2\pi mkT/h^2)^{\frac{1}{2}}\exp(\mu/kT). \qquad (3)$$

The quantities m, k, and h have the usual meanings,

The grand partition function of the gas in the volume V is

$$\mathcal{Q}_V=\sum_{N=0}^{M}\frac{Q_N}{N!}y^N, \qquad (4)$$

where M is the maximum number of atoms that can be crammed into V.

III. THE LIMIT OF INFINITE VOLUME

The average pressure and the average density of such a gas in V are calculable in terms of $\mathcal{Q}_V$ by the standard treatment of statistical mechanics, and are evidently dependent on V. In thermodynamics, however, one is only interested in an infinite sample and the thermodynamic functions are limits of these average quantities as $V\to\infty$. The pressure p and density ρ are accordingly given by

$$\frac{p}{kT}=\operatorname{Lim}_{V\to\infty}\frac{1}{V}\log\mathcal{Q}_V, \qquad (5)$$

$$\rho=\operatorname{Lim}_{V\to\infty}\frac{\partial}{\partial\log y}\frac{1}{V}\log\mathcal{Q}_V. \qquad (6)$$

The question of whether these limits do exist is usually not discussed.[4] It is, however, generally believed that in the gas phase such limits do exist and that (5) and (6) give the correct equation of state. At the point of condensation and in the liquid phase the situation has been extremely unclear. As a matter of fact doubts have been raised[5] as to whether the equation of state of both the liquid and the gas phase can be obtained from the same interaction (1) through the considerations of statistical mechanics.

We shall try to resolve these problems and prove that (5) and (6) do give a complete description of the equation of state of both the gas and condensed phases. In fact in paper II we shall give a concrete example in which it is seen how the same partition function describes both phases, and in which the two-phase-equilibrium region is exactly known.

We first state the following:

Theorem 1.—(Proved in Appendix I.) For all positive real values of y, $V^{-1}\log\mathcal{Q}_V$ approaches, as $V\to\infty$, a limit which is independent of the shape of V. Furthermore, this limit is a continuous, monotonically increasing function of y.

The assumption is made, of course, that the shape of V is not so queer that its surface area increases faster than $V^{\frac{3}{4}}$.

One might be tempted to conclude from the independence of the limit on the shape of V that the system under consideration exists only in fluid phases (i.e., gas and liquid) with no elastic resistance against shearing strain. It is to be emphasized that this is not the case. The independence of the limit on the shape of V is not due to the lack of elastic resistance against shearing strain, but rather due to the fact that for an infinite sample changing the shape of V does not produce a strain in the interior which might serve to differentiate between a fluid and a solid. This is so because the strain at the boundary only penetrates to a finite depth and is inconsequential for an infinite sample.

To study the limit of $(\partial/\partial\log y)V^{-1}\log\mathcal{Q}_V$ we notice that $\mathcal{Q}_V$ is a polynomial in y of finite degree M. This is a direct consequence of the assumed impenetrable core of the atoms. It is therefore possible to factorize $\mathcal{Q}_V$ and write

$$\mathcal{Q}_V=\prod_{i=1}^{M}\left(1-\frac{y}{y_i}\right), \qquad (7)$$

where $y_1,\cdots y_M$ are the roots of the algebraic equation

$$\mathcal{Q}_V(y)=0. \qquad (8)$$

Evidently none of these roots can be real and positive, since all the coefficients in the polynomial $\mathcal{Q}_V$ are positive.

As V increases these roots move about in the complex y plane and their number M increases (essentially) linearly with V. Their distribution in the limit $V\to\infty$ gives the complete analytic behavior of the thermodynamic functions in the y plane. In fact one can prove the following:

Theorem 2.—(Proved in Appendix II.) If in the complex y plane a region R containing a segment of the positive real axis is always free of roots, then in this region as $V\to\infty$ all the quantities:

$$\frac{1}{V}\log\mathcal{Q}_V,\quad\left(\frac{\partial}{\partial\log y}\right)\frac{1}{V}\log\mathcal{Q}_V,\quad\left(\frac{\partial}{\partial\log y}\right)^2\frac{1}{V}\log\mathcal{Q}_V\cdots,$$

[4] The behavior of the partition function Q_N as the volume approaches infinite was discussed by L. von Hove, Physica 15, 951 (1949), where it is proved that $N^{-1}\log Q_N$ approaches a limit as the volume approaches infinity at constant density. His proof is similar to our proof of theorem 1.

[5] There was apparently some discussion on this point at the International Conference held in Amsterdam, 26 November 1937. The doubts can perhaps be formulated in the form of the question: "How can the gas molecules know when they have to coagulate to form a liquid or solid?" See p. 391 of reference 3 (Born and Fuchs).

approach limits which are analytic with respect to y. Furthermore the operations $(\partial/\partial \log y)$ and $\lim_{V\to\infty}$ commute in R so that, e.g.,

$$\lim_{V\to\infty} \frac{\partial}{\partial \log y} \frac{1}{V} \log \mathcal{Q}_V = \frac{\partial}{\partial \log y} \lim_{V\to\infty} \frac{1}{V} \log \mathcal{Q}_V. \quad (9)$$

This gives, together with (5) and (6),

$$p = (\partial/\partial \log y)(p/kT). \quad (10)$$

IV. PHASE TRANSITIONS

The quantity $(\partial/\partial \log y) V^{-1} \log \mathcal{Q}_V$ does not, however, always approach a limit p for all values of y. Physically this must evidently be as the density of the system does not assume a single value at the point of condensation. It is clear therefore that the problem of phase transition is intrinsically related to the form of the regions R described in theorem 2. We discuss the following cases:

(1) The roots of $\mathcal{Q}_V(y) = 0$ do not close in onto the positive real axis of y as $V \to \infty$, or more exactly, there exists a region R which contains the whole positive real axis and is free of roots.

In this case from the two theorems one concludes that the pressure and density of the system are analytic functions of y (along the positive real axis). They are related by Eq. (10). Furthermore, p is an increasing function of y. We shall show in Appendix III that ρ is also an increasing function of y. Consequently in the $p-\rho$ diagram, on the isotherm p increases analytically as the specific volume v decreases. The system under consideration is thus a single phase system (see Fig. 1).

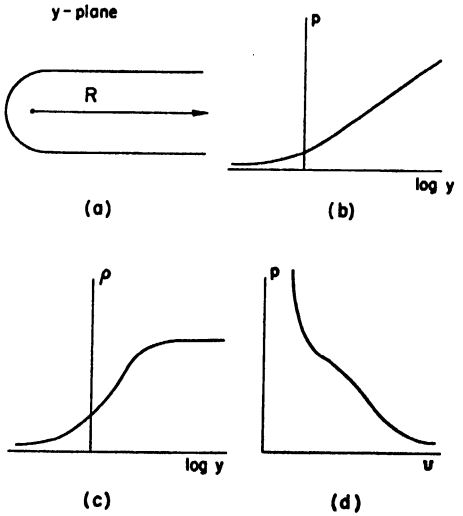


FIG. 1. Analytical behavior at a given temperature of thermodynamic functions for a single phase system. The quantity y is defined by Eq. (3) in the text. The region R is free of roots of Eq. (8). Notice that the density ρ of (c) is proportional to the slope of the $p-\log y$ curve in (b).

(2) The roots of $\mathcal{Q}_V(y) = 0$ do close in onto the real axis as $V \to \infty$, say at the points $y = l_1, l_2$; and regions R_1, R_2, and R_3 free of roots enclose, respectively, the three segments of the positive real axis as in Fig. 2(a).

By the same reasoning as in the previous case one concludes, within any one of these three segments, that the system exists in a single phase, that p and ρ are analytic and increasing functions of y, that ρ is $(kT)^{-1}(\partial p/\partial \log y)$, and that on the isotherm p increases analytically as v decreases.

At the points $y = l_1, l_2$ the pressure p is continuous (by theorem 1), but its derivative ρ has in general a discontinuity. By Appendix III one shows easily that ρ increases across the discontinuity. The functions p and ρ are schematically plotted in Figs. 2(b) and 2(c) which together give the isotherm in Fig. 2(d).

As the temperature varies the points l_1 and l_2 will in general move along the y axis. If at a certain temperature T_c the roots cease to close in onto one of the points, say l_1, then T_c is the critical temperature for the transition phase $1 \leftrightarrow$ phase 2. If, on the other hand, l_1 and l_2 merge together at a particular temperature T_0, we would then have a triple point at that temperature.

It may be remarked that at $y = l_1$ or l_2 the density ρ may in some cases be continuous (although its derivative will in general be discontinuous). At the critical temperature this will happen, but not at neighboring temperatures. If, however, this happens over an extended temperature range, one would have a transition of second (or higher) order.

It is clear therefore that phase transitions of the system occur only at the points on the positive real y axis onto which the roots of $\mathcal{Q}(V) = 0$ close in as $V \to \infty$. For other values of the fugacity y a single phase system obtains.

As mentioned before, the theory can be easily generalized to include many body forces and forces with a weak long tail. In fact, the generalization does not lead to any alterations of the conclusions reached above.

Generalization can also be made to other kinds of phase transitions such as order-disorder phenomena and ferromagnetism, as will be discussed in paper II. The study of the equations of state and phase transitions can thus be reduced to the investigation of the distribution of roots of the grand partition function. In many cases, as will be seen in paper II, such distributions turn out to have some surprisingly simple regularities.

V. COMPARISON WITH MAYER'S THEORY

We first notice that by expanding in powers of y one obtains from (7)

$$\frac{1}{V} \log \mathcal{Q}_V = \sum_{l=1}^{\infty} b_l(V) y^l, \quad (11)$$

where

$$b_l(V)=\frac{-1}{lV}\sum_{j=1}^{M}\left(\frac{1}{y_j}\right)^l. \tag{12}$$

Combining (11) and (3) we have

$$\frac{Q_N}{N!}=\text{coefficient of } y^N \text{ in } \exp[V\sum_{l=1}^{\infty} b_l y^l]. \tag{13}$$

Comparison of this equation with Mayer's theory shows that the b_l's defined by (12) are identical with the reducible cluster integrals defined[6] by Mayer. It is interesting to notice that these reducible cluster integrals are, according to (12), closely related to the moments of the roots y_j of Eq. (8). It should be emphasized that in both (12) and in Mayer's definition the b_l's are functions of the volume V. It is evident from Mayer's definition that they approach definite limits $b_l(\infty)$ as $V\to\infty$.

In Mayer's theory the cluster integrals b_l are replaced from the very beginning by their limiting values $b_l(\infty)$. He then considers the series

$$\chi(y)=\sum_{1}^{\infty} b_l(\infty)y^l \tag{14}$$

and its analytical continuation along the positive real axis. If one calls the first singularity of $\chi(y)$ along the positive real axis t_1, one shows in Mayer's theory that

(1) for densities ρ less than

$$\rho_1=\lim_{y\to t_1-} y\chi'(y), \tag{15}$$

the system exists in a single phase;

(2) for $\rho\geqq\rho_1$, the pressure p (at a given temperature) becomes independent of the density. Consequently, one identifies the density ρ_1 as the density of the gas at condensation.

An essential difficulty of Mayer's theory is that it does not admit of the existence of a liquid phase with finite density, since the isotherm remains horizontal for all specific volume less than ρ_1^{-1}. This is clearly due to the replacement of the volume dependent b_l's by their limiting values. The question is therefore often raised[7] as to exactly at what point on the isotherm Mayer's theory breaks down.

In the present theory by retaining the volume dependence of the partition function $\mathcal{Q}_V$ we do not encounter these difficulties. To clarify the relationship with Mayer's theory, we refer back to Fig. 2(a) and draw a small circle C within R_1 with the center at the origin. The series

$$\sum_{l=1}^{\infty} b_l(V)y^l$$

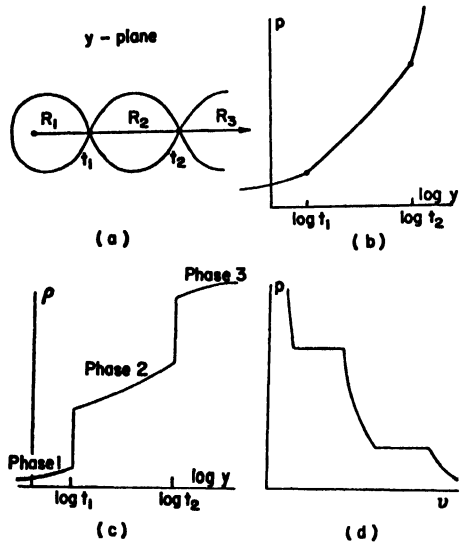


FIG. 2. Analytical behavior at a given temperature of thermodynamic functions for a system that undergoes two phase transitions. The transitions occur at t_1 and t_2 which are the points at which the roots of Eq. (8) close in onto the positive real y axis. The regions R_1, R_2, and R_3 are free of roots. The three phases 1, 2, and 3 are indicated in (c). The horizontal parts of (d) represent two-phase-equilibrium regions.

is easily shown[8] to converge uniformly in the circle C. By a well-known mathematical theorem on double limiting processes one concludes that in C

$$\lim_{V\to\infty}\sum_{1}^{\infty} b_l(V)y^l=\sum_{1}^{\infty} b_l(\infty)y^l. \tag{16}$$

The left-hand side of this equation is by definition

$$\lim_{V\to\infty} V^{-1}\log\mathcal{Q}_V,$$

and the right-hand side $\chi(y)$. Therefore within C the function $\chi(y)$ in Mayer's theory is indeed $(kT)^{-1}$ times the pressure p as defined by (5). By analytical continuation one concludes that this holds throughout the region R_1.

In the interval $0\leqq y<t_1$ it is evident that $\rho<\rho_1$ and Mayer's theory is seen to give a correct description of the system.

Beyond the point $y=t_1$ (i.e., $y\geqq t_1$) it is not possible in Mayer's theory to analytically continuate $\chi(y)$. The $p-\log y$ and $\rho-\log y$ diagrams [Figs. 2(b) and 2(c)] therefore exist in his theory only to the left of the first singularity. This explains the nonexistence of the liquid phase in Mayer's theory.

[6] See for example Eq. (13.5) in J. E. Mayer and Mayer, *Statistical Mechanics* (John Wiley and Sons, Inc., New York, 1946), p. 280.

[7] See, for example, reference 3 (Kahn and Uhlenbeck), p. 415.

[8] All roots y_j have absolute values larger than the radius σ of the circle C. By Eq. (12) we have $|b_l|\leqq(M/V)l^{-1}\sigma^{-l}$. But M/V is bounded. Hence the statement.

We thus remark:

(1) Throughout the gas phase (i.e., $\rho<\rho_1$) Mayer's theory gives correct results.

(2) For $\rho\geqq\rho_1$ Mayer's conclusion that the $p-v$ diagram becomes horizontal is, as already mentioned, incorrect for high densities due to the existence of the liquid phase. It is not even justified for densities immediately above ρ_1, as for transitions of high order the isotherm does not even have any horizontal part at all.

We are indebted to Professor J. R. Oppenheimer for criticism and comment.

APPENDIX I

To prove theorem 1 we first establish the following:

Lemma 1.—Let V and W be two cubes of linear dimensions L and $L+(b/2)$, respectively. Keeping b fixed one has as $L\to\infty$,

$$\operatorname{Lim} W^{-1}(\log\mathcal{Q}_W-\log\mathcal{Q}_V)=0. \tag{17}$$

Proof.—Put V completely inside W and write $\mathcal{Q}_W$ as the sum of contributions $A_0, A_1\cdots$ from configurations with zero, one . . . atoms outside of V. Now (a) since the interaction has finite range, each atom interacts with at most a finite and definite number of other atoms. Also (b) the available volume for the first atom outside of V is $\Delta=W-V$. If the volume of the impenetrable core of an atom is α, the available volume for the second atom outside of V is less than $\Delta-\alpha$, the third, $\Delta-2\alpha$, etc. Combining (a) and (b) one concludes that

$$A_m\leqq\beta^m[\Delta(\Delta-\alpha)(\Delta-2\alpha)\cdots(\Delta-m\alpha+\alpha)/m!]\mathcal{Q}_V, \tag{18}$$

where β is a constant. (This inequality is obtained by comparing the contributions to A_m and $\mathcal{Q}_V$ of a distribution of atoms with, say, N atoms inside of V and m atoms outside.) Adding all A's one obtains

$$\mathcal{Q}_W\leqq\mathcal{Q}_V(1+\beta\alpha)^{\Delta/\alpha}.$$

But $\Delta\sim L^2$, and clearly

$$\mathcal{Q}_V\leqq\mathcal{Q}_W.$$

Hence the lemma.

Lemma 2.—Let W_i be a cube of linear dimension 2^iL and $\mathcal{Q}_i$ an abbreviation for $\mathcal{Q}_{W_i}$. Then

$$\lim_{i\to\infty} W_i^{-1}\log\mathcal{Q}_i=K \text{ exists.}$$

Proof.—Consider W_j to be built up from 8^{j-i} smaller cubes $W_i(j>i)$. Evidently the number of atoms interacting across the boundaries of the small cubes is at most proportional to the area of such boundaries and hence is less than $8^i2^{-i}\gamma$, ($\gamma=$constant). Should one neglect these interactions, $\mathcal{Q}_j$ would become $\mathcal{Q}_i$ raised to the power 8^{j-i}. The inclusion of these terms violates this identity by not more than a constant factor β raised to the power 8^i2^{-i}, i.e.,

$$\log\mathcal{Q}_j\leqq 8^{j-i}\log\mathcal{Q}_i+8^i2^{-i}\gamma\log\beta. \tag{19}$$

Next draw within each small cube W_i a concentric cube V_i with linear dimension $2^iL-(b/2)$. Since b is the range of interaction, clearly atoms in different V_i's do not interact. Hence,

$$8^{j-i}\log\mathcal{Q}_{V_i}\leqq\log\mathcal{Q}_j. \tag{20}$$

Equations (19) and (20) give

$$W_i^{-1}\log\mathcal{Q}_{V_i}\leqq W_j^{-1}\log\mathcal{Q}_j\leqq W_i^{-1}\log\mathcal{Q}_i+2^{-i}\gamma L^{-2}\log\beta. \tag{21}$$

The last term approaches zero as $i\to\infty$. Also by Lemma 1, $W_i^{-1}(\log\mathcal{Q}_i-\log\mathcal{Q}_{V_i})\to 0$. Thus as $j>i\to\infty$ $W_j^{-1}\log\mathcal{Q}_j-W_i^{-1}\log\mathcal{Q}_i\to 0$. Hence the lemma.

Proof of the theorem.—Given any $\epsilon>0$ there exists by Lemma 2 a large enough box W such that

$$|K-W^{-1}\log\mathcal{Q}_W|<\epsilon.$$

In fact by the same reasoning as used in proving that lemma, one easily sees that this can also be made true of any box Ω which can by adding partitions be divided into cubes of size W:

$$|K-\Omega^{-1}\log\mathcal{Q}_\Omega|<\epsilon.$$

Now consider a volume V of arbitrary shape. For sufficiently large V one can build two Ω-type boxes Ω_1 and Ω_2 such that Ω_1 is contained in V and Ω_2 contains V and that

$$|(\Omega_1/\Omega_2)-1|<\epsilon.$$

Using

$$\mathcal{Q}_{\Omega_1}\leqq\mathcal{Q}_V\leqq\mathcal{Q}_{\Omega_2},$$

one proves easily that $V^{-1}\log\mathcal{Q}_V$ also approaches K.

That this limit monotonically increases with y follows from the same property of $V^{-1}\log\mathcal{Q}_V$. That it is also continuous follows from the observation that $(\partial/\partial\log y)V^{-1}\log\mathcal{Q}_V$ has a finite and definite upper bound (equal to the density of closest packing).

APPENDIX II

We first prove the following:

Lemma 3.—Consider the series

$$\sum_{l=0}^{\infty} b_l(V)z^l=S_V(z),$$

where

$$|b_l(V)|\leqq A\sigma^{-l},$$

A and σ being positive constants. For all real z between $-\sigma$ and $+\sigma$ assume $\operatorname{Lim} S_V(z)$ to exist as $V\to\infty$. Then (a) $\operatorname{Lim} b_l(V)$ as $V\to\infty$ exists and will be denoted by $b_l(\infty)$. (b) $S_V(z)$ approaches the limit

$$\sum_{l=0}^{\infty} b_l(\infty)z^l, \tag{22}$$

as $V\to\infty$ for all $|z|<\sigma$. The series (22) is convergent for all $|z|<\sigma$.

Proof.—(a) Evidently $b_0(\infty)$ exists and is equal to $\mathrm{Lim}S_V(0)$ as $V\to\infty$. To prove the existence of $\mathrm{Lim}b_1(V)$: Given any real ϵ between 0 and $\sigma/2$ consider the convergence of $S_V(\epsilon)$ and $b_0(V)$ as $V\to\infty$. There exists a volume V_0 such that for any volumes V and W greater than V_0 one has

$$|S_V(\epsilon)-S_W(\epsilon)|<\epsilon^2, \tag{23}$$

$$|b_0(V)-b_0(W)|<\epsilon^2. \tag{24}$$

But

$$|\sum_{l=2}^{\infty} b_l(V)\epsilon^l|\leqq\frac{A\epsilon^2}{1-\epsilon\sigma^{-1}}\leqq 2A\epsilon^2. \tag{25}$$

The same is true if one replaces V by W. Using

$$\epsilon b_1(V)=S_V(\epsilon)-b_0(V)-\sum_{l=2}^{\infty} b_l(V)\epsilon_l,$$

one proves easily with the aid of (23), (24), and (25) that

$$\epsilon|b_1(V)-b_1(W)|<(2+4A)\epsilon^2.$$

Hence $\mathrm{Lim}b_1(V)$ exists as $V\to\infty$. Similar proof holds for the other b_l's.

(b) The series $\sum b_l(V)z^l$ evidently converges uniformly in z for $|z|<\sigma$. The lemma follows from a well-known theorem on double limits.

Proof of theorem 2.—Consider first a circle C, lying inside R, with its center at the point $y=\eta$ along the positive real axis. We shall first prove the theorem inside this circle. Making the displacement $z=y-\eta$ we express (7) in the form

$$\mathcal{Q}_V=\prod_{i=1}^{M}\left(1-\frac{z}{z_i}\right)\left(\frac{z_i}{y_i}\right), \tag{26}$$

where $z_i=y_i-\eta$ are the roots of $\mathcal{Q}_V$. Expanding $V^{-1}\log\mathcal{Q}_V$ in powers of z one obtains

$$\frac{1}{V}\log\mathcal{Q}_V=\sum_{l=0}^{\infty}[b_l(V)z^l, \tag{27}$$

where

$$b_l(V)=\frac{-1}{Vl}\sum_{i=1}^{M}\left(\frac{1}{z_i}\right)^l \text{ for } l\geqq 1, \tag{28}$$

and

$$b_0(V)=\frac{1}{V}\sum_{i=1}^{M}\log\frac{z_i}{y_i}. \tag{29}$$

If σ is the radius of C, since C is free of roots we have $|z_i|\geqq\sigma$. Hence by (28)

$$|b_l(V)|\leqq(M/V)l^{-1}\sigma^{-l} \text{ for } l\geqq 1.$$

But M/V is bounded; hence we can use Lemma 3 and the theorem is proved in C.

By similar arguments we can extent the theorem into a circle C' lying inside R with its center inside C. One can easily prove the theorem in the whole region R by repeating this process.

APPENDIX III

To prove that ρ is an increasing function of y it is only necessary to show for any finite V the inequality

$$\frac{d^2}{(d\log y)^2}\log\mathcal{Q}_V>0.$$

Now $\mathcal{Q}_V$ is a polynomial in y with positive coefficients. Regarding the various terms of $\mathcal{Q}_V$ as relative probabilities we have obviously

$$\frac{d}{d\log y}\log\mathcal{Q}_V=\langle N\rangle,$$

where $\langle\ \rangle$ means "average." Also

$$\frac{d^2}{(d\log y)^2}\log\mathcal{Q}_V=\langle N^2\rangle-\langle N\rangle^2=\langle(\Delta N)^2\rangle,$$

which is always positive. Here ΔN is the deviation of N from the average

$$\Delta N=N-\langle N\rangle.$$

Comments

1. The complex plane behavior of the grand partition function for particles not having hard cores is discussed in the paper of Penrose.

2. For the extension of Yang and Lee's results in Appendix I to more general potentials see the previous paper by Ruelle and references quoted there. The equivalence of the canonical and grand canonical ensemble in the thermodynamic limit is also established there. A general discussion of the relations and proof of equivalence, between canonical, grand canonical and pressure ensembles is given By M. B. Lewis and J. F. Siegert, Extension of the Condensation Theory of Yang and Lee to the Pressure Ensemble, Phys. Rev. 101, 1227 (1956); and M. B. Lewis, Relation between the Canonical and Grand Canonical Ensemble, Phys. Rev. 105, 348 (1957).

3. A proof of the existence of "average" distribution functions in the thermodynamic limit has been given by M. Fisher, (to be published). This extends the results of Groeneveld, Ruelle, Penrose, Lebowitz and Penrose (see following papers) who prove the existence of these distributions ot low densities.

4. A discussion of the N-dependence of the distribution functions in the canonical and grand canonical ensemble is given by J. L. Lebowitz and J. K. Percus, Long-Range

Correlations in a Closed System with Applications to Non-uniform Fluids, Phys. Rev. 122, 1675 (1961). It was shown there that in a fluid having a "finite correlation length", the distribution function of $q + \ell$ particles $n_{q+\ell}(\underline{r}_1, \ldots, \underline{r}_{q+\ell})$ will behave asymptotically, when the set of q particles is far from the set of ℓ particles, as a product of n_q and n_ℓ in a grand ensemble while in a canonical ensemble there will be a N^{-1} correction term,

$$n_{q+\ell} \rightarrow n_q n_\ell - \frac{nkT\kappa}{N}\left(\frac{n\partial n_q}{\partial n}\right)\left(\frac{n\partial n_\ell}{\partial n}\right)$$

where n is the average particle density and κ is the isothermal compressibility.

Statistical Theory of Equations of State and Phase Transitions. II. Lattice Gas and Ising Model

T. D. LEE AND C. N. YANG
Institute for Advanced Study, Princeton, New Jersey
(Received March 31, 1952)

The problems of an Ising model in a magnetic field and a lattice gas are proved mathematically equivalent. From this equivalence an example of a two-dimensional lattice gas is given for which the phase transition regions in the $p-v$ diagram is exactly calculated.

A theorem is proved which states that under a class of general conditions the roots of the grand partition function always lie on a circle. Consequences of this theorem and its relation with practical approximation methods are discussed. All the known exact results about the two-dimensional square Ising lattice are summarized, and some new results are quoted.

INTRODUCTION

IN paper I[1] we have seen that the problem of a statistical theory of phase transitions and equations of state is closely connected with the distribution of roots of the grand partition function. It was shown there that the distribution of roots determines completely the equation of state, and in particular its behavior near the positive real axis prescribes the properties of the system in relation to phase transitions. It was also shown there that the equation of state of the condensed phases as well as the gas phase can be correctly obtained from a knowledge of the distribution of roots. While this general and abstract theory clarifies the problems underlying the statistical theory of phase transitions and condensed phases, it is natural to ask whether it also provides us with a means of obtaining practical approximation methods for calculating properties pertaining to phase transitions and condensed phases.

The problem is clearly that of seeking for the properties of the distribution of roots of the grand partition function. At first sight this appears to be a formidable problem, as the roots are in general complex and would naturally be expected to spread themselves for an infinite sample in the entire complex plane, or at least regions of the complex plane, and make it very difficult to calculate their distribution. We were quite surprised, therefore, to find that for a large class of problems of practical interest, the roots behave remarkably well in that they distribute themselves not all over the complex plane, but only on a fixed circle. This fact will be stated as a theorem in Sec. IV of the present paper and proved in the appendix. Implications of the theorem are discussed in Sec. V.

Also in this paper we shall give a proof (Sec. II) that the problem of an Ising model with a magnetic field is mathematically identical with that of a "lattice gas." From this identification we were able to trace exactly the transition region in the $p-v$ diagram of a two-dimensional lattice gas in detail. This will be presented in Sec. III and forms a clear illustration of the discussions of paper I and of Sec. V of the present paper.

At the end of Sec. V we give a summary of all the exact knowledge known to us about the two-dimensional Ising model in a magnetic field and its relationship with the distribution of roots of the partition function.

TABLE I. Identification of corresponding quantities in Ising model and lattice gas.[a]

Ising model		Lattice gas
No. of spins	=	volume
No. of ↓ spins	=	No. of atoms
$2/(1-I)$	=	specific volume v
$-F-H$	=	pressure p

[a] I, H, and F are respectively the intensity of magnetization, the magnetic field, and the free energy per spin in the Ising model problem.

II. ISING MODEL AND LATTICE GAS

We shall in this section show that the problems of an Ising model in a magnetic field and of a lattice gas are mathematically equivalent. In the former problem one considers a lattice of interacting spins each of which can assume two possible positions: ↑ and ↓. In the latter, one considers a corresponding lattice with each lattice point either vacant or occupied by an atom.[2] To each configuration of the lattice of spins there corresponds a configuration of the lattice gas in which a lattice point is vacant or occupied according as whether the corresponding spin is ↑ or ↓. Using this geometrical correspondence, one could establish the mathematical equivalence of the two problems.

For clarity of presentation we shall take as an example a simple cubic lattice and consider first the Ising model problem with nearest neighbor interaction. The same treatment can be applied to any lattice with arbitrary interaction between the spins. Denote by $[\uparrow\uparrow]$, $[\downarrow\downarrow]$, and $[\uparrow\downarrow]$ the total number of nearest neighboring spins that are respectively parallel and upward, parallel and downward, and antiparallel to each other. Also denote by $[\uparrow]$ and $[\downarrow]$ the total numbers of upward and downward spins in the lattice.

[1] C. N. Yang and T. D. Lee, Phys. Rev. 87, 404 (1952).

[2] Similar ideas have been used in the "hole theory of liquids." See, e.g., J. E. Lennard-Jones and A. F. Devonshire, Proc. Roy. Soc. (London) A169, 317 (1939); A170, 464 (1939). F. Cernuschi and H. Eyring, J. Chem. Phys. 7, 547 (1939).

Counting the number of nearest neighbors of all the upward spins, one arrives at the following identity:

$$2[\uparrow\uparrow]+[\uparrow\downarrow]=6[\uparrow]. \tag{1}$$

Similarly

$$2[\downarrow\downarrow]+[\uparrow\downarrow]=6[\downarrow]. \tag{2}$$

Evidently

$$\mathfrak{N}=[\uparrow]+[\downarrow] \tag{3}$$

is the total number of spins in the lattice.

The interaction energy of all the spins can be written as $[\uparrow\downarrow]\epsilon$ if the interaction energy between parallel spins is taken to be zero. Here ϵ is a constant and is positive for ferromagnetic interactions and negative for antiferromagnetic interactions. In an external magnetic field H (measured in proper units), there is an additional magnetic energy so that the total energy of the Ising lattice is

$$U_I=H([\downarrow]-[\uparrow])+[\uparrow\downarrow]\epsilon. \tag{4}$$

The partition function is therefore

$$\exp(-\mathfrak{N}F/kT)=\sum \exp(-U_I/kT), \tag{5}$$

where F is the free energy per spin and the summation extends over all arrangements of the spins. F as a function of the magnetic field H and the temperature T defines completely the thermodynamic behavior of the lattice. Its derivative with respect to H gives the intensity of magnetization I,

$$\partial F/\partial H=-I, \tag{6}$$

where I is defined to be

$$I=\mathfrak{N}^{-1}([\uparrow]-[\downarrow]). \tag{7}$$

Now consider a lattice gas on the same simple cubic lattice. According to the geometrical correspondence discussed before, each downward spin $\downarrow$ corresponds to one gas atom, hence $[\downarrow]$ is equal to the number of atoms in the gas. Also the "volume" of the gas (in proper units) is simply $\mathfrak{N}$. The specific volume v per atom is, by Eqs. (3) and (7), related to the intensity of magnetization I by

$$v=2/(1-I). \tag{8}$$

To prevent more than one atom from occupying the same lattice site and to correspond to the case of nearest neighbor interaction in the Ising model problem, we consider here the following potential energy u between two atoms:

$u=+\infty$ if the two atoms occupy the same lattice site,

$u=-2\epsilon$ if the two atoms are nearest neighbors, (9)

and

$u=0$ otherwise.

It should be remarked that this interaction closely simulates the actual interaction between atoms of a atonatomic gas. For $\epsilon>0$ (ferromagnetic case), the gas moms attract each other at intermediate distances and

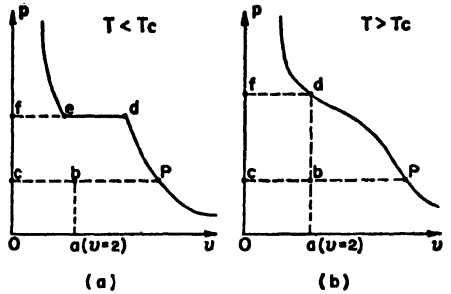


FIG. 1. Geometrical construction of the thermodynamic functions of the Ising model from the $p-v$ diagram of the corresponding lattice gas. Solid curves are isotherms. Given a point P the corresponding Ising model has a magnetic field $H=\frac{1}{2}$ (area $cPdf$), a free energy per spin $F=-\frac{1}{2}$ (area $OabPdf$).

for $\epsilon<0$ (antiferromagnetic case) they repel each other. Since $[\downarrow\downarrow]$ is by the geometrical correspondence the number of nearest neighboring pairs of atoms the energy of the gas is

$$U_G=-2[\downarrow\downarrow]\epsilon.$$

The grand partition function of the gas is

$$\exp(p\mathfrak{N}/kT)=\sum y^{[\downarrow]}\exp(2[\downarrow\downarrow]\epsilon/kT), \tag{10}$$

where p is the pressure of the gas and y the fugacity given by Eq. (3) in paper I. Now on using (2), (3), and (4), one can write Eq. (5) in the form

$$\exp\{-\mathfrak{N}(F+H)/kT\}=\sum \exp\{(-2H[\downarrow]-6\epsilon[\downarrow]+2\epsilon[\downarrow\downarrow])/kT\}.$$

Upon comparing this with Eq. (10), one concludes that for a value of the fugacity given by

$$y=\exp\{(-2H-6\epsilon)/kT\}, \tag{11}$$

the pressure p of the lattice gas is related to the free energy F per spin of the Ising lattice by

$$p=-F-H. \tag{12}$$

The same treatment can be easily applied to *any general lattice with arbitrary interactions, yielding results identical with Eqs.* (8) *and* (12). These results are compiled in Table I which lists the corresponding quantities for the problems of an Ising model with a magnetic field and of a lattice gas. The two problems are completely equivalent; the thermodynamical properties of one system can be derived with the aid of Table I from those of the other and vice versa. In particular, the isotherms in the $I-H$ and $p-v$ diagrams bear a very close relationsip to each other. We mention specifically that a discontinuity in I corresponds to a discontinuity in v. In the case of ferromagnetism, below the Curie temperature T_C, I has a discontinuity, and this corresponds to a phase transition at the same temperature in the lattice gas. Above T_C the isotherms in

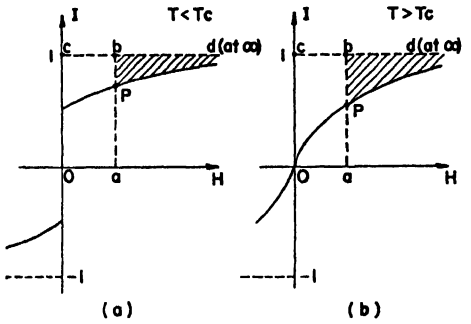


FIG. 2. Geometrical construction of the thermodynamic functions of the lattice gas from the $I-H$ diagram of the corresponding Ising model. Solid curves are isotherms. Given a point P the corresponding lattice gas had a density $\rho=\frac{1}{2}$ (length bP), a pressure $p=$area Pdb. p can be obtained from the free energy of the Ising model which is $F=-$area $OaPdc$.

both of the two diagrams become smooth, and we therefore identify the Curie temperature with the critical temperature of the lattice gas.

From the $p-v$ diagram one can construct the thermodynamic properties of the Ising model in a simple geometric manner and vice versa. These are illustrated in Figs. 1 and 2.

III. AN EXAMPLE OF A TWO-DIMENSIONAL LATTICE GAS

The question naturally arises as to the relationship between a lattice gas and a real gas in which the atoms are not confined to move on lattice points. If one replaces the configurational integral in the partition function of the real gas by a summation over lattice sites, one would obtain the partition function of the lattice gas. Theoretically speaking, by making the lattice constant smaller and smaller one could obtain successively better approximations to the partition function of the real gas.[3] In practice this is a very difficult procedure. However, referring back to the simple example discussed in detail in Sec. II one sees that if $\epsilon>0$ the interaction u there has all the characteristics of the interaction between gas atoms that are usually considered responsible for the phenomenon of condensation, namely, an attractive force with a finite range outside of a strongly repulsive core. Thus one would expect that the main features of the phenomenon of condensation should be revealed even in this simple example.

It happens fortunately that in two dimensions this example can be solved exactly in the transition region. To be more specific, the problem is that of a two-dimensional lattice gas with the interaction u specified in (9). (A square lattice is considered.) ϵ is assumed to be positive so that the gas atoms attract each other. The corresponding Ising model problem has been studied extensively. It is known that the model exhibits ferromagnetism so that according to the discussion of Sec. II the lattice gas has a phase transition. In the Ising model, for temperatures lower than the Curie temperature T_C which is given by

$$\exp(-\epsilon/kT_C)=\sqrt{2}-1, \tag{13}$$

the intensity of magnetization has a discontinuity at $H=0$. The corresponding phase transition of the lattice gas occurs at a value of the fugacity[4]

$$y=x^4,$$

where

$$x=\exp(-\epsilon/kT).$$

Since the free energy[5] and the spontaneous magnetization[6] of the Ising model problem are known exactly at zero magnetic field, the transition region in the $p-v$ diagram of the lattice gas can be mapped out completely with the aid of Table I. We list here the formulas used for the vapor pressure p and specific volumes v_g and v_l of the equilibrium gas and liquid phases:

$$\frac{p}{kT}=\log(1+x^2)+\frac{1}{2\pi}\int_0^{\pi}\times\log\{\tfrac{1}{2}[1+(1-k_1^2\sin^2\varphi)^{\frac{1}{2}}]\}d\varphi, \tag{14}$$

$$v_g^{-1}=\tfrac{1}{2}-\tfrac{1}{2}[(1+x^2)(1-6x^2+x^4)^{\frac{1}{2}}/(1-x^2)^2]^{\frac{1}{2}}, \tag{15}$$

and

$$v_g^{-1}+v_l^{-1}=1, \tag{16}$$

where

$$k_1=4x(1-x^2)(1+x^2)^{-2}. \tag{17}$$

In Fig. 3 the transition region in the $p-v$ diagram is plotted. The isotherms are calculated in the following way:

(1) For small values of the fugacity y, one uses Mayer's series expansion[7] in powers of y:

$$\frac{p}{kT}=y+y^2\left(\frac{2}{x^2}-\frac{5}{2}\right)+y^3\left(\frac{6}{x^4}-\frac{16}{x^2}+\frac{31}{3}\right)+y^4\left(\frac{1}{x^8}+\cdots\right)+\cdots, \tag{18}$$

[3] See Appendix I for a simple illustration on this point.

[4] The relation between the quantity $\exp(-2H/kT)$ of the two-dimensional Ising model in a magnetic field and the fugacity y of the corresponding lattice gas can be easily obtained by using similar reasonings as that used in deriving Eq. (11). One obtains

$$y=x^4\exp(-2H/kT). \tag{A}$$

The fourth power in x comes from the fact that each lattice point has four nearest neighbors. The corresponding equation for any general lattice with arbitrary interaction can be written as

$$y=\sigma\exp(-2H/kT), \tag{B}$$

where σ is a constant and is determined by the structure of the lattice and the interaction between the atoms.

[5] L. Onsager, Phys. Rev. 65, 117 (1944); B. Kaufman, Phys. Rev. 76, 1232 (1949).

[6] C. N. Yang, Phys. Rev. 85, 808 (1952).

[7] J. E. Mayer, J. Chem. Phys. 5, 67 (1937); J. E. Mayer and Ph. G. Ackermann, J. Chem. Phys. 5, 74 (1937); J. E. Mayer and S. F. Harrison, J. Chem. Phys. 6, 87, 101 (1938).

and

$$\frac{1}{v}=y\frac{\partial}{\partial y}\left(\frac{p}{kT}\right). \tag{19}$$

We shall see in the next section that (18) is convergent for all values of y less than x^4.

(2) When y is equal to x^4, the gas undergoes a phase transition at temperatures lower than the critical temperature. The vapor pressure and specific volumes of the gas and liquid at this value of y have been given before in Eqs. (14), (15), and (16).

(3) At $y=x^4$, but above the critical temperature, the pressure of the gas is still given by Eq. (14), but the specific volume v now is a constant

$$v=2. \tag{20}$$

This is so because the corresponding Ising model problem has now, according to Eq. (A),[4] zero magnetic field and consequently zero intensity of magnetization.

(4) For $y>x^4$, the Ising model problem has, according to Eq. (A),[4] a negative magnetic field. Changing the sign of the magnetic field corresponds to changing the fugacity to a value y' given by, according to Eq. (A),[4]

$$yy'=x^8. \tag{21}$$

Since the free energy F is even in H and the intensity of magnetization I is odd, one could express with the aid of Table I the pressure and density in terms of their values when the fugacity is equal to y'. But y' is less than x^4. On using (18) and Table I, one can therefore expand the pressure and density *in inverse powers of the fugacity* y:

$$\frac{p}{kT}=\log\frac{y}{x^4}+y^{-1}x^8+y^{-2}(2x^{14}-(5/2)x^{16})$$
$$+y^{-3}(6x^{20}-16x^{22}+(31/3)x^{24})$$
$$+y^{-4}(x^{24}+\cdots)+\cdots, \tag{22}$$

and

$$\frac{1}{v}=y\frac{\partial}{\partial y}\left(\frac{p}{kT}\right).$$

This series is convergent for all $y>x^4$, which holds everywhere in the liquid phase and also in part of the gas phase above the critical temperature.

It is really quite remarkable that a model with so many properties characteristic of a real gas should allow of a complete and exact solution in the transition region where the usual virial expansion does not apply. A complete solution outside of the transition region is related to the solution of an Ising model in a non-vanishing magnetic field, which still remains unknown. However, the problem can be reduced by the results of paper I, to that of the distribution of roots of the grand partition function; and although the complete distribution is not known, many of its properties have been obtained. We shall return to this problem in Sec. V.

Before closing this section we shall make the following remark. Equation (16) shows that the sum of the densities of the vapor and liquid in equilibrium is a constant independent of the temperature.[8] This closely resembles the behavior of a real gas where a law called "law of rectilinear diameter"[9] is known to hold. It states that the sum of the densities of the vapor and liquid in equilibrium increases linearly with decreasing temperature. This increment is, however, very slow, being not more than 10 percent for a temperature variation over a factor 2 for He near its critical temperature. Now our model obeys what may be called the "law of constant diameter," which provides no increment of the sum of densities with decreasing temperature and may be considered a first approximation to the law of rectilinear diameter. The difference lies, we believe, in the inadequacy of the lattice model (with a finite lattice constant) as an approximation to a true gas. One can formulate arguments which indicate that the correction to the lattice model is in the right direction (i.e., it tends to make the sum of the densities of the vapor and liquid in equilibrium increase with decreasing temperature).

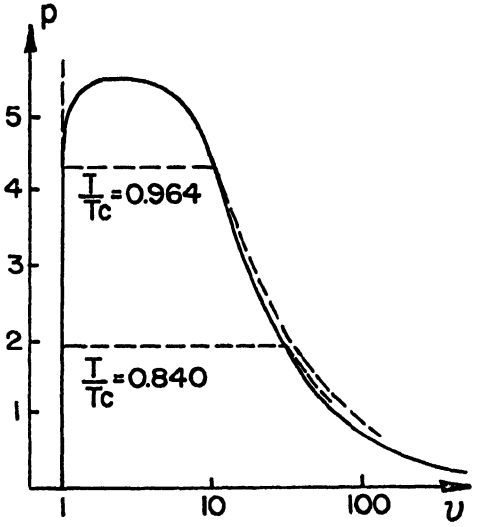


FIG. 3. $p-v$ diagrams for a two-dimensional lattice gas. The solid curve is the exact boundary of the two-phase region. The dotted curves are the isotherms.

[8] The constancy of the sum of the densities of the vapor and liquid in equilibrium is true even for any general lattice. This can be seen very easily since there exists a one to one correspondence between the configurations of the vapor phase and that of the liquid phase, which is obtained simply by replacing the lattice sites occupied by the atoms in the former by vacant sites in the latter and vice versa. Compare Cernuschi and Eyring's paper quoted in footnote 2.

[9] Mathias, Onnes, and Crommelin, Proc. Sect. Sci. Amsterdam 15, 960 (1913).

IV. A THEOREM ON THE DISTRIBUTION OF ROOTS OF THE GRAND PARTITION FUNCTION

We shall now return to the general problem of the condensation of gases, and shall in the following apply the results of paper I to the problem of a lattice gas. There is actually no loss of generality in confining our attention to a lattice gas, as a real continuum gas can be considered as the limit of a lattice gas as the lattice constant becomes infinitesimally small.

The equivalence proved in Sec. II states that the problem of a lattice gas is identical with that of an Ising model in a magnetic field, and that the grand partition function in the former problem is proportional to the partition function in the latter problem. It is convenient to introduce in the Ising model problem the variable

$$z=\exp(-2H/kT), \tag{23}$$

which is by Eq. (B),[4] proportional to the fugacity y of the lattice gas:

$$y=\sigma z, \tag{24}$$

where σ is a constant. In terms of z the partition function $\exp(-\mathfrak{N}F/kT)$ of the Ising lattice is equal to $\exp(\mathfrak{N}H/kT)$ times a polynomial $\mathcal{P}$ in z of degree $\mathfrak{N}$:

$$\exp(-\mathfrak{N}F/kT)=\mathcal{P}\exp(\mathfrak{N}H/kT), \tag{25}$$

where

$$\mathcal{P}=\sum P_n z^n, \quad (n=0, 1, \cdots \mathfrak{N}). \tag{26}$$

The coefficients P_n are the contribution to the partition function of the Ising lattice in zero external field from configurations with the number of $\downarrow$ spins equal to n. It should be noticed that

$$P_n=P_{n'} \text{ if } n+n'=\mathfrak{N}.$$

The P_n's are, of course, real and positive.

Evidently the roots of the polynomial $\mathcal{P}$ are never on the positive real z axis, and are in general complex.

The results of paper I show that if at a given temperature as $\mathfrak{N}$ approaches infinity, the roots of the polynomial $\mathcal{P}$ do not close in onto the positive real axis in the complex z plane, the free energy F is an analytic function of the positive real variable z. Physically this means that the Ising model has a smooth isotherm in the $I-H$ diagram and that the corresponding lattice gas undergoes no phase transition at the given temperature. If, on the other hand, the roots of the polynomial $\mathcal{P}$ do close in onto the positive real z axis at the points $z=l_1, l_2 \cdots$, each of these points would correspond to a discontinuity of the isotherm in the $I-H$ diagram of the Ising lattice and to a phase transition of the lattice gas.

To study the problem of phase transitions of the lattice gas (and of an Ising model), one therefore needs only to study the distribution in the complex z plane of the roots of the polynomial $\mathcal{P}$. The surprising thing is that under quite general conditions this distribution shows a remarkably simple regularity, which may be stated in the form of the following theorem:

Theorem 3. If the interaction u between two gas atoms is such that

$$u=+\infty \text{ if the two atoms occupy the same lattice}$$
$$\text{and } u\equiv 0 \text{ otherwise,} \tag{27}$$

then all the roots of the polynomial $\mathcal{P}$ lie on the unit circle in the complex z-plane.

This theorem will be proved in Appendix II. It should be noticed that in the theorem no assumptions are made about (1) the range of the interaction u, (2) the dimensionality of the lattice, and (3) the size and structure of the lattice. In fact, even the periodicity property of the lattice plays no part at all in the proof.

For the Ising model problem Eq. (27) means that the interaction between all pairs of spins (not limited to pairs of nearest neighbors) are ferromagnetic.

Some immediate consequences of theorem 3 may be enumerated as follows:

(1) The lattice gas cannot undergo more than one phase transition, which must occur, if at all, at a value of the fugacity equal to σ, which according to Eq. (24), corresponds to $z=1$. The isotherms in the $I-H$ diagram of the corresponding Ising model problem is smooth everywhere except possibly at zero magnetic field (which occurs at $z=1$). This is usually believed to be true but was not proved.

(2) Mayer's series expansions[7] of the pressure p and density $1/v$ in powers of the fugacity y [see, e.g., Eq. (18)] are convergent for all values of y less than σ. This is easily proved by the argument used in footnote 9 of paper I. On using the symmetry property of the Ising model problem with respect to a reversal of sign of the magnetic field, one obtains, for values of y greater than σ, convergent series expansions of p and $1/v$ in inverse powers of the fugacity [see, e.g., Eq. (22)].

V. DISTRIBUTION OF ROOTS ON THE UNIT CIRCLE

A. Distribution Function $g(\theta)$

We have seen in the last section that for the interaction (27) the roots of $\mathcal{P}$ lie on the unit circle. Its distribution as $\mathfrak{N}\to\infty$ may be described by a density function $g(\theta)$[10] so that $\mathfrak{N}g(\theta)d\theta$ is the number of roots with z between $e^{i\theta}$ and $e^{i(\theta+d\theta)}$. Evidently one has

$$g(\theta)=g(-\theta), \tag{28}$$

[10] The average density of a finite lattice gas is easily seen to be

$$\Sigma_k z/[z-\exp(i\theta_k)],$$

where $z=\exp(i\theta_k)$ are the zeros of the grand partition function. The results of paper I show that this average density converges to an analytic function in z both inside and outside of the unit circle as the size of the lattice approaches infinity. It seems intuitively clear from this that the distribution of these roots should also approach a limiting distribution on the unit circle for an infinite lattice. This is indeed the case and a rigorous mathematical proof exists in the literature. See A. Wintner, Monatsh. Math. Phys. 4, 1 (1934). We are indebted to Professor Kac for showing us the proof.

and

$$\int_0^\pi g(\theta)d\theta=\tfrac{1}{2}. \tag{29}$$

Taking the logarithm of (25) and factorizing $\mathcal{P}$ one obtains for the Ising model

$$\begin{aligned}\frac{-F}{kT}&=\frac{H}{kT}+\int_0^{2\pi} g(\theta)\log(z-e^{i\theta})d\theta\\&=\frac{H}{kT}+\int_0^{\pi} g(\theta)\log(z^2-2z\cos\theta+1)d\theta.\end{aligned} \tag{30}$$

The intensity of magnetization is obtained from Eqs. (6) and (30):

$$I=1-4z\int_0^\pi g(\theta)\frac{z-\cos\theta}{z^2-2z\cos\theta+1}d\theta. \tag{31}$$

For the lattice gas one has, on using Table I and Eqs. (30) and (31),

$$\frac{p}{kT}=\int_0^\pi g(\theta)\log(z^2-2z\cos\theta+1)d\theta, \tag{32}$$

and

$$\frac{1}{v}=2z\int_0^\pi g(\theta)\frac{z-\cos\theta}{z^2-2z\cos\theta+1}d\theta. \tag{33}$$

These equations enable one to calculate the isotherms in the $I-H$ and $p-v$ diagrams from the distribution function $g(\theta)$. The isotherms thus obtained extend to the condensed region, the two-phase region as well as the gas region. They approximate very realistically the isotherms of a real gas even for very simple distribution functions $g(\theta)$.

B. An Electrostatic Analog

A very simple analog of Eqs. (32) and (33) may be found in electrostatics in the following way: Consider a circular cylinder of unit radius perpendicular to the complex z plane discussed above, cutting it at the unit circle. Assume the cylinder to be charged with a surface charge density dependent only on the angle θ and equal to $g(\theta)$ per unit area. Denote the electrostatic potential and the field produced by this charge distribution at any point on the real axis in the z plane by ϕ and E. Since $g(\theta)=g(-\theta)$, the electric field E is evidently parallel to the real z axis. One can easily verify the following equations:

$$-2p/kT=\text{electrostatic potential } \phi, \tag{34}$$

$$2/vz=-\partial\phi/\partial z=\text{electric field } E. \tag{35}$$

If $g(\theta)$ vanishes in the neighborhood of $\theta=0$, the sheet of charge has the shape "C" and the electric potential and electric field are well behaved for all real positive values of z. Consequently the pressure and density of the lattice gas are analytic in the fugacity and the system undergoes no phase transition. If, on the other hand, $g(0)$ does not vanish, the electric field has a discontinuity at $z=1$ as one goes along the positive real z axis through the sheet of surface charge distribution. Evidently

$$E(z=1+)-E(z=1-)=4\pi g(0). \tag{36}$$

This means of course that the specific volume v of the lattice gas has a discontinuity at $z=1$, showing that the system undergoes a phase transition. The specific volumes v_g and v_l of the gas and liquid phases in equilibrium are related to $g(0)$ by

$$1/v_l-1/v_g=2\pi g(0). \tag{37}$$

Another relation between v_l and v_g is given by the "law of constant diameter" discussed in Sec. 3 (see especially footnote 8).

Equation (37) asserts that the value of the distribution function g at the transition point ($\theta=0$) is equal to $(2\pi)^{-1}$ times the difference of the densities of the liquid and gas phases in equilibrium, which is a directly physically observable quantity.

It should be emphasized that this relationship has actually a much wider range of validity and holds even when the roots are not necessarily on the unit circle.

It is evident that the variation of the pressure and density near the point $z=1$ depends very sensitively on the behavior of $g(\theta)$ near $\theta=0$. One would for example expect that the derivatives of $g(\theta)$ near $\theta=0$ should determine the derivatives of the isotherm in the $p-v$ diagram near the region of condensation. This is indeed the case, as can be seen in the example: If for small values of θ,

$$g(\theta)=g(0)+a|\theta|^n+\cdots, \quad (n>0)$$

the $g(0)$ term, (if nonvanishing), will give rise to a flat horizontal portion on the $p-v$ diagram. The next term will give rise to a discontinuity in the nth derivative of v with respect to p for integral values of n. If n is a fraction between the integers, say, m and $m-1$ then the mth derivative of v with respect to p is discontinuous.

For the Ising model problem, the relation corresponding to Eq. (37) can be obtained by using Table I. The intensity of spontaneous magnetization I is thus expressible in terms of $g(0)$:

$$I=2\pi g(0). \tag{38}$$

C. Relationship between the Cluster Integrals and $g(\theta)$

If one expands (32) in powers of z and identifies the coefficient with Mayer's expansion,[7]

$$\frac{p}{kT}=\sum_{l=1}^{\infty} b_l y^l=\sum_{l=1}^{\infty} b_l\sigma^l z^l,$$

where σ is given by (24), one obtains the following expression for the cluster integrals b_l:

$$b_l=\frac{-1}{l\sigma^l}\int_0^{2\pi} g(\theta)\cos l\theta d\theta, \quad l\geqq 1. \tag{39}$$

In other words the b_l's are the Fourier coefficients of the distribution function $g(\theta)$. One has, of course, the inverse relation,

$$g(\theta)=\frac{1}{2\pi}-\frac{1}{\pi}\sum_{l=1}^{\infty} l\sigma^l b_l\cos l\theta, \tag{40}$$

which expresses the distribution function $g(\theta)$ in terms of the cluster integrals b_l, which in turn are themselves calculable from the virial coefficients.

Equations (40) and (37) enable one to make practical approximate calculations of the distribution function $g(\theta)$ from the virial coefficients and the change in specific volume in the phase transition, both of which are experimentally osbervable quantities.

D. $g(\theta)$ and the Analytical Behavior of the Specific Volume

The density $1/v$ of the lattice gas is evidently an analytic function of z both inside and outside of the unit circle in the complex z plane. Its value is equal to 1 at $z=\infty$ where it is also analytic. All its singularities lie on the unit circle. The values of $1/v$ on the two sides of the unit circle are related to the distribution function $g(\theta)$ by

$$\lim_{r\to 1+}\frac{1}{v}\bigg|_{z=re^{i\theta}}-\lim_{r\to 1-}\frac{1}{v}\bigg|_{z=re^{i\theta}}=2\pi g(\theta). \tag{41}$$

This equation can be proved by extending the reasoning that led to Eq. (37) in the following way: Consider the electric field at a point not necessarily on the real z axis produced by the charge distribution discussed in subsection *B*. Let E_x and E_y be the components of the electric field $\parallel$ and $\perp$ to the real z axis. Defining

$$\mathcal{E}=E_x+iE_y, \quad \mathcal{E}^*=E_x-iE_y,$$

one can easily verify the identity

$$2/vz=\mathcal{E}^* \tag{42}$$

at every point in the complex z plane. The discontinuity of $\mathcal{E}$ across the unit circle at any point $z=e^{i\theta}$ is related to the charge density $g(\theta)$ by

$$\mathcal{E}|_{\text{outside}}-\mathcal{E}|_{\text{inside}}=4\pi g(\theta)z.$$

Taking the complex conjugate of this equation and using (42) one obtains Eq. (41).

The corresponding equation for the Ising model problem is

$$\lim_{r\to 1+} I\bigg|_{z=re^{i\theta}}-\lim_{r\to 1-} I\bigg|_{z=re^{i\theta}}=-4\pi g(\theta). \tag{41a}$$

E. Example of One- and Two-Dimensional Ferromagnetic Ising Model

We shall illustrate the above discussion with the problem of a one-dimensional ferromagnetic Ising model with nearest neighbor interaction in an external magnetic field. This problem can be rigorously solved by the matrix method.[11] We quote here only the results. For a closed chain of $\mathfrak{N}$ spins denote the roots of the partition function by $e^{\pm i\theta_1}$, $e^{\pm i\theta_2}$, $\cdots$. These roots are given by

$$\cos\theta_j=-x^2+(1-x^2)\cos[\pi(2j-1)/\mathfrak{N}], \tag{43}$$

where x is defined to be $\exp(-\epsilon/kT)$ and j runs through all integers 1, $2\cdots$ less than or equal to $\frac{1}{2}(\mathfrak{N}+1)$. As $\mathfrak{N}\to\infty$ these roots distribute themselves continuously on an arc of the unit circle lying to the left of the points

$$z=(1-2x^2)\pm i2x(1-x^2)^{\frac{1}{2}}. \tag{44}$$

For all values of $x>0$, the roots therefore do not close in onto the positive real axis, confirming the well-established fact that a one-dimensional Ising model does not exhibit ferromagnetism. The density of roots is given by the distribution function

$$\begin{aligned} g(\theta)&=\frac{1}{2\pi}\frac{\sin^2\frac{1}{2}\theta}{(\sin^2\frac{1}{2}\theta-x^2)^{\frac{1}{2}}}, && \text{for } \cos\theta<1-2x^2,\\ g(\theta)&=0, && \text{for } \cos\theta>1-2x^2. \end{aligned} \tag{45}$$

On substituting this into Eq. (30), one obtains the correct free energy. The intensity of magnetization is given by

$$I=\left[\frac{z^2-2z+1}{z^2-2z(1-2x^2)+1}\right]^{\frac{1}{2}}.$$

This is analytic everywhere except at the two points given in Eq. (44). The cut between these two points should be made along the unit circle toward the left of the two points. One easily verifies that the discontinuity across the cut is exactly $4\pi g(\theta)$, as given by Eq. (41a).

For a two-dimensional ferromagnetic Ising square lattice the problem with a finite magnetic field, to our knowledge, has not been solved. However, the following exact knowledge is available about the problem:

(1) The free energy F at zero magnetic field (i.e., $z=1$) was obtained by Onsager[5] as

$$F(z=1)=\frac{-kT}{2\pi^2}\int_0^{\pi}\int_0^{\pi}\log[1+2x(\cos\omega+\cos\omega')+2x^2-2x^2(\cos\omega+\cos\omega')+x^4]d\omega d\omega'. \tag{46}$$

(2) The intensity of spontaneous magnetization at

[11] H. A. Kramers and G. H. Wannier, Phys. Rev. 60, 252, 263 (1941).

zero magnetic field was obtained by one of us[6] as

$$I(z=1)=\left[\frac{1+x^2}{(1-x^2)^2}(1-6x^2+x^4)^{\frac{1}{2}}\right]^{\frac{1}{4}} \quad \text{for } x\leqq\sqrt{2}-1,$$
$$I(z=1)=0 \quad \text{for } x>\sqrt{2}-1. \tag{47}$$

(3) By an extension of a method due to Kac and Ward[12] for obtaining a combinatorial solution of the free energy of the Ising model problem in the absence of a magnetic field, one could obtain[13] the free energy at $z=-1$ which corresponds to a pure imaginary magnetic field equal to $i\pi/2$:

$$F(z=-1)=-\frac{i\pi}{2}-\frac{kT}{4\pi^2}\int_0^\pi\int_0^\pi \log\{(1-x^2)^2 \times[1+(6-4\cos^2\omega-4\cos^2\omega')x^2+x^4]\}d\omega d\omega'. \tag{48}$$

(4) The intensity of magnetization at this imaginary value[13] of the magnetic field is

$$I(z=-1)=\left[\frac{(1+x^2)^2}{1-x^2}(1+6x^2+x^4)^{-\frac{1}{2}}\right]^{\frac{1}{4}}. \tag{49}$$

This can be verified by series expansion of both sides in powers of x.

The results quoted above have direct bearing on the distribution function $g(\theta)$ of the zeros of the partition function on the unit circle. In particular (47) and (49) are precisely the value of $g(\theta)/2\pi$ at $\theta=0$ and $\theta=\pi$. It is interesting to notice that $g(0)/2\pi$ is always less than unity while $g(\pi)/2\pi$ is always greater than unity. Also $g(0)$ increases with decreasing temperature while $g(\pi)$ decreases with decreasing temperature. This shows the motion of the roots toward the right along the unit circle as the temperature decreases.

On the other hand, (46) and (48) give through Eq. (30) certain averages with respect to the distribution function $g(\theta)$. The form of these averages are extremely suggestive and we have tried to construct the distribution function from them, but without success.

VI. CONCLUDING REMARKS

The relation between the distribution of roots of a polynomial and its coefficients is mathematically a very complicated problem. It is therefore very surprising that the distribution should exhibit such simple regularities as proved in theorem 3 which applies, as remarked before, under very general conditions. One cannot escape the feeling that there is a very simple basis underlying the theorem, with much wider application, which still has to be discovered.

It is a great pleasure to thank Professor M. Kac for many stimulating and very pleasant discussions from which we learned much in mathematics.

[12] M. Kac and J. C. Ward (to be published).

[13] We hope to publish in a future communication the details of the steps that led to (48) and (49).

APPENDIX I. ONE-DIMENSIONAL HARD RODS AND LATTICE GAS

A simple example that illustrates clearly the relationship between a continuum gas and a lattice gas is the problem of one-dimensional hard rods and the corresponding problem of a one-dimensional lattice gas, as the problem can be exactly solved in both cases. To be specific, let us consider a lattice gas of n atoms distributed on a one-dimensional lattice of total length L and lattice constant δ. The total number of lattice sites is clearly

$$\mathfrak{N}=L/\delta. \tag{50}$$

The interaction potential $u(r)$ between the gas atoms is

$$u(r)=+\infty \text{ for } r<m\delta, \quad u(r)=0 \text{ for } r\geqq m\delta, \tag{51}$$

where r is the distance between atoms and m is a positive integer. Evidently, r can only be 0, δ, 2δ, etc.

The evaluation of the spatial partition function Q can be reduced to that of a simple combinatorial problem of distribution $(\mathfrak{N}-mn)$ identical pieces into $(n+1)$ different bags. One easily obtains

$$Q/n!=[\mathfrak{N}-(m-1)n]!/(\mathfrak{N}-mn)!n!. \tag{52}$$

The pressure p of the gas is related to Q through the relation

$$\frac{p}{kT}=\frac{\partial}{\partial L}\log\left(\frac{Q}{n!}\right). \tag{53}$$

On using (50) and differentiating (52) one obtains

$$\frac{p}{kT}=\frac{1}{\delta}\log\left(1+\frac{n\delta}{L-mn\delta}\right). \tag{54}$$

In the limit of an infinitesimal lattice constant δ but a finite value D for $m\delta$, the pressure approaches a limit given by

$$\lim_{\substack{\delta\to 0\\ m\delta=D}}\frac{p}{kT}=\frac{n}{L-nD}, \tag{55}$$

which is precisely the expression for the pressure of a system of n hard rods each of length D, enclosed in a one-dimensional box of total length L. Thus one sees that when the lattice constant δ approaches zero the thermodynamic functions of a lattice gas indeed approach that of a continuum gas.

APPENDIX II. PROOF OF THEOREM 3

Theorem 3 is a special case of the following more general theorem:

Theorem: Let $x_{\alpha\beta}=x_{\beta\alpha}$ $(\alpha\neq\beta,\ \alpha,\ \beta=1,\ 2,\ \cdots n)$ be real numbers *whose absolute values are less than or equal to* 1. Divide the integers $1, 2, \cdots n$ into two groups a and b so that there are γ integers in group a and $(n-\gamma)$ in group b. Consider the product of all $x_{\alpha\beta}$ where α belongs to group a and β to group b. We shall denote by P_γ the sum of all such products over all the

$n!/[\gamma!(n-\gamma)!]$ possible ways of dividing the n integers. In other words

$$P_\gamma=\sum[\gamma!(n-\gamma)!]^{-1}\prod_{j=1}^{n-\gamma}\prod_{i=1}^{\gamma}x_{a_ib_j}, \tag{56}$$

where $a_1\cdots a_\gamma$, $b_1\cdots b_{n-\gamma}$ is any permutation of the integers 1, 2, $\cdots n$ and the summation extends over all such permutations; e.g.,

$$P_0=P_n=1, \quad P_1=P_{n-1}=x_{12}x_{13}\cdots x_{1n}+x_{21}x_{23}\cdots x_{2n}+\cdots+x_{n1}x_{n2}\cdots x_{n(n-1)}.$$

It is easy to verify that $P_\alpha=P_{n-\alpha}$. Consider the polynomial

$$\mathcal{P}(z)=1+P_1z+P_2z^2+\cdots+P_{n-1}z^{n-1}+z^n. \tag{57}$$

The theorem asserts that all the roots of the equation

$$\mathcal{P}=0$$

are on the unit circle.

To prove this theorem it is convenient to introduce the following polynomials $\mathfrak{P}$ of the variables z_1, z_2, $\cdots z_n$:

$$\mathfrak{P}(z_1,\cdots z_n)=\sum\left\{[\gamma!(n-\gamma)!]^{-1}\left(\prod_{j=1}^{n-\gamma}\prod_{i=1}^{\gamma}x_{a_ib_j}\right)\prod_{i=1}^{\gamma}z_{a_i}\right\}, \tag{58}$$

where, as in the definition of P_γ, $a_1, a_2, \cdots a_\gamma, b_1, \cdots b_{n-\gamma}$ is any permutation of the integers 1, 2, $\cdots n$ and the summation extends over all such permutations and over all γ. It is clear that

$$\mathfrak{P}(z, z, z, \cdots z)=\mathcal{P}(z). \tag{58a}$$

To prove the theorem it is only necessary to prove that there are no roots of $\mathcal{P}=0$ that have an absolute value >1. Equation (58a) therefore shows that the theorem is an immediate consequence of the following lemma:

Lemma: If $\mathfrak{P}(z_1, z_2, \cdots z_n)\neq 0$ and none of the z's has an absolute value less than one, then

$$1=|z_1|=|z_2|=\cdots=|z_n|.$$

Proof: (1) We shall assume throughout the proof that all the x's are different from zero and ± 1. The proof can then be easily generalized to include the case when one or more of the x's either vanish or are equal to ± 1.

The lemma is clearly true when $n=1$, for which $\mathfrak{P}_1=1+z_1$. For $n=2$,

$$\mathfrak{P}_2=1+x_{12}(z_1+z_2)+z_1z_2.$$

$\mathfrak{P}_2=0$ therefore implies

$$z_1=-(1+x_{12}z_2)/(z_2+x_{12}).$$

It is easy to prove that for $|z_2|>1$ the right-hand side of this equation always has an absolute magnitude <1. Hence if $|z_1|\geqq 1$, $|z_2|\geqq 1$, they must both be equal to 1, proving the lemma for the case $n=2$.

(2) We shall prove the lemma for general values of n by induction. Assume it is true for $n=m-2$ and $n=m-1$, but not true for $n=m$. We shall call this assumption hypothesis A and prove that it is self-contradictory. Under hypothesis A there exists a set of z's equal to $z_1, z_2, \cdots z_m$ such that

$$\mathfrak{P}_m(z_1, z_2, \cdots z_m)=0 \tag{59a}$$

and

$$|z_1|>1, \quad \text{and} \quad |z_2|, |z_3|\cdots, |z_m| \text{ all } \geqq 1. \tag{59b}$$

The subscript m in $\mathfrak{P}_m$ is to indicate that it has m independent variables z.

(3) We now prove in this paragraph that keeping $z_3, z_4, \cdots z_m$ fixed and regarding z_2 as a function of z_1 defined by (59a) one obtains a limit δ_2 for z_2 as $z_1\to\infty$ and that $|\delta_2|<1$. To prove this we take, say, $m=3$ and notice that

$$\mathfrak{P}_3=1+z_1x_{12}x_{13}+z_2x_{21}x_{23}+z_3x_{31}x_{32}+z_1z_2x_{13}x_{23}+z_2z_3x_{21}x_{31}+z_3z_1x_{32}x_{12}+z_1z_2z_3. \tag{61}$$

As $z_1\to\infty$, $\mathfrak{P}_3=0$ gives

$$x_{12}x_{13}+\delta_2x_{13}x_{23}+z_3x_{32}x_{12}+\delta_2z_3=0, \tag{62}$$

i.e.,

$$\delta_2=-(x_{12}x_{13}+z_3x_{32}x_{12})/(z_3+x_{13}x_{23}). \tag{63}$$

It is easy to prove that unless the denominator vanishes this does give the correct limit for z_2. But the denominator cannot vanish, as will be seen later. Hence the limit δ_2 is correctly given by (62). Now write

$$\delta_2x_{12}^{-1}=\zeta_2, \quad z_3x_{13}^{-1}=\zeta_3. \tag{64}$$

Equation (62) reduces to

$$1+\zeta_2x_{23}+\zeta_3x_{23}+\zeta_2\zeta_3=0.$$

But this is exactly an equation of the form

$$\mathfrak{P}_2(\zeta_2\zeta_3)=0,$$

where $\mathfrak{P}_2$ is a polynomial of the general form (58) with $n=2$. Now by condition (59b)

$$|\zeta_3|>|z_3|\geqq 1.$$

Hence hypothesis A asserts that $|\zeta_2|<1$. Hence

$$|\delta_2|<|\zeta_2|<1.$$

We need now only prove that the denominator in (63) does not vanish. This is evident in the present case of $m=3$. We shall, however, give a formal proof which holds in general for any value of m. The denominator in (63) is clearly the coefficient of z_2z_1 in $\mathfrak{P}_3$ given in (61), just as the left-hand side of (62) is the coefficient of z_1 in $\mathfrak{P}_3$. We therefore make a similar transformation as (64):

$$z_3x_{13}^{-1}x_{23}^{-1}=\zeta_3',$$

and reduce the denominator in (63) into the form

$$x_{13}x_{23}\mathfrak{P}_1(\zeta_3'). \tag{65}$$

Clearly $|\zeta_2'|>1$ so that hypothesis A implies that (65) does not vanish. The above arguments evidently hold for any general value of m and this completes the proof that as $z_1\to\infty$, z_2 approaches a limit smaller than unity in absolute value.

(4) Keeping $z_3, z_4\cdots z_m$ fixed one can increase $|z_1|$ and define z_2 as a continuous function of z_1. Since by (59b), $|z_2|$ starts to be $\geqq 1$ in absolute magnitude and tends to a limit <1 in absolute magnitude as $z_1\to\infty$, there must be a value of z_1 equal to z_1' so that z_2 assumes a value z_2' equal to 1 in absolute magnitude, i.e.,

$$\mathfrak{P}_m(z_1', z_2', z_3, \cdots z_m)=0 \tag{66a}$$

and

$$|z_1'|>1, \quad |z_2'|=1, \quad |z_3|, \quad |z_4|, \quad \cdots |z_m|\geqq 1. \tag{66b}$$

We can fix $z_2', z_4, z_5, \cdots z_m$ and regard z_3 as a function of z_1' and follow the same procedure by increasing $|z_1'|$ till z_3 assumes a value equal to 1 in absolute magnitude. Continuing this way we finally get a set of values z_1'', z_2'', $\cdots z_m''$ such that

$$\mathfrak{P}_m(z_1'', z_2'', \cdots z_m'')=0, \tag{67a}$$

and

$$|z_1''|>1, \quad |z_2''|=|z_3''|=\cdots=|z_m''|=1. \tag{67b}$$

But $\mathfrak{P}_m$ is linear in z_1''. Writing $\mathfrak{P}_m=Bz_1''+C$ where B and C are independent of z_1'' one verifies easily that

$$B=z_2''z_3''\cdots z_m''\mathfrak{C},$$

where $\mathfrak{C}$ is the complex conjugate of C under the condition (67b). Hence

$$|z_1''|=|C|/|B|=1, \tag{68}$$

which contradicts (67b). (It is easy to show that B does not vanish by making a transformation similar to (64) and reduce B to products of some x's with $\mathfrak{P}_{m-1}$.)

This completes the proof by induction.

Comment

The distribution of zeros of the grand partition function for a one-dimensional system of hard rods is discussed by E. Hiis Hauge and P. C. Hemmer, Physica 29, 1338 (1963). They show that in the thermodynamic limit the zeros are located on the negative real z-axis. (A more rigorous proof of this was given by O. Penrose (unpublished).) The distribution of zeros for a gas obeying the van der Waals equation of state is discussed by P. C. Hemmer and E. Hiis Hauge in Phys. Rev. 133, A190 (1964).

Volume 3, number 1 PHYSICS LETTERS 15 November 1962

TWO THEOREMS ON CLASSICAL MANY-PARTICLE SYSTEMS

J. GROENEVELD
Instituut voor Theoretische Fysica, Universiteit van Amsterdam

Received 23 October 1962

The following two general theorems concerning the partition function and the distribution functions of a classical system of identical particles with additive, central forces can rigorously be proved:

Theorem 1. If the two-particle interaction potential $\varphi(i,j)$ $(= \varphi(r_i, r_j))$ satisfies the two conditions

a. ("Hard core") There exists a non-negative constant a (which may be zero) such that $\varphi(i,j) = +\infty$ whenever $|r_i - r_j| \leq a$; and

b. ("Convergence condition") There exists a non-negative constant E such that

$$\sum_{i \geq 2} \varphi(1,i) \geq -E$$

for all sets of coordinates $r_2, r_3, \ldots$ satisfying $|r_k - r_l| > a$ for all k and l $(2 \leq k < l)$, then the following relations hold true:

1. For all volumes V and all (positive) activities z_1 and z_2 the pressure $p(z,V) = (\beta V)^{-1} \ln Z_G(z,V)$ satisfies:

$$p(z_1 + z_2, V) \leq p(z_1 e^{\beta E}, V) + p(z_2, V) . \qquad (1)$$

2. For all V and all (positive) z the k-particle distribution function n_k in the grand canonical ensemble * satisfies:

$$n_k(1,\ldots,k;z,V) \leq z^k W_k(1,\ldots,k) (e^{\beta E})^k , \qquad (2)$$

where

$$W_k(1,\ldots,k) = \exp\{-\beta \sum_{1 \leq i<j \leq k} \varphi(i,j)\} .$$

Proof. From the conditions a and b can be deduced that the probability densities W satisfy:

$$W_{k+l}(1,\ldots,k,k+1,\ldots,k+l) \leq W_k(1,\ldots,k) \times W_l(k+1,\ldots,k+l) (e^{\beta E})^k . \qquad (3)$$

By integration over the coordinates, multiplication with $(k!\,l!)^{-1} z_1^k z_2^k$, summation over k and l and taking the logarithms one obtains the inequality (1). In a similar manner (2) can be proved.

Remark. The inequality (1) can be used for the calculation of a rigorous upper bound for the function $p(z,V)$ for all positive z, starting from either experimental or theoretical values of $p(z,V)$ which are given in a finite interval of the positive z-axis.

A second theorem will now be proved concerning the coefficients $b_l(V)$ in the well known Mayer expansion

$$p(z,V) = \sum_l b_l(V) z^l$$

and the radius of convergence of this series.

Theorem 2. If the two-particle interaction φ satisfies the conditions:

a. $\varphi(i,j)$ depends on $|r_i - r_j|$ only;

b. φ is non-negative $(\varphi \geq 0)$; and

c. the integral

$$\int [\exp\{-\beta\varphi(r)\} - 1] \, d\mathbf{r} = \int f(r) \, d\mathbf{r} = \tilde{f}(0)$$

converges,

then (if $b_l = \lim_{V \to \infty} b_l(V)$):

1. For all V and l $(l = 1,2,\ldots)$

$$0 \leq (-1)^{l-1} b_l(V) \leq (-1)^{l-1} b_l . \qquad (4)$$

2. For all l $(l = 1,2,\ldots)$

$$\frac{1}{l} \leq \frac{b_l}{(2b_2)^{l-1}} \leq \frac{l^{l-2}}{l!} . \qquad (5)$$

3. The radius of convergence R of the series $\Sigma_l b_l z^l$ satisfies the inequalities:

$$\frac{1}{2e|b_2|} \leq R \leq \frac{1}{2|b_2|} . \qquad (6)$$

4. In the region $|z| < R$ the functions $p(z,V)$ and $p(z) = \lim_{V \to \infty} p(z,V)$ are regular, and in this region

$$p(z) = \sum_l b_l z^l . \qquad (7)$$

5. The point $z = -R$ is a singular point of the function $p(z)$.

Proof. Only the general outline of a proof of this theorem will be presented here. As is well known, the expression for the coefficient

$$b_l(V) = (V\,l!)^{-1} \int \ldots \int U_l(1,\ldots,l) \, d\mathbf{r}_1 \ldots d\mathbf{r}_l$$

contains the "Ursell function" $U_l(1,\ldots,l)$, defined as the sum of all products of functions

* Cf. ref. 1), p. 334, eq. (37.3).

$$f(i,j) = \exp\{-\beta\varphi(i,j)\} - 1 ,$$

corresponding to "connected Mayer diagrams" of the points $1,2,\ldots,l$.

In addition we consider the function $T_l(1;2,\ldots,l+1)$ ($l = 1,2,\ldots$) defined here as the sum of all products of functions $f(i,j)$ corresponding to connected Mayer diagrams of the points $1;2,\ldots,l+1$ in which the point 1 is *no* articulation point. The function $T_l(1;2,\ldots,l+1)$ is symmetrical in the coordinates $r_2,\ldots,r_{l+1}$ only. The Ursell functions can then be expressed in the form *:

$$U_2(1,2) = T_1(1;2)\,U_1(1) ,$$

$$U_3(1,2,3) = T_1(1;2)\,U_2(1,3) + T_2(1;2,3)\,U_1(1) ,$$

$$U_l(1,\ldots,l) = \sum_{j=1}^{l-1} \mathop{S}_{(3,\ldots,l)} T_j(1;2,\ldots,j+1) \times U_{l-j}(1;j+2,\ldots,l) , \quad (8)$$

where $U_1(1) \equiv 1$ and where $\mathop{S}_{(3,\ldots,l)}$ represents a summation over all possible, $\binom{l-2}{j-1}$, divisions of the molecules $3,\ldots,l$ into two groups of $(j-1)$ and $(l-j-1)$ molecules respectively.

It can be proved that the T-functions can be written as:

$$T_l(1;2,\ldots,l+1) = f(1;2,\ldots,l+1)\,U_l(2,\ldots,l+1) , \quad (9)$$

where the "generalised f-function"

$$f(1;2,\ldots,l+1) = \exp\{-\beta \sum_{j=2}^{l+1} \varphi(1,j)\} - 1$$

can be written as:

$$f(1;2,\ldots,l+1) = f(1,2) + \sum_{j=3}^{l+1} f(1,j) \times \exp\{-\beta \sum_{k=2}^{j-1} \varphi(1,k)\} . \quad (10)$$

Because $\varphi(r) \geq 0$, all $f \leq 0$. From this and from (8) and (9) it follows by induction with respect to l that

$$(-1)^{l-1}\, U_l(1,\ldots,l) \geq 0 , \quad (11)$$

$$(-1)^{l}\, T_l(1;2,\ldots,l+1) \geq 0 . \quad (12)$$

The inequalities (4) follow immediately from (11). Introducing (for $V = \infty$) the coefficients

$$t_l = (l!)^{-1} \int \ldots \int T_l(1;2,\ldots,l+1)\, dr_2,\ldots,dr_{l+1}$$

it follows from (12) that also:

$$(-1)^{l}\, t_l \geq 0 . \quad (13)$$

Using the fact that $f(i,j) \leq 0$ and $0 \leq \exp\{-\beta\varphi(i,j)\} \leq 1$, one can derive from (10):

$$|f(1,2)| \leq |f(1;2,\ldots,l+1)| \leq \sum_{j=2}^{l+1} |f(1,j)| . \quad (14)$$

Integrating eq. (9), using (11), (12) and (14), one finds:

$$|\tilde{f}(0)| \cdot |b_l| \leq |t_l| \leq l\, |\tilde{f}(0)| \cdot |b_l| . \quad (15)$$

Integrating eq. (8), using (4) and (13) one finds:

$$l(l-1)\, |b_l| = \sum_{j=1}^{l-1} j\, |t_j|\, (l-j)\, |b_{l-j}| . \quad (16)$$

From (15) and (16) one can now deduce by means of a technique of generating functions that:

$$\frac{|\tilde{f}(0)|^{l-1}}{l} \leq |b_l| \leq |\tilde{f}(0)|^{l-1}\, \frac{l^{l-2}}{l!} . \quad (17)$$

From this, using $\tilde{f}(0) = 2b_2 < 0$ and (4), one obtains the desired result (5).

The statement 3 follows from eq. (5) using Cauchy's convergence criterion *. From (4) it follows, using Weierstrass's theorem **, that (for $|z| < R$) the series $\Sigma_l\, b_l(V)\, z^l$ converges uniformly with respect to V, and hence † that the limit $V \to \infty$ and the summation over l can be interchanged, as expressed by (7). Finally the statement 5 follows from the fact that the b_l's alternate in sign ††.

A more detailed account of this work, containing also the corresponding results for the pressure and the distribution functions will be published in Physica.

The author is much indebted to Prof. Dr. J.De Boer for his valuable comments and criticisms during the preparation of this letter.

References

1) T.L.Hill, Statistical mechanics (New York, Toronto, London, 1956).
2) J.De Boer, Repts. Progr. Phys. 12 (1959) 305.
3) E.C.Titchmarsh, The theory of functions, 2nd ed. (London, 1939).

* These relations are analogous to the relations for the W- and U-functions as given by De Boer [2], p. 338, eq. (7.12).

* Cf. ref. 3), par. 7.1, p. 213.
** Cf. ref. 3), par. 1.14, p. 3.
† Cf. ref. 3), par. 1.14, p. 7.
†† Cf. ref. 3), par. 7.21, p. 214.

Comment

Groeneweld (private communication) has recently extended his method to apply also to more general potentials, (see also the next paper by Penrose). He has also obtained, for small z, upper and lower bounds on the two-body Ursell function $\rho^2[g(r_{12}) - 1]$. The asymptotic form of these bounds, which approach zero for $r_{12} \to \infty$, was also found (to be published; see also Ref. 14 of the paper by D. Ruelle).

Reprinted from JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 10 *Printed in U.S.A.* OCTOBER 1963

Convergence of Fugacity Expansions for Fluids and Lattice Gases*

OLIVER PENROSE†
Belfer Graduate School of Science, Yeshiva University, New York, New York

Upper and lower bounds are obtained for $R(V)$, the radius of convergence of the Mayer expansion $V\sum_l b_l(V)z^l$ expressing the logarithm of the classical grand partition function for a finite volume V as a power series in the fugacity z. The particles in V interact only through two-body forces whose potential $\varphi(\mathbf{r})$ satisfies $s^{-1}\sum_{i<j\leq s}\varphi(\mathbf{x}_i - \mathbf{x}_j) \geq \text{const} \equiv -\Phi$ for all s, $\mathbf{x}_1 \cdots \mathbf{x}_s$. The bounds are

$$[e^{l+2\Phi/kT}\int |e^{-\varphi(\mathbf{r})/kT} - 1|\, d^3\mathbf{r}]^{-1} \leq R(V) \leq |e^{\Phi/kT}l/(l-1)\, b_l(V)|^{1/(l-1)}$$

for any $l \geq 2$. For lattice gases the integral becomes a sum. The upper bounds, obtained from the theory of entire functions, include a subsequence converging to $R(V)$ as $l \to \infty$. The lower bound is obtained by using the Kirkwood–Salsburg integral equation to calculate upper bounds on the $b_l(V)$'s and the coefficients in the fugacity expansions of the s-particle distribution functions. For hard-core potentials some of these bounds can be strengthened. For nonnegative potentials, $1/2|b_2(V)|$ is an extra upper bound on $R(V)$. The radius of convergence of the infinite-volume series $\sum b_l z^l$ is shown to be at least $\lim_{V\to\infty} R(V)$, with equality for nonnegative potentials.

1. INTRODUCTION

MANY formal power series expansions are used in the theories of classical fluids[1] and Ising lattices[2] but there are few theorems about their convergence properties. Rigorous results regarding the convergence of these series would be of interest for the theory of phase transitions. For example, let R be the radius of convergence of the Mayer expansion for the thermodynamic pressure p of a fluid in powers of fugacity z:

$$p(z)/kT = \sum_{l=1}^{\infty} b_l z^l; \tag{1.1}$$

and suppose that only a finite number of the b_l's are negative, so that the function $p(z)$ has a singularity at $z = R$. Then if we can show[3] that (1.1) correctly gives the thermodynamic pressure for small z, we can deduce[4] that there is a phase transition either at $z = R$ or at some smaller value of z where (1.1) ceases to give the thermodynamic pressure correctly.

For particles interacting by additive central forces with a nonnegative potential $\varphi(r)$, Groeneveld[5] has obtained upper and lower bounds for R by studying the Mayer graphs of the coefficients b_l. His formulas imply that the same lower bound also holds for $R(V)$, the radius of convergence of the finite-volume analogue of (1.1),

$$\mathcal{V}^{-1}\log \Xi(z, V) = \sum_l b_l(V)z^l, \tag{1.2}$$

where the $b_l(V)$'s are the Mayer cluster integrals for a finite region V, whose volume is $\mathcal{V}$, and $\Xi(z, V)$ is the grand partition function for this region. Moreover, he shows that (1.1) correctly represents the pressure if z is between 0 and the lower bound on R, so that there can be no phase transition for z within this range. Ruelle[6] has generalized Groeneveld's lower bounds to all $\varphi(r)$ with the property

$$\sum_{i<j\leq s}\varphi(|\mathbf{x}_i - \mathbf{x}_j|) \geq -s\Phi \text{ for all } s, \mathbf{x}_1 \cdots \mathbf{x}_s, \tag{1.3}$$

where Φ is a constant, and $\sum_{i<j\leq s}$ means $\sum_{i=1}^{i-1}\sum_{j=i+1}^{s}$ for $s \geq 2$, and 0 for $s = 1$. He used the Kirkwood–Salsburg integral equation[7] and the theory of linear operators in Banach space.

In the present paper we shall obtain a new set of upper bounds on $R(V)$, forming an infinite sequence which converges from above to the limit $R(V)$. We shall also rederive Ruelle's lower bound on $R(V)$ using a method based, like his, on the Kirkwood–Salsburg integral equation, but not

* Work supported by the U. S. Air Force Office of Scientific Research under grant No. 62-64.

† On leave of absence from Imperial College, London, England.

[1] See the review article by G. W. Ford and G. E. Uhlenbeck in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, 1962), Vol. I.

[2] See the review article by C. Domb, Phil. Mag. Suppl. 9, 149 (1960).

[3] A proof is needed since the derivation of (1.1) involves an interchange of the limiting processes $\lim_{V\to\infty}$ and $\sum_{l=1}^{\infty}$. See C. N. Yang and T. D. Lee, Phys. Rev. 87, 404 (1952).

[4] I owe this argument to Dr. J. Lebowitz (private communication).

[5] J. Groeneveld, Phys. Letters 3, 50 (1962).

[6] D. Ruelle, "Correlation Functions of Classical Gases," Institute for Advanced Study, Princeton, New Jersey, 1963 [to be published in Ann. Phys. (N. Y.)].

[7] J. G. Kirkwood and Z. W. Salsburg, Discussions Faraday Soc. 15, 28 (1953); T. L. Hill, *Statistical Mechanics* (McGraw-Hill Book Company, Inc., New York, 1956), p. 251. The equation is a special case of one due to J. E. Mayer [J. Chem Phys. 15, 187 (1947), Eq. (54′)].

requiring the theory of Banach spaces. Our method incidentally yields upper bounds on the $b_l(V)$'s and on the coefficients in the fugacity expansions of the s-particle distribution functions ($s = 1, 2, 3, \cdots$). These results apply to classes of potentials φ a little wider than those treated by Groeneveld and Ruelle, enough so to make the results applicable to lattice systems as well as to fluids. For potentials with hard cores, very similar methods yield additional upper and lower bounds which are sometimes stronger than those given by the more general methods.

2. PROPERTIES OF THE INTERACTION POTENTIAL

The results of this paper apply to systems of interacting particles whose positions $\mathbf{x}_1, \mathbf{x}_2, \cdots$ are either continuously variable or confined to a lattice (Ising model). Their potential energy is assumed to have the form[8] $\sum_{i<j} \varphi(\mathbf{x}_i - \mathbf{x}_j)$, where $\varphi(\mathbf{r})$ is symmetric under space inversion, though not necessarily spherically symmetric, and falls into one or more of the following classes:

I. *nonnegative potentials* satisfying

$$\varphi(\mathbf{r}) \geq 0 \quad \text{for all} \quad \mathbf{r}; \tag{2.1}$$

II. *hard-core potentials* satisfying

$$\varphi(\mathbf{r}) = +\infty \quad \text{if} \quad r < a, \tag{2.2}$$

$$|\varphi(\mathbf{r})| < Ar^{-\nu-\epsilon} \quad \text{if} \quad r \geq a, \tag{2.3}$$

where r is the length of $\mathbf{r}$, and a, A, and ϵ are positive constants, and ν is the number of space dimensions;

III. *potentials with finite binding energy per particle* satisfying

$$\sum_{i<j\leq s} \varphi(\mathbf{x}_i - \mathbf{x}_j) \geq -s\Phi \text{ for all } s, \mathbf{x}_1 \cdots \mathbf{x}_s, \tag{2.4}$$

where Φ is a (finite) constant depending on the form of $\varphi(\mathbf{r})$ only.

The conditions defining categories I and III are very much like the ones used by Groeneveld[5] and by Ruelle[6]. However, we do not require spherical symmetry of $\varphi(\mathbf{r})$, nor do we require absolute convergence of the integral (or sum) for b_2. Thus our results apply to lattice systems as well as to fluids, and also to long-range potentials such as $\varphi(\mathbf{r}) = \text{const}/r$. The condition (2.1) is equivalent to (2.4) with $\Phi = 0$, so that potentials in the class I are also in the class III. The remainder of this section is devoted to the proof of a theorem and corollary showing that the class II is also included in the class III.

Theorem. For any hard-core or nonnegative potential, there exists a constant Φ' such that

$$\sum_{i\leq l} \varphi(\mathbf{x}_i - \mathbf{y}) \geq -2\Phi' \tag{2.5}$$

for all l, $\mathbf{x}_1 \cdots \mathbf{x}_l$, $\mathbf{y}$ satisfying

$$\sum_{i<j\leq l} \varphi(\mathbf{x}_i - \mathbf{x}_j) < \infty. \tag{2.6}$$

Proof. For nonnegative potentials the result is trivial, with $\Phi' = 0$. For hard-core potentials, (2.2) and (2.6) imply that l spheres of radius $\frac{1}{2}a$ centered on the points $\mathbf{x}_1 \cdots \mathbf{x}_l$ do not intersect. An arbitrary sphere of radius $(r + a/2)$ has volume $(1 + 2r/a)^\nu$ times that of one of the l small spheres, and therefore cannot completely enclose more than $(1 + 2r/a)^\nu$ of them. The function $F(r)$, defined as the number of points from the set $\mathbf{x}_1 \cdots \mathbf{x}_l$ which lie inside a sphere of radius r centered on the point $\mathbf{y}$, is therefore bounded above and below by

$$0 \leq F(r) \leq (1 + 2r/a)^\nu. \tag{2.7}$$

Moreover, we may assume that

$$F(r) = 0 \quad \text{if} \quad r \leq a \tag{2.8}$$

since otherwise (2.5) is trivially satisfied, the left side being infinite. The left side of (2.5) with reversed sign has, by (2.3), an upper bound which may be written as a Stieltjes integral and transformed by partial integration:

$$-\sum_i \varphi(\mathbf{x}_i - \mathbf{y}) < A \sum_i |\mathbf{x}_i - \mathbf{y}|^{-\nu-\epsilon}$$

$$= A \int_a^\infty r^{-\nu-\epsilon}\, dF(r) = A[r^{-\nu-\epsilon}F(r)]_a^\infty$$

$$+ (\nu + \epsilon)A \int_a^\infty r^{-\nu-\epsilon-1}F(r)\, dr. \tag{2.9}$$

The integrated part vanishes at $r = a$ by (2.8), and as $r \to \infty$ by (2.7); and the integral converges, again by (2.7). Therefore, (2.5) can be satisfied, for example by taking

$$\Phi' = \tfrac{1}{2}(\nu + \epsilon)A \int_a^\infty r^{-\nu-\epsilon-1}(1 + 2r/a)^\nu\, dr < \infty,$$

Q.E.D.

Corollary. In proving (2.4) we may assume that all terms on the left are finite; otherwise (2.4) is satisfied trivially. Transforming the left side of (2.4) and using (2.5) with some changes of notation, we obtain[9]

[8] The results of Sec. (3), which do not depend on integral equations, are easily extended to systems whose potential energy cannot be written as a sum of pair potentials.

[9] I owe this step to Dr. D. Ruelle.

$$\sum_{i\le j<s} \varphi(\mathbf{x}_i - \mathbf{x}_j) = \tfrac{1}{2} \sum_{i=1}^{s} \sum_{j}{}' \varphi(\mathbf{x}_i - \mathbf{x}_j)$$

$$\geq \tfrac{1}{2} \sum_{i=1}^{s} (-2\Phi') = -s\Phi', \qquad (2.10)$$

where $\sum_j'$ means a sum from 1 to s without the $j = i$ term. Thus (2.4) can be satisfied by taking $\Phi = \Phi'$, and the best value of Φ is at most Φ'. This completes the proof that hard-core potentials (class II) have a finite binding energy per particle (class III).

3. UPPER BOUNDS ON $R(V)$

The grand partition function for a ν-dimensional region V at fugacity z is defined by

$$\Xi(z, V) = 1 + \sum_N (z^N/N!)$$

$$\times \int_V \cdots \int_V e_N(\mathbf{x}_1 \cdots \mathbf{x}_N)\, d^\nu\mathbf{x}_1 \cdots d^\nu\mathbf{x}_N, \qquad (3.1)$$

where the sum is from 1 to ∞ and

$$e_N(\mathbf{x}_1 \cdots \mathbf{x}_N) \equiv \exp\{-\sum_{i<j\le N} \varphi(\mathbf{x}_i - \mathbf{x}_j)/kT\}. \qquad (3.2)$$

We interpret $e_1(\mathbf{x}_1)$ as 1. For a lattice system we interpret the integrations over V in (3.1), and throughout this paper, as sums over all lattice sites contained within V.

For hard-core potentials, an inequality due to Yang and Lee[10] yields upper bounds to $R(V)$. Let M be the largest number of nonoverlapping hard spheres of diameter a whose centers can be fitted into the region V. Then (3.1) reduces to a polynomial of degree M, which may be factorized in the form

$$\Xi(z, V) = \prod_\alpha (1 - z/z_\alpha), \qquad (3.3)$$

where $z_1, z_2 \cdots$ are the zeros of $\Xi(z, V)$.

Expanding the logarithm of (3.3) in powers of z and comparing with (1.2) yields Yang and Lee's formula for the $b_l(V)$'s:

$$\mathcal{V}lb_l(V) = -\sum_\alpha z_\alpha^{-l} \qquad (l = 1, 2, \cdots), \qquad (3.4)$$

where

$$\mathcal{V} \equiv \int_V d^\nu\mathbf{x} \qquad (3.5)$$

is the volume of the region V, or the number of lattice sites within V. Equation (3.4) implies

$$|\mathcal{V}lb_l(V)| \le \sum_\alpha |z_\alpha|^{-l} \le M/R(V)^l \qquad (l = 1, 2, \cdots), \qquad (3.6)$$

since $\Xi(z, V)$ has just M zeros, none of them inside the circle of convergence of (1.2). Rearranging (3.6) gives a set of upper bounds[8] on $R(V)$ for hard-core potentials:

$$R(V) \le |M/\mathcal{V}lb_l(V)|^{1/l} \qquad (l = 1, 2, \cdots). \qquad (3.7)$$

By Cauchy's nth root convergence test,[11] the smallest limit point of the right side of (3.7) is precisely $R(V)$; therefore, the sequence of upper bounds includes a subsequence converging from above to the limit $R(V)$.

Even if the potential does not have a hard core, a similar argument can be used. By virtue of (2.4), the series (3.1) can be majorized by an exponential series, so that for any real or complex value of z,

$$|\Xi(z, V)| \le \exp(|z|\, e^{\Phi/kT}\mathcal{V}). \qquad (3.8)$$

This inequality shows that $\Xi(z, V)$, regarded as a function of the complex variable z, is[12] an entire function (integral function) of order ≤ 1. Therefore it may be factorized using Hadamard's theorem[13,14]

$$\Xi(z, V) = e^{G(z)} \prod_\alpha \{(1 - z/z_\alpha)e^{z/z_\alpha}\}, \qquad (3.9)$$

where $G(z)$ is a polynomial of degree 1. Expanding the logarithm of (3.9) in powers of z and comparing with (1.2) yields a set of formulas differing only slightly from (3.4):

$$G(z) = \mathcal{V}b_1(V)z = \mathcal{V}z, \qquad (3.10)$$

$$\mathcal{V}lb_l(V) = -\sum_\alpha z_\alpha^{-l} \qquad (l = 2, 3, \cdots). \qquad (3.11)$$

The upper bound to $b_l(V)$ obtained from (3.11) may be written as a Stieltjes integral

$$|\mathcal{V}lb_l(V)| \le \sum_\alpha |z_\alpha|^{-l} = \int_0^\infty r^{-l}\, d\mathfrak{N}(r), \qquad (3.12)$$

where $l \ge 2$ and $\mathfrak{N}(r)$ is defined as the number of zeros of $\Xi(z, V)$ in $|z| < r$. In particular, since there are no zeros of $\Xi(z, V)$ inside the circle of convergence of (1.2), we have

$$\mathfrak{N}(r) = 0 \quad \text{if} \quad r \le R(V). \qquad (3.13)$$

---

[10] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952).

[11] E. T. Copson, *Theory of Functions of a Complex Variable*, (Oxford University Press, London, 1935), Sec. 3.33; E. C. Titchmarsh, *The Theory of Functions* (Oxford University Press, London, 1939), Section 7.1.

[12] E. T. Copson, reference 11, Sec. 7.4; E. C. Titchmarsh, reference 11, Sec. 8.2.

[13] E. T. Copson, Reference 11, Sec. 7.6; E. C. Titchmarsh, Reference 11, Sec. 8.24.

[14] If the order is less than 1, Hadamard's factorization is $\Pi_\alpha(1 - z/z_\alpha)$, but since $\sum |z_\alpha|^{-1}$ converges in this case the factorization (3.9) is also valid [E. Borel, *Leçons sur les Fonctions Entières* (Gauthier-Villars, Paris, 1921), p. 32].

D. Ruelle [Helv. Phys. Acta **36**, 183 (1963)] shows that for most potentials the order is actually zero.

Further information about $\mathfrak{N}(r)$ is provided by Jensen's inequality[15] which shows, because of (3.8), that

$$\int_0^r \mathfrak{N}(q)\, dq/q \leq re^{\Phi/kT}\mathfrak{v}. \tag{3.14}$$

This has the corollary

$$\mathfrak{N}(r) = \mathfrak{N}(r)\int_r^{er} dq/q \leq \int_0^{er} \mathfrak{N}(q)\, dq/q \leq ere^{\Phi/kT}\mathfrak{v}. \tag{3.15}$$

Equation (3.13) justifies replacing $\int_0^\infty$ by $\int_{R(V)}^\infty$ in (3.12), and then two partial integrations give, if $l \geq 2$,

$$|\mathfrak{v}lb_l(V)| \leq l\int_{R(V)}^\infty \mathfrak{N}(r)r^{-l-1}\, dr = l^2\int_{R(V)}^\infty r^{-l-1}\, dr \int_{R(V)}^r \mathfrak{N}(q)\, dq/q, \tag{3.16}$$

where the integrated parts vanish by (3.13)-(3.15). The inequalities (3.14) and (3.16) combine to give

$$|\mathfrak{v}lb_l(V)| \leq \mathfrak{v}e^{\Phi/kT}l^2/(l-1)R(V)^{l-1} \qquad (l = 2, 3, \cdots), \tag{3.17}$$

which leads to the upper bounds on $R(V)$,

$$R(V) \leq |e^{\Phi/kT}l/(l-1)b_l(V)|^{1/(l-1)} \qquad (l = 2, 3, \cdots). \tag{3.18}$$

Like the upper bounds in (3.7), these include a subsequence converging from above to the limit $R(V)$. They are valid[8] for all potentials with finite binding energy, including hard-core potentials.

4. RELATION BETWEEN R AND $\lim_{V\to\infty} R(V)$

The inequality (3.18) also furnishes a comparison between R and $\lim_{V\to\infty} R(V)$. Let $V \to \infty$ on both sides of (3.18); this gives[16]

$$\lim_{V\to\infty} R(V) \leq |e^{\Phi/kT}l/(l-1)b_l|^{1/(l-1)} \quad \text{if} \quad l \geq 2, \tag{4.1}$$

since $b_l = \lim_{V\to\infty} b_l(V)$. Now let $l \to \infty$ in (4.1); this gives[11]

$$\lim_{V\to\infty} R(V) \leq R. \tag{4.2}$$

This result can also be deduced from Yang and Lee's results[10] on the uniform convergence of the sequence of functions (1.2).

For nonnegative potentials, Grocneveld[8] has shown that $b_l(V) \leq b_l$, which implies $R(V) \geq R$. Combined with (4.2), this yields the result, due to J. Lebowitz,[4]

$$R = \lim_{V\to\infty} R(V) \quad \text{for nonnegative potentials.} \tag{4.3}$$

5. DISTRIBUTION FUNCTIONS

The s-particle distribution function at fugacity z is defined,[17] for any positive integer s, as

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) \equiv \frac{\sum_{N=s}^\infty \frac{z^N}{(N-s)!}\int_V \cdots \int_V e_N(\mathbf{x}_1 \cdots \mathbf{x}_N)\, d^3\mathbf{x}_{s+1} \cdots d^3\mathbf{x}_N}{\Xi(z, V)}. \tag{5.1}$$

For hard-core potentials the series in the numerator, like the series (3.1) for the denominator, is[10] a polynomial of degree M in z, so that n_s is a rational function of z. Even if the potential has no hard core, the numerator and denominator are still entire functions, so that n_s is still a meromorphic function of z.

The radius of convergence of the series expansion of this meromorphic function in powers of z is at least $R(V)$, since the denominator has no zeros in $|z| < R(V)$. This series expansion may be written

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) = \sum_{l=0}^\infty n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)z^{s+l}. \tag{5.2}$$

The coefficients in the series for $s = 1$ are related to the Mayer coefficients $b_l(V)$; for (5.1), (3.1), and (1.2) imply

$$\int_V n_1(\mathbf{x} \mid z)\, d^3\mathbf{x} = z\frac{d(\log \Xi)}{dz} = \mathfrak{v}\sum lb_l(V)z^l, \tag{5.3}$$

whence

$$\mathfrak{v}lb_l(V) = \int_V n_{1,l-1}(\mathbf{x})\, d^3\mathbf{x} \quad (l = 1, 2, \cdots). \tag{5.4}$$

We shall use this formula to estimate $R(V)$ from information about the coefficient functions $n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)$.

[15] E. T. Copson, reference 11, Sec. 7.42; E. C. Titchmarsh, reference 11, Sec. 8.21.

[16] *Note added in proof*: To allow for the possibility that $\lim_{V\to\infty} R(V)$ may not exist, it would be better to write lim sup for lim in Eqs. (4.1), (4.2), and (9.2).

[17] T. L. Hill, reference 7, Sec. 37. To shorten the formulas a little, the fact that n_s also depends on the form of the region V is not shown explicitly in our notation.

The first term in the series (5.2) is easily obtained by division from (5.1) and (3.1). It is

$$n_{s,0}(\mathbf{x}_1 \cdots \mathbf{x}_s) = e_s(\mathbf{x}_1 \cdots \mathbf{x}_s). \tag{5.5}$$

Recurrence relations for the higher coefficients are conveniently obtained by substituting (5.2) either into the Kirkwood–Salsburg integral equation[7] or the Mayer–Montroll integral equation.[18]

The Kirkwood–Salsburg equation is[7]

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) = z^s\left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \times \sum_{k=0}^{\infty} \frac{1}{k!} \int \cdots \int n_{s-1+k}(\mathbf{x}_2 \cdots \mathbf{x}_{s+k} \mid z) \times \prod_{i=s+1}^{s+k} \{f_1(\mathbf{x}_1; \mathbf{x}_i)\, d^\nu \mathbf{x}_i\}, \tag{5.6}$$

for $s = 1, 2, \cdots$, where

$$f_1(\mathbf{x}; \mathbf{y}) \equiv e_2(\mathbf{x}, \mathbf{y}) - 1, \tag{5.7}$$

and

$$n_0 \equiv 1. \tag{5.8}$$

Substituting (5.2) into (5.6) and equating coefficients of like powers of z gives the recurrence relation

$$n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) = \left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \times \sum_{k=0}^{l} \frac{1}{k!} \int \cdots \int n_{s-1+k,\,l-k}(\mathbf{x}_2 \cdots \mathbf{x}_{s+k}) \times \prod_{i=s+1}^{s+k} \{f_1(\mathbf{x}_1, \mathbf{x}_i)\, d^\nu \mathbf{x}_i\}, \tag{5.9}$$

where

$$n_{0,l} \equiv \delta_{0,l}. \tag{5.10}$$

Equation (5.9) holds for all $s \geq 1$, $l \geq 0$.

The Mayer–Montroll equation is[18]

$$n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z) = z^s e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \times \sum_{k=0}^{\infty} \frac{1}{k!} \int \cdots \int n_k(\mathbf{x}_{s+1} \cdots \mathbf{x}_{s+k} \mid z) \times \prod_{i=s+1}^{s+k} \{f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{x}_i)\, d^\nu \mathbf{x}_i\}, \tag{5.11}$$

for $s = 1, 2, \cdots$, where

$$f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y}) \equiv \left\{\prod_{i=1}^{s} e_2(\mathbf{x}_i, \mathbf{y})\right\} - 1, \tag{5.12}$$

and $n_0 \equiv 1$ as before. Equation (5.11) leads to the recurrence relation

$$n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) = e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \times \sum_{k=1}^{l} \frac{1}{k!} \int \cdots \int n_{k,\,l-k}(\mathbf{x}_{s+1} \cdots \mathbf{x}_{s+k}) \times \prod_{i=s+1}^{s+k} \{f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{x}_i)\, d^\nu \mathbf{x}_i\}, \tag{5.13}$$

which holds for all $s \geq 1$, $l \geq 1$.

[18] J. E. Mayer and E. Montroll, J. Chem. Phys. 9, 2 (1941).

6. LOWER BOUND ON $R(V)$

One way to obtain a lower bound on $R(V)$ is to calculate upper bounds on the coefficient functions $n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ from the recurrence relation (5.9), which depends on the Kirkwood–Salsburg integral equation. These upper bounds may be written

$$|n_{s,m-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq K_{s,m-s} \quad (s = 1, 2 \cdots m). \tag{6.1}$$

In accordance with (5.5), we can satisfy this inequality for $m = 1$ by taking

$$K_{1,0} = 1. \tag{6.2}$$

Values of $K_{s,m-s}$ satisfying (6.1) for $m > 1$ may now be calculated by induction. If (6.1) is true for $m = M - 1$, where $M - 1$ is any positive integer, then (5.9) and (5.10) imply

$$|n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq \left\{\prod_{i=2}^{s} e_2(\mathbf{x}_1, \mathbf{x}_i)\right\} \times \sum_{k=0}^{M-s} \frac{1}{k!} K_{s-1+k,\,M-s-k} B(V)^k, \tag{6.3}$$

for $s = 1, 2, \cdots M$, where

$$B(V) \equiv \operatorname*{Max}_{\mathbf{x}} \int_V |f_1(\mathbf{x}; \mathbf{y})|\, d^\nu \mathbf{y} \tag{6.4}$$

and

$$K_{0,m} \equiv \delta_{0,m}. \tag{6.5}$$

For nonnegative potentials, the product in braces in (6.3) has the upper bound 1 and may therefore be omitted. Thus, if (6.1) holds for $m = M - 1$, it also holds for $m = M$ provided

$$K_{s,M-s} \geq \sum_{k=0}^{M-s} \frac{1}{k!} K_{s-1+k,\,M-s-k} B(V)^k \quad (s = 1, 2, \cdots M). \tag{6.6}$$

It follows by induction that (6.1) holds for all $m \geq 1$, if $K_{s,l}$ satisfies (6.5) for $s = 0$, (6.2) for $s = 1$, $l = 0$, and (6.6) for all other nonnegative values of s and l. These conditions can be satisfied by taking $K_{s,l} = s(s + l)^{l-1} B(V)^l / l!$; in fact, this is their "best possible" solution, which makes both sides of (6.6) equal. Substituting this formula

for $K_{s,l}$ into (6.1) we obtain for nonnegative potentials

$$|\hat{n}_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s+l)^{l-1}B(V)^l/l! \qquad (6.7)$$

for all $s \geq 1$, $l \geq 0$.

This result may be generalized to other potentials with finite binding energy per particle, using a resource due to Ruelle.[6] Since $n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ is symmetric in $\mathbf{x}_1 \cdots \mathbf{x}_s$, it satisfies not only (6.3) but $s-1$ further inequalities obtained by cyclic permutations of the indices:

$$|n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq \{\prod_i{}' e_2(\mathbf{x}_J, \mathbf{x}_i)\} \times \sum_k K_{s-1+k,M-s-k}B(V)^k/k! \qquad (6.8)$$

for $J = 1, 2 \cdots s$, where $\prod_i'$ means a product over $i = 1, 2, \cdots J-1, J+1, \cdots s$. Therefore $n_{s,M-s}(\mathbf{x}_1 \cdots \mathbf{x}_s)$ is also less than the geometric mean of the s different right-hand sides of (6.8). By (3.2), this geometric mean is a similar expression with $\exp[(-2/skT)\sum_{i<j\leq s}\varphi(\mathbf{x}_i - \mathbf{x}_j)]$ replacing the product in the braces. By (2.4), this exponential is at most $e^{2\Phi/kT}$; therefore (6.1) is satisfied for $m = M$ provided

$$K_{s,M-s} \geq e^{2\Phi/kT}\sum_{k=0}^{M-s}\frac{1}{k!}K_{s-1+k,M-s\ k}B(V)^k \qquad (s = 1, \cdots M). \qquad (6.9)$$

It follows, by a discussion parallel to that of (6.6), that

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq K_{s,l} = e^{2(s+l-1)\Phi/kT}s(s+l)^{l-1}B(V)^l/l! \qquad (6.10)$$

for all $s \geq 1$ and $l \geq 0$.

A stronger upper bound on $n_{1,l}(\mathbf{x}_1)$ can be obtained by setting $s = 1$ in (6.3) and substituting on the right from (6.10). Since the product over i in (6.3) now degenerates to 1, we obtain

$$|n_{1,l}(\mathbf{x}_1)| \leq e^{2(l-1)\Phi/kT}(1+l)^{l-1}B(V)^l/l! \qquad (6.11)$$

for all $l \geq 1$. Substituting this into (5.4) gives

$$|lb_l(V)| \leq e^{2(l-2)\Phi/kT}[lB(V)]^{l-1}/l! \qquad (6.12)$$

for all $l \geq 2$. The radius of convergence of the series (1.2) is therefore

$$R(V) = \liminf_{l\to\infty}|b_l(V)|^{-1/l} \geq 1/B(V)e^{1+2\Phi/kT} \qquad (6.13)$$

If $\Phi = 0$ and $V \to \infty$, (6.12) reduces to one of Groeneveld's results;[5] if $V \to \infty$, (6.13) reduces to one of Ruelle's.[6]

7. UPPER BOUNDS ON $n_{s,l}(x_1 \cdots x_s)$

For nonnegative and hard-core potentials the calculation in the previous section can be supplemented by a similar calculation based on the recurrence relation (5.13) which depends on the Mayer–Montroll equation.[18] The resulting inequalities have the advantage of indicating the dependence of $n_{s,l}$ on $\mathbf{x}_1 \cdots \mathbf{x}_s$. They also lead to a new lower bound on $R(V)$.

The form of Eq. (5.13) suggests looking for bounds of the form

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)M_{s,l} \qquad (s = 1, 2, \cdots). \qquad (7.1)$$

Eq. (5.5) shows that we may take

$$M_{s,0} = 1 \qquad (s = 1, 2, \cdots). \qquad (7.2)$$

As an inductive hypothesis, suppose that (7.1) holds for $l = 0, 1, \cdots L-1$ where L is a positive integer. Then (5.13), (3.2), and (2.4) imply

$$|n_{s,L}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) \times \sum_{k=1}^{L}\frac{1}{k!}M_{k,L-k}e^{k\Phi/kT} \times \left\{\int |f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})|\, d^r\mathbf{y}\right\}^k \qquad (7.3)$$

for all $s \geq 1$.

For nonnegative potentials, Φ vanishes and $|f_s|$ has the upper bound used by Groeneveld[5],

$$|f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})| \leq \sum_{i=1}^{s}|f_1(\mathbf{x}_i; \mathbf{y})|; \qquad (7.4)$$

therefore if (7.1) holds for $l = 0, 1, \cdots L-1$ it also holds for $l = L$, provided

$$M_{s,L} \geq \sum_{k=1}^{L}M_{k,L\ k}[sB(V)]^k/k! \qquad (7.5)$$

for all $s \geq 1$. It follows by induction that (7.1) holds for all $l \geq 0$ provided the numbers $M_{s,l}$ satisfy (7.2) for $l = 0$ and (7.5) for all $l \geq 1$. The "best possible" solution of these conditions is $M_{s,l} = s(s+l)^{l-1}B(V)^l/l!$. Substituted into (7.1), this gives (nonnegative potentials only)

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s+l)^{l-1}B(V)^l e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)/l! \qquad (7.6)$$

for all $s \geq 1$, $l \geq 0$, which is at least as strong as the corresponding inequality (6.7) obtained from the Kirkwood–Salsburg equation.

For hard-core potentials, a generalization of (7.4) is required. In obtaining this we may assume that

$e_s(\mathbf{x}_1 \cdots \mathbf{x}_s) > 0$, since if the factor e_s vanishes in (7.3) the value of the other factor is immaterial. Consequently we may write, using (5.12) and (2.5),

$$1 + f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y}) = \prod_{i=1}^{s} (1 + g_i) \leq e^{2\Phi'/kT}, \quad (7.7)$$

where

$$g_i \equiv \exp\left[-\varphi(\mathbf{x}_i - \mathbf{y})/kT\right] - 1. \quad (7.8)$$

Using the inequality in (7.7), and the fact that $g_i \geq -1$, it can be proved by induction on s that

$$1 - \sum_{i=1}^{s} g_{i-} \leq \prod_{i=1}^{s} (1 + g_i) \leq 1 + e^{2\Phi'/kT} \sum_{i=1}^{s} g_{i+}/(1 + g_{i+}), \quad (7.9)$$

where

$$g_{i\pm} \equiv \text{Max}\,(0, \pm g_i) \geq 0. \quad (7.10)$$

Combining the equality in (7.7) with (7.9) gives the required generalization of (7.4):

$$|f_s(\mathbf{x}_1 \cdots \mathbf{x}_s; \mathbf{y})| \leq \sum_{i=1}^{s} \text{Max}\,\{g_{i-}, e^{2\Phi'/kT} g_{i+}/(1 + g_{i+})\}. \quad (7.11)$$

Using (7.11) in (7.3) we find that if (7.1) holds for $l = 0, 1, \cdots L-1$, it also holds for $l = L$ provided

$$M_{s,L} \geq \sum_{k=1}^{L} M_{k,L-k} e^{k\Phi/kT} [sB'(V)]^k/k! \quad (7.12)$$

for all $s \geq 1$, where

$$B'(V) \equiv \text{Max}_{\mathbf{x}} \int_V \text{Max}\,\{(1 - e^{-\varphi(\mathbf{x}-\mathbf{y})/kT}), \times e^{2\Phi'/kT}(1 - e^{\varphi(\mathbf{x}-\mathbf{y})/kT})\}\, d^3\mathbf{y}. \quad (7.13)$$

It follows, by a discussion parallel to that following (7.5), that for hard-core potentials

$$|n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s)| \leq s(s + l)^{l-1} \times [B'(V)e^{\Phi/kT}]^l e_s(\mathbf{x}_1 \cdots \mathbf{x}_s)/l! \quad (7.14)$$

for all $s \geq 1$, $l \geq 0$. By (7.6), this inequality also holds for nonnegative potentials. Substituted into (5.4) it gives

$$|lb_l(V)| \leq [lB'(V)e^{\Phi/kT}]^{l-1}/l!, \quad (7.15)$$

and the new radius of convergence estimate

$$R(V) \geq 1/B'(V)e^{1+\Phi/kT}. \quad (7.16)$$

This inequality is stronger than (6.13) for some potentials and temperatures, but weaker for others; the lattice gas example considered in Sec. 9 illustrates both possibilities.

8. SPECIAL UPPER BOUNDS ON R AND $R(V)$

For nonnegative potentials, Groeneveld[8] showed that R has the upper bound $1/2|b_2|$. By virtue of (4.3), this is also an upper bound on $\lim_{V\to\infty} R(V)$. This bound is stronger than the first upper bound $2/|b_2|$ deducible from (3.18), presumably because the result (3.18) is not restricted to two-body interactions.[8] The method of integral equations provides an alternative proof of Groeneveld's upper bound on R and also yields the corresponding result for finite volumes.

For nonnegative potentials, the definition (5.12) implies $f_s(\mathbf{x}_1, \cdots \mathbf{x}_s; \mathbf{y}) \leq 0$; it follows inductively from (5.5) and (5.13) that

$$(-)^l n_{s,l}(\mathbf{x}_1 \cdots \mathbf{x}_s) \geq 0 \quad (8.1)$$

for all $s \geq 1$, $l \geq 0$, and that all terms in the sum (5.13) have the same sign. Consequently the $k = 1$ term in (5.13) gives a lower bound to $|n_{s,l}|$ if $l \geq 1$. For $s = 1$, this bound is

$$|n_{1,l}(\mathbf{x})| \geq \int_V |f_1(\mathbf{x}; \mathbf{y})|\, |n_{1,l-1}(\mathbf{y})|\, d^3\mathbf{y} \quad (8.2)$$

for $l \geq 1$, with equality when $l = 1$. Since $n_{1,0}(\mathbf{x}) = 1$ by (5.5), it follows that

$$|n_{1,l}(\mathbf{x})| \geq \int_V K^{(l)}(\mathbf{x}, \mathbf{y})\, d^3\mathbf{y}, \quad (8.3)$$

where $K^{(l)}(\mathbf{x}, \mathbf{y})$ is the lth iterate[19] of the symmetric kernel $K^{(1)}(\mathbf{x}, \mathbf{y}) \equiv |f_1(\mathbf{x}; \mathbf{y})|$. Expanding[18] $K^{(l)}(\mathbf{x},\mathbf{y})$ in (8.3) in terms of its (real) eigenvalues $\lambda_0, \lambda_1, \cdots$ and eigenfunctions $\psi_0(\mathbf{x})$, $\psi_1(\mathbf{x})$, $\cdots$, and substituting into (5.4), we obtain

$$|\mathcal{V}(l+1)b_{l+1}(V)| \geq \int_V \int_V K^{(l)}(\mathbf{x}, \mathbf{y})\, d^3\mathbf{x}\, d^3\mathbf{y} = \sum_p \left\{\int \psi_p(\mathbf{x})\, d^3\mathbf{x}\right\}^2 / \lambda_p^l. \quad (8.4)$$

Therefore $R(V)$ has the lower bound

$$R(V) = \liminf_{l\to\infty} |b_l(V)|^{-1/l} \leq |\lambda_0|, \quad (8.5)$$

where λ_0 is the eigenvalue of $K(\mathbf{x}, \mathbf{y})$ with smallest absolute value. Since (8.4) is an equality for $l = 1$, it also implies

$$2\mathcal{V}\,|b_2(V)| \leq \sum_p \left\{\int \psi_p(\mathbf{x})\, d^3\mathbf{x}\right\}^2 / |\lambda_0| \leq \mathcal{V}/|\lambda_0|, \quad (8.6)$$

[19] R. Courant and D. Hilbert, *Methods of Mathematical Physics* (Interscience Publishers, Inc., New York, 1953) Chap. 3, Sec. 5.

by virtue of Bessel's inequality.[20] Combining (8.5) with (8.6) gives the main result of this section,

$$R(V) \leq 1/2|b_2(V)|. \tag{8.7}$$

Taking the limit $V \to \infty$ and using (4.3), we recover Groeneveld's result,[5]

$$R \leq 1/2|b_2|, \tag{8.8}$$

which can also be proved more simply by letting $V \to \infty$ at the beginning of the calculation [Eq. (8.2)] instead of the end.

Unfortunately, it seems to be impossible to extend this result to more general potentials.

9. CONCLUSIONS

Our basic inequalities (6.13), (3.18), and (4.2) for the radii of convergence R and $R(V)$ of the series (1.1) and (1.2) may be summarized in the form[16]

$$1/B(V)e^{1+2\Phi/kT} \leq R(V)$$

$$\leq |e^{\Phi/kT}l/(l-1)b_l(V)|^{1/(l-1)} \quad (l = 2, 3, \cdots), \tag{9.1}$$

$$\lim_{V\to\infty} R(V) \leq R. \tag{9.2}$$

These are valid for any pair potential[8] with finite binding energy per particle. Hard-core potentials also satisfy (7.16) and (3.7),

$$1/B'(V)e^{1+\Phi/kT} \leq R(V) \leq |M/Vlb_l(V)|^{1/l} \quad (l = 1, 2, \cdots), \tag{9.3}$$

and if the potential is nonnegative, the relations

$$R \leq R(V) \leq |1/2b_2(V)|, \tag{9.4}$$

$$\lim_{V\to\infty} R(V) = R \tag{9.5}$$

supplement (9.1) and (9.2), and (9.3) if it applies.

The quantities $B(V)$, Φ, $B'(V)$, and Φ' are defined in (6.4), (2.4), (7.13), and (2.5); M is defined just before (3.3). By virtue of the inequality

$$B(V) \leq B \equiv \lim_{V\to\infty} B(V)$$

$$= \int_{\text{all space}} |e^{-\varphi(\mathbf{r})/kT} - 1|\, d^3\mathbf{r}, \tag{9.6}$$

we may replace $B(V)$ in (9.1) by B, whose definition is neater. This makes the lower bound in (9.1) identical with that given by Ruelle.[6] Similarly we may replace $B'(V)$ in (9.3) by its upper bound

$$B' \equiv \int_{\text{all space}} \text{Max}\,\{(1 - e^{-\varphi(\mathbf{r})/kT}), e^{2\Phi'/kT} \times (1 - e^{\varphi(\mathbf{r})/kT})\}\, d^3\mathbf{r}. \tag{9.7}$$

These simplifications destroy hardly any information unless V is small or B infinite. A case where the original form of (9.1) does give more information is the Coulomb potential $\varphi(\mathbf{r}) = \text{const}/r$, for which (9.1) shows that $R(V) = O(\mathcal{V}^{-1})$.

A system for which R can be calculated exactly is the 1-dimensional gas of hard rods of length a, for which $\varphi(\mathbf{r}) = +\infty$ for $r < a$ and $= 0$ for $r \geq a$. The equation of state is[21] $p = \rho kT/(1 - a\rho)$ from which it follows that $z = (p/kT)\exp(ap/kT)$ and hence, by (1.1) and Lagrange's inversion formula,[22] that

$$b_l = (-al)^{l-1}/l!. \tag{9.8}$$

The function $p(z) = kT\sum b_l z^l$ is analytic on the positive real axis, but it has a branch point at $z = -1/ea$, so that $R = 1/ea = 0.368a^{-1}$. This system shows no phase transition, and the divergence of $\sum b_l z^l$ for real $z > R$ has no physical significance.

Equation (9.5) shows that $\lim_{V\to\infty} R(V) = R = 0.368a^{-1}$. The lower bound on $\lim_{V\to\infty} R(V)$ computed from (9.1) is $1/eB = 1/2e|b_2| = 0.184a^{-1}$, just half the correct value. The upper bound on $\lim_{V\to\infty} R(V)$ given by (9.4) is $1/2a = 0.500a^{-1}$. The convergent sequence of upper bounds given by (9.1) starts $2.000a^{-1}$, $1.000a^{-1}$, $0.795a^{-1}$, $0.700a^{-1}$, $0.644a^{-1}$, etc., while the one given by (9.3) starts a^{-1}, $0.707a^{-1}$, $0.606a^{-1}$, $0.553a^{-1}$, $0.521a^{-1}$, $0.499a^{-1}$, $0.483a^{-1}$, etc. The convergence to the limit $0.368a^{-1}$ is a little faster in the second sequence, whose derivation takes advantage of the information that the potential has a hard core.

For an Ising lattice gas with coordination number q and negative interaction energy -2ϵ between nearest neighbours, Lee and Yang have shown[23] that $\lim_{V\to\infty} R(V) = 1/w^q$ where $w \equiv e^{\epsilon/kT}$. This is a hard-core potential, for which $\Phi' = \Phi = q\epsilon$, and our lower bounds on $\lim R(V)$ are

$$1/Be^{1+2\Phi/kT} = 1/[1 + q(w^2 - 1)]ew^{2q}, \tag{9.9}$$

$$1/B'e^{1+\Phi/kT} = 1/[1 + w^{2q}q(1 - w^{-2})]ew^q. \tag{9.10}$$

For a given lattice (fixed $q \geq 3$), the first of these lower bounds is the stronger when $w \gg 1$ (low temperatures) and the second when w is close to 1 (high temperatures). Our simplest upper bound on $\lim R(V)$ is 1 [from (9.3) with $l = 1$], which is actually attained when $\epsilon = 0$.

The physical significance of the lower bounds (9.1)

[20] Reference 18, Chap. 2, Sec. 1.

[21] L. Tonks, Phys. Rev. 50, 955 (1936).

[22] E. T. Copson, reference 11, Sec. 6.23.

[23] T. D. Lee and C. N. Yang, Phys. Rev. 87, 410 (1952). In the limit $V \to \infty$ their z becomes our $z\exp(q\epsilon/kT)$.

and (9.3) is[5,6] that they imply there can be no phase transition for fugacities less than $1/Be^{1+2\Phi/kT}$ (or $1/B'e^{1+\Phi/kT}$ for hard-core potentials) since the convergent power series (1.1) correctly gives the pressure for these fugacities. This is proved[5] most directly by using (6.12) and (9.6) [or (7.15) and (9.7)] together with Weierstrass' theorem[24] to show that the series (1.2) converges uniformly in the limit $V \to \infty$ for any z inside the disk $|z| < 1/Be^{1+2\Phi/kT}$ (or $|z| < 1/B'e^{1+\Phi'/kT}$). Alternatively, one may apply Yang and Lee's theory[10,25], since (6.13) and (9.6) [or 7.16 and (9.7)] show that this disk contains no zeros of $\Xi(z, V)$ for any V. Unfortunately the upper bounds on R and $R(V)$ cannot be interpreted so easily, since the singularity of (1.1) or (1.2) nearest to the origin need not be on the positive real axis; for nonnegative potentials this singularity is in fact on the negative real axis.[5] An example is the one-dimensional gas of hard rods, which has no phase transition although the fugacity series diverges for $z > 1/ea$.

[24] E. C. Titchmarsh, reference 11, 1.11 and 1.14.

[25] D. Ruelle, Helv. Phys. Acta **36**, 183 (1963) extends Yang and Lee's theorems to a more general class of potentials.

10. ACKNOWLEDGMENTS

I am indebted to J. Groeneveld, J. Lebowitz, E. Lieb, J. Percus, D. Ruelle, and G. Stell for helpful correspondence and discussions.

Convergence of Virial Expansions*

J. L. LEBOWITZ AND O. PENROSE†

Belfer Graduate School of Science, Yeshiva University, New York, New York

(Received 5 September 1963)

Some bounds are obtained on $\mathfrak{B}(V)$, the radius of convergence of the density expansion for the logarithm of the grand partition function of a system of interacting particles in a finite volume V, and on $\mathfrak{B}$, the radius of convergence of the corresponding infinite-volume expansion (the virial expansion). A common lower bound on $\mathfrak{B}(V)$ and $\mathfrak{B}$ is $0.28952/(u+1)B$, where $u \equiv \exp[-\text{Min}\, s^{-1}\sum_{i<j\leq s} 2\varphi(\mathbf{x}_i - \mathbf{x}_j)]/\kappa T$ [so that $u \geq 1$, with equality for nonnegative $\varphi(\mathbf{r})$], $B \equiv \int|e^{-\varphi(\mathbf{r})/\kappa T} - 1|\, d^3\mathbf{r}$, and $\varphi(\mathbf{r})$ is the binary interaction potential; the irreducible Mayer cluster integrals have the related upper bounds $\beta_s \leq [(u+1)B/0.28952]^s/s!u = 1$, when $\varphi(r) \geq 0$]. For potentials with hard cores the maximum density is an upper bound on $\mathfrak{B}(V)$, though possibly not on $\mathfrak{B}$; an example shows how both $\mathfrak{B}(V)$ and $\mathfrak{B}$ can be less than the maximum density, even if there is no phase transition. A theorem is proved, analogous to Yang and Lee's theorem on uniform convergence in the complex z plane, defining a class of domains in the complex ρ plane within which the operations $V \to \infty$ and $d/d\rho$ commute. This theorem is used to show that $\lim_{V\to\infty}\mathfrak{B}(V) \leq \mathfrak{B}$, and that there is no phase transition for $0 \leq \rho < 0.28952/(u+1)B$.

1. INTRODUCTION

RECENTLY several authors[1-3] have obtained upper and lower bounds for the radius of convergence $R(V)$ of the Mayer fugacity expansions.[4]

$$\mathcal{V}^{-1}\log \Xi(z, V) \equiv p(z, V)/\kappa T = \sum_\ell b_\ell(V) z^\ell, \quad (1.1)$$

$$\rho(z, V) \equiv (z/\kappa T)\, dp(z, V)/dz = \sum_\ell \ell b_\ell(V) z^\ell. \quad (1.2)$$

Here $\Xi(z, V)$ and $\rho(z, V)$ are the grand partition function and the mean number density at fugacity z and temperature T for a system of particles with two-body interactions, confined to a spatial region V whose volume is $\mathcal{V}$. Boltzmann's constant is denoted by κ. The coefficients $b_\ell(V)$ are the finite-volume Mayer[4] cluster integrals. The s-particle distribution functions $n_s(\mathbf{x}_1 \cdots \mathbf{x}_s \mid z, V)$ can also be expanded as power series in z, with radius of convergence at least[3] $R(V)$.

The thermodynamic pressure and density are given[1-3] for small z by

$$p(z) \equiv \lim_{V\to\infty} p(z, V) = \kappa T \sum_\ell b_\ell z^\ell, \quad (1.3)$$

$$\rho(z) \equiv \lim_{V\to\infty} \rho(z, V) = \sum_\ell \ell b_\ell z^\ell, \quad (1.4)$$

where

$$b_\ell \equiv \lim_{V\to\infty} b_\ell(V) \qquad (\ell = 1, 2, \cdots). \quad (1.5)$$

Moreover, the common radius of convergence R of these two series satisfies[3]

$$R \geq \liminf_{V\to\infty} R(V), \quad (1.6)$$

since any point $z = a$ with $|a| < \liminf_{V\to\infty} R(V)$ must be a regular point of $p(z)$. This follows from Yang and Lee's theory.[5]

* Supported by the Air Force Office of Scientific Research under Grant 62-64.

† Present address: Imperial College, London, England.

[1] J. Groeneveld, Phys. Letters **3**, 50 (1962).

[2] D. Ruelle, *Correlation Functions of Classical Gases* (Institute for Advanced Study, Princeton, 1963); Ann. Phys. (N. Y.) **25**, 109 (1963).

[3] O. Penrose, J. Math. Phys. 1312 (1963).

[4] Mayer describes his theory in *Handbuch der Physik* (Springer-Verlag, Berlin, 1958), Vol. 12.

[5] C. N. Yang and T. D. Lee, Phys. Rev. **87**, 404 (1952). The theory is generalized to a wider class of potentials by D. Ruelle, Helv. Phys. Acta **36**, 183 (1963).

The purpose of this paper is to make a similar study of the radii of convergence $\mathfrak{R}(V)$ and $\mathfrak{R}$ of the finite- and infinite-volume density expansions obtained by eliminating z from (1.1) and (1.2), and from (1.3) and (1.4). These expansions may be written

$$p(z, V) = P(\rho(z, V), V) \equiv \kappa T \rho(z, V) \times \left[1 - \sum_k \frac{k}{k+1} \beta_k(V) \rho(z, V)^k\right], \quad (1.7)$$

$$p(z) = P(\rho(z)) \equiv \kappa T \rho(z) \times \left[1 - \sum_k \frac{k}{k+1} \beta_k \rho(z)^k\right]. \quad (1.8)$$

The $\beta_k(V)$'s can be expressed in terms of the $b\ell(V)$'s by algebraic relations,[4] such as $\beta_1(V) = 2b_2(V)$, $\beta_2(V) = 3b_3(V) - 6b_2(V)^2$, etc., which do not involve V explicitly. It follows by (1.5) that

$$\beta_k = \lim_{V \to \infty} \beta_k(V). \quad (1.9)$$

The β_k's are the irreducible Mayer cluster integrals[4] and (1.8) is the virial expansion. We shall study $\mathfrak{R}(V)$ by a method based on Lagrange's theorem for the expansion of one function of z in powers of another. This method incidentally yields upper bounds on the absolute values of the β_ℓ's and $\beta_\ell(V)$'s. We shall study $\mathfrak{R}$ by means of a generalization to the complex ρ-plane of Yang and Lee's results[5] on uniform convergence in the z plane.

Our lower bounds on $\mathfrak{R}(V)$ and $\mathfrak{R}$ apply to systems of particles whose positions $\mathbf{x}_1, \mathbf{x}_2, \cdots$ are either continuously variable or confined to a lattice. Their interaction stems from a two-body interaction potential $\varphi(r)$ for which there exists a constant Φ such that

$$\sum_{i<j\le s} \varphi(\mathbf{x}_i - \mathbf{x}_j) \ge -s\Phi \quad \text{for all } s, \mathbf{x}_1 \cdots \mathbf{x}_s. \quad (1.10)$$

The circumstances under which (1.10) is satisfied have been discussed by Ruelle[2,6] and Penrose.[3] We shall also make the convergence assumption

$$B \equiv \int_{\text{all space}} |e^{-\varphi(\mathbf{r})/\kappa T} - 1|\, d^\nu \mathbf{r} < \infty, \quad (1.11)$$

where ν is the number of space dimensions (= 1, 2, or 3). In discussing the upper bounds on $\mathfrak{R}(V)$ and $\mathfrak{R}$, we shall further assume that the potential has a hard core, i.e., that a positive constant a exists such that

$$\varphi(\mathbf{r}) = +\infty \quad \text{if} \quad r < a, \quad (1.12)$$

but this assumption is unnecessary in the other parts of the discussion.

[6] D. Ruelle, Ref. 5.

2. LAGRANGE'S THEOREM

Lagrange's theorem,[7] adapted to the expansion of $p(z, V)$ in powers of $\rho(z, V)$ may be stated thus: let the function $z/\rho(z, V)$ be analytic within and on a closed contour C surrounding the origin of the z plane, and let ρ be a complex number satisfying

$$|\rho| < \mu \equiv \operatorname*{Min}_{z \text{ on } C} |\rho(z, V)|. \quad (2.1)$$

Then the equation $\rho(z, V) = \rho$ is satisfied by just one value of z inside C, which we denote by $z(\rho, V)$; further, if the function $p(z, V)$ is analytic within and on C, it has the convergent expansion

$$P(\rho, V) \equiv p(z(\rho, V), V) = \sum_{n=1}^{\infty} c_n \rho^n, \quad (2.2)$$

where

$$c_n \equiv \frac{1}{2\pi i} \oint_{C'} \frac{dp(z, V)}{dz} \frac{dz}{n\{\rho(z, V)\}^n} = \frac{1}{n!} \frac{d^{n-1}}{dz^{n-1}} \left[\frac{dp(z, V)}{dz} \left\{\frac{z}{\rho(z, V)}\right\}^n\right]_{z=0}. \quad (2.3)$$

The path of integration is any contour C' surrounding $z = 0$ such that $|\rho(z, V)| \le \mu$ for all z on C'. The uniqueness of $z(\rho, V)$ follows from Rouché's theorem,[8] which shows that the functions z/ρ and $z/\rho - z/\rho(z, V)$ have the same number (one) of zeros inside C. The formula for c_n is obtained by expanding in powers of ρ on both sides of the following equation derived from Cauchy's residue theorem:

$$P(\rho, V) = \frac{1}{2\pi i} \oint_{C'} p(z, V) \frac{d\rho(z, V)}{dz} \frac{dz}{\rho(z, V) - \rho}, \quad (2.4)$$

and then integrating the resulting formula for c_n by parts. By virtue of the relation (1.2) between $\rho(z, V)$ and $p(z, V)$, and the definition (1.7) of the $\beta_k(V)$, Eq. (2.3) for $n = 2, 3, \cdots$ is equivalent to

$$-k\beta_k(V) = \frac{1}{2\pi i} \oint \frac{dz}{z\{\rho(z, V)\}^k} \quad (2.5)$$

for $k = 1, 2, \cdots$. This formula is used in Sec. 3 to estimate the $\beta_k(V)$'s.

3. LOWER BOUNDS ON $\mathfrak{R}(V)$ AND $\mathfrak{R}$

According to Lagrange's theorem, the series (2.2) converges if $|\rho|$ is less than the lower bound μ of $|\rho(z, V)|$ on the contour C; that is,

$$\mathfrak{R}(V) \ge \mu \equiv \operatorname*{Min}_{z \text{ on } C} |\rho(z, V)|. \quad (3.1)$$

[7] E. T. Whittaker and G. N. Watson, *Modern Analysis* (Cambridge University Press, New York, 1927), Sec. 7.32.

[8] E. T. Copson, *Theory of Functions of a Complex Variable* (Oxford University Press, London, 1935), Sec. 6.21.

A suitable lower bound on $|\rho(z, V)|$ can be found from Penrose's generalization[3] of Groeneveld's estimates[1] of the $b_\ell(V)$'s

$$|\ell b_\ell(V)| \leq u^{\ell-2}[\ell B]^{\ell-1}/\ell! \quad (\ell = 2, 3, \cdots), \tag{3.2}$$

where

$$u \equiv e^{2\Phi/\kappa T} \geq 1, \tag{3.3}$$

and Φ and B are defined in (1.10) and (1.11). These estimates imply[1,3] that $R(V) \geq 1/euB$. In the rest of the section we ensure the convergence of (1.1) and (1.2) by requiring

$$|z| < 1/euB. \tag{3.4}$$

Since $b_1 = 1$ the series (1.1) now gives the inequality

$$\begin{aligned}|\rho(z, V) - z| &= \left|\sum_{\ell=2}^{\infty} \ell b_\ell z^\ell\right| \leq \frac{1}{u^2 B}\sum_{\ell=2}^{\infty} \ell^{\ell-1}(uB\,|z|)^\ell/\ell! \\ &= w/u^2B - |z|/u, \end{aligned} \tag{3.5}$$

where w is defined by

$$we^{-w} = uB\,|z|, \qquad 0 \leq w < 1. \tag{3.6}$$

Since the function we^{-w} increases monotonically from 0 to e^{-1} in the range $0 \leq w < 1$, the condition (3.4) guarantees that w exists and is unique. In deriving the last line of (3.5) we used Euler's expansion[9] for w in powers of we^{-w}:

$$w = \sum \ell^{\ell-1}(we^{-w})^\ell/\ell!. \tag{3.7}$$

From (3.5) and (3.6) we obtain a lower bound on $|\rho(z, V)|$,

$$\begin{aligned}|\rho(z, V)| &\geq (1 + 1/u)\,|z| - w/u^2B \\ &= \{(u + 1)e^{-w} - 1\}w/u^2B. \end{aligned} \tag{3.8}$$

As the contour C in (3.1), we may choose any circle $|z| = \text{const} < 1/euB$. By (3.6) the equation of this circle may be written $w = \text{const}$, and the corresponding value of μ is $\geq \{(u+1)e^{-w}-1\}w/u^2B$. Since (3.1) holds whatever value of w in the range $0 \leq w < 1$ is used to define C, we must have

$$\mathfrak{R}(V) \geq \operatorname*{Max}_{0\leq w<1} \{(u + 1)e^{-w} - 1\}w/u^2B. \tag{3.9}$$

To obtain a convenient estimate of $\mathfrak{R}(V)$, we use the identity

$$\begin{aligned}\{(u + 1)e^{-w} - 1\}w/u^2B &= \left[v - v^2 g\left(\frac{vu}{1+u}\right)\right]/(u + 1)B, \end{aligned} \tag{3.10}$$

where $v \equiv w(1 + u)/u$ and $g(w) \equiv (1 - e^{-w})/w$. Since $u/(1 + u) \geq \frac{1}{2}$ by (3.3) and $g(w)$ decreases monotonically, the right side of (3.10) is at least $[v - v^2 g(\frac{1}{2}v)]/(u + 1)B$. Hence (3.9) implies

$$\begin{aligned}\mathfrak{R}(V) &\geq \operatorname*{Max}_{v} [v - v^2 g(\tfrac{1}{2}v)]/(u + 1)B \\ &= 0.28952/(u + 1)B. \end{aligned} \tag{3.11}$$

The maximum is attained when $v = 0.62984$. If we had not replaced $g(vu/(1 + u))$ by $g(\frac{1}{2}v)$ the numerator in (3.11) would have been replaced by a function of u increasing monotonically from 0.28952 when $u = 1$ (nonnegative potentials) to $e^{-1} = 0.36788$ as $u \to \infty$.

These methods also yield upper bounds on the $\beta_k(V)$'s. Taking the contour in (2.5) to be a circle $|z| = \text{const}$, we obtain the estimate

$$\begin{aligned}k\,|\beta_k(V)| &\leq \frac{1}{2\pi}\oint_C d\,|z| \operatorname*{Max}_{z \text{ on } C}\left|\frac{1}{z\{\rho(z, V)\}^k}\right| \\ &= [\operatorname*{Min}_{z \text{ on } C} |\rho(z, V)|]^{-k}. \end{aligned} \tag{3.12}$$

Choosing the radius of the circle, as before, to maximize the quantity in square brackets, we find that

$$\begin{aligned}k\,|\beta_k(V)| &\leq [\operatorname*{Max}_{w} \{(u + 1)e^{-w} - 1\}w/u^2B]^{-k} \\ &\leq \left[\frac{(u + 1)B}{0.28952}\right]^k \qquad (k = 1, 2, \cdots). \end{aligned} \tag{3.13}$$

Combined with (1.5) this gives upper bounds on the irreducible Mayer cluster integrals

$$k\,|\beta_k| \leq [(u + 1)B/(0.28952)]^k. \tag{3.14}$$

This set of inequalities implies, by (1.8) and Cauchy's kth-root convergence test, that

$$\begin{aligned}\mathfrak{R} = \liminf_{k\to\infty}\left|\frac{k}{k+1}\beta_k\right|^{-1/k} \\ \geq 0.28952/(u + 1)B, \end{aligned} \tag{3.15}$$

so that $\mathfrak{R}$ and $\mathfrak{R}(V)$ have the same lower bound. The result (3.15) can also be obtained by applying to the function $\rho(z)$ the same arguments which when applied to $\rho(z, V)$ led to (3.11); or by using (6.1).

Reprinted from
Journal of Mathematical Physics
Volume 5, Number 7, July 1964

[9] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze der Analysis* (Springer-Verlag, Berlin, 1925), Vol. I, Part III, Chap. 5, No. 209.

CONFIGURATION INTEGRAL FOR ONE-DIMENSIONAL SYSTEMS OF PARTICLES

by

L. Van Hove
University of Brussels*; now at the Institute for Advanced Study
Princeton, New Jersey, U.S.A.

(translated from the French article in Physica XVI No. 2; 137-143 (February 1950))

Synopsis

The free energy of a one-dimensional system of particles is calculated for the case of non-vanishing hard-core radius of the particles and a finite range of the forces. It is shown quite generally that no phase transformation phenomena can occur under these circumstances. The method used is the reduction of the problem to an eigenvalue problem.

§1. Introduction

The partition function of a one-dimensional system has already been studied by various authors in cases where the interaction between particles takes place solely among nearest neighbors and is sufficiently simple to permit direct integration.[1] No phenomenon of change in phase, such as condensation, has ever been proved to occur by these calculations. We used two different methods for the case of an arbitrary interaction taking place between nearest neighbors

*Department of Physical Chemistry.

only.* Unknown to us, Takahasi[2] had already applied one of these methods to the same problem, but, again no condensation phenomena were observed.

In the following, we shall discuss a more general interaction, not limited to nearest neighbors. We shall assume that the particles have a nonvanishing hard-core radius and that the potential interaction between two particles has a finite range; barring that, we simply suppose it to be continuous. Under these conditions, the problem can be reduced to a search for the eigenvalue an integral equation. We have here, moreover, an example of statistical mechanics being reduced to the determination of the eigenvalue of a matrix.[3] Nevertheless, our case concerns an integral equation; its kernel is generally not symmetric and the integration interval is infinite. By changing the variables, we render the integration interval finite and can thus make use of Fredholm's theory. Usually, the appearance of phase transformations is linked to the multiplicity of the eigenvalue obtained and is often excluded following a known theorem by

---

*For the mathematical aspect of a method which is equally suitable for the calculation of the distribution function the reader is referred to a discussion by the author in the Bulletin de la Societé Mathématique de Belgique (1949), to be published. The other method has been developed in the present article.

Frobenius.[5] Our problem is a little different; nevertheless, an extension of the theorem to integral equations[6] shows that any phenomenon of phase transformation must still be excluded in our case. It should be borne in mind, however, that the simultaneous existence of a nonvanishing hard-core radius and a finite range of the forces is essential for the validity of our reasoning and conclusion.

§2. The configuration integral

Let there be a system of N identical particles situated on a rectilinear segment of length L; their positions are determined by their abscisses x_i ($i = 1,\ldots,N$; $0 \leq x_i \leq L$). The potential energy of the system can be expressed as:

$$V = \sum_{i<j} U(|x_i - x_j|), \tag{2.1}$$

where $U(\xi)$ is a function such that

$$\begin{aligned} U(\xi) &= +\infty \quad \text{for} \quad 0 \leq \xi < d_0, \\ U(\xi) &= 0 \quad \text{for} \quad \xi \geq d_1, \quad 0 < d_0 < d_1. \end{aligned} \tag{2.2}$$

In the interval $d_0 \leq \xi \leq d_1$, we suppose $U(\xi)$ to be continuous and bounded below. Moreover, $U(\xi) = +\infty$ can apply only to $\xi < 2d_0$. The length d_0 appears as the diameter of the hard-core of the particles, and d_1 is the maximum range of the forces of interaction.

Let us designate by m the mass of the particles, supposing them to be points, and by T the absolute temperature multiplied by the Boltzmann constant. The Helmholtz free energy F, per particles, is then given by

$$F/T = - f(N,L) - \log(d_0 h^{-1}\sqrt{2\pi mT}) \qquad (2.3)$$

with

$$\exp[Nf(N,L)] = Q(N,L) = \left(1/N!d_0^N\right)\int_0^L dx_1 \cdots \int_0^L dx_N . \exp(-V/T)$$

$$= d_0^{-N}\int_0^L dx_1 \int_{x_1}^L dx_2 \cdots \int_{x_{k-1}}^L dx_k \cdots \int_{x_{N-1}}^L dx_N \exp(-V/T) \; *$$

As is known, for the whole set of lengths L_N such that $\lim_{N\to\infty} L_N/N = \ell > d_0$, the expression $f(N,L_N)$ converges to a finite limit $f(\ell)$, an increasing function of ℓ, with non-increasing derivative.[7] We can reduce the determination of $f(\ell)$, and thus the free energy F, to the solution of an integral equation. It should be noted that as a result of (2.2), two nearest particles can never approach each other closer than d_0; consequently, the particles capable of

*The factor d_0 is introduced solely so that the quantity Q(N,L) will be dimensionless.

interaction with a given particle are limited to their ν neighbors on the left, ν being the highest whole number for which $\nu d_0 < d_1$. This limitation which results from the fact that $d_0 > 0$ and $d_1 < +\infty$, is essential for the developments discussed hereafter.

§3. Free energy at constant pressure

As has been pointed out by Takahasi in the case where $\nu = 1$ [2] in a system having one dimension, the free energy at constant pressure $G = F + pl$ is much easier to determine than F. It is directly linked to the Laplace transform of $Q(N,L)$:

$$I_N(\alpha) = \int_0^\infty Q(N,L)e^{-\alpha L}dL. \tag{3.1}$$

In effect, from the behavior of $Q(N,L)$ for $N \to \infty$, it follows that $(I_N(\alpha))^{1/N}$ coverges to a finite limit

$$\lim_{N\to\infty}\{I_N(\alpha)\}^{1/N} = e^{-g(\alpha)} \tag{3.2}$$

given by

$$e^{-g(\alpha)} = \lim_{N\to\infty}\left\{Q\left(N,L_N^{(\alpha)}\right)\exp\left(-\alpha L_N^{(\alpha)}\right)\right\}^{1/N}, \tag{3.3}$$

where $L_N^{(\alpha)}$ is the value of L which maximizes

$$Q(N,L)e^{-\alpha L} = \exp[Nf(N,L) - \alpha L].$$

As $f(N,L_N) \to f(\ell)$ when $N \to \infty$ and $N^{-1}L_N \to \ell$, the length $N^{-1}L_N^{(\alpha)}$ must converge to value of ℓ which maximizes $f(\ell) - \alpha$ and is thus defined by

$$df(\ell)/d\ell = \alpha. \tag{3.4}$$

The equation (3.3) gives, therefore,

$$g(\alpha) = a\ell - f(\ell). \tag{3.5}$$

Taking into consideration (3.4) and (2.3) one sees that the pressure p and the free energy at constant pressure G are related to α and $g(\alpha)$ by

$$p/T = \alpha, \quad G/T = g(\alpha) - \log\left((d_0/h)\sqrt{2\pi mT}\right).$$

In the following paragraph we shall determine $g(\alpha)$. From this we will deduce $f(\ell)$ by inverting the Legendre transformation contained in (3.4) and (3.5):

$$f(\ell) = \alpha\ell - g(\alpha), \quad \ell = dg(\alpha)/d\alpha. \tag{3.6}$$

In the special case where $\nu = 1$, the interaction energy (2.1) is $V = \sum_{i=1}^{N-1} U(\xi_i)$, where the ξ_i designate the distances between nearest particles. The Laplace transform (3.1) can be effortlessly calculated and (3.2) gives

$$g(\alpha) = - \log\left\{\frac{1}{d_0} \int_0^\infty \exp\left[- \frac{U(\xi)}{T} - \alpha\xi\right] d\xi\right\},$$

from which

$$f(\ell) = \alpha\ell + \log\left\{\frac{1}{d_0} \int_0^\infty \exp\left[- \frac{U(\xi)}{T} - \alpha\xi\right] d\xi\right\},$$

$$\ell = \frac{\int_0^\infty \xi \exp[- U(\xi)/T - \alpha\xi] d\xi}{\int_0^\infty \exp[- U(\xi)/T - \alpha\xi] d\xi}.$$

These are the equations given by Takahasi;[2] we have obtained them by another method.* Since they express ℓ as a single valued function of α, no condensation is possible.

---

*See footnote * on page 2.

§4. The problem of eigenvalue

Let us now determine $g(\alpha)$ for the general case where $\rho = \nu - 1 > 0$. In this case the interaction energy is:

$$V = \sum_{i=1}^{N-1} U(\xi_i) + \sum_{i=2}^{N-1} U(\xi_{i-1}+\xi_i) + \ldots \sum_{i=\nu}^{N-1} U(\xi_{i-\rho}+\ldots+\xi_{i-1}+\xi_i),$$

the ξ_i designating again the distances between adjacent particles. Now let us define the functions

$$h(\xi_1,\ldots,\xi_\rho)$$

$$= \exp\left\{-\frac{1}{2T}\sum_{i=2}^{\rho} U(\xi_{i-1}+\xi_i) - \frac{1}{2T}\sum_{i=3}^{\rho} U(\xi_{i-2}+\xi_{i-1}+\xi_i) \ldots - \frac{1}{2T}U(\xi_1+\ldots+\xi_\rho)\right\},$$

$$k\left(\xi_1,\ldots,\xi_\rho,\xi_1'\ldots\xi_\rho'\right)$$

$$= \exp\left\{-\frac{1}{T}\left[U\left(\xi_\rho+\xi_\rho'\right) + U\left(\xi_{\rho-1}+\xi_\rho+\xi_1'\right) + U\left(\xi_\rho+\xi_1'+\xi_2'\right) + \ldots\right.\right.$$

$$+ U\left(\xi_\rho+\ldots+\xi_\rho+\xi_1'\right) + U\left(\xi_2+\ldots+\xi_\rho+\xi_1'+\xi_2'\right) + \ldots$$

$$\left.\left. + U\left(\xi_\rho+\rho_1'+\ldots+\xi_\rho'\right)\right]\right\}$$

$$K\left(\xi_1,\ldots,\xi_\rho,\xi_1',\ldots,\xi_\rho'\right) = h(\xi_1,\ldots,\xi_\rho)k\left(\xi_1,\ldots,\xi_\rho,\xi_1',\ldots,\xi_\rho'\right)h\left(\xi_1',\ldots,\xi_\rho'\right),$$

$$K^{(n)}\left(\xi_1,\ldots,\xi_\rho,\xi_1',\ldots,\xi_\rho'\right)$$

$$\doteq \exp\left\{-\frac{1}{2}\sum_{i=1}^{\rho}\left[\frac{U(\xi_i)}{T}+\alpha\xi_i\right]\right\} K\left(\xi_1,\ldots,\xi_\rho,\xi_1',\ldots,\xi_\rho'\right)$$

$$\times\exp\left\{-\frac{1}{2}\sum_{i=1}^{\rho}\left[\frac{U\left(\xi_i'\right)}{T}+\alpha\xi_i'\right]\right\}. \qquad (4.1)$$

The iterated kernel $K_n^{(\alpha)}\left(\xi_1,\ldots,\xi_\rho,\xi_1',\ldots,\xi_\rho'\right)$ is determined by the recurrence relation:

$$K_{n-1}^{(\alpha)}\left(\xi_1,\xi_1'\right)=\int_0^\infty d\xi_1''\cdots\int_0^\infty d\xi_\rho'' K_n^{(\alpha)}\left(\xi_1,\xi_1''\right)K^{(\alpha)}\left(\xi_1'',\xi_1'\right);\quad K_1^{(\alpha)}=K^{(\alpha)}.$$

It permits us to express function $I_N(\alpha)$ given in (3.1) with the aid of an integral whose multiplicity does not depend anymore on N:

$$I_S(\alpha)=\int_0^\infty d\xi\int_0^\infty d\xi_1\cdots\int_0^\infty d\xi'\int_0^\infty d\xi_1'\cdots\int_0^\infty d\xi_\rho'\cdot K_n^{(\alpha)}\left(\xi_1,\xi_1'\right)h(\xi_1)h\left(\xi_1'\right)$$

$$\exp\left\{-2T^{-1}\sum_{i=1}^{\rho}\left[U(\xi_i)+U\left(\xi_i'\right)\right]\right\}\exp\left\{-a\left[\xi+\xi'+\frac{1}{2}\sum_{i=1}^{\rho}\left(\xi_i'+\xi_i'\right)\right]\right\}$$

if $N=\nu+n\rho$. From this equation and (3.2) one obtains

$$\lim_{n\to\infty}\left\{K_n^{(\alpha)}\left(\xi_1,\xi_1'\right)\right\}^{1/n}=\lim_{n\to\infty}\left\{I_N(\alpha)\right\}^{1/n}=e^{-\rho g(\alpha)}.$$

If ever $U(\xi) \neq +\infty$, $U\left(\xi_1'\right) \neq +\infty$, the first member of this equality will not depend anymore on the variables ξ_1, ξ_1'; of this one can make sure with the help of the considerations which establish the existence of the limit of $\{Q(N, L_N)\}^{1/N}$ when $N \to \infty$ and $N^{-1}L_N \to \ell$.[7]

As often happens with this type of problem,[3] we are led to consider the integral equation

$$\varphi(\xi_1) = \lambda \int_0^\infty d\xi_1' \cdots \int_0^\infty d\xi_\rho' K^{(\alpha)}\left(\xi_1, \xi_1'\right)\varphi\left(\xi_1'\right), \qquad (4.3)$$

where the $K_n^{(\alpha)}$ is the iterated kernel of order n. This allows the integrations to be limited to the interval (d_0, ∞). We shall transform this into an integral equation with a bounded region of integration and with a positive definite kernel. We shall use the change of variables,

$$\eta = q^{(\alpha)}(\xi) = \int_\xi^\infty \exp\left[-\left[T^{-1}U(\xi') + \alpha\xi'\right]\right] d\xi',$$

whose inverse we describe by $\xi = r^{(\alpha)}(\eta)$ and the change of the unknown function

$$\varphi(\eta_1) = \exp\left\{\frac{1}{2}\sum_1^\rho \left[T^{-1}U(\xi_1) + \alpha\xi_1\right]\right\}\varphi(\xi_1), \quad \text{with} \quad \eta_1 = q^{(\alpha)}(\xi_1).$$

According to (4.1), equation (4.3) becomes

$$\varphi(\eta_1) = \lambda \int_0^{\delta(\alpha)} d\eta_1' \cdots \int_0^{\delta(\alpha)} d\eta_\rho' H^{(\alpha)}\left(\eta_1, \eta_1'\right)\varphi\left(\eta_1'\right), \quad (4.4)$$

with

$$H^{(\alpha)}\left(\eta_1, \eta_1'\right) = K\left[r^{(\alpha)}(\eta_1), r^{(\alpha)}\left(\eta_1'\right)\right].^*$$

The iterated kernels $H_n^{(\alpha)}\left(\eta_1, \eta_1'\right)$ of (4.4) differ from the $K_n^{(\alpha)}\left(\xi_1, \xi_1'\right)$ only by a factor independent of n. The relation (4.2) becomes, therefore:

$$e^{-\rho g(\alpha)} = \lim_{n\to\infty}\left\{H_n^{(\alpha)}\left(\eta_1, \eta_1'\right)\right\}^{1/n},$$

where the limit of the second member is independent of η_1, η_1'. According to the Cauchy-Hadamard theorem the quantity $e^{\rho g(\alpha)}$ is therefore the radius of convergence of the resolvent series

---

*If $U(\xi)$ is bounded for $d_0 \leq \xi < d1$, it is possible to use the change of variables $\eta = 1/\alpha \exp(-\alpha\xi)$ and the change of unknown function $\psi(\eta_1) = \exp(\alpha/2\Sigma\rho\xi_1)\varphi(\xi_1)$; a positive kernel will also be obtained. Conversely, when for example $\lim U(\xi) = +\infty$ for $d_0 < \xi \to d_1$, it will be necessary to revert to the procedure described in the text.

$$H^{(\alpha)} + \lambda H_2^{(\alpha)} + \ldots \lambda^{n-1} H_n^{(\alpha)} + \ldots$$

According to the Fredholm theory which is applicable to equation (4.4), this radius is $|\lambda^{(\alpha)}|$, where $\lambda^{(\alpha)}$ is the eigenvalue of the absolute minimum of the equation.[8] Therefore, $g(\alpha) = \rho^{-1} \log|\lambda^{(\alpha)}|$. It should be noted that the kernels of the equations (4.2) and (4.4) are not symmetric if $\rho > 1$; one has

$$K^{(\alpha)}\left(\xi_1, \ldots, \xi_\rho, \xi_1', \ldots, \xi_\rho'\right) = K^{(\alpha)}\left(\xi_\rho', \ldots, \xi_1', \xi_\rho, \ldots, \xi_1\right),$$

and relation for $H^{(\alpha)}$.

§5. The impossibility of phase transformation

A phase transformation manifests itself by a constant pressure on a nonvanishing interval of ℓ values. With such an interval, α must remain constant when ℓ varies; according to equation (3.6) the derivative $dg(\alpha)/d\alpha$ must therefore possess a discontinuity. We shall see that such a condition is impossible.

According to the theory of Fredholm, the eigenvalues of the equation (4.4) are the zeros of an entire function $D^{(\alpha)}(\lambda)$. It can be easily proved that for $\mathrm{Re}(\alpha) > 0$, the function $D^{(\alpha)}(\lambda)$ is simultaneously holomorphic in α and λ. Since a theorem of Jentzsch[6] that $D^{(\alpha)}(\lambda)$ has a root of

minimum absolute value, which at the same time is unique, simple and real. It is this one which is designated by $\lambda^{(\alpha)}$ and $|\lambda^{(\alpha)}| = \lambda^{(\alpha)} > 0$. Given the analytic character of $D^{(\alpha)}(\lambda)$, the root $\lambda^{(\alpha)}$ is a holomorphic function of α, when $\mathrm{Re}(\alpha) > 0$. Consequently, the function $g(\alpha) = \rho^{-1}\log \lambda^{(\alpha)}$ is holomorphic. No condensation phenomenon can occur in the type of system studied.

Received December 19, 1949

LITERATURE

1. K. F. Herzfeld and M. Goeppert-Mayer, Journ. Chem. Phys. 2, 38, 1934. L. Tonks, Phys. Rev. 50, 955, 1936; T. Nagamiya, Proc. Phys. Math. Soc. Japan 22, 705, 1035, 1940.

2. H. Takahasi, Proc. Phys. Math. Soc. Japan 24, 60, 1942.

3. H. A. Kramers and G. H. Wannier, Phys. Rev. 60, 252, 1941; E. W. Montroll, Journ. Chem. Phys. 9, 706, 1941; E. N. Lassettre and J. P. Howe, Journ. Chem. Phys. 9, 747, 1941; G. S. Rushbrooke and H. D. Ursell, Proc. Cambr. phil. Soc. 44, 263, 1947.

4. E. W. Montroll, loc. cit.[3]

5. G. Frobenius, Sitzungsber. Preuss. Akad. Wiss. Berlin, 1908, 471, 1909, 514.

6. R. Jentzsch, Crelles Journ. 141, 235, 1912. E. Hopf, Sitzungsber. Preuss. Akad. Wiss. Berlin, 1928, 233.

7. I. Van Hove, Physica, in Press.

8. E. Goursat, Cours d'Analyse Mathématique, vol. III, Paris, 1927, chap. XXXI.

COMMENTS

(1) Van Hove's results have been generalized to systems having a finite extension in a second dimension by M. Kummer, Helv. Phys. Acta 35, 451 (1962).

(2) The Laplace transform of the radial distribution function of a one-dimensional system with nearest neighbor interactions was shown by Lebowitz, Percus and Zucker, Bull. Amer. Phys. Soc. 7, 415 (1962), to have the form

$$\sigma(s) = \ell[\exp[\mu(p+s) - \mu(p)] - 1]^{-1}$$

where μ, the chemical potential and p the pressure are measured in units of temperature. The radial distribution function is thus determined entirely by the thermodynamic functions for these systems.

(3) For the behavior of a one-dimensional system with long range forces, in which a phase transition can occur in a certain limit, see paper by M. Kac, G. Uhlenbeck and P. C. Hemmer in Section III. A thermodynamic discussion of the impossibility of a phase transitions in one dimension is given in Section 149 of Statistical Physics by Landau and Lifshitz (Pergamon, London, 1958).

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 3, NUMBER 4 JULY-AUGUST 1962

Exact Statistical Mechanics of a One-Dimensional System with Coulomb Forces. II. The Method of Functional Integration

S. F. EDWARDS
Department of Theoretical Physics, Manchester University, Manchester, England
AND
A. LENARD*
Plasma Physics Laboratory, Princeton University, Princeton, New Jersey
(Received January 11, 1962)

The statistical mechanics of a one-dimensional system of charged sheets is studied in the formalism of the grand canonical ensemble. It is shown that the grand partition function may be expressed as a Wiener integral, i. e., as an average of a certain functional of Brownian motion paths. This functional integral is then expressed in terms of the fundamental solution of a partial differential equation of diffusion type. This depends on a theorem of Kac whose proof is also given. The generality of this method is discussed. When all charges are integral multiples of a common unit the problem is reduced to the determination of the largest characteristic value of an ordinary differential operator with periodic coefficients. An invariance property of the thermodynamic potential is shown to imply charge neutrality in the infinite system limit. A theorem is proven which, in certain cases, excludes the possibility of a thermodynamic phase transition. The method is generalized to yield exact expressions for the n-particle reduced density functions. Some properties of the two-particle functions are discussed.

1. INTRODUCTION

IN two recent papers the thermodynamics of a one-dimensional "plasma" has been determined rigorously.[1,2] This system consists of a great number of uniformly charged plane sheets freely able to move and cross each other in the direction normal to the planes. It was shown in I that the thermodynamic functions are expressible in terms of a single function which may be defined by the Sturm-Liouville characteristic value problem of the Mathieu functions. An analysis of this result showed that the system behaves as a true plasma when its mean kinetic energy is much larger than its mean potential energy; in the opposite case it resembles a collection of infrequently colliding neutral molecules. The calculation was rendered possible by the observation that the partition function of the constant-pressure ensemble is a sum of simple algebraic expressions, and that the problem of summation over very many terms could be discussed systematically by using generating functions.

In the present paper we propose to discuss this problem with the aid of a more powerful method. It is a variant of a method used by one of us[3] for the calculation of the thermodynamic functions of a real (i.e., three-dimensional) plasma. Its essence is the expression of the grand partition function as the average of a certain functional over its argument function. An averaging over functions (or "functional integration"; we shall use the terms interchangeably) may be defined in a variety of ways. In probability theory a set of functions over which a method of averaging with non-negative weight is prescribed is termed a stochastic (or random) process. It turns out that for the one-dimensional plasma the appropriate random process is of a familiar and simple kind. It is the "Wiener process" which is the mathematical idealization of the Brownian motion of a particle.[4,5]

There are several motivations for this work. One of them is that in three dimensions the presence of the singularity of the interparticle force at small distances makes a theory based on a pure Coulomb force mathematically meaningless, and if this is remedied by introducing a repulsive hard core for the particles the resulting extra complication obscures the essentials of the functional integration method. There is no need for such a complication in the one-dimensional model since the interparticle potential remains regular with vanishing distance. Secondly, the present method offers greater insight

* Supported by the U. S. Atomic Energy Commission.
[1] A. Lenard, J. Math. Phys. 2, 682 (1961). This paper will be referred to as I.
[2] S. Prager, *Advances in Chemical Physics*, edited by I. Prigogine (Interscience Publishers, Inc., New York, 1961), Vol. IV.
[3] S. F. Edwards, Phil. Mag. 4, 1171 (1959). The author regrets that several calculational errors were inadvertently allowed to occur in this article.
[4] M. Kac, *Probability and Related Topics in Physical Sciences* (Interscience Publishers, Inc., New York, 1959), Chap. IV.
[5] J. M. Gel'fand and A. M. Yaglom, J. Math. Phys. 1, 48 (1960), where other references are to be found.

into some mathematical features of the one-dimensional model. In particular, the characteristic value problem of the Mathieu functions enters here in a quite natural way. This stands in contrast to the treatment in I where it appears as a trick whose origin remains somewhat of a mystery. The reduction to the characteristic value problem is accomplished by the use of a theorem due to Kac[6] which allows the expression of the average of a functional of a certain type in terms of the fundamental solution of a parabolic partial differential equation. This reduction has no analog in the three-dimensional case for reasons which will be pointed out below.

With the present method we also obtain some results that go far beyond those contained in I. For one, we are not restricted to the case of a two-component gas containing particles of equal and opposite charge, although this remains the simplest case in which concrete results can be obtained. More importantly, we obtain rigorous expressions for the reduced density functions involving an arbitrary number of particles, and, in particular, a fast converging expansion for the simplest of these, the two-particle density function.

Our plan is as follows. We first set up the machinery of statistical mechanics in the grand canonical ensemble formalism and explain the reason for the appropriate choice of the potential energy function (Sec. 2). Next, we give a brief resume of the Wiener process and reproduce a proof of a theorem due to Kac which is of decisive importance (Sec. 3). We then show how the grand partition function may be regarded as a certain functional integral (or average) over the Wiener process, and reduce its determination to an initial value problem for a partial differential equation of the diffusion type. We discuss the role that is played in this result by two properties of the Wiener process, its Markoffian and its Gaussian nature (Sec. 4). If all particles possess charges which are integral multiples of a common unit, the problem may be further reduced to a characteristic value problem of an ordinary differential operator with periodic coefficients. Two general questions are discussed, the question of charge neutrality and the question of the continuity of the thermodynamic quantities (Sec. 5). The dominant behavior in the "plasma limit" is derived by a simple heuristic approximation scheme on the appropriate partial differential equation (Sec. 6). Finally, we explain how the reduced density functions are determined and discuss, in particular, the properties of the two-particle function (Sec. 7).

[6] M. Kac, *Proceedings of the Second Berkeley Symposium on Probability and Statistics* (University of California Press, Berkeley, California, 1951), p. 189. See also reference 4.

2. THE GRAND PARTITION FUNCTION

The most convenient method from the present point of view is that based on the grand canonical ensemble. We suppose our system to consist of a great number of parallel charged sheets located in the region

$$0 \leq x \leq L. \tag{1}$$

These sheets (which in the following will be referred to as particles) carry surface charge densities which fall into a finite number of groups. We denote by N' the number of particles carrying charge σ', N'' carrying charge σ'', etc., and let $N = N' + N'' + \cdots$ stand for the total number. For the moment we make no assumptions regarding the charges σ', σ'', $\cdots$ which may be different or equal, positive, negative, or even zero, and need not be integral multiples of a common unit. A more restrictive hypothesis will be imposed later.

In the grand canonical ensemble the probability density that N arbitrary elements $dx_1, dx_2, \cdots, dx_N$ within the region (1) are occupied by particles of charges $\sigma_1, \sigma_2, \cdots, \sigma_N$, respectively, and that no other particles are in the region is

$$\frac{1}{\Omega} z'^{N'} z''^{N''} \cdots \times \exp\left\{-\frac{V_{N'N''\cdots}(x_1, x_2, \cdots, x_N)}{\theta}\right\}. \tag{2}$$

Here θ is the temperature; z', z'', $\cdots$ are positive parameters of the dimension of a number density; $V_{N'N''\cdots}$ is the total potential energy as a function of the coordinates; and $\Omega = \Omega(L, \theta, z', z'', \cdots)$ is the grand partition function. The latter is defined by

$$\Omega = \sum_{N'=0}^{\infty} \frac{z'^{N'}}{N'!} \sum_{N''=0}^{\infty} \frac{z''^{N''}}{N''!} \cdots \times \int_0^L dx_N \cdots \int_0^L dx_1 \exp\left\{-\frac{V_{N'N''\cdots}}{\theta}\right\}. \tag{3}$$

As is well known, the asymptotic behavior of Ω in the infinite system limit $L \to \infty$ specifies the thermodynamic properties of the system. In particular, the pressure is

$$P = \theta \lim_{L\to\infty} \frac{\ln \Omega}{L}. \tag{4}$$

Once $P = P(\theta, z', z'', \cdots)$ is determined, the densi-

ties of the several components may be derived by differentiation

$$n' = (z'/\theta)\,\partial P/\partial z',$$
$$n'' = (z''/\theta)\,\partial P/\partial z'', \qquad (5)$$
$$\text{etc.}$$

The internal energy per unit volume is

$$u = \tfrac{1}{2}z'\,\partial P/\partial z' + \tfrac{1}{2}z''\,\partial P/\partial z'' + \cdots + \theta\,\partial P/\partial\theta - P. \qquad (6)$$

We now turn to the consideration of the potential energy. If two particles, placed at x_1 and x_2, have charges σ_1 and σ_2, respectively, the potential energy of their mutual interaction is $-2\pi\sigma_1\sigma_2\,|x_1 - x_2|$. It would be incorrect, however, to take for the total potential energy merely the sum

$$-2\pi \sum_{1\le k<l\le N}\sum \sigma_k\sigma_l\,|x_k - x_l|. \qquad (7)$$

We must imagine that the system is in contact with an infinite reservoir (in the region $x \le 0$, say) which exchanges particles with it giving rise to the statistical fluctuations in particle numbers. Take the system *plus reservoir* electrically neutral as a whole and consider the state of affairs when some particular number of particles, N' of charge σ', N'' of charge σ'', etc., are in the system. Then there is a total charge

$$\sigma_{\text{res}} = -(N'\sigma' + N''\sigma'' + \cdots) \qquad (8)$$

in the reservior. This creates a constant electric field of magnitude $2\pi\sigma_{\text{res}}$ in the region $x \ge 0$. The particles of the system are subject to this electric field in addition to the forces they exert on each other. Thus, there is an additional contribution

$$-2\pi\sigma_{\text{res}} \sum_{1<k\le N} \sigma_k x_k \qquad (9)$$

to the potential energy. *We define $V_{N'N''\cdots}$ to be the sum of the functions* (7) *and* (9). The contribution (9) comes physically from the interaction of the particles inside the system with those inside the reservoir. That it must be properly attributed to the system as an additional potential energy and cannot be neglected as a "boundary effect" is due, of course, to the fact that the interparticle force does not tend to zero with increasing distance.

The expression for the potential energy may be conveniently simplified by the device made already use of in I. The grand partition function may be decomposed into a sum of integrals in each of which the integration variables are restricted by the inequalities

$$0 \le x_1 \le x_2 \le \cdots \le x_N \le L \qquad (10)$$

and each being associated with a definite succession of charges (a "configuration")

$$C\colon \sigma_1, \sigma_2, \cdots, \sigma_N. \qquad (11)$$

In addition to the sums over the numbers $N', N'', \cdots$ we now have also a sum over all sequences C restricted only by the requirement that N' of the σ_k have value σ', N'' the value σ'', etc. Let us consider now a definite configuration and introduce the quantities

$$E_0 = 4\pi\sigma_{\text{res}},$$
$$E_k = 4\pi\left(\sigma_{\text{res}} + \sum_{l=1}^{k}\sigma_l\right), \qquad (12)$$
$$(k = 1, 2, \cdots, N).$$

E_k is just the electric field in the space between the kth and $(k+1)$st particles. A simple calculation shows that the potential energy, as we have defined it, is equal to

$$V_{N',N'',\cdots} = \sum_{k=1}^{N} \frac{E_{k-1}^2}{8\pi}\,(x_k - x_{k-1}) \qquad (13)$$

($x_0 = 0$, by definition), i.e., *the total electrostatic field energy within the boundaries of the system.*

Making use of this we have the grand partition function in the form

$$\Omega = \sum_{N'=0}^{\infty} z'^{N'} \sum_{N''=0}^{\infty} z''^{N''} \cdots \sum_{C} \int_0^L dx_N \cdots \times \int_0^{x_3} dx_1 \exp\left\{-\frac{1}{8\pi\theta}\sum_{k=1}^{N} E_{k-1}^2(x_k - x_{k-1})\right\}. \qquad (14)$$

The symbol $\sum_C$ denotes the summation over the

$$\frac{N!}{N'!\,N''!\cdots}$$

configurations.

3. THE WIENER PROCESS AND THE THEOREM OF KAC

We shall now collect the mathematical apparatus needed for exhibiting the grand partition function as a functional integral. This material is not new and it is included here only in order to enable a reader not thoroughly familiar with it to follow the argument without interruption. If more detail and mathematical rigor than can be given here is desired, reference may be made to the review of Gel'fand and Yaglom,[5] or to the book of Kac.[4]

We consider the mathematical description of the Brownian motion of a particle. Let $\phi(x)$ stand for one of its coordinates as a function of time.[7] This function is specified only in a statistical sense, namely, by the following two postulates:

(a) The conditional probability that $\phi(x_1)$ is in an element $d\phi_1$ around ϕ_1 given that $\phi(x_0) = \phi_0$ is $P(\phi_1 - \phi_0, x_1 - x_0)\, d\phi_1$ for any $x_1 > x_0$. The function P is

$$P(\phi, x) = (4\pi Dx)^{-\frac{1}{2}} e^{-(\phi^2/4Dx)}, \tag{15}$$

D being a positive constant.

(b) Displacements in nonoverlapping time intervals are statistically independent quantities.

The two postulates are compatible on account of the identity

$$P(\phi_2 - \phi_0, x_2 - x_0) = \int_{-\infty}^{\infty} d\phi_1 P(\phi_2 - \phi_1, x_2 - x_1) \times P(\phi_1 - \phi_0, x_1 - x_0), \qquad (x_0 < x_1 < x_2), \tag{16}$$

whose meaning is that the displacement $\phi(x_2) - \phi(x_0)$ can be regarded as the sum of two statistically independent displacements $\phi(x_1) - \phi(x_0)$ and $\phi(x_2) - \phi(x_1)$. If $x_0, x_1, \cdots x_n$ in any increasing sequence of time values, the probability that $\phi(x_k)$ is in $d\phi_k (k = 1, 2, \cdots, n)$ given that $\phi(x_0) = x_0$ is the product

$$\prod_{k=1}^{n} P(\phi_k - \phi_{k-1}, x_k - x_{k-1})\, d\phi_k. \tag{17}$$

Any quantity which is an expression involving $\phi(x)$ is a random variable whose average value may be calculated by the above assignment of probabilities. It was Wiener[8] who first recognized that this probabilistic model of Brownian motion amounts to assigning a measure in the space of continuous functions $\phi(x)$, the "possible paths of the Brownian particle," and that statistical averaging corresponds to a theory of integration over this space.[4,5]

We shall be interested exclusively in averages of the form

$$\left\langle \exp \int_{x_0}^{x} dx' F(\phi(x'), x') \mid \phi(x_0) = \phi_0 \right\rangle, \tag{18}$$

where $F(\phi, x)$ is some given function. We shall always use this bracket notation to indicate averaging over $\phi(x)$ by the method just explained. The condition following the bar is to be satisfied by all Brownian paths over which the averaging is carried out. When misunderstanding is not likely we shall omit writing out this condition explicitly.

A method of calculating (18) will now be given.

Expand the exponential in a power series, interchange the averaging with the time integration, and make use of the symmetry of the integrand in the integration variables. The following series expansion results:

$$1 + \sum_{n=1}^{\infty} \int_{x_0}^{x} dx_n \int_{x_0}^{x_n} dx_{n-1} \cdots \int_{x_0}^{x_2} dx_1 \times \langle F(\phi_n, x_n) F(\phi_{n-1}, x_{n-1}) \cdots F(\phi_1, x_1) \rangle, \tag{19}$$

where we have set $\phi(x_k) = \phi_k$ as an abbreviation. The successive terms of this series will now be related to each other by an integral-recursion formula. We write the first term

$$1 = \int_{-\infty}^{\infty} d\phi Q_0(\phi, x; \phi_0, x_0) \tag{20}$$

with

$$Q_0 = P(\phi - \phi_0, x - x_0). \tag{21}$$

The next term ($n = 1$) may be written as

$$\int_{x_0}^{x} dx_1 \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\phi_1 P(\phi - \phi_1, x - x_1) \times F(\phi_1, x_1) P(\phi_1 - \phi_0, x_1 - x_0) = \int_{-\infty}^{\infty} d\phi Q_1(\phi, x; \phi_0, x_0). \tag{22}$$

The first factor P may appear superfluous inasmuch as the integration over ϕ can be trivially carried out. However, it is convenient to include it because this puts in evidence that

$$Q_1 = \int_{x_0}^{x} dx_1 \int_{-\infty}^{\infty} d\phi_1 P(\phi - \phi_1, x - x_1) \times F(\phi_1, x_1) Q_0(\phi_1, x_1; \phi_0, x_0). \tag{23}$$

The next term ($n = 2$) is

$$\int_{x_0}^{x} dx_2 \int_{x_0}^{x_2} dx_1 \int_{-\infty}^{\infty} d\phi \int_{-\infty}^{\infty} d\phi_2 \int_{-\infty}^{\infty} d\phi_1 \times P(\phi - \phi_2, x - x_2) F(\phi_2, x_2) P(\phi_2 - \phi_1, x_2 - x_1) \times F(\phi_1, x_1) P(\phi_1 - \phi_0, x_1 - x_0) = \int_{-\infty}^{\infty} d\phi Q_2(\phi, x; \phi_0, x_0), \tag{24}$$

where

$$Q_2 = \int_{x_0}^{x} dx_2 \int_{-\infty}^{\infty} d\phi_2 P(\phi - \phi_2, x - x_2) \times F(\phi_2, x_2) Q_1(\phi_2, x_2; \phi_0, x_0). \tag{25}$$

[7] The reason for this unconventional notation will become clear below.

[8] N. Wiener, J. Math. and Phys. 2, 131 (1923).

The structure of this formula is the same as (23), and this can be continued indefinitely. The general term of the sum (19) is

$$\int_{-\infty}^{\infty} d\phi Q_n(\phi, x; \phi_0, x_0), \tag{26}$$

and the successive integrands are related to each other by

$$Q_n(\phi, x; \phi_0, x_0) = \int_{x_0}^{x} dx' \int_{-\infty}^{\infty} d\phi'$$

$$\times P(\phi - \phi', x - x')F(\phi', x')Q_{n-1}(\phi', x'; \phi_0, x_0). \tag{27}$$

Let

$$Q = \sum_{n=0}^{\infty} Q_n. \tag{28}$$

Then we have shown that *the average value* (18) *can* be expressed as

$$\int_{-\infty}^{\infty} d\phi Q(\phi, x; \phi_0, x_0), \tag{29}$$

the function Q being the solution of the integral equation

$$Q(\phi, x; \phi_0, x_0) = P(\phi - \phi_0, x - x_0)$$

$$+ \int_{x_0}^{x} dx' \int_{-\infty}^{\infty} d\phi' P(\phi - \phi', x - x')$$

$$\times F(\phi', x')Q(\phi', x'; \phi_0, x_0). \tag{30}$$

This theorem and this derivation of it are due to Kac.[9,4] Other derivations have been given by Rosenblatt,[10] by Blanc-Lapierre and Fortet,[11] and by Darling and Siegert.[12]

The integral equation (30), in turn, is equivalent to a partial differential equation with an appropriate boundary condition. The function P, given explicitly in Eq. (15), satisfies the equation

$$\{\partial/\partial x - D\,\partial^2/\partial\phi^2\}P(\phi, x) = 0 \tag{31}$$

and the initial condition

$$P(\phi, 0) = \delta(\phi). \tag{32}$$

Therefore, applying the operator $\{\partial/\partial x - D\partial^2/\partial\phi^2\}$ on the right-hand side of (30) only one term survives; thus,

$$\{\partial/\partial x - D\,\partial^2/\partial\phi^2 - F(\phi, x)\}Q(\phi, x; \phi_0, x_0) = 0. \tag{33}$$

Equation (30) also shows that

$$Q(\phi, x_0; \phi_0, x_0) = \delta(\phi - \phi_0). \tag{34}$$

This means that Q is the fundamental solution of the partial differential equation (33).

The theorem of Kac can be interpreted physically in a variety of ways, but our concern with it is due only to the interesting fact that the grand partition function of the system under investigation can be identified as a functional average of the precise form (18). The next section is devoted to the demonstration of this fact and a discussion of its significance.

4. THE GRAND PARTITION FUNCTION AS A FUNCTIONAL AVERAGE

We intend to show that

$$\Omega = \left\langle \exp \int_0^L dx' F(\phi(x')) \mid \phi(0) = 0 \right\rangle, \tag{35}$$

where the function F is

$$F(\phi) = z'e^{i\sigma'\phi} + z''e^{i\sigma''\phi} + \cdots, \tag{36}$$

and the diffusion constant of the Wiener process [cf. Eq. (15)] is

$$D = 2\pi/\theta. \tag{37}$$

To show this we begin by expanding the exponential inside the bracket into a multiple power series of z', z'', etc., and then we interchange averaging with integrating over x. The series

$$\sum_{N'=0}^{\infty} \frac{z'^{N'}}{N'!} \sum_{N''=0}^{\infty} \frac{z''^{N''}}{N''!} \cdots \int_0^L dx_1 \cdots \int_0^L dx_N$$

$$\times \left\langle \exp\left\{ i \sum_{k=1}^{N} \sigma_k \phi_k \right\} \right\rangle \tag{38}$$

results, where again $\phi_k = \phi(x_k)$ is used as an abbreviation. The quantities $\sigma_1, \sigma_2 \cdots, \sigma_N$ are defined as follows: The first N' of them are equal to σ', the next N'' to σ'', etc., and $N = N' + N'' + \cdots$. The integration is over a domain symmetric in the N variables, so that the integrand, although not symmetric, may be replaced by its symmetric part

$$\frac{N'!\,N''!\cdots}{N!} \sum_{C} \left\langle \exp\left\{ i \sum_{k=1}^{N} \sigma_k \phi_k \right\} \right\rangle. \tag{39}$$

Here C is a symbol for an arbitrary succession of quantities $\sigma_1, \sigma_2 \cdots, \sigma_N$ with the property that *some* N' of them equal σ', *some* N'' equal σ'', etc., in other words it has the same significance as in Sec. 2. The summation goes over all such sequences, whose total number is just the inverse of the factor in front. Once (39) is substituted as the integrand of (38) it is permissible to integrate over the restricted domain (10) at the price of multiplying by

[9] Reference 6, Sec. 3.
[10] M. Rosenblatt, Trans. Am. Math. Soc. 77, 120 (1951).
[11] A. Blanc-Lapierre and R. Fortet, *Theorie des functions aleloires* (Masson & Cie, Paris, 1953), Chap. VII.
[12] D. A. Darling and A. J. F. Siegert, Proc. Natl. Acad. Sci. U. S. 42, 525 (1956).

a factor $N!$. Thus, (35) becomes

$$\sum_{N'=0}^{\infty} z'^{N'} \sum_{N''=0}^{\infty} z''^{N''} \cdots \sum_{C} \int_0^L dx_N \cdots \times \int_0^{s_1} dx_1 \left\langle \exp\left\{ i \sum_{k=1}^{N} \sigma_k \phi_k \right\} \right\rangle. \quad (40)$$

This is identical to Ω, in the form given by Eq. (14), if

$$\left\langle \exp\left\{ i \sum_{k=1}^{N} \sigma_k \phi_k \right\} \right\rangle = \exp\left\{ -\frac{1}{8\pi\theta} \sum_{k=1}^{N} E_{k-1}^2 (x_k - x_{k-1}) \right\}. \quad (41)$$

This identity is established as follows. We recall the definition of the quantities E_k given by Eqs. (12) and (8) which can also be written

$$(1/4\pi)(E_k - E_{k-1}) = \sigma_k \quad (k = 1, 2, \cdots, N), \quad (42)$$
$$E_N = 0.$$

Now multiply (42) by ϕ_k and sum over k from 1 to N. In view of $E_N = 0$ and $\phi_0 = 0$ this sum is

$$-\frac{1}{4\pi} \sum_{k=1}^{N} E_{k-1}(\phi_k - \phi_{k-1}) = \sum_{k=1}^{N} \sigma_k \phi_k. \quad (43)$$

Thus, the quantity to be averaged on the left-hand side of (41) is a product of N *statistically independent* quantities

$$\prod_{k=1}^{N} \exp\left\{ -\frac{i}{4\pi} E_{k-1}(\phi_k - \phi_{k-1}) \right\},$$

so that the product sign may be taken outside the bracket

$$\prod_{k=1}^{N} \left\langle \exp\left\{ -\frac{i}{4\pi} E_{k-1}(\phi_k - \phi_{k-1}) \right\} \right\rangle = \prod_{k=1}^{N} \int_{-\infty}^{\infty} d\phi P(\phi, x_k - x_{k-1}) \exp\left\{ -\frac{i}{4\pi} E_{k-1}\phi \right\} = \prod_{k=1}^{N} \exp\left\{ -\frac{D E_{k-1}^2 (x_k - x_{k-1})}{16\pi^2} \right\}. \quad (44)$$

This is identical to the right-hand side of (41) provided D is set equal to $2\pi/\theta$, as claimed.

The expression for Ω thus obtained is of the general form discussed in the previous section, therefore it may be expressed in yet another form, namely as an integral over the fundamental solution of a certain partial differential equation. Application of the theorem of Kac to Eq (35) establishes the following

Theorem 1. *The grand partition function is expressible as*

$$\Omega = \int_{-\infty}^{\infty} d\phi Q(\phi, L; 0, 0), \quad (45)$$

where $Q = Q(\phi, x; \phi_0, x_0)$ is the fundamental solution of the partial differential equation

$$\{\partial/\partial x - D\,\partial^2/\partial\phi^2 - F(\phi)\} Q = 0, \quad (46)$$

that is to say the solution satisfying the initial condition

$$Q(\phi, x_0; \phi_0, x_0) = \delta(\phi - \phi_0). \quad (47)$$

$F(\phi)$ is given by (36) *and the constant D by* (37). The function Q depends only on the difference $x - x_0$ because F does not depend explicitly on x. Thus, we shall more appropriately write $Q = Q(\phi, \phi_0, x)$ and for abbreviation often omit the argument ϕ_0 when it is understood to have the value zero.

Theorem 1 is the central result of our paper. Let us make some comments on its significance.

It is clear, first of all, that our result has its origin in two distinct calculations, the proof of the theorem of Kac and the representation of the grand partition function as a functional average. They represent two ways of "evaluating" the functional average and it is in the connection between the two end products that our interest lies. The details of the calculations are of a quite familiar kind, and it is apparent that the intermediate notion of functional average could be eliminated altogether. The really essential point is that in the iteration solution (Neumann series) of the integral equation (30) each term can be explicitly evaluated by a repeated application of the elementary formula

$$\int_{-\infty}^{\infty} d\phi \frac{1}{(4a\pi)^{\frac{1}{2}}} \exp\left\{ -\frac{\phi^2}{4a} + ib\phi \right\} = \exp\{-ab^2\}, \quad (48)$$
$$(a > 0)$$

and the terms turn out to be identical to the terms in the power series expansion (14) of the grand partition function.

Nevertheless it would be wrong to conceal the importance of the functional average concept with its probabilistic interpretation, because it sheds light on the question: To what extent is the success of our analysis due to the particular features of our model? The two aspects of the calculation are characterized by different assumptions. In the proof of the theorem of Kac—at least in its integral equation formulation—no use is made of the special Gaussian form (15) of the transition probability, but only that the joint probability for succeeding steps is the product (17). This is the assumption that the random process $\phi(x)$ is Markoffian. On the other hand, the explicit identification of the grand

partition function with the functional average is based on the Gaussian property, and on that alone.

In order to bring out this point we shall make a digression from the principal topic of this paper and consider the averaging over general Gaussian random processes in some detail. A general Gaussian random process $\phi(x)$ is defined by the postulate that for any finite number of points $x_1, x_2, \cdots, x_N$ the joint probability density for $\phi(x_k)$ in $d\phi_k$ is of the form

$$\frac{(\det B)^{\frac{1}{2}}}{(2\pi)^{N/2}} \exp\left\{-\tfrac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} B_{kl}\phi_k\phi_l\right\}, \qquad (49)$$

where the elements of the matrix B are functions of the x_k alone. The matrix B must be positive definite because otherwise the normalization integral would not converge, but this is not the only condition to which it is subject. Let $\alpha_k (k = 1, 2, \cdots, N)$ be arbitrary real numbers. Then

$$\left\langle \exp\left\{i\sum_{k=1}^{N}\alpha_k\phi(x_k)\right\}\right\rangle = \exp\left\{-\tfrac{1}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} C_{kl}\alpha_k\alpha_l\right\}, \qquad (C = B^{-1}) \quad (50)$$

because the left-hand side is just the N-dimensional Fourier transform of the function (49) which can be explicitly evaluated by the obvious generalization of the formula (48). If (50) is differentiated with respect to α_k and α_l (not excluding $k = l$) and then all α are set zero, one obtains

$$\langle\phi(x_k)\phi(x_l)\rangle = C_{kl} = C(x_k, x_l). \qquad (51)$$

Thus, B is necessarily the inverse of a matrix C whose elements are given by the second equation of (51) with $C(x_1, x_2)$ a single function (the "covariance function") of two variables. This function then completely characterizes the statistical nature of $\phi(x)$; however, it too is not arbitrary, but subject to the condition that for *any choice* of the points $x_k (k = 1, 2, \cdots, N$; N arbitrary) the matrix C_{kl} must be positive definite.

All this generalizes in an obvious manner to the case when the independent variable x ranges over a three-dimensional space, say, and also to the case when a certain finite number $\phi_1(x), \phi_2(x), \cdots \phi_k(x)$ of processes are involved. In the latter case the covariance function is replaced by the covariance matrix

$$\langle\phi_\alpha(x_1)\phi_\beta(x_2)\rangle = C_{\alpha\beta}(x_1, x_2). \qquad (52)$$

The formula (50) is the basis for the application to statistical mechanics. If all the α's are put equal to a common constant we get, integrating over the x's,

$$\left\langle\left[\int_0^L dx e^{i\alpha\phi(x)}\right]^N\right\rangle = \int_0^L dx_1 \cdots \int_0^L dx_N \times \exp\left\{-\frac{\alpha^2}{2}\sum_{k=1}^{N}\sum_{l=1}^{N} C(x_k, x_l)\right\}. \qquad (53)$$

If it were not for the terms $k = l$, this would be the partition function of a canonical ensemble for a system consisting of N particles between which there are two-body forces with a potential $C(x_1, x_2)$.[13] When several Gaussian functions $\phi_\alpha(x)$ are involved then the corresponding formula refers to the statistical mechanics of a system which contains different species of particles and (52) then just gives the interaction potential between the α and β species.

This is the point of contact of the present work with the paper of Edwards.[3] There the averaging over function space was done in a formal way; however, the method prescribed[14] shows that averaging over a Gaussian random process is involved. It is clear that no property of the interparticle potential needs to be used other than that it is the covariance function (or matrix, in the case of more than one species) of a Gaussian random process. This shows that the representation of the partition function as a functional average has great generality, although it should be said as a precaution that the positive definiteness condition on the covariance function is certainly not a natural condition if that function is to be interpreted as a potential.[15] We also see now that the really special feature of our model is the Markoffian nature of the random process associated with it in the sense just explained. This is a severe restriction. There is no possibility for application to three-dimensional systems, because the concept of a Markoffian process involves the idea of a succession in "time" and this is meaningless when there is more than one independent variable. More importantly, even a one-dimensional Gaussian process is not Markoffian in general. An interesting process, both Gaussian and Markoffian, is the Ornstein-Uhlenbeck process

[13] This observation lies at the root of a recent paper on the statistical mechanics of a one-dimensional system, M. Kac, Phys. Fluids **2**, 8 (1959).

[14] Reference 3, Eq. (1.7).

[15] If the positive definiteness condition is not respected then one deals with "integrals" of the type (48) with negative parameter a. This may not be fatal as long as the integral is only a symbolic expression which stands for the right-hand side. However, it is quite easy this way to get into contradictions because formal manipulations of such "integrals" may be illegitimate on account of their lack of convergence. There is a legitimate way of overcoming this difficulty due to A. J. F. Siegert, but we do not want to go into this subject here. One of us (A. L.) is indebted to Professor Siegert for letting him see his recent work before publication.

defined by the covariance function[16]

$$\langle\phi(x_1)\phi(x_2)\rangle = C_1 e^{-C_2|x_1-x_2|}, \qquad (C_1, C_2 > 0). \tag{54}$$

With an interparticle potential of this form it is then possible to investigate the statistical mechanics in much the same detail as for our model.[17] For the general case the representation of the partition function as a functional integral is a device which may or may not be a convenient aid in approximate calculations, but progress of the kind that we are reporting here on the one-dimensional plasma cannot be expected in general.

It remains to point out the specific connection between our work and the preceding considerations. Let $\phi(x)$ be the Wiener process as defined in Sec. 3 and normalized by the condition $\phi(0) = 0$, and let us associate with a particle species (having charge σ', say) the process which is just a numerical multiple of it

$$\phi'(x) = \sigma'\phi(x). \tag{55}$$

The processes $\phi'(x), \phi''(x), \cdots$ are Gaussian and their covariance matrix is given by the elements

$$C'''(x_1, x_2) = \langle\phi'(x_1)\phi''(x_2)\rangle = \sigma'\sigma''\langle\phi(x_1)\phi(x_2)\rangle = 2D\sigma'\sigma'' \min(x_1, x_2), \tag{56}$$

as one may verify easily by carrying out the relevant integrals. On the other hand, the total potential energy, i.e., the sum of (7) and (9), is

$$\begin{aligned} V &= -2\pi \sum_{k=2}^{N} \sum_{l=1}^{k-1} \sigma_k\sigma_l |x_k - x_l| + 2\pi \sum_{k=1}^{N} \sigma_k \sum_{l=1}^{N} \sigma_l x_l \\ &= 2\pi \sum_{k=1}^{N} \sum_{l=1}^{N} \sigma_k\sigma_l \left\{ -\frac{|x_k - x_l|}{2} + \frac{x_k + x_l}{2} \right\} \\ &= 2\pi \sum_{k=1}^{N} \sum_{l=1}^{N} \sigma_k\sigma_l \min(x_k, x_l). \end{aligned} \tag{57}$$

Thus, in view of (37)

$$\frac{V}{\theta} = \frac{1}{2} \sum_{k=1}^{N} \sum_{l=1}^{N} C_{kl}(x_k, x_l), \tag{58}$$

the subscripts being interpreted in the obvious sense as referring to the charges of the particles which label the rows and columns of the matrix C. Now the identification of the right-hand side of Eq. (35) with that of Eq. (3) is a simple matter of using the formula (53). This derivation of (35) is even simpler than the one presented at the beginning of this section. We have presented the latter because is presupposes less familiarity with the ideas of a Gaussian random process.[18]

One final remark. The Wiener process is not "stationary," i.e., explicitly independent of time, since the condition $\phi(0) = 0$ puts the Brownian particle at the origin at a given time thus distinguishing this instant from all others. This implies that the covariance function is not a function of the difference $x_1 - x_2$ alone [cf. Eq. (56)]. The *general* identification of the covariance with the interparticle potential demands that the process be stationary because the interparticle potential is a function of the difference of the two position variables. But the one-dimensional plasma is an exception. Here the total potential energy is not just the sum of pair potentials [cf. discussion in Sec. 2]. This explains how it comes about that the Gaussian random process appropriate to this problem is the nonstationary Wiener process.

5. THE CHARACTERISTIC VALUE PROBLEM

We now assume that all charges are rational multiples of each other, so that with a proper choice of the charge unit σ', σ'', etc., will be integers. For convenience we also choose the unit of length in such a way that $D = 1$.

The coefficient function $F(\phi)$ in Eq. (46) is now periodic with period 2π. This makes it possible to reduce the problem posed by Theorem 1 to the characteristic value problem of an ordinary differential operator on a finite interval of the independent varible. Let

$$\hat{Q}(\phi, x) = \sum_{n=-\infty}^{\infty} Q(\phi + 2\pi n, x). \tag{59}$$

This function is the "periodic fundamental solution" of the partial differential equation (46), i.e., for $x = 0$ it reduces to

$$\hat{Q}(\phi, 0) = \sum_{n=-\infty}^{\infty} \delta(\phi + 2\pi n). \tag{60}$$

It is clear that

$$\Omega = \int_{-\pi}^{\pi} d\phi \hat{Q}(\phi, L). \tag{61}$$

Since F does not depend on x, we may use the method of separation of variables. This leads to the charac-

[16] M. C. Wang, and G. E. Uhlenbeck, Revs. Modern Phys. 17, 323 (1945).

[17] M. Kac, "Probability in Classical Physics," *Proceedings of the Seventh Symposium in Applied Mathematics of the American Mathematical Society* (McGraw-Hill Book Company, Inc., New York, 1957). It appears to be the first published suggestion of using stochastic averaging as an aid in calculating partition functions.

[18] An excellent general introduction to this subject is contained in Wang and Uhlenbeck, reference 16.

teristic value problem

$$\{d^2/d\phi^2 + F(\phi)\}y(\phi) = \gamma y(\phi) \tag{62}$$

with the boundary condition

$$y(\phi + 2\pi) = y(\phi). \tag{63}$$

We shall label the characteristic values and functions with a subscript m ($m = 0, 1, 2, \cdots$) if the need arises. The functions are assumed to be normalized, orthogonal

$$\int_{-\pi}^{\pi} d\phi y_m(\phi) y_{m'}(\phi) = \delta_{mm'}, \tag{64}$$

and to form a complete set for functions with period 2π. The expansion of $\hat{Q}$ in terms of these functions is

$$\hat{Q}(\phi, x) = \sum_{m=0}^{\infty} e^{\gamma_m x} y_m(0) y_m(\phi). \tag{65}$$

Hence, the grand partition function appears exhibited in the following form

$$\Omega = \sum_{m=0}^{\infty} A_m e^{\gamma_m L}, \tag{66}$$

where the coefficients are

$$A_m = y_m(0) \int_{-\pi}^{\pi} d\phi y_m(\phi). \tag{67}$$

The γ_m and the y_m depend parametrically on z', z'', etc., which enter into the definition of $F(\phi)$.

Let us make some general remarks on the characteristic value problem. It is a generalization of the type known in the mathematical literature as Hill's problem.[19] The only difference is that our function $F(\phi)$ is complex in general while in the Hill problem the corresponding function is real. This has the consequence that our problem, while being self adjoint in the real sense, is not necessarily *Hermitian*. Thus, it cannot be excluded *a priori* that characteristic values and functions be complex, but it seems likely that the expansion theorem still holds.[20] The function F has the symmetry

$$F^*(-\phi) = F(\phi), \tag{68}$$

which does imply that together with any characteristic function $y(\phi)$ another characteristic function is $y^*(-\phi)$, the corresponding values of γ being complex conjugates. But from this it only follows that $\hat{Q}$ has the same symmetry (68) and that the terms of the expansion (66) are either real or occur in complex conjugate pairs.

Let us consider now the asymptotic behavior of Ω as $L \to \infty$. It is clear that those terms of (66) survive which have the largest real parts for the γ's. Thus,

$$\Omega(L) \sim \sum_{m}{}' A_m e^{\gamma_m L}, \qquad (L \to \infty), \tag{69}$$

where the dash indicates that the summation is extended only over those terms for which Re γ_m is the maximum.[21] As just remarked, the terms of this sum are either real or occur in complex conjugate pairs, but one at least must be real otherwise the sum could not be positive for all L which would preclude the asymptotic equality with the positive function $\Omega(L)$. Let us call this real value γ_0. By taking the logarithm and comparing with (4) we have obtained

Theorem 2. *The thermodynamic pressure is given by*

$$P = 2\pi\gamma_0, \tag{70}$$

where γ_0 is the largest real characteristic value of the problem (62) *with F being given by* (36).[22]

Regarding the magnitude of γ_0 we have the inequalities

$$0 < \gamma_0 \leq (z' + z'' + \cdots), \tag{71}$$

the equality being possible only in the trivial case of no interaction (all σ's zero). This follows from the corresponding inequalities

$$1 < \Omega(L) \leq \exp[(z' + z'' + \cdots)L] \tag{72}$$

for the grand partition function, the second of which depends on the non-negative nature of the potential energy [see Eq. (13)].

We shall now discuss two general questions concerning the thermodynamic properties of the model. The first of these is the question of charge neutrality. It is intuitively evident that in the infinite system limit the total charge density should be zero

$$\sigma' n' + \sigma'' n'' + \cdots = 0. \tag{73}$$

This is not in contradiction with the fact that for any fixed value of L the mean particle numbers can be freely adjusted by varying the parameters z', z'' etc. The point is that as L becomes larger it becomes

[19] E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis* (Cambridge University Press, New York, 1927), Chap. XIX.

[20] See E. Hill, Math. Ann. 71, 76 (1912) where a closely related problem is studied.

[21] If is not excluded, of course, that this sum contains a single term; indeed, this should be expected to be the case in general. However, for special values of the z's the possibility of a degeneracy cannot be excluded.

[22] This (in a more restricted case) is the content of Theorem 4 of I.

"more and more difficult" to do this, and as $L \to \infty$ for any fixed z values, (73) becomes established as a limiting relation. We shall now prove this.

Hitherto ϕ was always a real variable, but once the characteristic values γ for Eq. (62) are determined the corresponding functions $y(\phi)$ can be analytically continued into the complex ϕ plane. This follows from a general theorem on differential equations with analytic coefficients.[23] Indeed, our equation has no finite singular points so that its solutions are entire functions, i.e., regular in the whole open complex plane. Together with any characteristic function $y(\phi)$ with characteristic value γ, consider the new function $Y(\phi) = y(\phi - i\kappa)$ with κ an arbitrary real number. Evidently

$$\{d^2/d\phi^2 + F(\phi - i\kappa)\}\, Y(\phi) = \gamma Y(\phi), \qquad (74)$$

and furthermore Y has the period 2π because $Y(\phi + 2\pi) - Y(\phi)$ is an entire function which vanishes on the line Im $\phi = i\kappa$ and hence vanishes everywhere. Now $F(\phi - i\kappa)$, regarded as a function of the real variable ϕ, arises out of $F(\phi)$ by the substitutions

$$z' \to z'e^{\sigma' \kappa},$$
$$z'' \to z''e^{\sigma'' \kappa}, \qquad (75)$$
$$\text{etc.}$$

Thus, we have constructed a characteristic function with the same value of γ of the problem (62) transformed by the substitution (75). This means that we have proven the following.

Theorem 3. *The characteristic values γ of the problem* (62), *regarded as functions of the parameters z', z'', etc., are invariant under the group of substitutions* (75). This means that the quantities

$$\gamma(z'e^{\sigma' \kappa}, z''e^{\sigma'' \kappa}, \cdots) \qquad (76)$$

are in effect independent of κ for any values of the z's. Differentiating with respect to κ and setting it equal to zero it follows that

$$\{\sigma' z' \partial/\partial z' + \sigma'' z'' \partial/\partial z'' + \cdots\}\gamma(z', z'', \cdots) = 0. \qquad (77)$$

In particular, this is true for $\gamma = \gamma_0$ and now the stated conclusion follows from Eqs. (5) and Theorem 2.

The charge neutrality condition implies that if all nonvanishing[24] σ values are of the same sign the corresponding densities vanish. This is quite obvious physically: If only charges of one sign are admitted through the "membrane" separating the system from the reservoir [cf. discussion of Sec. 2], a "sheath" builds up which prevents particles to enter beyond the sheath region into the system. It is less obvious what mathematical property of the characteristic value problem is responsible for this radically different behavior of the case of all σ's having one sign. To clear this up it is convenient to reformulate the problem in terms of the complex variable $u = e^{i\phi}$. We then have

$$\{-u^2 d^2/du^2 - u\, d/du + z'u^{\sigma'} + z''u^{\sigma''} + \cdots\}y(u) = \gamma y(u), \qquad (78)$$

and the periodicity in ϕ is replaced by the demand that $y(u)$ be *single valued* in u. Now, the differential equation (78) has two singular points, $u = 0$ and $u = \infty$. They are of the irregular type when σ's with both signs occur; $u = 0$ is of the regular type, $u = \infty$ irregular, when all σ's are positive; and the other way round when all σ's are negative.[25] To be specific, let all of them be positive. Then two independent solutions of the equation exist[25] of the form

$$y(u) = u^{\sqrt{-\gamma}}\, f(u), \qquad (79)$$

where $f(u)$ is regular in the neighborhood of the origin. Evidently this is single valued only when

$$\gamma = -m^2, \quad (m = 0, 1, 2, \cdots). \qquad (80)$$

The remarkable fact is that the characteristic values are *independent* of the z's. In particular $\gamma_0 = 0$ which, as we have seen, implies the vanishing of the densities as well as the pressure. This analysis shows that certain striking features of the characteristic value problem, formulated at first in the real domain, are uncovered easiest by going into the complex plane. We have here the example of a problem where the "perturbation" does not affect the spectrum of characteristic values.

The second general question concerns the continuity of thermodynamic functions. We have seen how the latter are derivable from γ_0 given as a function of z', z'', etc. γ_0 has been defined as the largest real characteristic value for any given values of the parameters, but in some respect this is not a natural labeling. Characteristic values may be labeled alternatively by means of their association with the functions $y(\phi)$ which change smoothly as

[23] Reference 19, p. 194.

[24] Vanishing charge numbers mean added ideal gas components without interaction. This is easily seen to be compatible with Theorem 2.

[25] Reference 19, p. 197.

the z parameters are varied from zero up. In such a labeling one of the γ's tends to zero with decreasing z's which is thus γ_0 for sufficiently small z values. But it is in principle possible that a lower value may with increasing z's, actually "overtake" γ_0. It is evident that if this happens "the largest characteristic value" suffers a discontinuity in at least one of its z derivatives which leads to a discontinuity of some of the densities [cf. Eqs. (5)]. As the pressure is continuous, we get in the pressure-density diagram the horizontal portion typical of a phase transition.

Does this happen? We do not have a definite answer to this question. For a one-dimensional system of particles interacting with forces of finite range the answer is negative,[26] but our system is not of this type and a special investigation is needed. We have a partial result in this direction, and in order to formulate and prove it we need to broaden somewhat the characteristic value problem (62), namely, by replacing the condition of periodicity by

$$y(\phi + 2\pi) = \pm y(\phi). \tag{81}$$

Characteristic values where the plus sign applies will be called of type I, where the minus sign applies, of type II.

Theorem 4. *Let $\mathfrak{R}$ be a connected region in the space of z', z'', $\cdots$ which includes the point $z' = z'' = \cdots = 0$ and which is such that for all of its points all characteristic values (of both types) are real. Then within $\mathfrak{R}$ the largest characteristic value is not crossed by another one.*

When all z's vanish $F(\phi) = 0$ and the problem is trivial. The characteristic values of type I are given be (80) and those of type II by

$$\gamma = -(m + \tfrac{1}{2})^2, \qquad (m = 0, 1, 2, \cdots). \tag{82}$$

All of them are doubly degenerate except the single value $\gamma_0 = 0$ of type I. Moreover, those of type I alternate on the real line with those of type II. It is clear that when the z's are varied in $\mathfrak{R}$ the largest γ cannot be crossed by another one without this ordering being violated and if that is so there is a point in $\mathfrak{R}$ at which one of the characteristic values of type II coincides with one of type I (which may or may not be the largest one). But this is impossible because the two corresponding characteristic functions would have to satisfy

$$(d/d\phi)(y_{\mathrm{I}}\, dy_{\mathrm{II}}/d\phi - y_{\mathrm{II}}\, dy_{\mathrm{I}}/d\phi) = 0, \tag{83}$$

but the quantity inside the bracket is not a constant. It changes sign upon increasing the argument by 2π. This proves the theorem.

As a minimum the region $\mathfrak{R}$ includes those points where $z_1 = z_{-1}$, $z_2 = z_{-2}$, etc. (the subscripts are the values or the corresponding σ's). Then the problem (62) is real, thus also Hermitian, so that all γ's are real. Within this region then no phase transition can occur. It must be remembered that due to the invariance under the transformation (75) actually more points belong to $\mathfrak{R}$: all of those for which

$$z_1/z_{-1} = (z_2/z_{-2})^{\frac{1}{2}} = \cdots . \tag{84}$$

By a transformation of the form (75) the common value of these ratios may be brought to unity. In the case studied in I, namely a two-component system with charges equal in magnitude and opposite in sign, this condition is fulfilled trivially. We have excluded the possibility of a phase transition in those cases when a "detailed" balancing of charge obtains, i.e., a balancing for any pair of components carrying a charge of the same magnitude and opposite sign. Of course, the region $\mathfrak{R}$ may be larger, and it would be interesting to know how large, and also (if it is smaller than the totality of all z values) whether the possibility of a phase transition can be excluded on grounds independent of Theorem 4. It goes without saying that if a phase transition is possible,[27] a proof would be of great interest, as well as an analysis of its physical nature.

6. THE PLASMA LIMIT

In this section we restrict our considerations to the simplest case studied in I. Thus, we have only two kinds of particles with $\sigma = \pm 1$, and the corresponding values of z may be assumed equal without loss of generality. The problem is then to solve for the function $Q(\phi, x)$ which is the fundamental solution of the equation

$$\{\partial/\partial x - \partial^2/\partial\phi^2 - 2z\cos\phi\}Q(\phi, x) = 0. \tag{85}$$

The problem depends parametrically on the quantity z,[28] and we propose to investigate it in the limit of large z. As pointed out in I,[29] this limit corresponds to the true "plasma state," i.e., a state when the charges move independently of one another in the first approximation, and the kinetic energy is large compared to the potential energy.

[26] L. Van Hove, Physica 16, 137 (1950).

[27] We do not think this is likely.

[28] $z/D\sigma^2 = \theta z/2\pi\sigma^2$ in ordinary units.

[29] Section 8.

In I the treatment was based on the asymptotic analysis of the characteristic value problem (the Mathieu equation). Here we want to show that a more direct, even if somewhat heuristic, treatment of Eq. (85) gives easily the dominant behavior in the limit of large z.

It is convenient to split off a factor and write

$$Q(\phi, x) = e^{2zx}Q_1(\phi, x), \tag{86}$$

so that Q_1 is the fundamental solution of

$$\{\partial/\partial x - \partial^2/\partial\phi^2 + 4z\sin^2\tfrac{1}{2}\phi\}Q_1(\phi, x) = 0. \tag{87}$$

Now that the coefficient function is non-negative, the equation can be interpreted as a diffusion problem with absorption. The "absorption coefficient" $4z\sin^2\frac{1}{2}\phi$ is large everywhere except in the immediate neighborhoods of the points $\phi = 2\pi n$ (n integer). Hence, Q_1 can be expected to be very small everywhere except in these neighborhoods. But again, all of these except $\phi = 0$ can be ignored because initially Q_1 has a δ singularity at $\phi = 0$, and the other points with no absorption are separated from this by regions of very strong absorption. Thus, it seems reasonable that a good approximation to Q can be obtained by replacing (87) by the equation

$$\{\partial/\partial x - \partial^2/\partial\phi^2 + z\phi^2\}Q_1(\phi, x) = 0, \tag{88}$$

which represents correctly the behavior of the absorption coefficient near $\phi = 0$. But this problem has the exact analytic solution

$$Q_1(\phi, x) = \left(\frac{z^{1/2}}{2\pi\sinh(2z^{1/2}x)}\right)^{1/2}\exp\left\{-\frac{z^{1/2}\phi^2}{2\tanh(2z^{1/2}x)}\right\} \tag{89}$$

as may be verified by substitution. In addition, the integration over ϕ can be carried out explicitly and one obtains

$$\Omega = e^{2zL}/(\cosh 2z^{\frac{1}{2}}L)^{\frac{1}{2}}. \tag{90}$$

This is only an approximation, of course, the true Ω being larger because the "absorpton coefficient" has been increased in the transition from Eq. (87) to (88). Remembering the definition of γ_0 we get in the same approximation

$$\gamma_0 = 2z - z^{\frac{1}{2}}. \tag{91}$$

This is the beginning of the asymptotic development of γ_0 in inverse powers of $z^{1/2}$ quoted in I.[29] The first term leads to the ideal gas law, the next to the "Debye-Hückel correction."[30]

It does not seem easy to improve on the formula (91) with the present method. The root of the difficulty is that we are not looking for an approximation to $Q(\phi, x)$, nor even to $\Omega = \Omega(L)$, but rather to the ultimate exponential rate of growth of Ω as $L \to \infty$. That this should not be easy is not surprising because the asymptotic analysis of the periodic solution of the Mathieu equation is involved and that is quite an elaborate affair.[31] Nevertheless, it is gratifying that at least the dominant behavior can be understood with the heuristic discussion of this section. If Eq. (88) is attacked by the method of separation of variables the characteristic value problem for the Hermite functions results.[32] That the analysis in the limit of large z leads to this problem has already been observed by Prager.[2] An approach to the Mathieu problem based on the Hermite functions as a first approximation is given by Morse and Feshbach.[33]

7. THE REDUCED DENSITY FUNCTIONS

We come now to a subject which demonstrates forcefully the superiority of the functional integration method over the direct treatment of I. We shall construct explicit and exact expressions for the reduced density functions.

These functions form a hierarchy of ever increasing complexity and are defined as follows. $f_{\sigma_1}(x_1)\,dx_1$ is the probability that the element dx_1 is occupied by a particle of charge σ_1; $f_{\sigma_1\sigma_2}(x_1x_2)\,dx_1\,dx_2$ is the joint probability that the two elements dx_1 and dx_2 are occupied by particles of charge σ_1 and σ_2, respectively; and so on. In contrast to the probability densities (2) the function f refer to occupation probabilities *regardless of the presence of other particles* in the system volume. There is a connection between the f's and the probability densities (2) which can be written down in the most concise form by making use of generating functionals introduced by Bogoliubov[34] and recently discussed by Green.[35] Let us introduce arbitrary functions $\zeta_\sigma(x)$, $\zeta_\sigma(x)$, etc., one associated

[29] For a discussion of the validity of the Debye-Hückel treatment see N. Balazs, Phys. Fluids 4, 1259 (1961).

[31] E. L. Ince, Proc. Roy. Soc. Edinburgh 46, 316 (1926). Ince wrote "If anyone had the courage to push the development a stage or two further he would greatly enhance the value of an important expansion. But any reader who attempts to verify [my] results given above will realize that the work involved would be tremendous."

[32] The equation differs only in a factor i in front of the first term from the Schrödinger equation for a harmonic oscillator.

[33] P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill Book Company, Inc., New York, 1953), p. 1416.

[34] N. N. Bogoliubov, *Problems of a Dynamical Theory in Statistical Physics* (in Russian). Also available in English translation by E. K. Gora, *Studies in Statistical Mechanics* (Interscience Publishers, Inc., New York, 1962), Vol. I.

[35] M. S. Green, "Some Applications of the Generating Functional of the Molecular Distribution Functions" in Vol. III of *Lectures in Theoretical Physics* held at the University of Colorado in 1960 (Interscience Publishers, Inc., New York, 1961).

with each species of particles, and form the quantity

$$\Phi = \frac{1}{\Omega}\sum_{N'=0}^{\infty}\sum_{N''=0}^{\infty}\cdots\frac{1}{N'!\,N''!\cdots}\int_0^L dx_1\cdots \times\int_0^L dx_N\prod_{k=1}^{N}[z_k+\zeta_{\sigma_k}(x_k)]\exp\left\{-\frac{V_{N'N''\cdots}}{\theta}\right\}. \quad (92)$$

It is to be regarded as a functional of the ζ's. We have then the following formulas for the reduced density functions

$$f_{\sigma_1}(x_1) = z_1\,\delta\Phi/\delta\zeta_{\sigma_1}(x_1),$$
$$f_{\sigma_1\sigma_2}(x_1x_2) = z_1z_2\frac{\delta^2\Phi}{\delta\zeta_{\sigma_1}(x_1)\,\delta\zeta_{\sigma_2}(x_2)}, \quad (93)$$

and so on. It is understood that after the functional differentiation all ζ's are set identically equal zero.[34,35]

We now make the important observation that Φ (apart from the factor Ω^{-1}) has a structure very closely resembling that of the partition function itself. In fact, the difference is only that with each integration variable x_k we now have a factor $z_k+\zeta_{\sigma_k}(x_k)$ instead of just z_k as before. Therefore the analysis of Sec. 4 which led us to the formula (35) can be duplicated with these obvious and simple changes. The analog of Eq. (35) for Φ is then as follows

$$\Phi = \frac{1}{\Omega}\left\langle\exp\int_0^L dxF(\phi(x),x)\mid\phi(0)=0\right\rangle \quad (94)$$

where the function F is now

$$F(\phi,x) = [z'+\zeta_{\sigma'}(x)]e^{i\sigma'\phi} + [z''+\zeta_{\sigma''}(x)]e^{i\sigma''\phi}+\cdots. \quad (95)$$

The functional dependence of Φ on $\zeta_{\sigma'}$, $\zeta_{\sigma''}$, etc., comes through the functional dependence of F. The formula (94) together with (95) represents a *closed, exact expression* for the Bogoliubov functional of the reduced density functions in the one-dimensional plasma.

The theorem of Kac [cf. Sec. 3] could now be applied thus expressing Φ in terms of the fundamental solution of a partial differential equation. The method of separation of variables is inapplicable, however, because the coefficient function (35) is explicitly dependent on x through the functions $\zeta_{\sigma'}$, $\zeta_{\sigma''}$, etc. Fortunately this causes no difficulty because we are really not interested in Φ but rather in its functional derivatives with respect to the ζ's. For these a different method of evaluation suggests itself.

Since the integral over F occurs as the argument of the exponential function, every functional differentiation with respect to

$$\zeta_{\sigma_1}(x_1),\qquad \zeta_{\sigma_2}(x_2),\quad \text{etc.},$$

brings down factors

$$e^{i\sigma_1\phi(x_1)},\qquad e^{i\sigma_2\phi(x_2)},\quad \text{etc.} \quad (96)$$

Thus, we have

$$f_{\sigma_1\sigma_2\cdots}(x_1x_2\cdots) = \frac{z_1z_2\cdots}{\Omega}\left\langle\exp\left\{i\sigma_1(x_1)+i\sigma_2(x_2)+\cdots\int_0^L dxF(\phi(x))\right\}\,\middle|\,\phi(0)=0\right\rangle, \quad (97)$$

where F is now just (36) since the ζ's are to be put zero after the functional differentiation with respect to them. It is clearly possible to arrange the notation so that the successive arguments of f form an increasing sequence, provided we leave the corresponding charges arbitrary. Thus, it will be understood that

$$0 < x_1 < x_2 < \cdots < L. \quad (98)$$

We now break up the integral in the exponent into a sum

$$\int_0^L dx = \int_0^{x_1} dx + \int_{x_1}^{x_2} dx + \cdots. \quad (99)$$

The Wiener average is over all continuous functions such that $\phi(0)=0$. We carry out this averaging by making, initially, a number of additional restrictions $\phi(x_1)$ in $d\phi_1$, $\phi(x_2)$ in $d\phi_2$, etc., and afterward integrating over all ϕ_1, ϕ_2, etc. In the restricted Wiener integrals the factors (96) are just constants, and according to (99) we just have product of factors of the form

$$\left\langle\exp\left\{\int_{x_{k-1}}^{x_k} dxF(\phi(x))\right\}\,\middle|\,\phi(x_{k-1}) = \phi_{k-1},\ \phi(x_k)=\phi_k\right\rangle. \quad (100)$$

We have seen in Sec. 3 that this is just $Q(\phi_k,\phi_{k-1},x_k-x_{k-1})$, the function Q being defined in Theorem 1. Thus, we obtain the following results

$$f_{\sigma_1}(x_1) = \frac{z_1}{\Omega}\int_{-\infty}^{\infty}d\phi\int_{-\infty}^{\infty}d\phi_1 \times Q(\phi,\phi_1,L-x_1)e^{i\sigma_1\phi_1}Q(\phi_1,0,x_1), \quad (101)$$

$$f_{\sigma_1\sigma_2}(x_1x_2) = \frac{z_1z_2}{\Omega}\int_{-\infty}^{\infty}d\phi\int_{-\infty}^{\infty}d\phi_2\int_{-\infty}^{\infty}d\phi_1 \times Q(\phi,\phi_2,L-x_2)e^{i\sigma_2\phi_2}Q(\phi_2,\phi_1,x_2-x_1) \times e^{i\sigma_1\phi_1}Q(\phi_1,0,x_1), \quad (102)$$

and so on. When the σ's are integers (which will be assumed in the following) the function Q may be replaced in these formulas by

$$\hat{Q}(\phi, \phi', x) = \sum_{n=-\infty}^{\infty} Q(\phi + 2\pi n, \phi', x) \tag{103}$$

and the integration limits changed to $-\pi$ and π. $\hat{Q}$ is periodic in *both* variables ϕ and ϕ', and its expansion in terms of the characteristic functions of the problem (62) is

$$\hat{Q} = \sum_{m=0}^{\infty} e^{\gamma_m x} y_m(\phi) y_m(\phi'). \tag{104}$$

The results we have obtained are exact but they do not quite correspond to the customary definitions of the density functions. The reason is that our functions depend not only on all position variables, but even on L, the size of the system. In the customary definition the limiting form of these functions is meant, namely, the limit

$$\begin{aligned} &x_1 \to \infty, \\ &L - x_n \to \infty, \\ &x_2 - x_1, \quad x_3 - x_2, \cdots, x_n - x_{n-1} \quad \text{fixed.} \end{aligned} \tag{105}$$

These limiting functions truly express the bulk properties of the model inasmuch as "surface effects" have been removed. It is quite easy to carry out the limit (105) provided one notes that as $x \to \infty$

$$\hat{Q}(\phi, \phi', x) \sim e^{\gamma_0 x} y_0(\phi) y_0(\phi') \tag{106}$$

and of course the corresponding limiting form of Ω.[36] The results are

$$f_{\sigma_1} = z_1 \int_{-\pi}^{\pi} d\phi_1 y_0(\phi_1) e^{i\sigma_1\phi_1} y_0(\phi_1), \tag{107}$$

$$f_{\sigma_1\sigma_2} = z_1 z_2 e^{-\gamma_0(x_2 - x_1)} \int_{-\pi}^{\pi} d\phi_2 \int_{-\pi}^{\pi} d\phi_1 y_0(\phi_2) e^{i\sigma_2\phi_2} \times \hat{Q}(\phi_2, \phi_1, x_2 - x_1) e^{i\sigma_1\phi_1} y_0(\phi_1), \tag{108}$$

$$\begin{aligned} f_{\sigma_1\sigma_2\sigma_3} = z_1 z_2 z_3 e^{-\gamma_0(x_3 - x_1)} &\int_{-\pi}^{\pi} d\phi_3 \int_{-\pi}^{\pi} d\phi_2 \int_{-\pi}^{\pi} d\phi_1 \\ &\times y_0(\phi_3) e^{i\sigma_3\phi_3} \hat{Q}(\phi_3, \phi_2, x_3 - x_2) \\ &\times e^{i\sigma_2\phi_2} \hat{Q}(\phi_2, \phi_1, x_2 - x_1) e^{i\sigma_1\phi_1} y_0(\phi_1), \end{aligned} \tag{109}$$

and so on. These functions are manifestly translation invariant. Furthermore, they satisfy the "product condition." This means that if a group of arguments is far removed from another group than the f corresponding to the whole group is the product of the two f's corresponding to the two separated groups. This is shown by making use of the asymptotic relation (106) for that factor $\hat{Q}$ whose x argument is the separation between the two groups.

Let us consider the simplest of these functions. $f_\sigma \cdot dx$ is the probability that an element dx is occupied by a particle of charge σ', therefore, f_σ should also be the mean number density for this species. This, we have seen, is

$$n' = z' \, \partial\gamma_0/\partial z'. \tag{110}$$

Let us show that the right-hand sides of Eqs. (107) and (110) are identical. We have

$$d^2 y_0/d\phi^2 + F y_0 = \gamma_0 y_0. \tag{111}$$

Differentiate this with respect to the parameter z'.

$$\frac{d^2}{d\phi^2}\left(\frac{\partial y_0}{\partial z'}\right) + \frac{\partial F}{\partial z'} y_0 + F \frac{\partial y_0}{\partial z'} = \frac{\partial \gamma_0}{\partial z'} y_0 + \gamma \frac{\partial y_0}{\partial z'} \tag{112}$$

Now multiply (111) by $\partial y_0/\partial z'$, (112) by y_0 and integrate them between $-\pi$ and π. The difference of the two equations so resulting is just

$$\int_{-\pi}^{\pi} \frac{\partial F}{\partial z'} y_0^2 \, d\phi = \frac{\partial \gamma_0}{\partial z'} \int_{-\pi}^{\pi} y_0^2 \, d\phi \tag{113}$$

which, in view of the normalization of y_0, is the identity desired.

Let us consider now the two particle density functions (108). By means of the expansion (104) we obtain[37]

$$f_{\sigma_1,\sigma_2}(x) = \sum_{m=0}^{\infty} B_{m,\sigma_1} B_{m,\sigma_2} e^{-(\gamma_0 - \gamma_m)x}. \tag{114}$$

The coefficients are

$$B_{m,\sigma} = z_\sigma \int_{-\pi}^{\pi} d\phi y_0(\phi) y_m(\phi) e^{i\sigma\phi}. \tag{115}$$

Clearly f_{σ_1,σ_2} is symmetric in the two subscripts which run over all particle species.

Certain general properties of these functions will now be shown under the hypothesis of "detailed charge neutrality." This means that the species occur in pairs of equal and opposite charge and the corresponding values of z are equal

$$z_\sigma = z_{-\sigma}. \tag{116}$$

This implies that the corresponding densities are also equal

$$n_\sigma = n_{-\sigma}. \tag{117}$$

In more generality, all reduced density functions are invariant under a simultaneous change of sign of all their charge subscripts.

[36] We ignore here the possibility of degeneracy for the largest characteristic value.

[37] The variable x is here always taken positive. This is no restriction on generality.

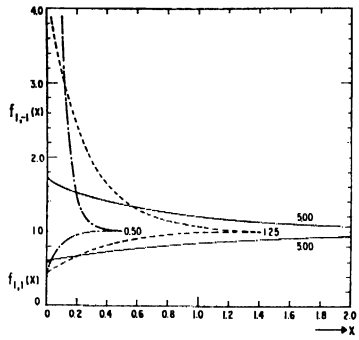


FIG. 1. Two particle density functions.

Theorem 5. *Suppose there is detailed charge neutrality. Let ξ_σ be arbitrary real numbers, one associated with each particle species. Then the function*

$$\sum_{\sigma_1} \sum_{\sigma_2} \xi_{\sigma_1}\xi_{\sigma_2} f_{\sigma_1,-\sigma_2}(x) \tag{118}$$

is positive and monotone decreasing. To prove this, note that in the case of detailed charge neutrality

$$F(\phi) = 2z_1 \cos\phi + 2z_2 \cos 2\phi + \cdots, \tag{119}$$

so that the basic characteristic value problem (62) is real. Hence, in this case, all γ_m are real as well as all $y_m(\phi)$. But then

$$B_{m,-\sigma} = B^*_{m,\sigma}, \tag{120}$$

and therefore the function (117) is

$$\sum_{m=0}^{\infty} |\sum_{\sigma} \xi_\sigma B_{m,\sigma}|^2 e^{-(\gamma_0-\gamma_m)x}. \tag{121}$$

The conclusion of the theorem is hereby exhibited.

The physical significance is best appreciated by a choice of suitable values for the ξ_σ. It is easy to see, for instance, that the functions $f_{\sigma,-\sigma}(x)$ and $f_{\sigma,\sigma}(x) + f_{\sigma,-\sigma}(x)$ are monotone decreasing. Also $f_{\sigma,-\sigma}(x) - f_{\sigma,\sigma}(x)$ are positive and monotone decreasing. The latter fact means that *it is more likely for two particles, a given distance apart, to have equal and opposite charges than the same charge, and the difference decreases with distance.* This is quite reasonable. It is a manifestion of the attraction (repulsion) between charges of opposite (same) sign.

There is a more subtle effect connected with the tendency for the two particles in dx_1 and dx_2 to have the same *magnitude* of their charge. To show this, we consider any even function of the charge σ, such as $|\sigma|$ for instance, and calculate its covariance at the two points x_1 and $x_2 = x_1 + x$. Let ξ_σ be this function, and let us denote by $\langle\cdots\rangle_{\text{av}}$ averaging over charges. The covariance is

$$\begin{aligned}\langle(\xi_{\sigma_1} - \langle\xi_{\sigma_1}\rangle_{\text{av}})(\xi_{\sigma_2} - \langle\xi_{\sigma_2}\rangle_{\text{av}})\rangle_{\text{av}} \\ = \langle\xi_{\sigma_1}\xi_{\sigma_2}\rangle_{\text{av}} - \langle\xi_{\sigma_1}\rangle_{\text{av}}\langle\xi_{\sigma_2}\rangle_{\text{av}} \\ = \sum_{\sigma_1}\sum_{\sigma_2} \xi_{\sigma_1}\xi_{\sigma_2}[f_{\sigma_1,\sigma_2}(x) - f_{\sigma_1}f_{\sigma_2}].\end{aligned} \tag{122}$$

Since $\xi_\sigma = \xi_{-\sigma}$ we can ignore the minus sign in (118), and now the theorem implies that the expression (122) is positive, since it obviously tends to zero as $x \to \infty$. Thus ξ_σ *at the point* x_1 *and* ξ_σ *at the point* x_2 *are positively correlated random variables.*

Certain rather plausible properties of the two particle density functions are not implied by our theorem. In particular, no assertion is made about the functions $f_{\sigma,\sigma}(x)$ although it seems likely on physical grounds that they are monotone increasing functions of x. To check this, we have performed numerical computations in the simplest case of a two-component plasma with $\sigma = \pm 1$ and $z_1 = z_{-1}$. In this case the characteristic value problem is just the Mathieu equation

$$d^2y(\phi)/d\phi^2 + 2z\cos\phi y(\phi) = \gamma y(\phi), \tag{123}$$

and detailed tables are available on the periodic solutions as well as the corresponding values of γ.[38] The Mathieu functions are given in terms of the coefficients of their Fourier series which converge rapidly. Therefore the integrals $B_{m,\sigma}$ of Eq. (115) can be also expressed as fast converging series with tabulated numerical coefficients. The decrease with m is also fairly rapid, although this depends on the value of z,[28] and becomes worse with increasing z. We present the results on the accompanying figure for three representative values of z. For convenience of physical interpretation the unit of distance is the mean interparticle distance n^{-1}. The decreasing functions are $f_{1,-1}(x)$, the increasing ones $f_{1,1}(x)$. We note that for larger z the approach to the asymptotic value is slower and the deviation from it smaller at small distances. For very small values of z the approach to the asymptotic value takes place in a distance small compared to the mean interparticle distance. This is a manifestation of the "condensation" into pairs of oppositely charged particles in this limit. The distance between members of a pair becomes small compared to the distance between pairs. This is a confirmation of the picture of the system obtained in I on thermodynamic grounds.[39]

[38] National Bureau of Standards, *Tables Relating to Mathieu Functions* (Columbia University Press, New York, 1951).

[39] Section 9.

Reprinted from *Printed in USA.*
JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 4 APRIL 1963

Exact Statistical Mechanics of a One-Dimensional System with Coulomb Forces. III. Statistics of the Electric Field*

ANDREW LENARD
Plasma Physics Laboratory Princeton University, Princeton, New Jersey
(Received 19 November 1962)

The statistics of the electric field in a one-dimensional "sheet" plasma is studied. It is shown that the electric field as a function of the space coordinate is a Markov process, whose basic probabilities are related to the Fourier coefficients of certain functions that play an important role in the previously found development of the grand partition function. Some implications regarding the statistics of particle locations are explored. The one-point probability distribution of the electric field is found to be asymptotically Gaussian in the plasma limit. It is pointed out that the one-dimensional analog of the Holtsmark calculation leads to an incorrect conclusion because electrostatic shielding is not properly taken into account.

Comment

For further discussion and utilization of the technique of functional integration see paper by E. Helfand

Comment on the Indirect Problem of Statistical Physics

The general problem of statistical physics, which is discussed in this book, has been formulated as:[1] "Given the structure and the laws of interaction of the molecules, what are the properties of the matter composed of these molecules?". The starting point of a particular investigation is a molecular model subject to certain microscopic force laws. The final aim is the specification of the thermodynamic properties and distribution functions of the macroscopic system. We shall call this the direct problem.

In distinction to this we can formulate indirect problems. Roughly speaking, we can state these as follows: Given, as a starting point, a model class of functional relations between the values of the macroscopic properties of a hypothetical Newtonian molecular system, what is the functional class of Hamiltonians which is consistent with the given information? Accepting the usual expression for the kinetic energy of the molecules and restricting ourselves to the case where the only external forces acting on the molecules are such as to keep the molecules inside the volume

---

[1] G. E. Uhlenbeck and G. W. Ford, "Lectures in Statistical Mechanics", American Mathematical Society, Providence, Rhode Island, 1963.

of their thermostatted container, the final aim of such an investigation is the specification of the intermolecular potential energy. Loosely speaking the explicit solution of the indirect problem is tantamount to the solution of the direct problem for the found potential.

To give an illustration, given the pressure $p = p(\rho,T)$ as function of ρ and T of a monoatomic molecular system can one find the class of intermolecular potentials consistent with this information? In particular, one can ask whether on restricting the class of potentials to pairwise additive, central potentials of the van der Waals type, there exists a unique intermolecular potential function, $\varphi = \varphi(r)$, r the distance between a pair of molecules, which corresponds to the given equation of state $p = p(\rho,T)$. An affirmative answer to this question can be given[2,3] in the special case that a single-valued inverse function $r = \tilde{r}(\varphi)$ exists (i.e., $\varphi(r)$ is purely repulsive or purely attractive) whose cube is Laplace transformable.

One requires in this case only the second virial coefficient as a function of $\beta = 1/kT$, where[3]

$$B(\beta) = \lim_{\rho\to 0} [p\beta/\rho - 1]$$

$$= -(2\pi\beta/3)\int_0^\infty dr\, r^3 \exp(-\beta\varphi)(d\varphi/dr), \qquad (1)$$

---

[2] J. B. Keller and B. Zumino, J. Chem. Phys. 30, 1351 (1959).

[3] H. L. Frisch and E. Helfand, J. Chem. Phys. 32, 269 (1960).

as can be easily verified from the virial equation of state. Introducing the inverse function $r = \tilde{r}(\varphi)$, we find

$$-\frac{3\beta B(\beta)}{2\pi} = \int_0^\infty \tilde{r}^3(\varphi) e^{-\beta\varphi} d\varphi = \mathfrak{L}\{\tilde{r}^3(\varphi)\} \tag{2}$$

and thus we obtain the explicit result,

$$r = \tilde{r}(\varphi) = \sqrt[3]{-\frac{3}{2\pi}\,\mathfrak{L}^{-1}\{\beta B(\beta)\}}, \tag{3}$$

with $\mathfrak{L}^{-1}\{\beta B(\beta)\}$ the inverse Laplace transform of $\beta B(\beta)$. An interesting application of these results is to van der Waals like equations of state since one can thus easily demonstrate that they correspond to long-range potentials.

If φ is of the van der Waals type[1], and has one repulsive portion $\varphi_+(r)$, $d\varphi_+(r)/dr < 0$, and one attractive portion $\varphi_-(r)$, $d\varphi_-(r)/dr$, with single-valued inverses $r = \tilde{r}_+(\varphi)$ and $r = \tilde{r}_-(\varphi)$, where $\tilde{r}_-(\varphi) = 0$ for $\varphi \geq -\varepsilon$, (2) is replaced by

$$-\frac{3\beta B(\beta)}{2\pi} = \int_0^\infty \left[\tilde{r}_+^3(\varphi) - \tilde{r}_-^3(\varphi)\right] e^{-\beta\varphi} d\varphi. \tag{4}$$

An inversion gives the volume contained between the points of equal potential on the attractive and repulsive branch, and gives the complete potential for values greater than ε.[2,3]

It is apparently not known whether a finite number of virial coefficients, the complete virial series or knowledge of the equation of state $p = p(\rho,T)$ suffices to uniquely determine a van der Waals type of φ. Certain other conditions imposed by gross properties (nonequilibrium as well as equilibrium) are discussed in Ref. 3.

An equally open indirect question concerns what class of intermolecular potentials is consistent with given values of the algebraic geometric invariants of the phase diagram (e.g., the number of coexistence lines, the number of triple or critical points, etc.), and what can be said about the stability of the invariants with respect to parameter variations in the intermolecular potential. Such questions may not be easy to answer. Yet the sometimes successful heuristic device of working backwards from a possible answer suggests that some attention to these problems may be useful. We should also mention here that explicit inversion formulas for the one-and two-body potentials in terms of the one- and two-body distribution functions have been given by Morita and Hiroiki and De Dominicis (cf. article by G. Stell). Their expressions are formal infinite series whose convergence is not established (c.f. Sec. II of paper by Percus).

SECTION II

MAY, 1935 JOURNAL OF CHEMICAL PHYSICS VOLUME 3

Statistical Mechanics of Fluid Mixtures*

JOHN G. KIRKWOOD, *Department of Chemistry, Cornell University*

(Received March 1, 1935)

Expressions for the chemical potentials of the components of gas mixtures and liquid solutions are obtained in terms of relatively simple integrals in the configuration spaces of molecular pairs. The molecular pair distribution functions appearing in these integrals are investigated in some detail, in their dependence upon the composition and density of the fluid. The equation of state of a real gas mixture is discussed, and an approximate molecular pair distribution function, typical of dense fluids, is calculated. Applications of the method to the theory of solutions will be the subject of a later article.

I.

ALTHOUGH statistical mechanics has met with considerable success in the interpretation of the properties of rarefied gases as well as those of crystalline solids, serious practical difficulties have retarded its application to dense gases and to liquids. In the present article, we shall describe a statistical treatment of fluid mixtures, which is formally applicable both to real gases and to liquids. It has the advantage of yielding expressions which lend themselves more readily than the usual statistical formulas, to approximations suitable for condensed systems. While we make no pretension to a resolution of the difficulties presented by the liquid state, the more serious of which resist more than a superficial treatment, we believe that the method to be described throws interesting light upon the problem, and that it may prove useful in opening up new lines of attack, particularly in the field of liquid solutions.

We must first ask whether classical statistical mechanics is competent to describe the dependence of the thermodynamic functions of fluids upon intermolecular forces. At ordinary temperatures, there is good reason to believe that the classical approximation is adequate for most gases and liquids consisting of chemically saturated molecules. When certain conditions involving the molecular masses, moments of inertia, and the intermolecular forces are satisfied, the partition function of a system of molecules is closely approximated by the product of a function of the temperature alone, and the Gibbs phase integral in the phase space associated with the low frequency translational and rotational motion of the molecules.[1] The validity of the assumptions involved in this approximation is of course more questionable for liquids than for gases, particularly when there is strong coupling between the rotational motion of neighboring molecules as in polar liquids. We shall provisionally assume the adequacy of the classical approximation, remembering that our results are applicable only when this is true.

When the potential of the intermolecular force is of special form, for example a sum of terms depending upon the relative coordinates of molecular pairs, the explicit use of unwieldy integrals in a many dimensional phase space may be avoided. We shall show how the statistical formulas for certain important thermodynamic functions of a fluid mixture may be recast in terms of comparatively simple integrals in the configuration space associated with the relative motion of a molecular pair. The molecular pair distribution functions, which appear in these integrals are, to be sure, formally defined as ratios of many dimensional phase integrals. However, by certain transformations, they may be expressed in a form which leads to reasonable approximations both in condensed and rarefied systems. The approximation for rarefied systems leads to a novel approach to the problem of the equation of state of a real gas, while the approximation for condensed system promises to be of use in the theory of liquid solutions. On the other hand, the approximate distribution functions which have thus far been obtained do not

* The material of Section V of this article was presented at the eighty-ninth meeting of the American Chemical Society held in New York City during the week of April 22, 1935.

[1] E. Wigner, Phys. Rev. **40**, 749 (1932); Uhlenbeck and Gropper, Phys. Rev. **41**, 79 (1932); J. G. Kirkwood, Phys. Rev. **44**, 31 (1933); J. Chem. Phys. **1**, 597 (1933).

appear to be sufficiently exact for the calculation of the equations of state of liquids.

The method of the present treatment is based upon the same principle as the charging processes employed in the theory of electrolyte solutions. As Onsager has clearly pointed out, parameters which appear in the potential of the intermolecular forces have essentially the same status as the parameters of external force,[2] which determine the mechanical interaction between a thermodynamic system and its surroundings, and may be manipulated in the same manner. The application of this principle is by no means restricted to intermolecular forces of the Coulomb type, in the potential of which ionic charges enter as natural parameters. Arbitrary parameters may be introduced into the potential of intermolecular force in any system of molecules, by means of which the coupling between the molecules may be continuously varied in any convenient manner. This device allows one to express the chemical potential of any component of a fluid mixture in a particularly simple form. It also proves to be very useful in investigating the dependence of molecular pair distribution functions upon the composition and density of the fluid

II.

The following analysis is formally applicable to all homogeneous fluids, whether real gas mixtures or liquid solutions. Differences arise only in the method of approximating distribution functions. We therefore consider a homogeneous fluid of ν substances, in thermodynamic equilibrium. By suitably restricting the ranges of composition and pressure so as to exclude coexistent phases containing any of its components, we may suppose that the fluid completely fills a vessel of volume v, which confines the motion of the system to a finite region of configuration space.[3] We denote the absolute temperature by T, the numbers of molecules of the several components by $N_1, \cdots, N_\nu$, and the total number of molecules of all species, $\sum_{s=1}^{\nu} N_s$, by N. According to classical statistical mechanics, we represent the system by a canonically distributed petit ensemble, and obtain the usual expression for the Helmholtz free energy, F_N.

$$e^{-\beta F_N} = \left[\prod_{s=1}^{\nu} \frac{f_s(T)^{N_s}}{N_s!}\right] Z_N$$

$$Z_N = \int \cdots \int e^{-\beta V_N} d\omega_1 \cdots d\omega_N, \qquad (1)$$

where β is equal to $1/kT$ and $d\omega_i$ is a differential element of the configuration space associated with the low frequency rotational and translational degrees of freedom of the molecule i. The integration in the sub-space of each molecule is extended over a region bounded by the volume of the vessel containing the fluid. In general

$$\int d\omega_i = \sigma_i v \qquad (2)$$

where σ_i is equal to 4π for diatomic molecules which have only two low frequency rotational degrees of freedom and to $8\pi^2$ for polyatomic molecules which in general have three. The potential of the intermolecular forces, V_N, is a function of the configuration coordinates of all N molecules. The functions $f_s(T)$, depending upon the temperature alone, are products of certain momentum integrals associated with the low frequency degrees of freedom of a molecule of type s and the partition function associated with its high frequency vibrational degrees of freedom. For our present purposes we do not need to characterize these functions more closely.

The chemical potential per molecule of a component i of the fluid is thermodynamically defined by the relation

$$\mu_i = (\partial F/\partial N_i)_{T, v, N_1, \ldots N_{\nu-1}}. \qquad (3)$$

It is statistically defined by Gibbs[4] in terms of the grand ensemble. For our purposes, it is more convenient to define μ_i in terms of the petit ensemble upon which Eq. (1) is based. This may

[1] L. Onsager, Chem. Rev. **13**, 73 (1933); J. W. Gibbs, *Elementary Principles in Statistical Mechanics, Collected Works II*, Chap. VII (Longmans).

[2] We postulate the absence of coexistent phases containing any of the components of the fluid, since we wish to represent the system by a petit ensemble, each example of which contains the same number of molecules. Actually this is a trivial restriction, for Gibbs has shown that although an open system is represented by a grand ensemble, its equilibrium properties do not differ sensibly from those of a closed system represented by a petit ensemble, each example of which is made up of the average number of molecules of a system of the grand ensemble. In case coexistent phases exist, we simply regard v as the volume of the homogeneous part in which we are interested.

[3] J. W. Gibbs, *Elementary Principles in Statistical Mechanics, Collected Works II*, p. 187, Longmans.

be done in a very simple manner. We first observe that the treatment of N as a continuous variable in the thermodynamic definition (3) is simply a convenient mathematical device. A physical definition of chemical potential requires the use of finite differences, ΔF and ΔN_i, in the ratio $\Delta F/\Delta N_i$. For values of ΔN_i which do not sensibly affect the composition of the fluid, this ratio will not differ sensibly from the derivative on the right-hand side of Eq. (3). We are therefore justified in defining μ_i as the change in free energy produced by the subtraction or addition of a single molecule of type i from the macroscopic system, while maintaining the temperature, volume and the numbers of molecules of the other species constant.

$$\mu_i = F(N_1, \cdots, N_i, \cdots N_r, T, v) - F(N_1, \cdots, N_i - 1, \cdots N_r, T, v). \quad (4)$$

The fact that the removal of a single molecule is scarcely a physically realizable process does not alter the value of the definition (4) when used in conjunction with the petit ensemble, each system of which is supposed to comprise exactly the same number of molecules.

From Eqs. (1) and (4), we obtain the following expression for μ_i:

$$\mu_i = -kT\log\left[\frac{f_i(T)}{N_i}\,\frac{Z(N_1, \cdots, N_i, \cdots N_r, T, v)}{Z(N_1, \cdots, N_i - 1, \cdots N_r, T, v)}\right]. \quad (5)$$

Further simplification depends upon the form of the potential of the intermolecular forces, V_N. We shall treat here only systems, for which V_N may be expressed as a sum of terms V_{kl}, each depending only upon the relative coordinates of a molecular pair (kl).[5]

$$V_N = \sum_{k<l} V_{kl}. \quad (6)$$

We now define a fictitious potential of the following type

$$V_N(\lambda_1, \cdots \lambda_N) = \sum_{k<l} \lambda_k \lambda_l V_{kl}, \quad (7)$$

where $\lambda_1 \cdots \lambda_N$ are arbitrary parameters. By varying these parameters between zero and unity, we can continuously vary the coupling between the N molecules from zero to its full value. In particular, we notice that when a value unity is assigned to each of the N parameters λ, we have

$$V_N(1, \cdots, 1) = V_N, \quad (8)$$

where V_N is given by Eq. (6) and represents the actual coupling between the molecules. If, on the other hand, a value unity is assigned to $N-1$ of the parameters λ and a value zero to a single λ_i, we have

$$V_N(1, \cdots, 0, \cdots 1) = V_{N-1}, \quad (9)$$

where V_{N-1} is the potential in a system of $N-1$ molecules formed by the removal of a single molecule i. Corresponding to $V_N(\lambda_1, \cdots \lambda_N)$, we define a phase integral,

$$Z_N(\lambda_1, \cdots \lambda_N) = \int \cdots \int e^{-\beta V_N(\lambda_1 \cdots \lambda_N)} d\omega_1 \cdots d\omega_N. \quad (10)$$

When each of the parameters, λ, is assigned the value unity, we have the following identity:

$$Z_N(1, \cdots 1) = Z(N_1, \cdots, N_i, \cdots N_r). \quad (11)$$

When $N-1$ of the parameters λ are equal to unity, while a single λ, corresponding, let us say,

[5] It is to be remarked that the potential of Eq. (6) in no way excludes the simultaneous interaction of groups of more than two molecules. It simply states that in such a group, the mutual potential energy of any pair is independent of the presence of the other molecules. For example the electrostatic energy of a system of point charges is of this type. London has shown that the potential of the attractive van der Waals forces between chemically saturated molecules satisfies the requirements of Eq. (6). [F. London, Zeits. f. physik. Chemie **11**, 222 (1930).]

No such clear cut statement can be made regarding the repulsive forces, which operate at small distances, determining the effective collision diameters of molecules. However, it is plausible to assume that the potentials of these forces are approximately additive, and even if they are not exactly so, it is only necessary for our purposes that their narrow range of action is not sensibly displaced. In general, we should expect terms depending upon the relative coordinates of three or more molecules to be important only if the total intermolecular field is strong enough to produce a large perturbation in the electronic structure of a molecule in the fluid. Theoretical considerations suggest that not only in gases but also in liquids, this is not the case. Empirical arguments against any appreciable distortion of the electronic structure of molecules in most liquids are furnished by spectroscopic and optical evidence. For example, the molecular refractivity of most liquids is substantially the same as that of their vapor.

There are to be sure some types of intermolecular potential which do not satisfy Eq. (6), for example those of chemical forces of the homeopolar type. Another example is furnished by the energy of polarization of a molecule by a group of ions.

to a molecule of type i is equal to zero, we may write

$$Z_N(1, \cdots, 0, \cdots 1) = \int \cdots \int e^{-\beta V_{N-1}} d\omega_1 \cdots d\omega_N. \quad (12)$$

Now since V_{N-1} contains no terms depending upon the coordinates of the molecule i, the integration over its configuration space in Eq. (12) may be carried out immediately. Performing this integration and making use of Eqs. (1) and (2), we obtain

$$Z_N(1, \cdots, 0, \cdots 1) = \sigma_i v Z(N_1, \cdots, N_i - 1, \cdots N_s). \quad (13)$$

From Eqs. (11) and (13) we may write

$$\frac{Z(N_1, \cdots, N_i, \cdots N_s)}{Z(N_1, \cdots, N_i - 1, \cdots N_s)} = \sigma_i v \frac{Z_N(1, \cdots, 1, \cdots 1)}{Z_N(1, \cdots, 0, \cdots 1)}. \quad (14)$$

Further we have the identity,

$$\log \frac{Z_N(1, \cdots, 1, \cdots 1)}{Z_N(1, \cdots, 0, \cdots 1)} = \int_0^1 \frac{\partial \log Z_N(\lambda_i)}{\partial \lambda_i} d\lambda_i. \quad (15)$$

To simplify notation, we write $Z_N(\lambda_i)$ for $Z_N(1, \cdots, \lambda_i, \cdots 1)$. In the future this convention will be adopted for all functions of the parameters λ, only those differing from unity being explicitly indicated. Returning to Eq. (10), we observe that

$$Z_N(\lambda_i) = \int \cdots \int e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N, \quad (16)$$

$$V_i = \sum_{k=1}^{N} V_{ik}.$$

Differentiation of (16) with respect to λ_i yields

$$\partial \log Z_N(\lambda_i)/\partial \lambda_i = -\beta \overline{V_i(\lambda_i)},$$

$$\overline{V_i(\lambda_i)} = \frac{\int \cdots \int V_i e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N}{\int \cdots \int e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N}. \quad (17)$$

Combining Eqs. (5), (14), (15) and (17), we obtain for the chemical potential:

$$\mu_i = kT \log N_i/v + \int_0^1 \overline{V_i(\lambda_i)} d\lambda_i + \varphi_i(T), \quad (18)$$

$$\varphi_i(T) = -kT \log \sigma_i f_i(T).$$

The mean potential energy $\overline{V_i(\lambda_i)}$ can be expressed as follows:

$$\overline{V_i(\lambda_i)} = \sum_{k=1}^{N} \overline{V_{ik}(\lambda_i)},$$

$$\overline{V_{ik}(\lambda_i)} = \frac{\int \cdots \int V_{ik} e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N}{\int \cdots \int e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N}. \quad (19)$$

Since V_{ik} depends only on the coordinates of the molecules i and k, we may write,

$$\overline{V_{ik}(\lambda_i)} = \frac{1}{\sigma_i \sigma_k v^2} \int \int V_{ik} e^{-\beta W_i^k(\lambda_i)} d\omega_i d\omega_k, \quad (a) \quad (20)$$

$$e^{-\beta W_i^k(\lambda_i)} = \frac{\sigma_i \sigma_k v^2 \int \cdots \int e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_{N-2}}{\int \cdots \int e^{-\beta[V_{N-1} + \lambda_i V_i]} d\omega_1 \cdots d\omega_N} \quad (b)$$

The function $W_i^k(\lambda_i)$, as we shall presently show, is the potential of the mean force acting between the molecules i and k. At present we merely observe that in the absence of external fields of force $W_i^k(\lambda_i)$ can depend only upon the relative coordinate of the molecule pair i and k, except in a region of negligible volume near the surface of the fluid.

We may therefore refer the position of the center of gravity and the orientation of the molecule k to a coordinate system fixed in molecule i. Both V_{ik} and $W_i^k(\lambda_i)$ will depend upon these relative coordinates alone. Moreover, the Jacobian of this transformation is unity. If we then integrate over the configuration space of molecule i, the expression for $\overline{V_{ik}(\lambda_i)}$, Eq. (20), becomes:

$$\overline{V_{ik}(\lambda_i)} = (1/\sigma_k v) \int V_{ik} e^{-\beta W_i^k(\lambda_i)} d\omega_k, \quad (21)$$

where $d\omega_k$ is understood to refer to the coordinates of molecule k relative to a system of axes fixed in molecule i. Introducing Eq. (21) into Eq. (18), and remembering that $\overline{V_{ik}(\lambda_i)}$

is the same function for all molecules k of the same type, we may write:

$$\mu_i = kT\left\{ \log N_i/v + \sum_{k=1}^{\nu} \frac{N_k}{v}\int_0^1 B_{ik}(\lambda_i)d\lambda_i \right\} + \varphi_i(T), \tag{22}$$

$$B_{ik}(\lambda_i) = (1/\sigma_k kT)\int V_{ik}e^{-\beta W_i^k(\lambda_i)}d\omega_k,$$

where the summation extends over all molecular types k present in the fluid, and the function B_{ik} refers to a single pair of molecules of types i and k, respectively. Each B_{ik} is a function of temperature, volume and composition, since $W_i^k(\lambda_i)$ is a function of all of these variables. The notation $W_i^k(\lambda_i)$ means of course that all the parameters λ have the value unity except λ_i corresponding to a *single* molecule of type i.

The mean energy $\bar{E}$ of the fluid may be expressed quite simply, if we assume classical equipartition with respect to the low frequency rotational and translational degrees of freedom

$$\bar{E} = E_0(T) + \bar{V}_N, \tag{23}$$

where $\bar{V}_N$ is the mean potential energy associated with the intermolecular forces and $E_0(T)$, a function of the temperature alone, is sum of the internal energies of the molecules and their mean translational and rotational kinetic energies. Writing V_N in the form of Eq. (6), and remembering that V_{ik} depends only on the relative coordinates of molecules k and i, we may write*

$$\bar{V}_N = \sum_{i<k} \bar{V}_{ik} = \tfrac{1}{2}\sum_{\substack{i=1\\ i\neq k}}^{\nu} (N_i N_k/v)\bar{V}_{ik},$$

$$\bar{V}_{ik} = (1/\sigma_k v)\int V_{ik}e^{-\beta W_i^k(1)}d\omega_k. \tag{24}$$

We note that W_i^k appearing in Eq. (24) is $W_i^k(1)$, the value of $W_i^k(\lambda_i)$ when molecule i has its full coupling with the other molecules of the system. Finally we may write,

$$\bar{E} = E_0(T) + \tfrac{1}{2}kT\sum_{\substack{i=1\\ i\neq k}}^{\nu} (N_i N_k/v)B_{ik}. \tag{25}$$

The quantity B_{ik} in Eq. (25) is equal to $B_{ik}(1)$ of Eq. (22) when λ_i is unity. $E_0(T)$ may be explicitly calculated if the internal partition functions of the molecules are known. We shall not be interested in this point here.

III.

The use of Eqs. (22) and (25) for purposes of calculation depends upon an investigation of the functions $W_i^k(\lambda_i)$ which determine the probability distribution of the pair of molecules i and k relative to each other. It may readily be shown that $W_i^k(\lambda_i)$ is the potential of the mean force acting between the molecules i and k, averaged over all configurations of the remaining $N-2$ molecules of the system. The force acting at the center of gravity of molecule i for a fixed configuration of the other molecules, is

$$F_i(\lambda_i) = -\nabla_i V_N(\lambda_i), \tag{26}$$

where ∇_i is the gradient operator associated with the coordinates of the center of gravity of molecule i. If we take a mean value with the pair of molecules i and k held fixed, we obtain

$$\overline{F_i(\lambda_i)}^{ik} = \frac{\int\cdots\int \nabla_i V_N(\lambda_i)e^{-\beta V_N(\lambda_i)}d\omega_1\cdots d\omega_{N-2}}{\int\cdots\int e^{-\beta V_N(\lambda_i)}d\omega_1\cdots d\omega_{N-2}}. \tag{27}$$

Differentiation of both sides of Eq. (20b), which defines $W_i^k(\lambda_i)$ and comparison with Eq. (27) yields

$$\overline{F_i(\lambda_i)}^{ik} = -\nabla_i W_i^k(\lambda_i). \tag{28}$$

The function $W_i^k(\lambda_i)$ is therefore the potential of the mean force acting at the center of gravity of molecule i when k is held fixed at some point in its vicinity. Similarly, it may be shown that $W_i^k(\lambda_i)$ is the potential of the mean generalized forces associated with the relative orientation of i and k. If the relative orientation is specified by the Eulerian angles ϑ, φ and ψ, the mean torque tending to produce rotation through the angle ϑ is $-\partial W_i^k(\lambda_i)/\partial\vartheta$. These relations have been used by Einstein, Schmoluchowski and others,[6] who showed that the simple Boltzmann distribution is valid provided the potential of the mean force is employed in the exponential factor.

It is interesting to remark that the potential $W_i^k(1)$, corresponding to full coupling between molecule i and the other molecules, is susceptible

* Eq. (24) has been given by Hildebrand and Wood, J. Chem. Phys. 1, 817 (1933).

[6] For a discussion of this point, see L. Onsager, Chem. Rev. 13, 73 (1933).

of experimental determination, at least in its dependence on the relative coordinates of the molecular centers of gravity. According to the theory of x-ray scattering by liquids, developed by Prins and Debye,[7] the angular distribution of intensity of the scattered radiation is a sum of mean values of functions depending only upon the relative coordinates of pairs of molecules. Mean values of such functions are determined by the probability distribution function $e^{-\beta W_\lambda{}^k}$ in a system in thermodynamic equilibrium. By a Fourier integral inversion of the theoretical scattering formula, $(e^{-\beta W_\lambda{}^k}-1)$ may be computed from the observed intensity measured as a function of the scattering angle. By Eq. (25), we see that a knowledge of $W_\lambda{}^k(1)$ suffices for the calculation of the mean energy of the fluid from the potentials V_{ik}. However, it does not suffice for the calculations of chemical potentials or the Gibbs free energy, ζ, which is equal to $\sum_{i=1}^{r} N_i\mu_i$. For the latter purpose we must know $W_\lambda{}^k(\lambda_i)$ for all values of λ_i between zero and unity. This information is not furnished by x-ray scattering experiments. Thus, while they yield valuable information concerning the statistical structure of liquids, they do not provide sufficient data for the calculation of entropy and free energy.

We must consequently depend upon theoretical information about the potentials of mean force $W_\lambda{}^k(\lambda_i)$. Since the many dimensional phase integrals of Eq. (20b), which defines $W_\lambda{}^k(\lambda_i)$, are inconvenient to handle, we follow an indirect method of attack, leading to reasonable approximations both for condensed and rarefied systems. Logarithmic differentiation of both sides of Eq. (20b) yields

$$\partial W_\lambda{}^k(\lambda_i)/\partial\lambda_i = \overline{{}^{ik}V_i(\lambda_i)} - \overline{V_i(\lambda_i)}, \qquad (29)$$

where

$$\overline{{}^{ik}V_i(\lambda_i)} = \frac{\int\cdots\int V_i e^{-\beta[V_{N-1}+\lambda_i V_{ik}]}d\omega_1\cdots d\omega_{N-2}}{\int\cdots\int e^{-\beta[V_{N-1}+\lambda_i V_{ik}]}d\omega_1\cdots d\omega_{N-2}}, \qquad (30)$$

$\overline{{}^{ik}V_i(\lambda_i)}$ is the mean value of V_i for a fixed configuration of the molecular pair i and k, while $\overline{V_i(\lambda_i)}$ is the mean value of the same function taken over all configurations of all molecules of the system. By Eq. (10), we may write

$$\overline{{}^{ik}V_i(\lambda_i)} = V_{ik} + \sum_{\substack{s=1\\ s\neq i\\ s\neq k}}^{N} \overline{{}^{ik}V_{is}(\lambda_i)}, \qquad (31)$$

since the term V_{ik} is constant for fixed configurations of molecules i and k and is unaffected by a mean value operation in which these are held fixed. Because V_{is} depends only on the relative coordinates of i and s, we have

$$\overline{{}^{ik}V_{is}(\lambda_i)} = (1/\sigma_s v)\int V_{is}e^{-\beta W_{ik^s}(\lambda_i)}d\omega_s, \qquad (a)$$

(32)

$$e^{-\beta W_{ik^s}(\lambda_i)} = \frac{\sigma_s v\int\cdots\int e^{-\beta[V_{N-1}+\lambda_i V_{ik}]}d\omega_1\cdots d\omega_{N-3}}{\int\cdots\int e^{-\beta[V_{N-1}+\lambda_i V_{ik}]}d\omega_1\cdots d\omega_{N-2}}, \qquad (b)$$

where the integration in the numerator of (32b) is to extend over the coordinates of all molecules except i, k and s and in the denominator over all but those of i and k. It will be seen at once that $W_{ik^s}(\lambda_i)$ is the potential of the mean force acting upon molecule s, averaged over the configurations of all molecules except i, k and s. We note that $W_{ik^s}(\lambda_i)$ is not invariant under a permutation of upper and lower indices, although $W_\lambda{}^k(\lambda_i)$ is. It is possible to define a potential of mean force which is symmetric in the coordinates of i, k and s, differing from $W_{ik^s}(\lambda_i)$ by a term independent of the coordinates of s, but the latter form is not the most convenient for the present purposes. The second term of Eq. (29) is the sum $\sum_{s=1}^{N}\overline{V_{is}(\lambda_i)}$, each member of which is given by an expression of the form (21). Moreover, each of the mean values $\overline{{}^{ik}V_{is}}$ and $\overline{V_{is}}$ are identical molecules of the same type s, so that we may write,

$$\partial W_\lambda{}^k(\lambda_i)/\partial\lambda_i = V_{ik} + \sum_{s=1}^{r}\frac{N_s}{v}\frac{1}{\sigma_s}\int V_{is}\left[e^{-\beta W_{ik^s}(\lambda_i)} - e^{-\beta W_\lambda{}^s(\lambda_i)}\right]d\omega_s, \qquad (33)$$

where the sum is extended over all types s. Each of the integrals refers to a single group of three molecules, i, k, and a single molecule of type s. Properly a term, $(v\sigma_k)^{-1}\int V_{ik'}e^{-\beta W^{k'}{}_{ik}(\lambda_i)}d\omega_{k'}$, where k' is a single molecule of the same type as k, should be omitted from the sum in Eq. (33). Since, however, this term bears a ratio of the

[7] Zernicke and Prins, Zeits. f. Physik **41**, 184 (1927); Debye and Mencke, Physik. Zeits. **31**, 797 (1930).

order of $1/N_k$ to terms retained, it can be ignored altogether. This is not true of V_{ik}, which is of dominant importance when k is situated in the immediate neighborhood of i.

Before integrating Eq. (33) we shall find it convenient to investigate the behavior of $W_i^k(\lambda_i)$ in the neighborhood of $\lambda_i=0$. First we note that $W_i^k(0)$ must be independent of the coordinates of i, since the mean force on i vanishes when λ_i is equal to zero. Moreover, it must also be independent of the coordinates of k except in the surface region, since the mean force on k must also vanish when λ_i is equal to zero. From these facts we conclude that $W_i^k(0)$ must be a constant. Making use of Eq. (20b), we write

$$e^{-\beta W_i^k(0)}=\frac{\sigma_i\sigma_k v^2\int\cdots\int e^{-\beta V_N}\,d\omega_1\cdots d\omega_{N-2}}{\int\cdots\int e^{-\beta V_N}\,d\omega_1\cdots d\omega_N}. \tag{34}$$

Now the integral in the denominator of (34) is obviously independent of the coordinate of i and k, since it includes integrations over these coordinates. Since $W_i^k(0)$ is independent of the coordinates of i and k, we must conclude that the integral in the numerator is also independent of them. The factor $\sigma_i\sigma_k v^2$ may therefore be written as an integration over the configuration spaces of molecules i and k,

$$\sigma_i\sigma_k v^2\int\cdots\int e^{-\beta V_N}\,d\omega_1\cdots d\omega_{N-2}=\int\cdots\int e^{-\beta V_N}\,d\omega_1\cdots d\omega_N. \tag{35}$$

From Eqs. (34) and (35) we conclude that $e^{-\beta W_i^k(0)}$ has the value unity and that $W_i^k(0)$ vanishes. Bearing this fact in mind, we integrate Eq. (33) between zero and λ_i with the result:

$$W_i^k(\lambda_i)=\lambda_i V_{ik}+\sum_{s=1}^{r}(N_s/v)(1/\sigma_s)\int_0^{\lambda_i}\int V_{is}[e^{-\beta W_{ik}^s(\lambda_i)}-e^{-\beta W_i^s(\lambda_i)}]\,d\omega_s d\lambda_i. \tag{36}$$

From Eq. (36), we observe that $W_i^k(\lambda_i)$ approaches $\lambda_i V_{ik}$ as the densities of the several components of the fluid approach zero. This approximation is also obtained by neglecting the influence of the fixed molecule k on the distribution of a third molecule s in the vicinity of molecule i, that is by ignoring the difference between $W_{ik}^s(\lambda_i)$ and $W_i^s(\lambda_i)$. In the theory of the equation of state of a rarefied gas, this is essentially the approximation which is made. To obtain a more accurate expression for $W_i^k(\lambda_i)$ we must know something about the potential $W_{ik}^s(\lambda_i)$ involving the group of three molecules i, k and s. Proceeding from Eq. (32b), which defines $W_{ik}^s(\lambda_i)$, we obtain by differentiation with respect to λ_i

$$\partial W_{ik}^s(\lambda_i)/\partial\lambda_i=\lambda_i V_{is}+\sum_{l=1}^{r}(N_l/v\sigma_l)\int V_{il}[e^{-\beta W_{ik,s}^l(\lambda_i)}-e^{-\beta W_{ik}^l(\lambda_i)}]\,d\omega_l, \tag{37}$$

$$e^{-\beta W_{ik,s}^l(\lambda_i)}=\frac{\sigma_l v\int\cdots\int e^{-\beta(V_N+\lambda_i V_{il})}\,d\omega_1\cdots d\omega_{N-4}}{\int\cdots\int e^{-\beta(V_N+\lambda_i V_{il})}\,d\omega_1\cdots d\omega_{N-3}}.$$

Examination of (32b) shows at once that $W_{ik}^s(0)$ is merely W_k^s in a system containing N_i-1 molecules of type i. It is apparent that this function can differ from W_k^s in the original system only by a quantity of completely negligible order, $O(1/N_i)$. Thus when Eq. (37) is integrated with respect to λ_i, there is obtained

$$W_{ik}^s(\lambda_i)=W_k^s+\lambda_i V_{is}+\sum_{l=1}^{r}(N_l/v\sigma_l)\int_0^{\lambda_i}\int V_{il}[e^{-\beta W_{ik,s}^l(\lambda_i)}-e^{-\beta W_{ik}^l(\lambda_i)}]\,d\omega_l d\lambda_i. \tag{38}$$

This equation is entirely analogous to Eq. (36). Since W_k^s is given by an equation of the form of (36), a first approximation to $W_{ik}^s(\lambda_i)$ is evidently $V_{ks}+\lambda_i V_{is}$. Similarly a first approximation to $W_i^s(\lambda_i)$ is $\lambda_i V_{is}$. If these approximations are employed in the integrals of Eq. (36), we obtain a second approximation to $W_i^k(\lambda_i)$, namely

$$W_i^k(\lambda_i)=\lambda_i V_{ik}-kT\sum_{s=1}^{r}(N_s/v)(1/\sigma_s)\int(e^{-\beta\lambda_i V_{is}}-1)(e^{-\beta V_{ks}}-1)\,d\omega_s. \tag{39}$$

We should expect this approximation to be adequate in gases of moderate density. The method of successive approximation employed to obtain Eq. (39), can of course be extended without great difficulty, by transforming the potentials $W_{ik\lambda}{}^{l}(\lambda_i)$, involving groups of four molecules, in exactly the same manner as $W_i{}^{k}(\lambda_i)$ and $W_{ik}{}^{s}(\lambda_i)$. Presumably any desired degree of approximation could be attained in this fashion, but the expressions become very cumbersome. It should also be remarked that the method of approximation just outlined can only be employed when no long range forces, such as the Coulomb forces, act between the molecules, for in such cases the integral of Eq. (39) diverges. It converges whenever the individual potentials decrease asymptotically at least as rapidly as r_{ik}^{-4}.

For an exact calculation of $W_i{}^{k}(\lambda_i)$, it would be necessary to solve a set of $N-1$ simultaneous integral equations of the type of Eqs. (36) and (38), involving potentials $W^{s}{}_{ik\cdots l}$, in which the number of fixed molecules i, k, $\cdots l$, ranges from one to $N-1$. Moreover, when the number of fixed molecules becomes comparable with the total number of molecules, N, the equations become more complicated than (36) and (38), due to growth of terms, which are of zero order in the latter equations. Since this is not a practicable procedure, we must resort to some means of approximation. Eq. (39) represents a possible method of approximation in which $W_i{}^{k}(\lambda_i)$ is developed in powers of the densities of the components of the fluid. However, we should like to find an approximation more suitable in condensed systems than Eq. (39) is likely to be. An obvious and simple approximation may be obtained by assuming superposition in the potentials of mean force,

$$W_{ik}{}^{s}(\lambda_i) = W_k{}^{s} + W_i{}^{s}(\lambda_i). \tag{40}$$

Although the additivity expressed in Eq. (40) holds for the direct action of i and k on s through the terms $V_{ks}+\lambda_i V_{is}$, it cannot hold exactly for their indirect action upon s through their effect on the distribution of the remaining $N-3$ molecules. The magnitude of this discrepancy will be discussed later. If we introduce the approximation (40) into Eq. (36), we obtain the following integral equation:

$$W_i{}^{k}(\lambda_i) = \lambda_i V_{ik} + \sum_{s=1}^{\nu} (N_s/v)(1/\sigma_{ik}) \int_0^{\lambda_i}\int V_{is} e^{-\beta W_i{}^{s}(\lambda_i)} [e^{-\beta W_k{}^{s}(1)} - 1] d\omega_s d\lambda_i. \tag{41}$$

The potentials $W_i{}^{k}(\lambda_i)$ will have the same form for any pair of molecules of the same type. From the ν types of molecules present in the fluid, $\nu(\nu+1)/2$ types of molecular pairs can be formed. There will, therefore, be $\nu(\nu+1)/2$ types of functions $W_i{}^{k}(\lambda_i)$, each satisfying an equation like (41). Simultaneous solution of these integral equations will yield values of the functions $W_i{}^{k}(\lambda_i)$ consistent with the approximation expressed in Eq. (40). Of course for a pure fluid, having only one component, only one integral equation has to be solved.

$$W(q_{12}, \lambda) = \lambda V(q_{12}) + (N\rho/M\sigma) \int_0^{\lambda}\int V(q_{13}) e^{-\beta W(q_{13},\ \lambda)} [e^{-\beta W(q_{23},\ 1)} - 1] d\omega_3 d\lambda, \tag{42}$$

where q_{12}, q_{13} and q_{23} are the relative coordinates of a set of three molecules, N is Avogadro's number, M the molecular weight of the substance, and ρ its density. We note that the approximation (39) is obtainable from Eq. (41) as the first step in its solution by successive substitution. In general, it will not be feasible to seek analytic solutions of Eqs. (41) or (42). However, by graphical or mechanical quadrature, it should be possible to determine self-consistent functions satisfying these equations. We remark that this process provides a theoretical means of calculating the x-ray scattering curves of liquids, since they are determined by $W_i{}^{k}(1)$. In some cases the potentials of intermolecular force can be calculated from atomic structure. Although this can be done in relatively few cases, quantum mechanics prescribes an approximate functional form for the V_{ik}. With this functional form, the numerical values of V_{ik} may be deduced from the temperature dependence of the second virial coefficient of the substance in the gaseous state.

While Eq. (40) represents an approximation which is intuitively attractive, it is of course not exact. Although a numerical estimate of the error involved is difficult, we can say something quite

definite about the nature of the approximation. To facilitate discussion, we shall introduce the symbol ϵ_{ik}^{s} to denote the deviation from additivity

$$\epsilon_{ik}^{s} = W_{ik}^{s}(\lambda_i) - [W_k^{s} + W_i^{s}(\lambda_i)]. \tag{43}$$

It is possible to express ϵ_{ik}^{s} in terms of a third order fluctuation in the potential energy of molecule i. Without entering into the details of the calculation, we state the following result.

$$\epsilon_{ik}^{s} = \beta^2 \int_0^{\lambda_i}\int_0^{1}\int_0^{1} \overline{(V_i - \overline{V_i})(V_k - \overline{V_k})(V_s - \overline{V_s})}^{ik s} d\lambda_k d\lambda_s d\lambda_i, \tag{44}$$

where the fluctuation is understood to be a function of the variables λ_i, λ_k, λ_s, mean values being taken with the distribution function $e^{-\beta V_N(\lambda_i, \lambda_k, \lambda_s)}$. A more useful expression for ϵ_{ik}^{s} is obtained simply by subtracting the equation for $W_i^{s}(\lambda_i)$ corresponding to (36) from Eq. (38). The result is,

$$\epsilon_{ik}^{s} = \sum_{l=1}^{r} (N_l/v\sigma_l) \int_0^{\lambda_i}\int V_{il}\{e^{-\beta W_{iks}^{l}(\lambda_i)} + e^{-\beta W_i^{l}(\lambda_i)} - e^{-\beta W_{ik}^{l}(\lambda_i)} - e^{-\beta W_{is}^{l}(\lambda_i)}\} d\omega_l d\lambda_i. \tag{45}$$

An exact calculation of ϵ_{ik}^{s} from Eq. (45) is of course impossible without a knowledge of the potential $W_{iks}^{l}(\lambda_i)$. However, we should expect to obtain at least a rough estimate of ϵ_{ik}^{s} by using the additivity approximation for the potentials $W_{iks}^{l}(\lambda_i)$, $W_{ik}^{l}(\lambda_i)$, and $W_{is}^{l}(\lambda_i)$. Introduction of this approximation yields

$$\epsilon_{ik}^{s} = \sum_{l=1}^{r} (N_l/v\sigma_l) \int_0^{\lambda_i}\int V_{il} e^{-\beta W_i^{l}(\lambda_i)} (e^{-\beta W_k^{l}} - 1)(e^{-\beta W_s^{l}} - 1) d\omega_l d\lambda_i. \tag{46}$$

If we rely on Eq. (46) for the order of magnitude of ϵ_{ik}^{s}, we see that the integrand will be very small except for configurations of molecule such that W_k^{l} and W_s^{l} are simultaneously large. If s and k are situated at some distance from each other so that their spheres of influence on the distribution of a third molecule l, do not overlap appreciably, ϵ_{ik}^{s} will be small. Returning to Eq. (36) we remark that for the calculation of $W_i^{k}(\lambda_i)$, it is only necessary to know $W_{ik}^{s}(\lambda_i)$ accurately in a region in the vicinity of molecule i, where V_{is} is sensibly different from zero. If the molecules i and k are separated by an appreciable distance, s will be distant from k in the greater part of this region, and we may conclude that ϵ_{ik}^{s} is small. Thus we may expect Eq. (41) to provide a good approximation to $W_i^{k}(\lambda_i)$ at least when the molecules i and k are separated by an appreciable distance, and that $W_i^{k}(\lambda_i)$ will approach the solution of Eq. (41) asymptotically as the distance between i and k is increased. This point is of some importance in the case of long range intermolecular forces, such as those between electrolyte ions. At small distances between i and k the dominant term in $W_i^{k}(\lambda_i)$ becomes V_{ik}, and although the terms depending on the density may still be large, it is not necessary to know them with such great accuracy. Once solutions of Eq. (41) have been obtained, they may be used in Eq. (46) to obtain a rough estimate of ϵ_{ik}^{s}. If this turns out to be small, the solutions of Eq. (41) can be accepted with some degree of assurance as reasonable approximations to the potentials $W_i^{k}(\lambda_i)$. It would be desirable to have a more rigid criterion for judging the adequacy of these approximations, but unfortunately this cannot be obtained without introducing distribution functions involving larger and larger groups of fixed molecules. Unless some approximation such as superposition is introduced at some early stage in this sequence of distribution functions, we are driven back to the many dimensional phase integrals with which we started.

Having obtained approximate expressions for $W_i^{k}(\lambda_i)$ by solution of Eqs. (41), we can calculate the chemical potential of any component of the fluid by means of Eq. (22), as well as the total energy of the fluid by Eq. (25). From the chemical potentials, we may obtain the Gibbs free energy, ζ, of the system.

$$\zeta = E + pv - TS = \sum_{i=1}^{r} N_i\mu_i = kT\left\{\sum_{i=1}^{r} N_i \log N_i/v + \sum_{i,k=1}^{r} \frac{N_iN_k}{v}\int_0^{1} B_{ik}(\lambda_i)d\lambda_i\right\} + \sum_{i=1}^{r} N_i\varphi_i(T). \tag{47}$$

In principle one can obtain the compressibility, κ, of the fluid by means of the thermodynamic formula,

$$1/\kappa = -v(\partial p/\partial v)_T = -(\partial \zeta/\partial v)_{T,\, N_1,\, \cdots, N_r}. \tag{48}$$

However the $B_{ik}(\lambda_i)$ depend upon v through the $W_i^k(\lambda_i)$. Even though Eq. (41) may provide good approximations to the $W_i^k(\lambda_i)$, it is by no means certain that the derivatives of these solutions with respect to v will be good approximations to $\partial W_i^k(\lambda_i)/\partial v$. Thus while Eqs. (22), (35) and (47) may express the dependence of the μ_i, E and ζ upon the composition of the fluid with some accuracy, the compressibility calculated from Eqs. (47) and (48) may be a much poorer approximation. Quite formally, we may write

$$\left(\frac{\partial p}{\partial v}\right)_T = -kT\left\{\sum_{i=1}^{r}\frac{N_i}{v^2} + \sum_{i,\,k=1}^{r}\frac{N_iN_k}{v^3}\int_0^1\left[B_{ik}(\lambda_i) - v\frac{\partial B_{ik}(\lambda_i)}{\partial v}\right]d\lambda_i\right\}. \tag{49}$$

In case the fluid under consideration is a gas, we know that as the volume is increased to infinity, the pressure decreases continuously to zero. In this case we may integrate Eq. (49) between the limits v and infinity. After a simple partial integration, the following equation of state is obtained.

$$\frac{p\bar{v}}{RT} = 1 + \sum_{i,\,k=1}^{r}\frac{x_ix_k}{\bar{v}}\left\{\int_0^1 N\left[B_{ik}(\lambda_i) + \bar{v}^2\int_\infty^{\bar{v}}\frac{B_{ik}(\lambda_i)}{\bar{v}^3}d\bar{v}\right]d\lambda_i\right\} \tag{50}$$

$$n_i = N_i/N; \qquad x_i = n_i/\sum n_k; \qquad \bar{v} = v/\sum n_k,$$

where N is Avogadro's number, n_i the number of moles of component i present in the gas and x_i the mole fraction of that component. Eq. (50) may be made the basis for a statistical calculation of the equation of state of a real gas mixture. When the potentials $W_i^k(\lambda_i)$ in the $B_{ik}(\lambda_i)$ are approximated simply by λ_iV_{ik} the usual expressions for the second virial coefficients are obtained.

IV.

The present method provides a particularly simple means of determining the equation of state of a real gas mixture. We shall derive expressions for the second and third coefficients in the virial expansion of the equation of state, indicating how the method may be extended to higher coefficients. Results are obtained which are in agreement with those of Ursell, who has treated the same problem from a somewhat different point of view.[8]

If the function $W_i^k(\lambda_i)$ is approximated by Eq. (39), we obtain the following expression for $B_{ik}(\lambda_i)$.

$$B_{ik}(\lambda_i) = \frac{1}{\sigma_kkT}\int V_{ik}\exp\left\{-\beta\lambda_iV_{ik} + \sum_{s=1}^{r}\frac{N_s}{v\sigma_s}\int(1-e^{-\beta\lambda_iV_{is}})(1-e^{-\beta V_{ks}})d\omega_s\right\}d\omega_k. \tag{51}$$

Substitution of this expression in Eqs. (22) and (50) would yield the chemical potentials of the several components and the equation of state of the gas mixture. We should expect the equation of state, although approximate, to be much more satisfactory than a virial expansion. However, the integrals involved cannot be expressed in terms of elementary functions, and we shall not concern ourselves with their simplification here. We note that Eq. (39) is exact to terms in $1/v^2$, so that the first two terms of an expansion of the right-hand side of Eq. (51) are exact. The expansion is

$$B_{ik}(\lambda_i) = \frac{1}{kT}\left\{\frac{1}{\sigma_k}\int V_{ik}e^{-\beta\lambda_iV_{ik}}d\omega_k + \sum_{s=1}^{r}\frac{N_s}{v}\frac{1}{\sigma_s\sigma_k}\int\int V_{ik}e^{-\beta\lambda_iV_{ik}}(1-e^{-\beta\lambda_iV_{is}})(1-e^{-\beta V_{ks}})d\omega_kd\omega_s + O(1/v^2)\right\}. \tag{52}$$

[8] Ursell, Proc. Camb. Phil. Soc. **23**, 685 (1927); R. H. Fowler, *Statistical Mechanics*, Cambridge University Press (1929), p. 173.

Substitution of this expansion in Eq. (22), integration with respect to λ_i and a rearrangement of terms yields[9]

$$\mu_i = kT\Big\{ \log N_i/v + \sum_{k=1}^{r} \frac{N_k}{v\sigma_k} \int (1-e^{-\beta V_{ik}}) d\omega_k + \tfrac{1}{2} \sum_{k,s=1}^{r} \frac{N_k N_s}{v^2} \frac{1}{\sigma_s \sigma_k} \int\int (1-e^{-\beta V_{ik}})(1-e^{-\beta V_{is}})(1-e^{-\beta V_{ks}}) d\omega_k d\omega_s + O(1/v^3) \Big\}. \quad (53)$$

When the approximate expressions for the μ_i given by Eq. (53) are introduced into Eqs. (47) and (48), we obtain

$$\left(\frac{\partial p}{\partial v}\right)_T = -kT\Big\{ \sum_{i=1}^{r} \frac{N_i}{v^2} + \sum_{i,k=1}^{r} \frac{N_i N_k}{v^3} \frac{1}{\sigma_k} \int (1-e^{-\beta V_{ik}}) d\omega_k + \sum_{i,k,s=1}^{r} \frac{N_i N_k N_s}{v^4} \frac{1}{\sigma_k \sigma_s} \int\int (1-e^{-\beta V_{ik}})(1-e^{-\beta V_{is}})(1-e^{-\beta V_{ks}}) d\omega_s d\omega_k + O(1/v^5) \Big\}. \quad (54)$$

Integration between the volume limits ∞ and v leads to the following equation of state

$$\frac{p\bar{v}}{RT} = 1 + \sum_{i,k=1}^{r} \frac{x_i x_k}{\bar{v}} B_{ik}{}^0 + \sum_{i,k,s=1}^{r} \frac{x_i x_k x_s}{\bar{v}^2} C_{iks}{}^0 + O(1/\bar{v}^3)$$

$$B_{ik}{}^0 = \frac{N}{2\sigma_k} \int (1-e^{-\beta V_{ik}}) d\omega_k \quad (55)$$

$$C_{iks}{}^0 = \frac{N^2}{3\sigma_k \sigma_s} \int\int (1-e^{-\beta V_{ik}})(1-e^{-\beta V_{is}})(1-e^{-\beta V_{ks}}) d\omega_k d\omega_s.$$

If the method of approximating $W_i{}^k(\lambda_i)$ embodied in Eq. (39) is carried one stage farther, one more term may be retained in the expansion of $B_{ik}(\lambda_i)$ and the fourth virial coefficient may be calculated. The process can be continued indefinitely, but of course the algebraic operations become tedious, and the final expressions unwieldy. The method consists simply of continued substitution and

[9] After integration with respect to λ_i, the terms in μ_i involving $1/v^2$ have the form

$$\sum_{\substack{k,s=1 \\ k\neq s\neq i}}^{N} \frac{1}{v^2 \sigma_k \sigma_s} \int\int (1-e^{-\beta V_{ks}}) \left\{ \frac{V_{ik}}{V_{ik}+V_{is}} (e^{-\beta(V_{ik}+V_{is})}-1) - (e^{-\beta V_{ik}}-1) \right\} d\omega_i d\omega_s,$$

where the original summation indices, referring to the individual molecules, have been restored for clearness. For every term in the sum with the integrand

$$\left\{ \frac{V_{ik}}{V_{ik}+V_{is}} (e^{-\beta(V_{ik}+V_{is})}-1) - (e^{-\beta V_{ik}}-1) \right\} (1-e^{-\beta V_{ks}}),$$

corresponding to any specified pair of molecules k and s, there is also a term involving the same pair, having the form

$$\left\{ \frac{V_{is}}{V_{is}+V_{ik}} (e^{-\beta(V_{ik}+V_{is})}-1) - (e^{-\beta V_{is}}-1) \right\} (1-e^{-\beta V_{ks}}).$$

The sum of this pair of terms is readily seen to be

$$(1-e^{-\beta V_{ik}})(1-e^{-\beta V_{is}})(1-e^{-\beta V_{ks}})$$

so that the double sum may be rearranged as follows:

$$\tfrac{1}{2} \sum_{\substack{k,s=1 \\ k\neq s\neq i}}^{N} \frac{1}{v^2 \sigma_k \sigma_s} \int\int (1-e^{-\beta V_{ks}})(1-e^{-\beta V_{is}})(1-e^{-\beta V_{ik}}) d\omega_i d\omega_s.$$

This sum contains $(N-1)(N-2)$ terms of similar magnitude. It differs from the sum over molecular types appearing in Eq. (55) by $(3N-2)$ terms of the same magnitude. The aggregate of these terms bears a ratio of the order of $1/N$ to either sum and is consequently entirely negligible.

expansion. If one wishes a virial expansion including terms of the order of $1/v^n$, one begins by approximating the $W^s_{ik\cdots l}$, involving n fixed molecules, by sums of the form $\lambda_i V_{is}+\sum_{t=1}^{n-1} V_{ts}$. These expressions, substituted in an equation similar to (38), yield an approximate $W^s_{ik\cdots l}$ involving $(n-1)$ fixed molecules. Continued substitution finally leads to an expression for $W_i{}^k(\lambda_i)$, which when expanded in powers of $1/v$, is exact to terms involving $1/v^n$. It is only necessary to have the expression for $W^s_{ik\cdots l}$ involving $(n-1)$ molecules, exact to terms in $1/v$, since in the final expansion, it is multiplied by a factor $1/v^{n-1}$.

V.

It is interesting to investigate the form of the molecular pair distribution function, $e^{-\beta W_i{}^k}$ by means of the approximation, Eq. (39). Although Eq. (39) can scarcely be accurate at liquid densities, we might expect it to be adequate up to moderately high densities. Moreover, even where inexact, it should serve to illustrate the form of the distribution function.

We shall restrict our attention to a pure fluid, consisting of spherically symmetrical molecules. For such a fluid, Eq. (39) reduces to the following expression when λ_i is equal to unity

$$W(r)=V(r)-(NkT/v)\times\int(1-e^{-\beta V(r_{13})})(1-e^{-\beta V(r_{23})})dv_3,$$

$$r=|\mathbf{r}_1-\mathbf{r}_2|\ ;\quad r_{13}=|\mathbf{r}_1-\mathbf{r}_3|\ ;\quad r_{23}=|\mathbf{r}_2-\mathbf{r}_3|. \tag{56}$$

Since all molecules are of the same type, $W_i{}^k$ and V_{ik} have the same functional form for every molecular pair, and we may omit the indices i and k. We shall suppose the molecules to be rigid spheres of diameter b, so that when r is less than b, $V(r)$ becomes positively infinite. Attractive forces outside of the sphere a should properly be represented by a potential of the form, $-\gamma/r^6$. However to simplify calculations, we shall represent this potential by a trough of depth u_0 and width, $a-b$. Between the concentric spheres of radii b and a, $V(r)$ is thus equal to $-u_0$. The potential $V(r)$, upon which the following calculations are based is represented by the following step function.

$$\begin{aligned} V(r)&=\infty, && 0\leqslant r\leqslant b\\ &=-u_0, && b\leqslant r\leqslant a\\ &=0, && a<r. \end{aligned} \tag{57}$$

With this potential, the integral on the left-hand side of Eq. (56) may be expressed in terms of the volumes of intersection of spheres of radii a and b.

$$\begin{aligned} W(r)/kT&=V(r)/kT-(N\omega_0/v)J(r),\\ \omega_0&=4\pi b^3/3,\\ J(r)&=\frac{1}{\omega_0}\{\omega_{bb}+2(\omega_{ab}-\omega_{bb})(1-e^{\beta u_0})\\ &\quad+(\omega_{aa}-2\omega_{ab}+\omega_{bb})(1-e^{\beta u_0})^2\}, \end{aligned} \tag{58}$$

where ω_{bb} is volume of intersection of two spheres of equal radii b, ω_{aa} that of two spheres of equal radii a, and ω_{ab} that of two spheres of unequal radii a and b, the distance between the centers of each pair of spheres being r. A simple calculation yields

$$\begin{aligned} \frac{\omega_{bb}}{\omega_0}&=1-3r/4b+\tfrac{1}{16}(r/b)^3\ ;\quad 0<r\leqslant 2b\\ &=0,\quad r>2b\\ \frac{\omega_{aa}}{\omega_0}&=(a/b)^3[1-3r/4a+\tfrac{1}{16}(r/a)^3]\ ;\quad 0<r\leqslant 2a\\ &=0,\quad r>2a\\ \frac{\omega_{ab}}{\omega_0}&=\tfrac{1}{2}\{[1-\tfrac{3}{4}(1-(\alpha/r^2))(r/b)\\ &\quad+\tfrac{1}{16}(1-(\alpha/r^2)^3(r/b)^3]+(a/b)^3\\ &\quad\times[1-\tfrac{3}{4}(1+(\alpha/r^2))(r/a)\\ &\quad+\tfrac{1}{16}(1+(\alpha/r^2))^3(r/a)^3]\}\ ;\\ &\qquad\qquad (a^2-b^2)^{\frac{1}{2}}<r\leqslant a+b\\ &=0,\quad r>a+b\ ;\quad \alpha=a^2-b^2. \end{aligned} \tag{59}$$

By means of Eqs. (58) and (59), $W(r)/kT$ may be computed as a function of the ratios, r/b, a/b, $N\omega_0/v$ and u_0/kT. The probability distribution function, which we shall designate by $G(r)$ may then be calculated from the relation $G(r)=e^{-W(r)/kT}$. We note that $N\omega_0$ is equal to eight times the total volume of the molecules,

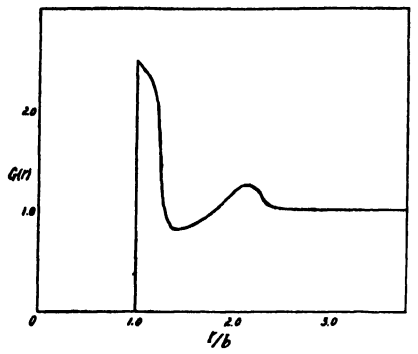


FIG. 1. Molecular pair distribution function based upon Eq. (58).

their radii being $b/2$. For most liquids the ratio $N\omega_0/v$ would have a value between 4 and 6. Moreover, we should expect u_0/kT to be at least unity in liquids. The ratio a/b is somewhat arbitrary. However, in order to make the rectangular trough approximate the actual potential of the attractive forces, which diminishes rapidly with increasing distance, a/b cannot differ much from unity. In order to approach conditions in the liquid state, we assign the following values to the constants in question: $N\omega_0/v=2$; $u_0/kT=1$; $a/b=1.25$. The above value of $N\omega_0/v$ corresponds to a density about a third of ordinary liquid densities. With the above values of the constants, $G(r)$ has been calculated with the aid of Eqs. (58) and (59). It is plotted as a function of r/b in Fig. 1, and tabulated in Table I. Actually there is a discontinuity in the curve at $r/b=1.25$, arising from the artificial nature of the step function potential. In the figure, this discontinuity has been eliminated by drawing a smooth curve through points calculated at r/b intervals of 0.2. It will be observed that the calculated $G(r)$ bears a close resemblance to the molecular distribution functions in liquids derived from measurements of x-ray scattering, exhibiting the characteristic maxima and minima of the empirically determined functions. While we realize that at the temperature and density of the above calculation, the approximation (39) is being pushed to the limit of its applicability, we believe that it adequately illustrates the mechanism by which the intermolecular forces operate to establish molecular pair distribution functions in dense fluids.

TABLE I.

r/b	$G(r)$	r/b	$G(r)$
0.0	0.00	1.6	0.86
0.8	0.00	1.8	0.98
1.0	2.48	2.0	1.20
1.2	2.23	2.2	1.25
1.4	0.81	2.4	1.03
		2.5	1.00

It is interesting to observe that the attractive van der Waals forces play a dominant role in the production of the secondary maximum and minimum in the $G(r)$ of Fig. 1. On the basis of Eq. (39), the distribution function in a fluid of rigid spheres with no attractive forces would exhibit a single maximum at $r=b$. The absence of secondary maxima and minima for rigid spheres is of course the fault of Eq. (39). If the approximation were carried one stage farther, they would begin to appear. However, the point to be emphasized is that probability distribution functions derived from the packing of rigid spheres can scarcely be applied to actual fluids, where the attractive forces play a dominant role in determining the form of the distribution function.

A further limitation of the approximation (39) is that it fails to produce more than a single pair of secondary maxima and minima. For a more exact calculation, it is suggested that the distribution function derived from Eq. (39) be used as a starting point for the self-consistent solution of the more exact integral Eqs. (41) and (42). Thus, it will be observed that tertiary maxima and minima would doubtless arise from the secondary maximum and minimum of the approximate $W(r, \lambda)$ based on Eq. (39), if the latter function were substituted in the integral of Eq. (42). The computational difficulties involved in the self-consistent solution of Eq. (42) are unfortunately rather serious, since even for spherically symmetrical molecules, a triple numerical or mechanical integration is required at each stage of the process. However, these difficulties are not insurmountable.

The formulation of a theory of solutions on the basis of the present ideas offers intriguing possibilities. Several methods of approach are open,

the direct use of our expressions for the chemical potentials being the most obvious. An alternative and more phenomenological treatment applicable to dilute solutions has also been undertaken. These investigations will be described in a future article.

Note added in proof:

(a) A remark about the normalization of the molecular pair distribution functions is perhaps worth while. In defining the potential of mean force, $W_i^k(\lambda_i)$, an arbitrary constant is at our disposal. This constant has been so chosen that

$$(1/v\sigma_k)\int e^{-\beta W_i^k(\lambda_i)} d\omega_k = 1$$

a relation which may be verified by reference to Eq. (20b). The approximate forms for $W_i^k(\lambda_i)$ of course do not exactly satisfy this condition.

(b) At sufficiently high temperatures, the exponentials of Eq. (41) may be expanded and the following approximate linear integral equation is obtained

$$W_i^k(\lambda_i) = \lambda_i V_{ik} - \frac{\lambda_i}{kT}\sum_s \frac{N_s}{v}\frac{1}{\sigma_s}\int V_{is} W_k^s(1) d\omega_s$$

where regions in which W_i^s and W_k^s become infinite due to repulsive forces, are excluded from the integration. If we apply the equation to the ions of an electrolyte solution, the solvent being treated as a continuous medium of dielectric constant D, we know that V_{ik} is equal $e_i e_k / D r_{ik}$ where e_i and e_k are the ionic charges. For this potential it may be verified without difficulty that the above set of linear integral equations has the following solutions

$$W_i^k(\lambda_i) = \lambda_i \frac{e_i e_k}{D r_{ik}} e^{-\kappa r_{ik}}; \qquad \kappa^2 = \frac{4\pi}{DkT}\sum_s \frac{N_s}{v} e_s^2$$

if molecular size is neglected. These solutions, together with Eq. (22) lead to the Debye-Hückel limiting law for the chemical potentials of the ions.

Comments

1. For further development of particle coupling methods see H. L. Frisch, J. Chem. Phys. 22, 1713 (1954), J. G. Kirkwood and J. C. Poirier, J Phys. Chem. 58, 591 (1954) as well as the paper by G. Stell and references cited there.

2. The superposition approximation (SA) has been tested directly numerically recently by B. J. Alder, Phys. Rev. Letters 12, 317 (1964) for the hard sphere fluid and by A. Rahman, Phys. Rev. Letters 12, 575 (1964) for a Lennard-Jones liquid with parameters appropriate for argon. Alder finds that the hard sphere triplet distribution function in the dense fluid and solid region, for spheres in contact, is given by SA to within a few per cent accuracy but the SA is less adequate at low densities.

Reprinted from THE JOURNAL OF CHEMICAL PHYSICS, Vol. 15, No. 4, 187–201, April, 1947
Printed in U. S. A.

Integral Equations between Distribution Functions of Molecules

JOSEPH E. MAYER
Institute of Nuclear Studies, University of Chicago, Chicago, Illinois

(Received February 18, 1947)

Integral equations are derived that relate variations in the potentials of average force between molecules of a system at two different densities or activities. These permit the calculation of the change in thermodynamic properties, or of the change in the distribution of molecules in space, in a liquid or crystalline phase, if either the temperature is varied, or if the the mutual potentials between the molecules is assumed to change. The equations are in a somewhat complex, but still distinguishable, matrix form. A matrix operates on the variations in potential occurring at one activity and transforms them into those occurring at a second activity. The matrix elements are combinations of the distribution functions at the second activity, to which the transformation is made, multiplied by powers of the difference of the two activities. The matrix approaches the unit matrix in value as this activity difference approaches zero. The product of the two matrices, one which transforms from activity α to activity β, with that which transforms from activity β to α, is the unit matrix. This leads to an integral (matrix) equation between the distribution functions at any two activities.

The calculation of one element of this matrix product leads to a cell type equation for the computation of the activity, or free energy, of a system in terms of the potential between a single molecule and the molecules that immediately surround it in the system.

The transformation matrices have solutions, at certain values of the activity difference y, corresponding to variations in the potentials of average force, which are transformed into zero variations at the new activity. The activities at which such variations can occur are those of the phase transitions in the systems. The solutions at these special values of the activity are differences of combinations of the distribution functions in the two pure phases that can coexist at these activities.

INTRODUCTION

ONE may define "distribution functions" for a given number, n, of molecules as being proportional to the probability density of finding n molecules in a certain configurational position, averaged over all positions of the other molecules of the system. These functions are the exponent of minus the potential of average force between the molecules, divided by kT. The distribution functions of a given system at constant tem-

perature, but at two different densities, are then related to each other by equations which can be obtained by a simple extension of the concept of the Gibbs grand canonical ensemble. The special case of most general interest is that for which one of the two densities is the limit of zero density. At zero density the distribution function for a classical system is the exponent of minus the potential energy divided by kT, and may be regarded as known.

Although the equations relating the distribution functions of a liquid or solid to those at infinite dilution are rigorous and exact, they are not susceptible of numerical solution, since they involve integration over the coordinates of the number of molecules in a macroscopic portion of the system. It is, however, possible to consider a variational method which leads to certain integral equations involving integrations over fewer numbers of molecules.

The variational method discussed in this paper is essentially as follows: We consider the potentials of average force at one density, say those at infinite dilution, to be capable of some continuous variation. This, of course, causes a variation in the potentials at some other density, say that of the liquid. By assuming the variation in the potentials of zero density to be of the sort induced by a change in the mutual potentials between pairs, or between some relatively small number of molecules, the equations for the variation in the potentials of average force in the liquid can first be integrated over all but the molecules for which the change in potential is being computed, and the few with which these are interacting. One thus obtains a set of integral equations giving the variations in the liquid potentials as the sum of integrals of kernels times the variations in the potentials of zero density. The kernels are combinations of the unvaried distribution functions in the liquid.

One of the purposes of this paper is merely the formal attainment of these equations. Their use is not further discussed in the body of this article, but will be dealt with in a subsequent paper. By considering variations due to changes in temperature, or other physically meaningful changes, one has a tool for the calculation of the thermodynamic properties of the system in terms of the distribution functions.

However, certain formal mathematical relationships that follow from these equations appear to be of interest. The discussion of these relationships occupies the main body of the paper.

NOTATION

Owing to the relatively complex equations involved, it is desirable to use a special symbolism in order to attain some brevity of expression. The final equations are in no way more complicated if a system of many components is considered than if a single component system is handled. We deal with an infinite system containing σ types of molecular species, and use the subscript s (or r) to refer to a special species of type s (or r). Symbols ρ, z, without subscripts refer to a set of quantities:

$$\rho = \rho_1, \rho_2, \cdots, \rho_s, \cdots \rho_\sigma, \tag{1}$$

$$z = z_1, z_2, \cdots, z_s, \cdots z_\sigma. \tag{2}$$

The symbol ρ_s indicates the average concentration of molecules of species s, in molecules per unit volume, and z_s is the activity or fugacity, defined by the equation:

$$z_s = \lim_{\rho_0 \to 0} [\rho_{0s} \exp(\mu_s - \mu_{0s})/kT], \tag{3}$$

in which μ_s is the chemical potential of species s in the system, and the limit is taken as all concentrations, ρ_0, approach zero, with μ_{0s} the chemical potential, and ρ_{0s} the concentration, of species s, in this limit at which the system approaches a perfect gas in properties. Thus z_s is proportional to the exponent of the chemical potential μ_s, divided by kT, with the proportionality constant so chosen that z_s/ρ_s becomes unity in the perfect gas at infinite dilution of the system:

$$\lim \rho \to 0 [z_s/\rho_s] = 1. \tag{4}$$

The letters k, m, n, κ, μ, ν, λ, etc., when written without subscripts, will be used to indicate sets of integers:

$$k = k_1, k_2, \cdots, k_s, \cdots k_\sigma, \tag{5}$$

of k_s molecules of type s. The shorthand notation:

$$k! = k_1! k_2! \cdots k_s! \cdots k_\sigma!, \tag{6}$$

$$\rho^k = \rho_1^{k_1} \rho_2^{k_2} \cdots \rho_s^{k_s} \cdots \rho_\sigma^{k_\sigma},$$
$$z^k = z_1^{k_1} z_2^{k_2} \cdots z_s^{k_s} \cdots z_\sigma^{k_\sigma}, \tag{7}$$

will be employed. With this notation the equations for the σ component system become identical with those for a single component system of activity z, or concentration ρ, in which k, m, n, etc., is used to indicate the number of molecules of the single component. The reader will probably find the ensuing discussion simpler if the one component case only is kept in mind.

The coordinates of the ith molecule of type s are indicated by

$$(i_s) = x_{i_s}, y_{i_s}, z_{i_s}, q_{1i_s}, \cdots q_{fi_s}, \tag{8}$$

including any internal coordinates, q, and the volume element, calculated in the Cartesian space of the constituent atoms by

$$d(i_s) = dx_{i_s} dy_{i_s} dz_{i_s} \begin{bmatrix} q \\ x \end{bmatrix} dq_{i_s} \cdots dq_{fi_s}, \tag{9}$$

with $\begin{bmatrix} q \\ x \end{bmatrix}$ the Jacobian transformation to the Cartesian volume. The introduction of the internal coordinates causes no change in the ensuing equations. The reader will probably find the visualization of the operations easier if the case that no internal coordinates are present is kept in mind. The coordinates of k_s molecules of type s are indicated by

$$\{k_s\} = (1_s), (2_s), \cdots, (k_s), \tag{10}$$

and the volume element by

$$d\{k_s\} = d(1_s)d(2_s)\cdots d(k_s), \tag{11}$$

whereas the coordinates of the set k of molecules are,

$$\{k\} = \{k_1\}, \{k_2\}, \cdots, \{k_s\}, \cdots \{k_\sigma\}, \tag{12}$$

and the volume element is

$$d\{k\} = d\{k_1\}d\{k_2\}\cdots d\{k_s\}\cdots d\{k_\sigma\}. \tag{13}$$

A function, $\Psi_n\{n\}$, of the coordinates of the set $\{n\}$ of molecules may happen to be simply expressible as a sum of functions of the coordinates $\{\nu\}_n$ of small subsets of molecules out of the set n. The typical component function, $\psi_\nu\{\nu\}_n$, composing a member of this sum, say for example, $\psi_2((i_s)(j_s))$, of the coordinates of the i_sth and j_sth molecule, will, if Ψ_n is symmetrical in permutations of identical molecules, have the same functional dependence on the coordinates that the function $\psi_2((j_s)(k_s))$ has upon the coordinates of the j_sth and k_sth molecule, but the two functions are distinct in the sum and have, in general, different values for some arbitrary value $\{n\}$ of the coordinates of the set n of molecules. The summation symbol, $(\sum\{\nu\}_n)$, will be used to indicate summation of functions of such subsets, $\{\nu\}_n$, over all subsets, for all values of ν, $0 \leq \nu_s \leq n_s$, including a single (constant) term for which all $\nu_s = 0$. It will be noted that there are $n!/(n-\nu)!\nu!$ different distinct subsets, $\{\nu\}_n$, for each value of the set of numbers, $\nu = \nu_1, \cdots, \nu_s, \cdots$, so that in the trivial case that the function is a constant, $(-1)^\nu$, for all νs, then

$$(\sum\{\nu\}_n)(-)^\nu = 0 \text{ if } n > 0, \tag{14}$$

$$(\sum\{\nu\}_n)(-)^\nu = 1 \text{ if } n = 0. \tag{15}$$

Thus if Ψ_n is one of a set of functions of the coordinates $\{n\}$, symmetrical in permutations of identical molecules, defined for all values of n including $n=0$, for which the function, Ψ_0, is a constant,

$$\Psi_n = \Psi_n\{n\}, \tag{16}$$

then we may define functions ψ_ν by

$$\psi_\nu = (\sum\{n\}_\nu)(-)^{\nu-n}\Psi_n, \tag{17}$$

and, in view of (14) and (15) we have that,

$$\Psi_n = (\sum\{\nu\}_n)\psi_\nu, \tag{18}$$

since then on inserting (17) in (18) we obtain the identity:

$$\begin{aligned} \Psi_n &= (\sum\{\nu\}_n)(\sum\{\kappa\}_\nu)(-)^{\nu-\kappa}\Psi_\kappa \\ &= (\sum\{\kappa\}_n)(\sum\{\nu-\kappa\}_{n-\kappa}(-)^{\nu-\kappa})\Psi_\kappa \\ &= \Psi_n. \end{aligned}$$

THE DISTRIBUTION FUNCTION EQUATION

With this notation we now define distribution functions $F_n(\rho, T, \{n\})$, of the coordinates of a set n of molecules, by the statement that in a system of infinite extent ($V \gg n_s/\rho_s$ for all s) at the temperature T, and concentration ρ, the probability that there are an appropriate set n of molecules at the coordinates $\{n\}$ within the volume element $d\{n\}$ is equal to $\rho^n F_n(\rho, T, \{n\})d\{n\}$. It is to be

noted that with this definition the functions F_n are so normalized that,

$$\lim_{V\to\infty}\left[\frac{1}{V^n}\int\int\cdots_V\cdot F_n d\{n\}\right]=1. \tag{19}$$

We arbitrarily include $F_0\equiv 1$ in the set of functions so defined.

In a previous paper[1] it has been shown that for two sets of concentrations ρ_α and ρ_β, for which the pressures are P_α and P_β, and for which the fugacity sets are z_α and z_β, respectively, the distribution functions are related by the equation set:

$$[\exp(P_\alpha V/kT)][\rho_\alpha/z_\alpha]^n F_n(\rho_\alpha, T, \{n\})$$
$$=\sum_{m\geq 0}\frac{(z_\alpha-z_\beta)^m}{m!}\int\int\cdots_V\int e^{P_\beta V/kT}\left[\frac{\rho_\beta}{z_\beta}\right]^{n+m}F_{n+m}(\rho_\beta, T, \{n+m\})d\{m\}. \tag{20}$$

This actually represents two sets of equations for each pair of concentrations ρ_α and ρ_β, since obviously the subscripts on the right and left can be interchanged. Thus if we define:

$$G_n=e^{P_\alpha V/kT}\left[\frac{\rho_\alpha}{z_\alpha}\right]^n F_n(\rho_\alpha, T, \{n\}), \tag{21}$$

$$G_n^*=e^{P_\beta V/kT}\left[\frac{\rho_\beta}{z_\beta}\right]^n F_n(\rho_\beta, T, \{n\}), \tag{21'}$$

$$z=z_\alpha-z_\beta, \tag{22}$$

we may write (20) as,

$$G_n=\sum_{m\geq 0}\frac{z^m}{m!}\int\int\cdots_V\int G^*_{n+m}d\{m\}, \tag{23}$$

or as,

$$G_n^*=\sum_{m\geq 0}\frac{(-z)^m}{m!}\int\int\cdots_V\int G_{n+m}d\{m\}. \tag{23'}$$

THE VARIATIONAL METHOD

Without regard, at first, to the physical interpretation of the functions G_n and G_n^* we shall now proceed to carry out some mathematical manipulations which lead us to a set of integral equations that are of considerable interest in interpreting the relationships between these two sets of functions. The only mathematical limitations which must be imposed upon the functions G_n and G_n^* are to require that they be symmetrical in permutations of identical molecules, and of such a form that the series (23) and (23') converge absolutely. These conditions are assuredly fulfilled for the physical case. The symmetry in permutations of the coordinates of identical molecules is obvious. In addition we have that G_n and G_n^* are positive definite, from their definitions in terms of probability functions, and for repelling molecules G_n and G_n^* approach zero in value for all values of the coordinates $\{n\}$ within a finite volume V as n/V approaches infinity.

We now assume that G_n, and therefore through (23') G_n^*, depend on a parameter y, defining:

$$\Psi_n=d\ln G_n/dy, \tag{24}$$

$$\Psi_n^*=d\ln G_n^*/dy. \tag{24'}$$

For instance, we may take y to be the temperature T, in which case, from (21) and (24) one could relate the Ψ's to the various temperature derivatives of the system. The function $F_n(0, T, \{n\})$, for $\rho\equiv 0$, is $\exp-(U_n/kT)$, with U_n the potential energy, if the system is classical. One might, for

[1] W. G. McMillan, Jr. and J. E. Mayer, J. Chem. Phys. 13, 276 (1945) and for component systems, J. E. Mayer, J. Chem. Phys. 10, 629 (1942).

instance, consider that U_n is $U_n^0 - yU_n^{(1)}$ with $U_n^{(1)}$ some sort of perturbation potential. In this case, if G_n^* referred to $\rho \equiv 0$, the functions Ψ_n^* would be $1/kT$ times the perturbation potential for the set n of molecules. The functions Ψ_n could then be related to the derivatives of the system with respect to this perturbation.

We may now define the functions ψ_ν and ψ_ν^* by (17) and by,

$$\psi_\nu^* = (\sum\{n\}_\nu)(-)^{\nu-n}\Psi_n^*, \tag{25}$$

so that (18), and,

$$\Psi_n^* = (\sum\{\nu\}_n)\psi_\nu^*, \tag{26}$$

are valid.

We now proceed to derive Eq. (29) by which the function ψ can be calculated as a sum of integrals of kernels multiplied by the functions ψ^*. The kernels are given by Eq. (30) as combinations of the functions G. A similar reciprocal set of equations, (29′) and (30′), can be used to compute the functions ψ^* by integration over kernels formed from the G^*'s multiplied by the functions ψ. These equations are obtained by differentiation of both sides of (23) with respect to y, obtaining:

$$\begin{aligned} G_n\Psi_n &= \frac{d}{dy}G_n = \frac{d}{dy}\sum_{m\geq 0}\frac{z^m}{m}\int\int\cdots\int G_{n+m}^* d\{m\} \\ &= \sum_{m\geq 0}\frac{z^m}{m!}(\sum\{\nu\}_n, \{\mu\}_m)\int\int\cdots\int \psi_{\nu+\mu}^* G_{n+m}^* d\{m\} \\ &= \sum_{\mu\geq 0}\frac{z^\mu}{\mu!}(\sum\{\nu\}_n)\int\int\cdots\int \psi_{\nu+\mu}^*\left[\sum_{k\geq 0}\frac{z^k}{k!}\int\int\cdots\int G_{n+\mu+k}^* d\{k\}\right]d\{\mu\} \\ &= \sum_{\mu\geq 0}\frac{z^\mu}{\mu!}(\sum\{\nu\}_n)\int\int\cdots\int \psi_{\nu+\mu}^* G_{n+\mu} d\{\mu\}. \end{aligned} \tag{27}$$

In this the second line follows from the first by the use of (24′) and (26) and by considering that

$$(\sum\{\nu\}_{n+m}) = (\sum\{\nu\}_n)(\sum\{\mu\}_m) = (\sum\{\nu\}_n\{\mu\}_m),$$

where the last notation means summation over all subsets $\{\nu\}_n$ and $\{\mu\}_m$ including $\nu \equiv 0$, and $\mu \equiv 0$. For any given value of the set m one may arbitrarily call $m-\mu = k$, and, remembering that there are $m!/\mu!k!$ different subsets $\{\mu\}_m$ for any given *value* of the set m and μ, one may multiply by this and sum over μ and k rather than m and $\{\mu\}_m$. This leads to the third line. The fourth line follows by using equation (23).

We thus have:

$$\Psi_n = \sum_{\mu\geq 0}\frac{z^\mu}{\mu!}(\sum\{\nu\}_n)\int\int\cdots\int \psi_{\nu+\mu}^*\frac{G_{n+\mu}}{G_n}d\{\mu\}. \tag{28}$$

To this equation we apply the operation of equation (17) to both sides, and obtain,

$$\psi_k = \sum_{m\geq 0}\frac{z^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int \psi_{\kappa+m}^* K_{k,\kappa,m} d\{m\}, \tag{29}$$

in which the kernels, $K_{k,\kappa,m}$, are defined by the equation:

$$K_{k+n,n,m} = (\sum\{\kappa\}_k)(-)^{k-\kappa}\frac{G_{\kappa+n+m}}{G_{\kappa+n}}. \tag{30}$$

The application of the equivalent series of operations to Eq. (23′), instead of (23), leads to the equations:

$$\psi_k^* = \sum_{m\geq 0}\frac{(-z)^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int \psi_{\kappa+m} K_{k,\kappa,m}^* d\{m\}, \tag{29′}$$

with,

$$K^*_{k+n,\,n,\,m}=(\sum\{\kappa\}_k)(-)^{k-\kappa}\frac{G^*_{k+n+m}}{G^*_{k+n}}. \tag{30'}$$

Special values of the kernels follow from (30) and (30') to be,

$$K_{0,\,0,\,0}=K^*_{0,\,0,\,0}=1, \tag{31a}$$

$$K_{0,\,0,\,m}=G_m/G_0;\quad K^*_{0,\,0,\,m}=G^*_m/G^*_0, \tag{31b}$$

$$K_{k+n,\,n,\,0}=K^*_{k+n,\,n,\,0}=0 \text{ if } k>0, \tag{31c}$$

$$K_{n,\,n,\,0}=K^*_{n,\,n,\,0}=1. \tag{31d}$$

AN OVERSIMPLIFICATION

Equations (29) and (29') apply for all values of $k\geq 0$. We may choose either the set ψ_k of functions *or* the set ψ_k^* as arbitrary, limited only by the condition that they be symmetrical in permutations of the coordinates of identical molecules.

A special case of the integral equation (29) (or (29') but not both) has been used by Kirkwood and Monroe[2] to determine distribution functions for liquids. The special case chosen by them may be constructed from our equations as follows. We use (23) with the asterisk, G_n^*, referring to infinite dilution of a one-component system, and G_n referring to the liquid state. Differentiation is then performed with respect to the coordinate of a molecule at a particular position over which the integration is not extended, that is, one of the coordinates of one of the molecules, (i), of the set n in (23). The functions $\Psi_n^* kT$ are then the true forces, and are zero unless n contains the molecule (i). They are symmetrical in the coordinates of all identical molecules except the molecule i. The ψ_n^* are assumed to be zero unless $n=2$. The functions $\Psi_n kT$ are the average forces in the liquid, which are related, by a well-known relationship to the unknown distribution functions G_n, in the liquid. Equation (29) then has unknown ψ_n and kernel K, but a subsidiary relationship between them is known. It then represents a non-linear integral equation, which Kirkwood and Monroe solved approximately with certain assumptions.

The form (29') for their problem would have presented them with a linear (non-homogeneous) integral equation with known kernel K^* and known function ψ^*, but unknown average force ψ. The possible advantages of this form have not been investigated.

The standard form of a non-homogeneous linear integral equation of the Fredholm type may be written as

$$\psi(x)=\psi^*(x)-z\int K(x,y)\psi^*(y)dy, \tag{32}$$

in which ψ and the kernel $K(x,y)$ are known functions, and ψ^* is the unknown function to be determined from the equation. There exists a reciprocal kernel $K_z^*(x,y)$ depending on the value of the parameter z, such that, independent of ψ,

$$\psi^*(x)=\psi(x)-z\int K_z^*(x,y)\psi(y)dy. \tag{32'}$$

It is, of course, quite unimportant whether the variables x and y are one dimensional or a set of coordinates in multidimensional space.

The reciprocal kernel, K^*, varies continuously with the parameter z between the value $z=0$ (where $K^*=K$) and the first eigenvalue, z_γ, for which the homogeneous equation,

$$\psi^*(x)=z_\gamma\int K(xy)\psi^*(y)dy, \tag{33}$$

has a solution. For values of z between z_γ and z_δ, the second eigenvalue, K^* again varies continuously with z. In general the functions which K^* approach as z approaches z_γ from above and below are different.

Now let us consider a special case of (29), namely that there is only one component kind of molecule, and that ψ_k^*, which we can choose arbitrarily, is identically zero except for $k=1$.

[2] John G. Kirkwood and Elizabeth Monroe, J. Chem. Phys. 9, 514 (1941).

There are, then, only two non-zero terms in the sum on the right, that with $\kappa=1, m=0$, and that with $\kappa=0, m=1$. We consider the case $k=1$, for which the kernel $K_{1,1,0}$ when $\kappa=1, m=0$ is unity from (31d). The equation for this case is,

$$\psi_1(i)=\psi_1^*(i)+z\int K_{1,0,1}((i),(j))\psi_1^*(j)d(j), \quad (34)$$

identical in form to (32). *Were we justified in assuming* that $\psi_k=0$ for all values of k except $k=1$ *then* (29′) for $k=1$ *would* be,

$$\psi_1^*(i)=\psi_1(i)-z\int K^*_{1,0,1}((i)(j))\psi_1(j)d(j)! \quad (34')$$

of the same form as (32′). That is *were the assumption correct* then $K_{1,0,1}$ and $K^*_{1,0,1}$ would be simple reciprocal kernels for the parameter z.

We may now examine more closely what the kernels $K_{1,0,1}$ and $K^*_{1,0,1}$ are. Choosing $\rho_\beta=0$, $z_\beta/\rho_\beta=1$, in Eqs. (21) and (21′) defining G and G^*, so that $z_\alpha=z$ from (22), and using (30) and (30′) defining the kernels we have,

$$K^*_{1,0,1}((i),(j))=\frac{F_2^*((i),(j))}{F_1^*(i)}-F_1^*(j), \quad (35)$$

$$K_{1,0,1}((i),(j))=\frac{\rho}{z}\left[\frac{F_2((i),(j))}{F_2(i)}-F_1(j)\right], \quad (35')$$

where the asterisks refer to infinite dilution and the functions F are the distribution functions at concentration ρ. Now if there are no internal coordinates then $F_1^*\equiv 1$, so that,

$$K^*_{1,0,1}=F_2^*-1, \quad (36)$$

which is a function symmetrical in permutations of the coordinates of the two molecules, approaching zero in value as the distance between them becomes large.

At least in the gaseous and liquid phases we also have $F_1=1$, so that for these phases,

$$K_{1,0,1}=\frac{\rho}{z}[F_2-1], \quad (36')$$

which is also symmetrical and approaches zeros for large distances.

The simplicity of the result obtained from this unjustified assumption of identifying $\psi_k=0$ for $k>1$ when $\psi_k^*=0$ for $k>1$ is made particularly intriguing when we consider that the function (36′) will show discontinuities in form as z changes through the eigenvalues of the kernel (36), so that these eigenvalues would be identified with the values of the activity at which phase transitions occur.

The actual situation is far more complicated due to the occurrence of the summations in the original Eqs. (29) and (29′). We have discussed this simple case at some length because of its suggestiveness, and because it appears to represent a simplified caricature of the true case.

THE KERNELS AS OPERATORS

We shall show that Eq. (29) may be regarded as a generalized matrix equation, in which a matrix, of which combinations of the kernels, K, constitute the elements, operate on the set of functions ψ_k^* to transform them to the set ψ_k. The equation (29′) is the reciprocal equation, in which the reciprocal matrix, of elements formed from the kernels K^*, operate on the set of functions ψ_k, transforming them to the set ψ_k^*.

These matrices, which transform from the activity z_β to the activity z_α, are functions of the distribution functions at the activity z_α, and of the difference, $z=z_\alpha-z_\beta$, Eq. (22), in the activity. We can thus consider the matrices, $A_z(y)$, as depending on two parameters, z and y, the matrix consisting of elements composed of the distribution functions at activity z, and transforming the set of functions ψ_k^* corresponding to the activity $z-y$ into the set ψ_k at z. We then obtain the matrix Eqs. (46) between the matrices for different values of z and y.

Let us use the symbol $a(\{k\},\{m\})$ for a function of the coordinates $\{k\}$ and $\{m\}$, defined by,

$$a(\{k\},\{m\})=(\sum\{\kappa\}_k\{\mu\}_m)\,\delta(\{\kappa\}_k,\{\mu\}_m)\frac{z^{m-\mu}}{(m-\mu)!}K_{k,\kappa,m-\mu}, \quad (37)$$

where the summation runs over all subsets, $\{\kappa\}_k$, of the set $\{k\}$, and all subsets, $\{\mu\}_m$, of $\{m\}$, but $\delta(\{\kappa\}_k,\{\mu\}_m)=0$ unless the two subsets contain the same numbers of the same types of molecules, and unless the coordinates of all corresponding molecules of the two sets are identical; and in the latter case the integral of $\delta(\{\kappa\}_k,\{\mu\}_m)$ about the position of identity of the coordinates

is unity. We similarly define:

$$a^*(\{k\},\{m\}) = (\sum\{\kappa\}_k, \{\mu\}_m) \delta(\{\kappa\}_k,\{\mu\}_m)\frac{(-z)^{m-\mu}}{(m-\mu)!}K^*_{k,\kappa,m-\mu}. \quad (37')$$

With these two quantities we may write Eqs. (29) and (29′) as,

$$\psi_k = \sum_{m\geq 0}\int a(\{k\},\{m\})\psi_m^* d\{m\}, \quad (38)$$

$$\psi_k^* = \sum_{m\geq 0}\int a^*(\{k\},\{m\})\psi_m d\{m\}. \quad (38')$$

These equations are now in a more conventional form than that of (29) and (29′). If Φ and Φ^* are the sets of functions:

$$\Phi = \psi_0, \psi_1, \cdots, \psi_k, \cdots, \quad (39)$$

$$\Phi^* = \psi_0^*, \psi_1^*, \cdots, \psi_k^*, \cdots; \quad (39')$$

and A and A^* are the matrices of elements $a(\{k\},\{m\})$ and $a^*(\{k\},\{m\})$, respectively; then (38) and (38′) may be written as:

$$\Phi = A\Phi^*, \quad \Phi^* = A^*\Phi, \quad (40)$$

$$AA^* = A^*A = E, \quad (41)$$

With E the unit matrix. The matrix multiplication of the two elements involves *both* summation over the middle index *m and* integration over the coordinates $\{m\}$ for each value of the index m. The matrix A is thus a generalized type of matrix with continuous and discrete ranges. Thus we see that the Eqs. (29) and (29′) do correspond to the equations for two reciprocal commutative linear operators whose matrix elements can be represented by the a and a^* of Eqs. (37) and (37′).

Written in terms of these matrix elements equation (41) takes the form:

$$\sum_{m\geq 0}\int\int\cdots\int_V a(\{k\},\{m\})a^*(\{m\},\{n\})d\{m\} = \sum_{m\geq 0}\int\int\cdots\int_V a^*(\{k\},\{m\})a(\{m\},\{n\})d\{m\} = \delta(\{k\}\{n\}). \quad (42)$$

This equation could be used, in principle at least, as a means of determination of the functions G from known functions G^*.

The matrices A and A^* which we have defined by means of their elements in Eqs. (37) and (37′) depend both on the activity set z at which the kernels are determined (z_α for A and z_β for A^* in Eqs. (21) and (21′)), and upon the difference of activity, $z = z_\alpha - z_\beta$, between the two systems. Let us change our nomenclature at this place, and write the matrices as $A_z(y)$ of element,

$$a_z(y,\{k\},\{m\}) = (\sum\{\kappa\}_k\{\mu\}_m)\delta(\{\kappa\}_k\{\mu\}_m)\frac{y^{m-\mu}}{(m-\mu)!}K_{k,\kappa,m-\mu}(z), \quad (43)$$

in which the kernels are made up according to (30) from the distribution functions at the activity set z.

If y becomes identically zero there remain in (43) only the terms for which $\mu = m$, and from (31c) and (31d) these are zero unless $k = \kappa = \mu = m$, and are then unity. We thus have,

$$\lim_{y\to 0}[A_z(y)] = E, \quad (44)$$

is the unit operator.

Now let $\Phi(z)$ represent the set of functions $\psi_0(z), \psi_1(z), \cdots, \psi_k(z)$ giving the variations in the potentials at the activity z. The transformation equations, (40), then take the form:

$$\Phi(z) = A_z(y)\Phi(z-y). \quad (45)$$

If we operate on this with $A_{z+x}(x)$ we find:

$$\Phi(z+x) = A_{z+x}(x)\Phi(z) = A_{z+x}(x)A_z(y)\Phi(z-y) = A_{z+x}(x+y)\Phi(z-y),$$

or

$$A_{z+x}(x+y) = A_{z+x}(x)A_z(y). \quad (46)$$

Equation (41) is seen to be a special case of (46), namely that for which $x = -y$, in which, from (44), $A_{z+x}(0) = E$, and,

$$A_{z+x}(x)A_z(-x) = E. \quad (46')$$

A SPECIAL EQUATION

Equation (42), which represnts a set of equations, one for each value of k and n, is comparatively awkward to handle for the case of general values of k and n. It can be seen that the equation is trivially satisfied if $n \equiv 0$, as we shall show in this section. If $k \equiv 0$ and $n = 1$, one obtains an expression which leads to an equation

for the activity, z_s, of the single molecule represented in the set $n=1$.

We use the symbol i_s or j_s to represent a single molecule of type s, and (i_s) or (j_s) to represent its coordinates, but use the subscript s to indicate a kernel in which the corresponding set n is a set consisting of one molecule of type s, and $m-s$ when the set is some general set, m, $(m_s \geq 1)$ less one molecule of type s. If n represents a set consisting of a single molecule of type s then $z^n = z_s$ in our notation. Using (37) and (37′), and Eqs. (31a) to (31d) inclusive, for special values of the kernels, we may write:

$$\left.\begin{aligned} a(0, \{m\}) &= (z^m/m!)K_{0,0,m} \\ &= (z^m/m!)G_m/G_0, \\ a^*(\{m\}, 0) &= a(\{m\}, 0) = \delta(m, 0), \\ a^*(\{m\}, (i_s)) &= -z_s K^*_{m,0,s} + \delta(\{m\}, (i_s)). \end{aligned}\right\} \quad (47)$$

Using these, the sum $\sum a(0, \{m\})a^*(\{m\}, 0)$ has only the single term $G_0/G_0 = 1$ when $m = 0$, so that Eq. (42) for the case $k \equiv 0$, $n \equiv 0$, is trivially satisfied. Similarly, all sums

$$\sum a(\{k\}\{m\})a^*(\{m\}, 0)$$

contain no non-zero terms if $k > 0$, from the second of these equations, so that (42) for this case is also trivially satisfied.

The sum $\sum a(0, \{m\})a^*(\{m\}, (i_s) = 0$, leads to the equation:

$$-z_s \frac{1}{G_0} \sum_{m \geq 0} \frac{z^m}{m!} \int\int \cdots \int G_m K^*_{m,0,s} d\{m\} + z_s \frac{G_s}{G_0} = 0. \quad (48)$$

If, in this, one sets the special case that the functions with asterisks refer to infinite dilution, see Eqs. (21), (21′) and (22), then $z \equiv z_\alpha$, and

$$z^m G_m/G_0 = \rho^m F_m(\rho, T, \{m\}),$$

is the probability density that the set m of molecules will occur at the position $\{m\}$. The functions $K^*_{m,0,s}$ are now, from (30′),

$$K^*_{m,0,s} = (\sum \{\mu\}_m)(-)^{m-\mu} \frac{F^*_{\mu+s}(0, T, \{\{\mu\}_m + (i_s)\})}{F^*_\mu(0, T, \{\mu\}_m)},$$

where the distribution function at zero concentration is:

$$F^*_\mu(0, T, \{\mu\}) = \exp -(U_\mu\{\mu\}/kT), \quad (49)$$

and U_μ is the potential energy of the μ molecules, if the system is classical.

Equation (48) now becomes,

$$(\rho_s/z_s)F_s(\rho, T, (i_s)) = \sum_{m \geq 0} \frac{\rho^m}{m!} \int\int \cdots \int F_m(\rho, T, \{m\}) \times \{\exp[(-1/kT)(U_{\mu+s}\{\{\mu\} + (i_s)\} - U_\mu\{\mu\})]\} d\{\mu\}, \quad (50)$$

as a general equation for any system, if U is understood to be defined by (49) when the system is non-classical, in which case it is not a potential energy.

If Eq. (50) is divided by V, and integrated over the coordinates $d(i_s)$, due to the normalization condition (19), one obtains ρ_s/z_s, the inverse of the activity coefficient, on the left-hand side.

Let us assume a more specialized case, and one that is particularly simple. We assume that there are no internal coordinates, or that they have been integrated out of the distribution functions at zero density so that the potential U_μ is a function of the centers of mass, averaged over the internal coordinates. We further consider a single component system, and one existing in the liquid state so that F_s is unity. If, now, it is legitimate to express U_μ as a sum of terms, u_{jk}, each dependent only on the coordinates of the pair of molecules, j and k, then if,

$$f_{ij} = [\exp -(u_{ij}/kT)] - 1, \quad (51)$$

we have that,

$$K^*_{m,0,s} = \prod_{j=1}^{j=m} f_{ij}. \quad (52)$$

The functions f_{ij} approach zero in value if the distance between the molecules i and j become large, since their mutual potential is then zero. Equation (50) now takes the form,

$$\frac{\rho}{z} = \sum_{m \geq 0} \frac{\rho^m}{m!} \int\int \cdots \int F_m(\rho, T, \{m\}) \times \prod_{j=1}^{j=m} f_{ij} d\{m\}. \quad (53)$$

The distribution functions F_m become zero if too many molecules are crowded into too small a space. The product of f_{ij}'s is zero unless *all* of the m molecules are within interacting distance of the molecule i. As m becomes large there are no positions of the coordinates for which the integrand is non-zero, so the sum converges, and, indeed, terms with $m>13$ must be very small.

This equation is essentially a cell type equation in that one computes the activity, and therefore the free energy, by considering only the interactions of one molecule with those others that can approach it at any one time. The equation is exact in the form (50). It is not essentially different from that derived previously by the author with Montroll,[3] although it is cast in an entirely different appearing form.

EQUATIONS BETWEEN THE KERNELS

To use Eq. (42) in the case of general values of k and n, and convert it to a direct summation of integrals of products of the kernels K and K^* is comparatively difficult. One runs into rather awkward summations, which then simplify with considerable labor. The simplest form of these equations can be found more readily by another method. However, we wish to emphasize that the equations derived in this section are actually another form of (42).

We use Eq. (23′), write it for a set $k+n$ of numbers, sum over μ, and divide and multiply each term in the sum by $G_{k+\mu}$, using, however, (23) for $G_{k+\mu}$ in the numerator:

$$G^*_{k+n}=\sum_{\mu\geq 0}\frac{(-z)^\mu}{\mu!}\int\int\cdots\int_V G_{k+n+\mu}d\{\mu\}$$

$$=\sum_{\mu\geq 0}\frac{(-z)^\mu}{\mu!}\int\int\cdots\int_V\frac{G_{k+n+\mu}}{G_{k+\mu}}\left[\sum_{\nu\geq 0}\frac{z^\nu}{\nu!}\int\int\cdots\int G^*_{k+\mu+\nu}d\{\nu\}\right]d\{\mu\}$$

$$=\sum_{\mu\geq 0}\sum_{\nu\geq 0}\frac{z^{\nu+\mu}}{\nu!\mu!}(-)^\mu\int\int\cdots\int\frac{G_{k+n+\mu}G^*_{k+\mu+\nu}}{G_{k+\mu}}d\{\nu\}d\{\mu\}.$$

In this equation we write $\nu+\mu=m$, and sum over all sets of numbers m. Since there are $m!/\nu!\mu!$ different subsets, $\{\mu\}_m$, for every value of the set of numbers m and μ, we may divide by this and sum over all $\{\mu\}_m$. One obtains,

$$G^*_{k+n}=\sum_{m\geq 0}\frac{(-z)^m}{m!}\int\int\cdots\int_V G^*_{k+m}\left[(\sum\{\mu\}_m)(-)^{m-\mu}\frac{G_{k+\mu+n}}{G_{k+\mu}}\right]d\{m\},$$

or with Eq. (30),

$$G^*_{k+n}=\sum_{m\geq 0}\frac{(-z)^m}{m!}\int\int\cdots\int_V G^*_{k+m}K_{m+k,k,n}d\{m\}. \tag{54}$$

If instead we start with (23), the equation,

$$G_{k+n}=\sum_{m\geq 0}\frac{z^m}{m!}\int\int\cdots\int G_{k+m}K^*_{m+k,k,n}d\{m\}, \tag{54′}$$

may be obtained by the corresponding manipulations.

At first sight Eqs. (54) and (54′) appear to be only complications of the relatively simple forms (23) and (23′), but the forms (54) and (54′) have one enormous advantage, namely that the summations converge far more rapidly in these equations than in (23) and (23′). This superiority of convergence will be discussed in greater detail later, and is merely remarked here in passing.

[3] J. E. Mayer and E. Montroll, J. Chem. Phys. 9, 2 (1941), Eq. (70).

We now use (54) to obtain an equation for $K^*_{k,0,m}$, using (30′) defining this kernel as:

$$K^*_{k,,0n}=(\sum\{\kappa\}_k)(-)^{k-\kappa}G^*_{\kappa+n}/G^*_\kappa=\sum_{m\geq 0}\frac{(-z)^m}{m!}\int\int\cdots\int(\sum\{\kappa\}_k)(-)^{k-\kappa}K^*_{\kappa,\kappa,m}K_{m+\kappa,\kappa,n}d\{m\}.$$

However from (30′) we may write,

$$K^*_{\kappa,\kappa,m}=(\sum\{\lambda\}_{\kappa-\mu})K^*_{\lambda+\mu,\mu,m},$$

where $\mu=\{\mu\}_\kappa$ is any subset of κ, and from (30),

$$K_{m+\kappa,\kappa,n}=(\sum\{\mu\}_\kappa)K_{m+\mu,0,n},$$

so that,

$$(\sum\{\kappa\}_k)(-)^{k-\kappa}K^*_{\kappa,\kappa,m}K_{m+\kappa,\kappa,n}=(\sum\{\kappa\}_k\{\mu\}_k\{\lambda\}_{k-\mu})(-)^{k-\kappa}K^*_{\lambda+\mu,\mu,m}K_{m+\mu,0,n}$$

$$=(\sum\{\mu\}_k\{\lambda\}_{k-\mu})(-)^{k-\mu-\lambda}(\sum\{\kappa-\mu-\lambda\}_{k-\lambda-\mu}(-)^{\kappa-\mu-\lambda})K_{\lambda+\mu,\mu,m}K^*_{m+\mu,0,n}$$

$$=(\sum\{\mu\}_k)K^*_{k,\mu,m}K_{m+\mu,0,n},$$

where the last line follows from the fact that the summation in the brackets of the preceding expression is zero unless $\lambda+\mu=k$. One therefore has,

$$K^*_{k,0,n}=\sum\frac{(-z)^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int_V K_{k,\kappa,m}K_{m+\kappa,0,n}d\{m\}. \tag{55′}$$

If, instead, (54′) is used, one finds,

$$K_{k,0,n}=\sum_{m\geq 0}\frac{z^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int_V K_{k,\kappa,m}K^*_{m+\kappa,0,n}d\{m\}. \tag{55}$$

Now for fixed values of the coordinates $\{n\}$ the kernels $K^*_{k,0,n}$ and $K_{k,0,n}$ are functions of the coordinates $\{k\}$, symmetrical in permutations of the coordinates of identical molecules. It is seen that Eqs. (55) and (55′) are identical to Eqs. (29) and (29′) with $K^*_{k,0,n}$ replaces ψ_k^* and $K_{k,0,n}$ replacing ψ_k. That is, the functions $K_{k,0,n}$ and $K^*_{k,0,n}$ for any set of numbers n and any value of the coordinates $\{n\}$, regarded as functions of the coordinates $\{k\}$, form a suitable set ψ_k and the corresponding solutions ψ_k^*, respectively, to Eqs. (29) and (29′) or (38) and (38′).

THE CONVERGENCE OF THE EQUATIONS

Whether or not the sum in Eqs. (29) or (29′) converge depend on the variation functions ψ_k and ψ_k^*. We shall assume that the functions ψ_k^* are of the type which one might actually wish to handle if one considers variations in the potential energy of the molecules, or variations in temperature. In particular we assume that ψ_k^* approaches zero if the coordinates of any pair of the molecules composing the set k are sufficiently distant in the three-dimensional space of their mass centers, and that this approach to zero is more rapid than r^{-2}, so that,

$$\lim_{V\to\infty}\left[V^{-1}\int\int\cdots\int_V\psi_k^*d\{k\}\right]=\text{finite}. \tag{56}$$

We shall also assume that the functions ψ_k approach zero in value as $k\to\infty$, or at least that the integral of Eq. (56) approaches zero as $k\to\infty$.

We shall show that the form of the Eq. (30) defining the kernels is such, that, if the functions ψ_k^* obey the conditions above, then if the kernels are formed at an activity for which the system can only exist in one phase, the Eq. (29) will converge and the functions ψ_k will also obey the above conditions. If, however, the kernels are formed at an activity for which two phases are stable no such convergence can be expected.

For such values of the activity set z that the system exists in a single phase, we may expect the distribution functions to show certain characteristics. These functions may be written

$F_n=\exp(-W_n/kT)$ where W_n is the potential of average force. We would expect W_n in single phase systems to have no terms corresponding to long range forces. That is, we expect that for two sets of molecules, k and m, if the distances, $r_{k,m}$, between all molecules of the set k and all of the set m are large, then $W_{k+m}=W_k+W_m$. This means that:

$$\lim_{r_{m,k}\to\infty}[F_{k+m}]=F_kF_m. \tag{57}$$

Let us assume that Eq. (57) applies, and that the difference $F_{k+m}-F_kF_m$ approaches zero more rapidly than $r_{k,m}^{-2}$, an assumption that corresponds to having all forces approach zero more rapidly than r^{-3}.

Now let us examine the kernels,

$$K_{k+n,\,n\,m}=\left(\frac{\rho}{z}\right)^m(\sum\{\kappa\}_k)(-)^{k-\kappa}\frac{F_{\kappa+n+m}}{F_{\kappa+n}}, \tag{58}$$

which equation is equivalent to (30) when ρ and z refer to the concentration and activity set for the functions without asterisks, see Eq. (21). Suppose the distances, $r_{k,n+m}$, between all molecules of the set k, and all molecules of $n+m$ are large. Then from (57)

$$F_{\kappa+n+m}=F_\kappa F_{n+m},\quad F_{\kappa+n}=F_\kappa F_n,$$

and

$$F_{\kappa+n+m}/F_{\kappa+n}=F_{n+m}/F_n.$$

But the sum $(\sum\{\kappa\}_k)(-)^{k-\kappa}$ is zero if $k>0$, Eq. (15). It follows that,

$$\lim_{r_{k,\,n+m}\to\infty}[K_{k+n,\,n,\,m}]=0,\quad k>0, \tag{59}$$

and the approach to zero is more rapid than $r_{k,\,n+m}^{-2}$ according to our assumptions.

Let us now look at the product:

$$K_{k+n,\,n,\,m}\psi^*_{n+m} \tag{60}$$

which composes the integrand in the general term of the sum (29). If ψ^*_{n+m} approaches zero as any of the distances of the set $n+m$ becomes large, that is it obeys (56), then the integrand is zero unless all molecules of the set $n+m$ are close together. But if the set $n+m$ is far from the set k, then from (59) the kernel is zero, and the integrand is zero. Furthermore, the argument leading to Eq. (59) applies if *any one* of the molecules of the set k is far from all of the set $n+m$. It follows that, within our assumptions, the integrand (60) is zero unless *all* of the molecules of the complete set $k+n+m$ are reasonably close together (within range of molecular forces one to the other) in the three-dimensional space of their centers of mass. It also follows, since we have assumed conditions leading to a sufficiently rapid approach to zero, that the integral of (60) over the coordinates $\{m\}$ leads to a finite result independent of the volume V over which the range of integration is extended.

If we also assume that $\psi_k^*\to 0$ as $k\to\infty$, then we have assured the convergence of (29) since only a finite number of terms have non-zero integrands, and these each converge upon integration. But we have also shown that, within our assumptions, each integrand is zero in (29) unless all the molecules of the set k are close together, so the ψ_k's also obey the condition (56).

Now at such values of the activity set, z, that two phases can coexist in the system, Eq. (57) is not correct. For instance, in the absence of a gravitational field (which would make z a function of position in the system) the distribution function for the system would be:

$$\rho^nF_n=x\rho_a{}^nF_n{}^{(a)}+(1-x)\rho_b{}^nF_n{}^{(b)},$$

where ρ_a,ρ_b are the densities of the two phases, and $F_n{}^{(a)}$ and $F_n{}^{(b)}$ the distribution functions in the pure phases. The number x is the volume fraction of the system in phase a, and is variable at the given z. Functions constructed according to this equation do not obey (57) even if $F_n{}^{(a)}$ and $F_n{}^{(b)}$ do satisfy it singly. We may conclude that, in general, solutions to (29) are not obtainable at such values of the set z that two phases can coexist in the system at the activities at which the kernels are determined. However, satisfactory solutions should be obtainable as z is varied to approach the value of the transition, but the solutions will, in general, be different as this value is approached from the side of one phase or the other.

EIGENVALUES AS THE ACTIVITIES OF PHASE TRANSITION

The first terms in Eqs. (29) and (29′) are those for which $m=0$. From (31c) and (31d) we see that only the term with $\kappa=k$ in those of $m=0$ has a non-zero kernel, and for this term the

kernel is unity. We may consequently write:

$$\psi_k=\psi_k^*+\sum_{m\geq 1}\frac{z^m}{m!}(\sum\{\kappa\}_k)\times\int\int\cdots\int_V K_{k,\kappa,m}\psi^*_{\kappa+m}d\{m\}, \quad (61)$$

$$\psi_k^*=\psi_k+\sum_{m\geq 1}\frac{(-z)^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int_V K^*_{k,\kappa,m}\psi_{\kappa+m}d\{m\}, \quad (61')$$

as modified forms of (29) and (29′). The fact that the terms independent of z in the matrix element were $\delta(\{k\}\{m\})$ has already been noted in the discussion under Eq. (43), and Eq. (61) and (61′) merely repeat this.

If we define a matrix $B_z(y)$ by the equation:

$$B_z(y)=A_z(y)-E, \quad (62)$$

so that, from (44),

$$\lim_{y\to 0}[B_z(y)]=0, \quad (63)$$

then the transformation equation (45) becomes:

$$\Phi(z)=\Phi(z-y)+B_z(y)\Phi(z-y). \quad (64)$$

Let us write this for $z\equiv y$,

$$\Phi(z)=\Phi(0)+B_z(z)\Phi(0), \quad (65)$$

and the reciprocal equation,

$$\Phi(0)=\Phi(z)+B_0(-z)\Phi(z). \quad (65')$$

Suppose that there exists a value of the set z, namely z_γ, for which there is a solution, Φ_γ, to the homogeneous equation obtained by setting $\Phi(0)=0$ in (65′), namely:

$$\Phi_\gamma+B_0(-z_\gamma)\Phi_\gamma=0. \quad (66)$$

This would mean that there would exist a possible variation of the potentials of average force at the activity z_γ which would not be reflected in a change of the distribution functions at zero density. Without regard to the physical interpretation of this let us first examine its implication in our equations. With this Φ_γ for $\Phi(z_\gamma)$ in (65′) we find $\Phi(0)=0$. However, if we insert $\Phi(0)=0$ in (65) we cannot obtain $\Phi(z_\gamma)=\Phi_\gamma\neq 0$ if the matrix Eqs. (61) (or (38)) converge.

Now at such values of the activity set z that the system exists in a single phase only, one expects that the distribution functions be uniquely determined, and that no stable variations would exist that did not require alterations in the potentials at zero density. One also expects that no long range average forces be present, and that, therefore, as discussed in the preceding section, the matrix Eq. (61) would converge.

However, at values of the activity set for which two or more phases are present the situation is entirely altered. The distribution functions are *not* uniquely determined by the activity, since the relative fractions of the different phases vary at the constant activity. There *are*, then, variations possible which do not require changes in the potentials at zero density. But also the existence of more than one phase introduces a pseudo long range average force which completely spoils the convergence of the matrix Eq. (61), corresponding to (65).

We may therefore identify the values of z for which the homogeneous Eq. (66) has a solution with the values of the activity set for which more than one phase can coexist in the system.

We have previously[1] identified the occurrence of a phase transition with the divergence of a power series in the activity. That is, one could start with a system of any activity, say activity z, and compute the thermodynamic properties such as pressure, density, or the distribution functions, F_n, at some other activity, say $z+y$, in terms of the same property at z and a power series in y. If y were sufficiently small the power series converges, but if, as y increases, the activity $z+y$ approaches the activity, z_γ, at which a phase transition occurs, the power series diverges.

Now the kernel, reciprocal to that in a Fredholm type integral equation, can be calculated by an iterative process as a power series in the characteristic parameter, y, of the equation, provided the value of y is less than the first eigenvalue of the corresponding homogeneous equation. As y approaches this value, the series diverges.

One may regard Eq. (61′) as an equation for unknown functions ψ_k, and attempt to solve it by using $\psi_k^*=\psi_k$ under the integral as a first approximation, and use the resulting equation as a

second approximation. By reiterating this process one obtains an equation involving integrations of products of the kernels K^* multiplied by ψ^*, and by comparison with (61) one obtains an equation for K as a sum of integrals of products of the kernels K^*. The sum is a power series in z. These equations, so obtained, can be transformed into the equations formerly obtained for the thermodynamic properties and distribution functions as a power series in z.

But the iterative process can only lead to a convergent series for those values of the parameter z which are less than the first eigenvalue of the corresponding homogeneous equation. The two methods of obtaining the phase transitions are thus formally equivalent.

THE SOLUTIONS OF THE HOMOGENEOUS EQUATIONS

It is possible to find the significance of the solutions of the homogeneous Eq. (66) at the values, z_γ, of the activity coefficient for which two phases coexist.

We know, from experience, that the solutions G_n, to Eq. (23), for $n>0$, undergo discontinuous changes as z crosses the values at which phase transitions occur. It is possible to disentangle the mathematical reason for this, and to see that the value of m for the term of maximum value undergoes discontinuous transitions at these values of z. In general, then, there exist two values, $G_n^{(a)}$, and $G_n^{(b)}$, which the sum (23) approaches as z approaches z_γ, depending on whether the approach to the activity of the phase transition occurs from the activities for which phase (a) or phase (b) are stable. At $z=z_\gamma$ any value, $G_n=xG_n^{(a)}+(1-x)G_n^{(b)}$ for $0\leq x\leq 1$ is a physically meaningful solution to (23), corresponding to having the volume fraction x of the system in the volume V in the form of phase (1). We might expect to use this for G_n in (24) and take derivatives with respect to x to determine the functions Ψ_n for which no change in G_n^* is required. This procedure bogs down in mathematical difficulties.

Let us instead turn to equations (55′). We again interpret the functions with asterisks to refer to infinite dilution. As z approaches z_γ from the two sides there are two solutions, $K^{(a)}_{m,0,n}$, and $K^{(b)}_{m,0,n}$, for every m and n, that satisfy (55′) with the same values of K^*. It follows that:

$$\sum_{m\geq 0}\frac{(-z)^m}{m!}(\sum\{\kappa\}_k)\int\int\cdots\int_V K^*_{k,\kappa,m}$$
$$\times[K^{(a)}_{m+\kappa,0,n}-K^{(b)}_{m+k,0,n}]d\{m\}=0. \quad (67)$$

The functions,

$$\psi_k(\gamma,\{n\})=K^{(a)}_{k,0,n}-K^{(b)}_{k,0,n}, \quad (68)$$

and the sets;

$$\Phi_\gamma(\{n\})=\psi_0(\gamma\{n\}),\ \cdots,\ \psi_k(\gamma,\{n\}), \quad (69)$$

for any $\{n\}$ are therefore solutions to the homogeneous Eq. (66) at the activity z_γ at which phases (1) and (2) can coexist in the system.

CONCLUSION

The essential results may be summarized as follows. One may define distribution functions, F_n, as being proportional to the probability density that a set, n, of molecules are found at the coordinate postion $\{n\}$. The logarithms of these functions are $-W_n\{n\}/kT$, in which $W_n\{n\}$ is the potential of average force between the molecules at the position $\{n\}$. These potentials depend on the activity set z of the different molecular species present, and vary with z continuously from the true potential energy, at zero activity, but with discontinuities at the phase transitions.

The function, $W_n(z,\{n\})$ may be represented as a sum of terms, each depending on the coordinates, $\{\nu\}_n$, of a subset ν of molecules of the set n, summed over all subsets:

$$W_n=(\sum\{\nu\}_n)w_\nu(z,\{\nu\}_n). \quad (70)$$

At zero density, at least, one usually assumes that only the functions w_ν for $\nu\leq 2$, are non zero, corresponding to having the forces between molecules the sum of those between pairs only.

One considers a possible variation in the various functions, $w_k(z,\{k\})/kT$. A set, $\Phi(z)$, of functions, $\psi_0(z)$, $\psi_1(z,(i))$, $\cdots$, $\psi_k(z,\{k\})$, $\cdots$ may be defined by

$$\psi_k(z,\{k\})=\delta[w_k(\rho,\{k\})/kT], \quad (71)$$

where $\delta(w_k/kT)$ is the variation in w_k/kT. This set of functions, depending on the activity parameter z, represents a self-consistent set of

variations of w_k/kT, of which the set at one particular activity, most naturally that at zero activity, may be regarded as arbitrary.

In order to compute the set $\Phi(z)$, from a knowledge of the functions at some other activity, say $\Phi(z-y)$, a transformation matrix $A_z(y)$ may be employed which transforms $\Phi(z-y)$ into $\Phi(z)$. The elements, $a_z(y, \{k\}, \{m\})$, of this matrix, given by Eqs. (43), (30), and (21), consist of combinations of the distribution functions at the activity z, to which the transformation is made, multiplied by powers of y, the difference of activities. The limit of the transformation matrix, as the activity difference y becomes zero, is, of course, the unit matrix. The matrix multiplication involves both summation over all sets of molecules, and integration over the coordinates of each set.

The product of the two matrices $A_z(y)$ and $A_{z-y}(-y)$ is the unit matrix. This is a special case of the more general relation that

$$A_z(y)A_{z+z}(x) = A_{z+z}(x+y).$$

These matrix equations are equivalent to a set of sums of integral equations, (55) and (55′), relating the distribution functions at different activities of the system. The equations are convergent if the activities are such that only single phases are present in the system, even if the activity pairs are those at which different phases are stable.

One member of these equations, (53), permits a computation of the activity or free energy of the system in terms of the interactions of one molecule with its neighbors in the system.

If z is an activity for which the system exists in a single phase only, and z_γ an activity at which two phases, (a), and (b), can coexist, then there are function sets Φ_γ characteristic of z_γ which satisfy the transformation equation

$$A_z(z-z_\gamma)\Phi_\gamma = 0.$$

In particular if $z=0$, $A_0(-z_\gamma)\Phi_\gamma = 0$, Eqs. (66) and (62), show that there exist a set, Φ_γ, of variations of the potential of average force in the two phase system which can occur without variations in the temperature or potentials between the molecules at infinite dilution. The transformation matrix, $A_{z_\gamma}(y)$, to this activity, z_γ, of the phase transition, has elements such that the summations and integrations diverge. The solutions, Φ_γ, to the homogeneous equation, $A_0(-z_\gamma)\Phi_\gamma = 0$, are differences of certain combinations of the distribution functions of the two phases (a) and (b) which are at equilibrium at the activity z_γ, Eqs. (68), (30), and (21).

COMMENTS

For an alternate derivation and generalization of Eq. (54′) of this paper see Sec. VII of article by Percus and "Integral Equations and Inequalities in the Theory of Fluids" by J. L. Lebowitz and J. K. Percus, J. Math. Phys. 4, 1495 (1963).

THE PAIR DISTRIBUTION FUNCTION IN CLASSICAL STATISTICAL MECHANICS*†

by

J. K. Percus
Courant Institute of Mathematical Sciences
New York University
New York

I. INTRODUCTION

Much of physics deals with systems of enormously many degrees of freedom. The investigation of such systems is greatly facilitated if attention can be restricted to entities of direct physical significance. While there is little difficulty in selecting physical parameters which encompass all information which can reasonably be required, it is far more demanding to formulate a theory, even an approximate theory, in terms of these parameters alone. For a classical fluid in thermal equilibrium, the pair distribution function both characterizes the interparticle forces and predicts the results of the bulk of single-time observations which can be made. It is therefore the object of this paper to examine in some detail the dependence of the pair distribution function upon the forces which are acting, and to determine to what extent the pair distribution may be found without implicitly taking into account higher order distributions.

---

*Supported in part by the United States Atomic Energy Commission under Contract AT(30-1)-1480.

†This represents an elaboration of a talk given at Evanston in 1962: Bull. Am. Phys. Soc. 7, 407 (1962).

In summary, after defining the pertinent distributions in Section II, we consider the realizability problem for uniform systems in Section III. This is followed in Section IV by a discussion of the properties of nonuniform systems and of several formulations of use in their analysis. The developments of Section IV are strengthened by the introduction of the direct correlation function in Section V. In Section VI, we establish the relation between the pair distribution of a uniform system and the density of a nonuniform one, and from this develop a number of approximate integral equations for the pair distribution function. The nature of the approximations is examined more carefully in Sections VII and VIII. Finally, in Sections IX and X, the P.Y. approximate integral equation is considered in the special cases of lattice gases and hard core fluids.

II. THE PAIR DISTRIBUTION

Our primary concern will be the N-body problem of classical equilibrium statistical mechanics, in which basically all is known. That is, we can write down at sight the Gibbs phase space distribution function

$$\rho_N(\cdots \underset{\sim}{r}_i \underset{\sim}{p}_i \cdots) = e^{-\beta H(\cdots \underset{\sim}{r}_i \underset{\sim}{p}_i \cdots)}/Q_N \qquad (2.1)$$

$$\beta = 1/kT,$$

where H is the system Hamiltonian, T the temperature, and the normalizing factor Q_N the partition function.

Both in principle and in practice however, matters are not this simple. In principle, any underlying assumption of quasi-ergodicity is made tenuous on the basis of a number of regularities recently found[1] in the general structure. In practice, ρ_N is insufficient, for we are interested in average values and are thus called upon as well to perform an N-fold integration. Now most quantities of physical interest in single-component systems will involve no more than two particles at a time. Since the classical momentum distribution is trivially Maxwell-Boltzmann, we thus require the one and two-body coordinate distribution functions n_1 and n_2, defined more generally in terms of the expectation of a function of s particle coordinates:

$$(N!/N-s!)\langle f(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)\rangle \equiv \int n_s\left(\underset{\sim}{r}_1',\ldots,\underset{\sim}{r}_s'\right) f\left(\underset{\sim}{r}_1',\ldots,\underset{\sim}{r}_s'\right) d\underset{\sim}{r}_1'\ldots d\underset{\sim}{r}_s', \tag{2.2}$$

or more briefly

$$n_s\left(\underset{\sim}{r}_1',\ldots,\underset{\sim}{r}_s'\right) = (N!/N-s!)\left\langle \delta\left(\underset{\sim}{r}_1 - \underset{\sim}{r}_1'\right)\ldots\delta\left(\underset{\sim}{r}_s - \underset{\sim}{r}_s'\right)\right\rangle. \tag{2.3}$$

The normalization is such that n_s integrates to the number of s-tuplets: N particles, N(N-1) pairs, etc.

In a noninteracting system of N particles, the pair distribution attains the form $n_2(\underset{\sim}{r},\underset{\sim}{r}') = n_1(\underset{\sim}{r})n_1(\underset{\sim}{r}')(1-1/N)$. The deviation or Ursell function

$$\mathfrak{J}_2(\underset{\sim}{r},\underset{\sim}{r}') \equiv n_2(\underset{\sim}{r},\underset{\sim}{r}') - n_1(\underset{\sim}{r})n_1(\underset{\sim}{r}') \tag{2.4}$$

is then a measure of the strength of interaction, and it is this quantity which appears naturally in experimental observations. The principal source of information on the form of the Ursell function is that of light (or other wave) scattering, in which the scattered intensity at wave vector $\Delta\underset{\sim}{k}$ from incidence may be shown to have the form

$$I(\Delta\underset{\sim}{k})/I_o = \left| \int n_1(\underset{\sim}{r}) e^{i\Delta\underset{\sim}{k}\cdot\underset{\sim}{r}} d\underset{\sim}{r} \right|^2 + \iint [\mathfrak{J}_2(\underset{\sim}{r},\underset{\sim}{r}') + n_1(\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}')] e^{i\Delta\underset{\sim}{k}\cdot(\underset{\sim}{r}-\underset{\sim}{r}')} d\underset{\sim}{r}\, d\underset{\sim}{r}' \tag{2.5}$$

for a suitable constant of proportionality I_o. Closely related to this, but less directly accessible, is the degree of density fluctuation. One easily shows that the relative mean square deviation of particle number in a specified volume Ω is given by

$$\frac{\langle N(\Omega)^2\rangle - \langle N(\Omega)\rangle^2}{\langle N(\Omega)\rangle} = 1 + \frac{1}{\langle N(\Omega)\rangle} \iint_{\Omega} \mathfrak{J}_2(\underset{\sim}{r},\underset{\sim}{r}') d\underset{\sim}{r}\, d\underset{\sim}{r}'. \tag{2.6}$$

For a uniform system, one whose properties are invariant to any over-all translation in space, the one-body density is itself a constant $n_1(\underset{\sim}{r}) = n$, and one may define a dimensionless pair distribution g, a function of relative coordinate alone,

$$\begin{aligned} n_2(\underset{\sim}{r},\underset{\sim}{r}') &= n^2 g(\underset{\sim}{r}-\underset{\sim}{r}') \\ n_1(\underset{\sim}{r}) &= n. \end{aligned} \tag{2.7}$$

Rather restricted but extremely important information concerning uniform systems is that which expresses the thermodynamic equation of state. Producing uniformity by allowing the boundaries to go to infinity at fixed mean density, one has both the virial pressure relation[2]

$$\frac{\beta p}{n} = 1 + \frac{n}{6}\int e^{\beta v(\underset{\sim}{r})} g(\underset{\sim}{r})\underset{\sim}{r}\cdot\nabla e^{-\beta v(\underset{\sim}{r})} d\underset{\sim}{r} \tag{2.8a}$$

for interparticle potential $v(\underset{\sim}{r}_i-\underset{\sim}{r}_j)$, and the Ornstein-Zernike compressibility formula[3]

$$\frac{\partial n}{\partial \beta p} = 1 + n\int (g(\underset{\sim}{r}) - 1)d\underset{\sim}{r} \tag{2.8b}$$

Of course, to the same extent that the over-all distribution is given explicitly by (2.1), the s-particle distribution defined by (2.2) is also known explicitly. If

the N-body Hamiltonian splits into coordinate-independent kinetic and momentum-independent potential energies of the usual form:

$$H_N = T_N + V_N$$

where

$$T_N = \sum p_i^2/2m \tag{2.9}$$

$$V_N = \frac{1}{2} \sum_{i \neq j} v(\underset{\sim}{r}_i, \underset{\sim}{r}_j) + \sum u(\underset{\sim}{r}_i),$$

it follows at once from (2.1) and (2.2) that

$$n_s(\underset{\sim}{r}_1, \cdots, \underset{\sim}{r}_s) \frac{N!}{N-s!} \int \cdots \int e^{-\beta V_N(\underset{\sim}{r}_1, \cdots, \underset{\sim}{r}_N)} d\underset{\sim}{r}_{s+1} \cdots d\underset{\sim}{r}_N \Big/ \int \cdots \int e^{-\beta V_N(\underset{\sim}{r}_1, \cdots, \underset{\sim}{r}_N)} d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_N, \tag{2.10}$$

and our problem will be to determine n_s for s = 1 or 2 without literally performing the N - s fold integration. It will actually be more convenient for our remaining considerations to work with a grand canonical ensemble, in which each N-particle system of the ensemble is weighted by[4]

$$w(N) = (z^N/N!)Q_N, \tag{2.11}$$

z being a fugacity parameter related to the chemical potential μ (the Gibbs free energy per particle in a uniform system), by

$$\ln z = \beta\mu + \frac{3}{2} \ln(2\pi m/\beta h^2) \tag{2.12}$$

where m is the particle mass and h Planck's constant. For such an ensemble, (2.10) goes over to

$$n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) = \sum_{N=s}^{\infty} \frac{z^N}{N-s!} \int \cdots \int e^{-\beta V_N(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_N)} d\underset{\sim}{r}_{s+1}\cdots d\underset{\sim}{r}_N/\Xi \tag{2.13}$$

where

$$\Xi = \sum_{0}^{\infty} (z^N/N!) \int \cdots \int e^{-\beta V_N(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_N)} d\underset{\sim}{r}_1\cdots d\underset{\sim}{r}_N.$$

In terms of the distributions n_1 and n_2 of (2.13), the Ursell function of (2.4) now literally vanishes for a noninteracting system. Conversely, since N can fluctuate in the grand ensemble, the normalization of (2.3) is maintained only to order 1/N, e.g., it is no longer true for a finite system that $\int n_2(\underset{\sim}{r},\underset{\sim}{r}')d\underset{\sim}{r}' = \left(\int n_1(\underset{\sim}{r})d\underset{\sim}{r} - 1\right)n_1(\underset{\sim}{r}')$. Rather, since $\langle N^2\rangle \geq \langle N\rangle^2$ for $\langle N\rangle = \int n_1(\underset{\sim}{r})d\underset{\sim}{r}$, one has $\iint [n_2(\underset{\sim}{r},\underset{\sim}{r}') - n_1(\underset{\sim}{r})n_1(\underset{\sim}{r}') + n_1(\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}')]d\underset{\sim}{r}\, d\underset{\sim}{r}' \geq 0$ and the deviation from 0 is, according to (2.9), a measure of the compressibility. But it is true that at asymptotic separation, where $\underset{\sim}{r}$ and $\underset{\sim}{r}'$ refer to sensibly independent particles,[5]

$$n_2(\underset{\sim}{r},\underset{\sim}{r}') \to n_1(\underset{\sim}{r})n_1(\underset{\sim}{r}') \quad \text{as} \quad |\underset{\sim}{r}-\underset{\sim}{r}'| \to \infty. \tag{2.14}$$

The validity of (2.14) beyond some effective correlation length (except in singular states such as two-phase regions) will be of importance in the ensuing discussion.

III. KINEMATIC RESTRICTIONS

It is time to examine the one and two-body distributions in some detail. Consider first a uniform fluid, so that the relations (2.7) hold. By using periodic boundary conditions, we can imagine the system as contained in a finite box of volume V, but we will generally have in mind the limit $V \to \infty$ at constant mean density. Now aside from the conditions of positivity and normalization

$$g(\underset{\sim}{r}) \geq 0, \quad \frac{1}{V}\int g(\underset{\sim}{r})d\underset{\sim}{r} \to 1 \quad \text{as} \quad V \to \infty \tag{3.1}$$

(for a single-phase fluid), are there any restrictions on the form of $g(\underset{\sim}{r})$? Experimentally, e.g., from X-ray scattering

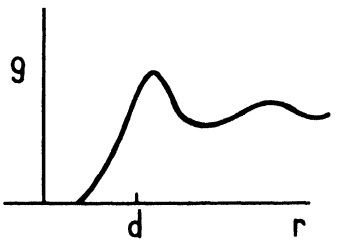


Fig. 1

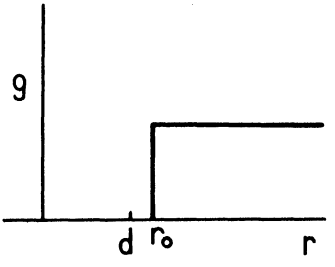


Fig. 2

g resembles Fig. 1, not dissimilar to its familiar zero density value

$$g(r) \to e^{-\beta v(\underset{\sim}{r})} \quad \text{as} \quad n \to 0 \tag{3.2}$$

for relative-coordinate two-body interactions alone: $u = 0$, $v(\underset{\sim}{r}_i, \underset{\sim}{r}_j) = v(\underset{\sim}{r}_i - \underset{\sim}{r}_j)$ in (2.9). But as an average (grand canonical) of N-2 fold integrals, of positive symmetric coordinate functions ρ_N, g is not at all arbitrary,[6] e.g., it cannot vanish identically (Fig. 2) for greater than mean interparticle spacing d, since the minimum interparticle spacing would then exceed its average.

Thus, there are kinematic restrictions on the form of g in order that it be possible to obtain g by integration of a many-body distribution, irrespective of dynamics. Two such restrictions have already appeared. First is the non-negativity of the scattering intensity (2.5) which, for a uniform system, attains the form

$$(g-1)_{\underset{\sim}{k}} \geq -\frac{1}{n}, \tag{3.3}$$

where $(g-1)_{\underset{\sim}{k}}$ denotes the Fourier coefficient $\int (g(\underset{\sim}{r}) - 1)\exp i\underset{\sim}{k}\cdot\underset{\sim}{r}\, d\underset{\sim}{r}$. For a grand ensemble, $(g-1)_{\underset{\sim}{k}}$ is continuous at $\underset{\sim}{k} = 0$, and (3.3) is a generalization of the nonnegativity of the Ornstein-Zernike compressibility (2.9); it implies qualitatively that g cannot depart too long and too far from its asymptotic value of unity (unity because distant particles do not influence each other's distribution - see (2.14) or (3.1)).

Equation (3.3) serves to eliminate a number of impossible forms for g; applied to Fig. 2, say in a one-dimensional fluid, it immediately yields the result $r_o \leq d/2$. Another obvious restriction is that the right-hand side of (2.6) be nonnegative. However, this is not independent of (3.3): defining $Z(\underset{\sim}{r})$ as the unit step function which restricts $\underset{\sim}{r}$ to volume Ω, (2.6) becomes $(n/\Omega)\iint Z(\underset{\sim}{r})(g(\underset{\sim}{r}-\underset{\sim}{r}') - 1 + \delta(\underset{\sim}{r}-\underset{\sim}{r}')/n)Z(\underset{\sim}{r}')d\underset{\sim}{r}\, d\underset{\sim}{r}'$ $= (n/\Omega)(2\pi)^{-3}\int Z_{\underset{\sim}{k}}^2[(g-1)_{\underset{\sim}{k}} + 1/n]d\underset{\sim}{k}$, which indeed is nonnegative according to (3.3).

The realizability problem for the pair distribution is in fact completely solvable, in principle.[7] Employing once more the condition that the average must exceed the minimum, it is clear that if we choose an interparticle "model potential" $\varphi(\underset{\sim}{r})$, then the total potential energy (including self-interaction for convenience) resulting from a proposed $g(\underset{\sim}{r})$ cannot be less than the pointwise minimum which the model system can attain: if

$$\Phi_\mu = \mathrm{Min}_N \mathrm{Min}_{(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_N)}\left[\frac{1}{2}\sum_{i,j}\varphi(\underset{\sim}{r}_i-\underset{\sim}{r}_j) + \mu(nV-N)\right]$$

Then

$$\frac{1}{2}\int \varphi(\underset{\sim}{r})\left(g(\underset{\sim}{r}) + \frac{1}{n}\delta(\underset{\sim}{r})\right)d\underset{\sim}{r} \geq \Phi_\mu/n^2V. \qquad (3.4)$$

Furthermore, it can be shown by a theorem on convex cones[8] that the validity of (3.4) for all model φ's is sufficient to ensure that $g(\underset{\sim}{r})$ be obtainable from some N-body distribution. Equation (3.3) is an example of (3.4) obtained from the special choice $\varphi(\underset{\sim}{r}) = \cos \underset{\sim}{k}\cdot\underset{\sim}{r}$, $\mu = 0$ and $\Phi_o = 0$ $\left(\text{since } \sum_{i,j} \cos \underset{\sim}{k}\cdot(\underset{\sim}{r}_i - \underset{\sim}{r}_j) = \left|\Sigma \exp i\underset{\sim}{k}\cdot\underset{\sim}{r}_i\right|^2\right)$ On the other hand, from $\varphi(r) = \delta(\underset{\sim}{r}-\underset{\sim}{x})$, $\Phi_o = 0$, there results the trivial but highly necessary

$$g(\underset{\sim}{x}) \geq 0. \tag{3.5}$$

The necessary and sufficient conditions (3.4) may be arrived at gradually by replacing Φ_μ by a smaller quantity (a necessary condition still resulting) and then successively raising its value. For this purpose, we rewrite

$$\Phi_\mu = \text{Min}_\rho \left[\frac{1}{2}\iint \rho(\underset{\sim}{r})\rho(\underset{\sim}{r}')\varphi(\underset{\sim}{r}-\underset{\sim}{r}')d\underset{\sim}{r}\, d\underset{\sim}{r}' + \mu\left(nV - \int \rho(\underset{\sim}{r})d\underset{\sim}{r}\right)\right] \tag{3.6}$$

over all

$$\rho(\underset{\sim}{r}) = \sum \delta(\underset{\sim}{r}-\underset{\sim}{r}_i), \tag{3.6'}$$

and then transform (3.4) into Fourier language,

$$\sum \varphi_{\underset{\sim}{k}}\left[(g-1)_{\underset{\sim}{k}} + \frac{1}{n}\right] \geq \frac{1}{n^2}\,\text{Min}_\rho\left[\frac{1}{V}\sum_{\underset{\sim}{k}\neq 0} \varphi_{\underset{\sim}{k}}\rho_{\underset{\sim}{k}}\rho_{-\underset{\sim}{k}} + 2\mu(nV-\rho_o)\right] \equiv \Phi_\mu^o \tag{3.7}$$

If ρ is varied all over all possible functions instead of the restricted class of (3.7), the minimum $\Phi_\mu^1 \leq \Phi_\mu^0$. In fact, clearly $\Phi_\mu^1 > -\infty$ only if all $\varphi_{\underset{\sim}{k}} \geq 0$, $\mu = 0$; then $\Phi_\mu^0 = 0$ and (3.7) reproduces (3.3). If Φ_μ^2 is obtained by further restricting $\rho(\underset{\sim}{r})$ to be nonnegative, $\Phi_\mu^1 \leq \Phi_\mu^2 \leq \Phi_\mu^0$; when all $\varphi_{\underset{\sim}{k}} \geq 0$, the minimizing $\rho(r) = N/V$ for Φ_μ^0 is already nonnegative and nothing is gained by the restriction. But if not all $\varphi_{\underset{\sim}{k}} \geq 0$, the nonnegativity of $\rho(\underset{\sim}{r})$ certainly provides a floor to Φ_μ^2. One consequence is (3.5) but it is simply not known whether other physically transparent conditions result. Finally, if ρ is compelled to be a true microscopic density as in (3.6), by e.g., imposing the condition

$$\int_\Omega \rho(\underset{\sim}{r}) d\underset{\sim}{r} \text{ is an integer} \tag{3.8}$$

for any volume Ω, we return to (3.4).

The realizability conditions (3.4) will be used as checks on approximate theories. They do not however significantly aid the determination of the pair distribution. The situation is otherwise in the case of the ground state energy of a quantum mechanical many-body system, for the Rayleigh-Ritz minimum principle is expressible completely in terms of the two-body reduced density matrix (when at most pair forces are present). Thus, a complete catalog of conditions[9] which permit a two-body object to be a reduced density matrix of an N-body system is in principle sufficient to allow the computation of the N-body ground state energy as effectively a two-body problem. Applications of this technique will be reported elsewhere.[8]

IV. NONUNIFORM SYSTEMS

We proceed to nonuniform systems, due to e.g., imposition of an external potential

$$U = \sum U(\underset{\sim}{r}_i) \tag{4.1}$$

upon a system with translation-invariant internal potential. Now even the computation of the one-body density $n_1(\underset{\sim}{r}) \equiv n(\underset{\sim}{r})$ is a nontrivial problem and will later provide immediate entree to the study of pair distribution functions. There are two limits in which the determination of $n(\underset{\sim}{r})$ is not difficult. If the change of external potential in a typical fluid correlation length λ (of the order of interparticle spacing) is small compared with the chemical potential: $|\lambda \nabla U| << |\mu|$; we would expect $n(\underset{\sim}{r})$ to satisfy the condition of hydrostatic equilibrium[10]

$$\nabla p_o(n(\underset{\sim}{r})) + n(\underset{\sim}{r}) \nabla U(\underset{\sim}{r}) = 0, \tag{4.2}$$

where $p_o(n)$ is the pressure for a uniform system at density n. Equally well, recalling that $dp/d\mu = n$, Eq. (4.2) can be integrated to give

$$\mu_o(n(\underset{\sim}{r})) + U(\underset{\sim}{r}) = \text{const.} \tag{4.3}$$

where $\mu_o(n)$ is the Gibbs free energy per particle in a uniform system.

The opposite limit, $|\lambda \nabla U| >> |\mu|$, is that in which the external potential changes so rapidly it may be discontinuous or even have infinite discontinuities. In the presence of a very large external force, the internal particle-particle forces are ineffective, so that one would expect $n(\underset{\sim}{r})$ to vary as the free particle Boltzmann factor in the external field, exp $- \beta U(\underset{\sim}{r})$. In other words, it is anticipated that $n(\underset{\sim}{r}) \exp \beta U(\underset{\sim}{r})$, which according to (2.10) for s = 1 and potential of (3.2), may be written as [using canonical ensemble; identical considerations apply for the grand]

$$n(\underset{\sim}{r}) e^{\beta U(\underset{\sim}{r})} = \frac{N}{Q_N} \int e^{-\beta V_{N-1}(\underset{\sim}{r}_2, \cdots, \underset{\sim}{r}_N)} e^{-\beta \sum_2^N [U(\underset{\sim}{r}_1) + v(\underset{\sim}{r} - \underset{\sim}{r}_1)]} d\underset{\sim}{r}_2 \cdots d\underset{\sim}{r}_N, \tag{4.4}$$

and its first few derivatives will be continuous even across a singularity in U. Indeed, since $\nabla_{\underset{\sim}{r}} \exp \left[- \beta \sum_2^N v(\underset{\sim}{r} - \underset{\sim}{r}_1) \right] = - \left(\sum_2^N \nabla_{\underset{\sim}{r}_1} \right) \exp \left[- \beta \sum_2^N v(\underset{\sim}{r} - \underset{\sim}{r}_1) \right]$, $ne^{\beta U}$ can have discontinuous derivatives only if the internal potential $v(\underset{\sim}{r} - \underset{\sim}{r}_1)$ has them. The discontinuities which do occur are geometrical in origin. E. g., in the prototype and worst

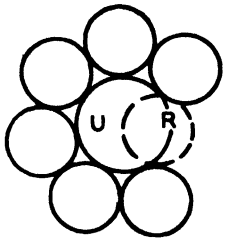


Fig. 3

possible case of hard sphere external potential, the distinguished particle R, of center $\underset{\sim}{r}$, can penetrate the potential U without causing $ne^{\beta U}$ to vanish; each time a new integer number of particles can surround the U,R pair, an analytic break will occur in the function $ne^{\beta U}$. However, it is easily verified that the discontinuity even in this case is first in the fourth derivative (or the second for the corresponding one-dimensional system).

It should be mentioned that the relations (4.2, 4.3) and the continuity of (4.4) do have a common region of validity. This is seen by first rewriting the local constancy of $ne^{\beta U}$ in the alternative forms

$$\begin{aligned} &\frac{1}{\beta} \ln n(\underset{\sim}{r}) + U(\underset{\sim}{r}) = \text{const.} \\ &\nabla \frac{1}{\beta} n(\underset{\sim}{r}) + n(\underset{\sim}{r}) \nabla U(\underset{\sim}{r}) = 0. \end{aligned} \tag{4.5}$$

Then in any region in which the density is rising from zero, we have $p \sim \frac{1}{\beta} n$, as well as (see (2.13)) $\beta\mu = \ln z + \text{const.} \sim \ln n$, so that (4.5) coincides with (4.2) and (4.3).

Ursell Distributions

In the limiting cases, we are able to write down approximations for the total effect of appending an external potential. For intermediate types of external potential, a more systematic procedure is required, one for example in which we observe at each stage the effect of a small change of potential. Let us then consider the process of starting with a uniform system and gradually turning on the external potential, or since only the more regular combination $\exp - \beta U(\underset{\sim}{r})$ occurs anyway, of turning this on instead. It is simplest to first compute the change in the grand partition function

$$\Xi[U] = \sum_0^\infty (z^N/N!) \int e^{-\beta V_N(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_N)} e^{-\beta \Sigma U(\underset{\sim}{r}_1)} d\underset{\sim}{r}_1 \ldots d\underset{\sim}{r}_N, \tag{4.6}$$

and then use this result to see how the distributions alter. When we change $e^{-\beta U(\underset{\sim}{r})}$ by the infinitesimal $\delta e^{-\beta U(\underset{\sim}{r})}$, it is clear from (2.13) that, if the applied potential U is set off by a bar in distribution functions,

$$\delta\Xi[U] = \Xi[U]\int e^{\beta U(\underset{\sim}{r}')} n_1(\underset{\sim}{r}'|U)\, e^{-\beta U(\underset{\sim}{r}')} d\underset{\sim}{r}'. \qquad (4.7)$$

In the convenient notation of functional differentiation, according to which

$$\frac{\delta F}{\delta y(\underset{\sim}{r})} = G(\underset{\sim}{r}) \qquad (4.8)$$

and

$$\delta F = \int G(\underset{\sim}{r})\delta y(\underset{\sim}{r}) d\underset{\sim}{r}$$

are equivalent statements, (4.7) now reads

$$\frac{\delta\Xi[U]}{\delta e^{-\beta U(\underset{\sim}{r})}} = \Xi[U] e^{\beta U(\underset{\sim}{r})} n(\underset{\sim}{r}|U). \qquad (4.9)$$

Further application of the operation $\delta/\delta e^{-\beta U}$ replaces successive factors $e^{-\beta U}$ in (4.6) by $\delta e^{-\beta U}$, with appropriate multiplicity, and thus (4.9) generalizes at once to[11]

$$\frac{\delta^s \Xi[U]}{\delta e^{-\beta U(\underset{\sim}{r}_1)} \cdots \delta e^{-\beta U(\underset{\sim}{r}_s)}} = \Xi[U] e^{\beta U(\underset{\sim}{r}_1)} \cdots e^{\beta U(\underset{\sim}{r}_s)} n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U), \quad (4.10)$$

a basic relation.

The explicit appearance of the grand partition function in (4.10) can be avoided if we consider, instead of the n_s, the corresponding Ursell distributions $\mathfrak{F}_s$. These are defined as[12] the portions of n_s not contained in any lower order distributions, e.g.

$$\begin{aligned}
\mathfrak{F}_1(\underset{\sim}{r}_1) &= n(\underset{\sim}{r}_1) \\
\mathfrak{F}_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) &= n_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) - n(\underset{\sim}{r}_1)n(\underset{\sim}{r}_2) \\
\mathfrak{F}_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3) &= n_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_2) - n_2(\underset{\sim}{r}_2,\underset{\sim}{r}_3)n(\underset{\sim}{r}_1) - n_2(\underset{\sim}{r}_3,\underset{\sim}{r}_1)n(\underset{\sim}{r}_2) \\
&\quad - n_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)n(\underset{\sim}{r}_3) + 2n(\underset{\sim}{r}_1)n(\underset{\sim}{r}_2)n(\underset{\sim}{r}_3)\cdots,
\end{aligned} \tag{4.11}$$

and may be obtained recursively from

$$\mathfrak{F}_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s) = n_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s) - \sum_{t=1}^{s-1}\binom{s-1}{t}\{n_t(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_t)\mathfrak{F}_{s-t}(\underset{\sim}{r}_{t+1},\cdots,\underset{\sim}{r}_s)\}_{\text{Perm.}}, \tag{4.12}$$

where $\{\}$ indicates an average over particle permutations. An alternative concise form derivable from (4.12) is that if

$$\mathfrak{F}[\varphi] \equiv \sum_1^\infty \frac{1}{s!}\int\cdots\int \mathfrak{F}_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s)\varphi(\underset{\sim}{r}_1)\cdots\varphi(\underset{\sim}{r}_s)d\underset{\sim}{r}_1\cdots d\underset{\sim}{r}_s \tag{4.12'}$$

and

$$n[\varphi] \equiv \sum_0^\infty \frac{1}{s!} \int \cdots \int n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)\varphi(\underset{\sim}{r}_1)\cdots\varphi(\underset{\sim}{r}_s)d\underset{\sim}{r}_1\cdots d\underset{\sim}{r}_s,$$

then

$$\mathcal{F}[\varphi] = \ln n[\varphi].$$

One of the major properties of $\mathcal{F}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)$, that of vanishing in a region in which a specified set of t particles $\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_t$ is independent of the remaining s-t: $n_{t+n}\left(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_t,\underset{\sim}{r}'_1,\ldots,\underset{\sim}{r}'_n\right) = n_t(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_t)n_n\left(\underset{\sim}{r}'_1,\ldots,\underset{\sim}{r}'_n\right)$ for $\underset{\sim}{r}'_1,\ldots,\underset{\sim}{r}'_n$ in the set $\underset{\sim}{r}_{t+1},\ldots,\underset{\sim}{r}_s$, can be verified from (4.12) or (4.12′). In fact, a sharper statement can be made. Suppose that $\underset{\sim}{r}_s$ is far enough removed from the set $\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}$ that it can be regarded as having no terms in common with it in the system potential energy. Then (2.14) extends, by virtue of (2.13), to $n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) = n_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1})n_1(\underset{\sim}{r}_s)$. Thus $\mathcal{F}_s$ vanishes, whenever $\underset{\sim}{r}_s$ is not near one of $\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}$, which we can indicate symbolically by $\mathcal{F}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)$ $\sim C(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1})\delta(\underset{\sim}{r}_s - \{\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}\})$; to determine the amplitude of this "smeared δ-function", we must integrate over $\underset{\sim}{r}_s$. Now a direct result of the definitions, (2.12) and (2.13) is that

$$\int n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)d\underset{\sim}{r}_s = \left(\frac{\partial}{\partial\beta\mu} + (N+1-s)\right)n_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}), \tag{4.13}$$

and one finds then from (4.12) that

$$\int \mathcal{F}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)d\underset{\sim}{r}_s = \left(\frac{\partial}{\partial\beta\mu} - (s-1)\right)\mathcal{F}_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}), \tag{4.14}$$

leading to the conclusion

$$\mathcal{F}_s(r_1,\ldots,r_s) \sim \left(\frac{\partial}{\partial\beta\mu} - (s-1)\right)\mathcal{F}_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1})\delta(\underset{\sim}{r}_s - \{\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1}\}). \tag{4.15}$$

Perhaps the most important formal consequence of (4.12) is the analog of (4.10). It can be shown inductively, and is easily obtained by direct computation for small s - that in terms of the "grand potential"

$$\frac{1}{\beta}\ln \Xi \tag{4.16}$$

(the pV product for a uniform system), (4.10) and (4.12') imply as well[13]

$$\frac{\delta^s \ln \Xi[U]}{\delta e^{-\beta U(\underset{\sim}{r}_1)}\ldots\delta e^{-\beta U(\underset{\sim}{r}_s)}} = e^{\beta U(\underset{\sim}{r}_1)}\ldots e^{\beta U(\underset{\sim}{r}_s)}\mathcal{F}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s|U). \tag{4.17}$$

As a consequence of (4.17), one also has

$$\delta^{s}\left(e^{\beta U(\underset{\sim}{r}_1')}\cdots e^{\beta U(\underset{\sim}{r}_t')}\mathcal{J}_t(\underset{\sim}{r}_1',\ldots,\underset{\sim}{r}_t')\right)/\delta e^{-\beta U(\underset{\sim}{r}_1)}\cdots e^{-\beta U(\underset{\sim}{r}_s)}$$

$$= e^{\beta U(\underset{\sim}{r}_1')}\cdots e^{\beta U(\underset{\sim}{r}_t')}e^{\beta U(\underset{\sim}{r}_1)}\cdots e^{\beta U(\underset{\sim}{r}_s)}\mathcal{J}_{s+t}(\underset{\sim}{r}_1',\ldots,\underset{\sim}{r}_t',\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) \tag{4.18}$$

Functional Expansion

By virtue of (4.18), one can now turn on an external potential in a uniform system and thereby correct the approximations of (4.4) and (4.5) which are specifically relevant to rapidly changing potentials. The point is this. Since we are applying a field which changes from 1 to $e^{-\beta U(\underset{\sim}{r})}$, any quantity F may be developed in a series expansion in

$$e^{-\beta U(\underset{\sim}{r})} - 1 \equiv f(\underset{\sim}{r}) \tag{4.19}$$

namely[14]

$$F[U] = F[0] + \int F(\underset{\sim}{r}_1)f(\underset{\sim}{r}_1)d\underset{\sim}{r}_1 + \frac{1}{2}\iint F(\underset{\sim}{r}_1,\underset{\sim}{r}_2)f(\underset{\sim}{r}_1)f(\underset{\sim}{r}_2)d\underset{\sim}{r}_1 d\underset{\sim}{r}_2$$

$$+ \cdots + \frac{1}{s!}\int\cdots\int F(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)f(\underset{\sim}{r}_1)\cdots f(\underset{\sim}{r}_s)d\underset{\sim}{r}_1\cdots d\underset{\sim}{r}_s + \cdots . \tag{4.20}$$

But on varying $e^{-\beta U}$ in (4.20) according to (4.8), we readily find

$$\frac{\delta^t F[U]}{\delta e^{-\beta U(\underset{\sim}{r}_1')} \ldots \delta e^{-\beta U(\underset{\sim}{r}_t')}} = \sum_{s=0}^{\infty} \frac{1}{s!} \int \cdots \int F\left(\underset{\sim}{r}_1', \ldots, \underset{\sim}{r}_t', \underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_s\right) f(\underset{\sim}{r}_1) \ldots f(\underset{\sim}{r}_s) d\underset{\sim}{r}_1 \ldots d\underset{\sim}{r}_s ; \tag{4.21}$$

choosing f = 0 or U = 0 hence converts the right-hand side of (4.21) to the Mac Laurin coefficients of (4.20), and we have

$$F\left(\underset{\sim}{r}_1', \ldots, \underset{\sim}{r}_t'\right) = \delta^t F[U] / \delta e^{-\beta U(\underset{\sim}{r}_1')} \ldots \delta e^{-\beta U(\underset{\sim}{r}_t')} \Bigg|_{U=0} . \tag{4.22}$$

In particular, if we select $F[U] = n(\underset{\sim}{r}|U) e^{\beta U(\underset{\sim}{r})}$ and take advantage of (4.18) with t = 1, there follows the desired expansion for the density $n(\underset{\sim}{r}|U)$ in the presence of U:

$$n(\underset{\sim}{r}|U) e^{\beta U(\underset{\sim}{r})} = n(\underset{\sim}{r}) + \int \mathcal{J}_2\left(\underset{\sim}{r}, \underset{\sim}{r}_1\right)\left(e^{-\beta U(\underset{\sim}{r}_1)} - 1\right) d\underset{\sim}{r}_1$$

$$+ \tfrac{1}{2} \iint \mathcal{J}_3\left(\underset{\sim}{r}, \underset{\sim}{r}_1, \underset{\sim}{r}_2\right)\left(e^{-\beta U(\underset{\sim}{r}_1)} - 1\right)\left(e^{-\beta U(\underset{\sim}{r}_2)} - 1\right) d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \cdots , \tag{4.23}$$

where the distributions on the right-hand side are those for the uniform system without external potential (but apply as well if the field is imposed upon an already nonuniform system.) Similarly, by choosing F appropriately, or also by directly taking logarithm and then gradient of (4.23), we find

$$\ln[n(\underset{\sim}{r}|U)/n(\underset{\sim}{r})] + \beta U(\underset{\sim}{r}) = \int \left(e^{-\beta U(\underset{\sim}{r}_1)} - 1\right)\mathfrak{F}_2(\underset{\sim}{r},\underset{\sim}{r}_1)/n(\underset{\sim}{r})d\underset{\sim}{r}_1 + \cdots \tag{4.24}$$

$$\nabla n(\underset{\sim}{r}|U) + n(\underset{\sim}{r}|U)\nabla\beta U(\underset{\sim}{r}) = e^{-\beta U(\underset{\sim}{r})}\nabla\left[n(\underset{\sim}{r}) + \int \left(e^{-\beta U(\underset{\sim}{r}_1)} - 1\right)\mathfrak{F}_2(\underset{\sim}{r},\underset{\sim}{r}_1)d\underset{\sim}{r}_1 + \cdots\right]. \tag{4.25}$$

By the method of construction, $n(\underset{\sim}{r}|U)e^{\beta U(\underset{\sim}{r})}$ of (4.23-25) is guaranteed to vary slowly at a singularity of U, i.e., for a variation within a region small compared with a typical scale length of $\mathfrak{F}_2$ - which is qualitatively to be identified with λ of the discussion following (4.1). But unless βU is itself small, the actual value of $ne^{\beta U}$ will not be obtained correctly from just the beginning of the series. On the other hand, if βU applied to a uniform system both changes slowly within the range of nonvanishing of $\mathfrak{F}_2$,

and is small as well, one can show that (4.2) and (4.3) are obtained from the first order terms of (4.24,4.25). Now (4.2) and (4.3) are not very difficult to obtain from our "turning-on" process, no matter how large βU becomes. Indeed, if we consider the basic variational expression

$$\delta\left(n(\underset{\sim}{r}|U)e^{\beta U(\underset{\sim}{r})}\right) = \int \mathfrak{F}_2(\underset{\sim}{r},\underset{\sim}{r}_1|U)e^{\beta U(\underset{\sim}{r})}e^{\beta U(\underset{\sim}{r}_1)}\delta e^{-\beta U(\underset{\sim}{r}_1)}d\underset{\sim}{r}_1 \tag{4.26}$$

which leads to (4.23) and suppose that $e^{\beta U}$ varies slowly in the range of nonvanishing of $\mathfrak{F}_2(\underset{\sim}{r},\underset{\sim}{r}_1|U)$, then the latter simply acts as a δ-function whose integral is slowly varying in space and may presumably be given its uniform system value of (2.9), $n\left(\frac{\partial n}{\partial \beta p} - 1\right)$. In other words, (4.26) reduces to

$$\delta(ne^{\beta U}) = -n\left(\frac{\partial n}{\partial \beta p} - 1\right)\delta e^{\beta U}, \tag{4.27}$$

and this, by means of the relation $n\frac{\partial n}{\partial \beta p} = \frac{\partial n}{\partial \mu}$, integrates to (4.3). Thus the relation between the two limits appropriate to (4.23) and (4.3) is a real one, but our expansion parameter $e^{-\beta U}$ is simply not one which leads to a series which converges rapidly enough to be useful for a large class of inhomogeneous systems. We shall now see how this situation may be improved.

V. THE DIRECT CORRELATION FUNCTION

Although inhomogeneity arises from an external potential, the strength of the potential does not necessarily give a simple indication of the degree of inhomogeneity, but by definition the density itself certainly does. This suggests that in some fashion changes in thermodynamic and other parameters of the system be assessed directly in terms of the change or nonuniformity of the density. For example, for weak nonuniformity, one would expect (and we have made preliminary use of this in (4.26,27)) $n_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)$ to be given by its uniform system value at some mean density, e.g., that at the average of $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_2$, irrespective of the strength of the potential which produced it. Since we already know how the system parameters vary with external potential, there remains the determination of the change in potential required by a given density alteration, i.e., we would like to evaluate $\delta\beta U(\underset{\sim}{r})/\delta n(\underset{\sim}{r}')$. Of course, we cannot a priori find this quantity; we can only express it in terms of others.

According to (4.15), we have $\delta n(\underset{\sim}{r}')/\delta(-\beta U(\underset{\sim}{r}))$

$$= \delta\left[\left(n(\underset{\sim}{r}')e^{\beta U(\underset{\sim}{r}')}\right)\left(e^{-\beta U(\underset{\sim}{r}')}\right)\right]/\delta(-\beta U(\underset{\sim}{r}))$$

$$= \left(n(\underset{\sim}{r}')e^{\beta U(\underset{\sim}{r}')}\right)\delta(\underset{\sim}{r}-\underset{\sim}{r}')e^{-\beta U(\underset{\sim}{r}')} + \delta\left(n(\underset{\sim}{r}')e^{\beta U(\underset{\sim}{r}')}\right)/\delta(-\beta U(\underset{\sim}{r}))e^{-\beta U(\underset{\sim}{r}')}$$

The explicit specification $|U)$ is, and will be, omitted when the meaning is clear without it. Thus

$$\frac{\delta n(\underset{\sim}{r}')}{\delta(-\beta U(\underset{\sim}{r}))} = n(\underset{\sim}{r}')\delta(\underset{\sim}{r}-\underset{\sim}{r}') + \mathfrak{F}_2(\underset{\sim}{r}',\underset{\sim}{r}), \tag{5.1}$$

a combination of a direct effect at the point $\underset{\sim}{r}$ due to a potential at $\underset{\sim}{r}$ plus a residual or longer (but still short, macroscopically) range reaction. Consequently, it is reasonable to separate $\delta(-\beta U(\underset{\sim}{r}))/\delta n(\underset{\sim}{r})$ as well into its singular and nonlocal parts:

$$\frac{\delta(-\beta U(\underset{\sim}{r}))}{\delta n(\underset{\sim}{r}')} \equiv \frac{1}{n(\underset{\sim}{r}')}\,\delta(\underset{\sim}{r}-\underset{\sim}{r}') - c(\underset{\sim}{r},\underset{\sim}{r}'). \tag{5.2}$$

We are now definitely operating in a grand ensemble, for only then can an arbitrary density change (not necessarily conserving particle number) be envisaged. The function $c(\underset{\sim}{r},\underset{\sim}{r}')$ thus introduced, known as the direct correlation function, solves in principle the question of the effect of a density change - due to a change in external potential, - on any quantity in the system, for by means of the chain rule for some intermediate function $y(\underset{\sim}{r})$,

$$\frac{\delta F}{\delta n(\underset{\sim}{r})} = \int \frac{\delta F}{\delta y(\underset{\sim}{r}')}\,\frac{\delta y(\underset{\sim}{r}')}{\delta n(\underset{\sim}{r})}\,d\underset{\sim}{r}', \tag{5.3}$$

we have from (5.2)

$$\frac{\delta F}{\delta n(\underset{\sim}{r})} = \frac{1}{n(\underset{\sim}{r})} \frac{\delta F}{\delta(-\beta U(\underset{\sim}{r}))} - \int \frac{\delta F}{\delta(-\beta U(\underset{\sim}{r}'))} c(\underset{\sim}{r}', \underset{\sim}{r}) d\underset{\sim}{r}', \tag{5.4}$$

and $\delta F/\delta(-\beta U(\underset{\sim}{r}'))$ is known e.g., from (4.10).

The direct correlation function first occurred in the work of Ornstein and Zernike.[3] Since $\delta n(\underset{\sim}{r}')/\delta(-\beta U(\underset{\sim}{r}))$ and $\delta(-\beta U(\underset{\sim}{r}))/\delta n(\underset{\sim}{r}))$ are matrix inverses of each other, by which we mean - from (5.3) - that

$$\int \frac{\delta(-\beta U(\underset{\sim}{r}))}{\delta n(\underset{\sim}{r}')} \frac{\delta n(\underset{\sim}{r}')}{\delta(-\beta U(\underset{\sim}{r}''))} d\underset{\sim}{r}' = \delta(\underset{\sim}{r}, \underset{\sim}{r}''), \tag{5.5}$$

direct substitution of (5.1) and (5.2) into (5.5) yields

$$g(\underset{\sim}{r}, \underset{\sim}{r}'') - 1 = c(\underset{\sim}{r}, \underset{\sim}{r}'') + \int c(\underset{\sim}{r}, \underset{\sim}{r}') n(\underset{\sim}{r}') \left(g(\underset{\sim}{r}', \underset{\sim}{r}'') - 1 \right) d\underset{\sim}{r}'$$

where

$$g(\underset{\sim}{r}, \underset{\sim}{r}'') \equiv n_2(\underset{\sim}{r}, \underset{\sim}{r}'')/n(\underset{\sim}{r}) n(\underset{\sim}{r}''). \tag{5.6}$$

In other words, in the form $g(\underset{\sim}{r},\underset{\sim}{r}') - 1 = c(\underset{\sim}{r},\underset{\sim}{r}')$ $+ \int c(\underset{\sim}{r},\underset{\sim}{r}'')n(\underset{\sim}{r}'')c(\underset{\sim}{r}'',\underset{\sim}{r}')d\underset{\sim}{r}'' + \iint c(\underset{\sim}{r},\underset{\sim}{r}'')n(\underset{\sim}{r}'')c(\underset{\sim}{r}'',\underset{\sim}{r}''')n(\underset{\sim}{r}''')$ $\times c(\underset{\sim}{r}''',\underset{\sim}{r}')d\underset{\sim}{r}''\, d\underset{\sim}{r}''' + \ldots$, the deviation of g from its non-interaction value of unity - the correlation function g-1 - for a particle at $\underset{\sim}{r}'$ due to one at $\underset{\sim}{r}$ is the sum of all chains of direct correlations which can occur between $\underset{\sim}{r}$ and $\underset{\sim}{r}'$. Ornstein and Zernike reasoned that c should be shorter-range than g-1, more like the interaction potential itself, and from this fact together with the normalization given by (2.9) [now equivalent to

$$\frac{\partial\beta p}{\partial n} = 1 - n\int c(\underset{\sim}{r})d\underset{\sim}{r} \tag{5.7}$$

for uniform systems] were able to estimate the asymptotic form of g-1.

A somewhat sharper specification of the direct correlation function may be obtained as follows. When (5.1) is written in the form

$$\frac{\delta^2 \ln \Xi}{\delta\beta U(\underset{\sim}{r})\delta\beta U(\underset{\sim}{r}')} = n(\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}') + \mathfrak{F}_2(\underset{\sim}{r},\underset{\sim}{r}'), \tag{5.8}$$

it states that, except for the singularity at $\underset{\sim}{r} = \underset{\sim}{r}'$, the change in grand potential due to changes in external potential at $\underset{\sim}{r}$ and $\underset{\sim}{r}'$ simultaneously exceeds that due to separate changes at $\underset{\sim}{r}$ and $\underset{\sim}{r}'$ by the Ursell distribution $\mathcal{U}_2$, which then clearly has the aspect of a correlation of fluctuations. On the other hand, consider the Legendre transformation

$$\ln \Xi - \int (-\beta U(\underset{\sim}{r}))\delta \ln \Xi/\delta(-\beta U(\underset{\sim}{r}))d\underset{\sim}{r} = \ln \Xi + \int \beta U(\underset{\sim}{r})n(\underset{\sim}{r})d\underset{\sim}{r} \equiv \beta W \tag{5.9}$$

where, since $\delta W = \int U(\underset{\sim}{r})\delta n(\underset{\sim}{r})d\underset{\sim}{r}$, W denotes the work done on the system during an isothermal change of external potential. But then directly from the relation $\delta\beta W/\delta n(\underset{\sim}{r}) = \beta U(\underset{\sim}{r})$, it follows that

$$\frac{\delta^2 \beta W}{\delta n(\underset{\sim}{r})\delta n(\underset{\sim}{r}')} = c(\underset{\sim}{r},\underset{\sim}{r}') - \frac{1}{n(\underset{\sim}{r})}\,\delta(\underset{\sim}{r}-\underset{\sim}{r}'), \tag{5.10}$$

so that $c(\underset{\sim}{r},\underset{\sim}{r}')$ is precisely the correlation contribution to the work of creating simultaneous density variations at $\underset{\sim}{r}$ and $\underset{\sim}{r}'$. This of course is also the change in Helmholtz free energy, arising formally from replacing $-\beta U$ by $\ln z - \beta U$ in (5.9).

Expansion in Density

We are now prepared to find the effect of "turning on" a density inhomogeneity (via an external potential which does not however have to be made explicit) for parameters and distributions of the system. Just as an expansion in $f = e^{-\beta U} - 1$ required successive functional derivatives with respect to $e^{-\beta U}$, giving rise to the $\mathcal{J}_s$ series (4.18), so we shall now need the sequence of higher derivatives with respect to n. From (5.2) we have

$$c(\underset{\sim}{r},\underset{\sim}{r}') = \frac{\delta(\beta U(\underset{\sim}{r}) + \ln n(\underset{\sim}{r}))}{\delta n(\underset{\sim}{r}')} \tag{5.11}$$

which therefore extends to the (symmetric) higher direct correlations

$$c_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s) \equiv \frac{\delta^{s-1}(\beta U(\underset{\sim}{r}_1) + \ln n(\underset{\sim}{r}_1))}{\delta n(\underset{\sim}{r}_2)\cdots\delta n(\underset{\sim}{r}_s)}$$

$$\equiv \delta^s\Big(\beta W + \int (n(\underset{\sim}{r})\ln n(\underset{\sim}{r}) - n(\underset{\sim}{r}))d\underset{\sim}{r}\Big)/\delta n(\underset{\sim}{r}_1)\cdots\delta n(\underset{\sim}{r}_s). \tag{5.12}$$

The relation of the c_s to the ordinary distributions follows from differentiation of the equation (5.6). From a comparison of (5.1) and (5.2), we see that the range of $c(\underset{\sim}{r},\underset{\sim}{r}')$ is expected to be of the order of the range of $\mathcal{F}_2(\underset{\sim}{r},\underset{\sim}{r}')$. Since the vanishing of $c(\underset{\sim}{r},\underset{\sim}{r}')$ at sufficient separation will not be affected by successive differentiation with respect to density, and since c_s is symmetric, we conclude that c_s approximately vanishes unless all arguments are within some correlation length of each other. As in (4.15), the magnitude of the resulting smeared δ-function is determined by integrating over the arguments. Clearly

$$\int c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)d\underset{\sim}{r}_s = \int \delta c_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1})/\delta n(\underset{\sim}{r}_s)d\underset{\sim}{r}_s, \text{ or}$$

$$\int c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)d\underset{\sim}{r}_s = \partial c_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s-1})/\partial n. \qquad (5.13)$$

For a uniform system, observing that (5.7) may be written as $\int c(\underset{\sim}{r},\underset{\sim}{r}')dr' = (1-\partial\beta p/\partial n)/n = (\partial/\partial n)(\ln n - \beta\mu)$, we thus have

$$c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) \sim \frac{\partial^{s-1}(\ln n - \beta\mu)}{\partial n^{s-1}}\,\delta(\underset{\sim}{r}_2-\underset{\sim}{r}_1)\ldots\delta(\underset{\sim}{r}_s-\underset{\sim}{r}_1). \qquad (5.14)$$

Suppose then that an initial external potential $U_0(\underset{\sim}{r})$ (not necessarily zero), with resulting density $n_0(\underset{\sim}{r})$, is altered to $U(\underset{\sim}{r})$, producing thereby a density $n(\underset{\sim}{r})$. Any quantity can be expanded, as in (4.20,4.22), in the difference $n(\underset{\sim}{r}) - n_0(\underset{\sim}{r})$. In particular, for $\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r})$ of (4.5), we have immediately from (5.12)[11]

$$\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r}) = \ln n_0(\underset{\sim}{r}) + \beta U_0(\underset{\sim}{r}) + \int c_2(\underset{\sim}{r},\underset{\sim}{r}_1)\Big|_0 (n(\underset{\sim}{r}_1) - n_0(\underset{\sim}{r}_1))d\underset{\sim}{r}_1$$

$$+ \frac{1}{2}\iint c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2)\Big|_0 (n(\underset{\sim}{r}_1) - n_0(\underset{\sim}{r}_1))(n(\underset{\sim}{r}_2) - n_0(\underset{\sim}{r}_2))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \cdots \tag{5.15}$$

Similarly, to expand $\nabla n(\underset{\sim}{r}) + n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r}) = n(\underset{\sim}{r})\nabla(\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r}))$, of (4.5), we obtain from (5.12)

$$\frac{\delta^s n(\underset{\sim}{r})\nabla(\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r}))}{\delta n(\underset{\sim}{r}_1)\cdots\delta n(\underset{\sim}{r}_s)} = n(\underset{\sim}{r})\nabla_{\underset{\sim}{r}} c_{s+1}(\underset{\sim}{r},\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s)$$

$$+ \sum \delta(\underset{\sim}{r}-\underset{\sim}{r}_i)\nabla_{\underset{\sim}{r}} c_s(\underset{\sim}{r},\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_{i-1},\underset{\sim}{r}_{i+1},\cdots,\underset{\sim}{r}_s), \tag{5.16}$$

which leads directly to

$$\nabla n(\underset{\sim}{r}) + n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r}) = \nabla n_o(\underset{\sim}{r}) + n_o(\underset{\sim}{r})\nabla\beta U_o(\underset{\sim}{r})$$

$$+\left[n_o(\underset{\sim}{r})\int \nabla_r c_2(\underset{\sim}{r},\underset{\sim}{r}_1)\Big|_0 (n(\underset{\sim}{r}_1) - n_o(\underset{\sim}{r}_1))d\underset{\sim}{r}_1 \right.$$

$$\left. + \nabla_{\underset{\sim}{r}} \ln n_o(\underset{\sim}{r})(n(\underset{\sim}{r}) - n_o(\underset{\sim}{r})) \right]$$

$$+\left[\tfrac{1}{2} n_o(\underset{\sim}{r})\iint \nabla_{\underset{\sim}{r}} c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2)\Big|_0 (n(\underset{\sim}{r}_1) - n_o(\underset{\sim}{r}_1)) \right.$$

$$\times(n(\underset{\sim}{r}_2) - n_o(\underset{\sim}{r}_2))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \int \nabla_r c_2(\underset{\sim}{r},\underset{\sim}{r}_1)\Big|_0$$

$$\left. \times(n(\underset{\sim}{r}_1) - n_o(\underset{\sim}{r}_1))d\underset{\sim}{r}_1(n(\underset{\sim}{r}) - n_o(\underset{\sim}{r})) \right] + \cdots$$

(5.17)

We can compute $n(\underset{\sim}{r})e^{\beta U(\underset{\sim}{r})}$ in the same fashion or simply exponentiate (5.15):

$$n(\underset{\sim}{r})e^{\beta U(\underset{\sim}{r})} = n_0(\underset{\sim}{r})e^{\beta U_0(\underset{\sim}{r})}$$

$$\times\left[1 + \int c_2(\underset{\sim}{r},\underset{\sim}{r}_1)\Big|_0 (n(\underset{\sim}{r}_1) - n_0(\underset{\sim}{r}_1))d\underset{\sim}{r}_1\right.$$

$$+ \frac{1}{2}\iint [c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2) - c_2(\underset{\sim}{r},\underset{\sim}{r}_1)c_2(\underset{\sim}{r},\underset{\sim}{r}_2)]_0(n(\underset{\sim}{r}_1) - n_0(\underset{\sim}{r}_1))$$

$$\left.\times(n(\underset{\sim}{r}_2) - n_0(\underset{\sim}{r}_2))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \cdots\right] \tag{5.18}$$

Finally, the higher distributions may be generated by differentiating (5.18) directly, or by merely expanding them in precisely the manner we have been employing, using (5.4) to accomplish the n-differentiation. Thus, we find

$$n_2(\underset{\sim}{r},\underset{\sim}{r}') = n_2(\underset{\sim}{r},\underset{\sim}{r}')\Big|_0 + \int \left(\frac{1}{n_o(\underset{\sim}{r}_1)}\,\delta(\underset{\sim}{r}_1-\underset{\sim}{r}_2) - c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)\right)$$

$$\times\left(n_3(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}_2) + n_2(\underset{\sim}{r},\underset{\sim}{r}')[\delta(\underset{\sim}{r}_2-\underset{\sim}{r}) + \delta(\underset{\sim}{r}_2-\underset{\sim}{r}') - n(\underset{\sim}{r}_2)]\right)\Big|_0$$

$$\times(n(\underset{\sim}{r}_1) - n_o(\underset{\sim}{r}_1))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \cdots . \qquad (5.19)$$

The expressions (5.15-5.19) may be used in two essentially different ways to analyze inhomogeneous systems. In each case $n_o(\underset{\sim}{r})$ is a constant. In the first interpretation, if one is looking at the point $\underset{\sim}{r}$, the constant n_o is taken as the value of the density at that point: $n_o(\underset{\sim}{r}') = n(\underset{\sim}{r})$, thereby in a sense minimizing the local effect of the potential. In this case, the external potential must start not at zero, but rather at that uniform value which changes the uniform system density to $n(\underset{\sim}{r})$. According to (2.12) and (2.13), this means that

$$\begin{aligned} U_o(\underset{\sim}{r}') &= \mu - \mu(n(\underset{\sim}{r})) \\ n_o(\underset{\sim}{r}') &= n(\underset{\sim}{r}). \end{aligned} \qquad (5.20)$$

For example, (5.17) may then be rewritten, after brief manipulation, as

$$\nabla n(\underset{\sim}{r}) + n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r}) = n(\underset{\sim}{r}) \int \nabla n(\underset{\sim}{r}_1) c_2(\underset{\sim}{r}-\underset{\sim}{r}_1;\ n(\underset{\sim}{r}))d\underset{\sim}{r}_1$$

$$+ n(\underset{\sim}{r}) \iint \nabla n(\underset{\sim}{r}_1)(n(\underset{\sim}{r}_2) - n(\underset{\sim}{r}))c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2;\ n(\underset{\sim}{r}))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 + \cdots, \tag{5.17'}$$

which has the desirable aspect of reducing properly in limiting cases. Thus, for a slowly varying $n(\underset{\sim}{r})$ in the domain of c, (5.17') reduces through terms in $(\nabla n)^2$, by virtue of (5.14), to

$$\nabla n(\underset{\sim}{r}) + n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r}) = \nabla n(\underset{\sim}{r})(1 - \partial\beta p(n(\underset{\sim}{r}))/\partial n(\underset{\sim}{r}))$$

$$+ n(\underset{\sim}{r})\nabla n(\underset{\sim}{r})\underset{\sim}{\lambda}(\underset{\sim}{r})\cdot\nabla n(\underset{\sim}{r})\partial^2[\ln n(\underset{\sim}{r}) - \beta\mu(n(\underset{\sim}{r}))]/\partial n(\underset{\sim}{r})^2 + \cdots \tag{5.21}$$

where

$$\underset{\sim}{\lambda}(\underset{\sim}{r}) = \iint (\underset{\sim}{r}_2-\underset{\sim}{r})c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2;\ n(\underset{\sim}{r}))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 / \iint c_3(\underset{\sim}{r},\underset{\sim}{r}_1,\underset{\sim}{r}_2;\ n(\underset{\sim}{r}))d\underset{\sim}{r}_1 d\underset{\sim}{r}_2,$$

or precisely (4.2) if $\underset{\sim}{\lambda}\cdot\nabla n \ll [\partial(\ln n - \beta\mu)/\partial n] \times[\partial^2(\ln n - \beta\mu)/\partial n^2]^{-1}$. In the same way, (5.15) and (5.18) reduce to their hydrostatic approximants, (4.3) and its exponential, the first correction term in these cases already being higher order. For (5.19), there are two distinguished points $\underset{\sim}{r}$ and $\underset{\sim}{r}'$; taking $n_o(\underset{\sim}{r}'') = n(\frac{1}{2}(\underset{\sim}{r}+\underset{\sim}{r}'))$, one can then show that the first correction again disappears for slow variations, so that the inhomogeneous system pair distribution is that for a homogeneous system at the midpoint density to terms of order $(\underset{\sim}{\lambda}\cdot\nabla n)^2$.

For the opposite regime, that of a density discontinuity due to a potential singularity, each of the three representations indeed consists of an expansion all of whose terms are continuous, but at least a priori, the actual value would appear to require the full series. Thus little is to be gained by the involved self-consistency of (5.17′) and its analogs if the problem at hand involves primarily a rapidly changing potential. Instead, the second interpretation of (5.15-5.19) is relevant: one will do just as well by choosing $n_o(\underset{\sim}{r})$ as precisely the uniform density with the external field turned off, $U_o = 0$ and $n_o = n$. This is the state of affairs when the external potential has the range of interparticle spacing, precisely the case which will next be discussed.

VI. INTEGRAL EQUATIONS

Analysis of inhomogeneous systems is not only of intrinsic interest, but also offers a systematic procedure for detailed determination of the higher distribution functions. This is related to the fact that in classical thermal equilibrium, momentum and coordinate portions of the N-body distribution or Boltzmann factor are independent. Thus one can obtain the s+t particle distribution by fixing s particles at the desired positions and then asking for the t particle distribution, i.e. recoil effects are not required. In fact, in a grand ensemble, we shall show in a moment that

$$n_{s+t}(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s,\underset{\sim}{x}_1,\cdots,\underset{\sim}{x}_t)/n_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s) = n_t(\underset{\sim}{x}_1,\cdots,\underset{\sim}{x}_t|U), \qquad (6.1)$$

where the internal and imposed external potentials are given by

$$V = \frac{1}{2}\sum_{i\neq j} v(\underset{\sim}{r}_i,\underset{\sim}{r}_j)$$
$$U(\underset{\sim}{x}) = v(\underset{\sim}{x},\underset{\sim}{r}_1) + \cdots + v(\underset{\sim}{x},\underset{\sim}{r}_s). \qquad (6.2)$$

Here U is precisely the potential seen by the other particles when the specified particles $\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s$ are fixed. Now to prove (5.1), we have first from (2.13), using a potential $U(\underset{\sim}{x}) + u(\underset{\sim}{x})$,

$$z^{-s} n_s(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_s | u) \left(\exp \beta \sum_{i<j\leq s} v(\underset{\sim}{r}_i, \underset{\sim}{r}_j) \right) e^{\beta \Sigma u(\underset{\sim}{r}_i)} \Xi[u] = \Xi[U+u]. \tag{6.3}$$

Differentiating t times with respect to $e^{-\beta U(\underset{\sim}{x})}$ and making use of (3.10), (6.3) generalizes immediately to

$$\left(z^{-s} \exp \beta \sum_{i<j\leq s} v(\underset{\sim}{r}_i, \underset{\sim}{r}_j) \right) \left(n_{s+t}(\underset{\sim}{r}_1, \ldots \underset{\sim}{r}_s, \underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_t | u) e^{\beta(\Sigma u(\underset{\sim}{r}_i) + \Sigma u(\underset{\sim}{x}_j)} \Xi[u] \right)$$

$$= n_t(\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_t | U+u) e^{\beta \Sigma u(\underset{\sim}{x}_j)} \Xi[U+u], \tag{6.3'}$$

and (6.1) is obtained on setting u = 0 in (6.3') divided by (6.3).

The relations (6.1) and (6.2) offer, in principle, a "bootstrap" operation. For example, we can determine the pair distribution in a uniform system by observing the one-body density on imposition of an external potential which mimics the internal potential:[13]

$$\begin{aligned} &\text{As} \quad U(\underset{\sim}{r}) \text{ changes from 0 to } v(\underset{\sim}{r}-\underset{\sim}{x}), \\ &\quad n_1(\underset{\sim}{r}) \text{ changes from n to } n_2(\underset{\sim}{r}-\underset{\sim}{x})/n. \end{aligned} \tag{6.4}$$

This then gives a self-consistent relation for the pair distribution if the turning on of $n_1(\underset{\sim}{r})$ can be deduced from knowledge of the uniform pair distribution alone. Although this is not truly the case, it does obtain for first order truncation of any of the expansions (e.g. 5.15, 5.17, or 5.18) we have already obtained, each one of which will then give a first order approximate theory of the pair distribution. For a second order theory in the same sense, one may use the same technique to get an additional relation for the three-body distribution, and then carry each relation through three-body terms. Thus, a hierarchy is produced by each choice of dependent and independent function in the expansion (not uniquely, although one is generally most natural) - whose truncation gives rise to successive approximations.

Potential Expansions

As prototype, albeit naive, we may literally expand n as U is turned on. Then according to (4.18, 4.20, 4.22), with the identification (6.4)

$$n_2(\underset{\sim}{r}-\underset{\sim}{x})/n = n + \int \delta n(\underset{\sim}{r})/\delta\beta U(\underset{\sim}{y}) \Big|_0 \beta v(\underset{\sim}{y}-\underset{\sim}{x})d\underset{\sim}{y} + \ldots$$

$$= n - \int [n_2(\underset{\sim}{r}-\underset{\sim}{y}) - n^2 + n\delta(\underset{\sim}{r}-\underset{\sim}{y})]\beta v(\underset{\sim}{y}-\underset{\sim}{x})d\underset{\sim}{y} + \ldots, \tag{6.5}$$

or on taking Fourier transforms to convert the convolution in (6.5) to a product,

$$(g-1)_{\underset{\sim}{k}} = -\beta v_{\underset{\sim}{k}}/(1+n\beta v_{\underset{\sim}{k}}) \qquad (6.6)$$

to first order. Equation (6.6) is a primitive collective[16] or Debye-Huckel[17] result, yielding e.g., $g(r) = 1 + (\beta e^2/r)\exp[-(4\pi\beta n e^2)^{\frac{1}{2}}r]$, for a Coulomb $v = e^2/r$ with neutralizing background. The result (6.6) can legitimately be regarded as a first approximation for sufficiently weak $v(\underset{\sim}{r})$, and then differs essentially from a straight virial (density) expansion only if v is sufficiently long range that $v_{\underset{\sim}{k}} \to \infty$ as $k \to 0$, e.g. if $v \sim r^{-(3-\varepsilon)}$, $\varepsilon > 0$.

The result (6.6) suffers from the well-known drawback that the resulting $g(\underset{\sim}{r})$ may become negative, even negative infinite. Amelioration is clear: one instead expands ln n as U is turned on, so that (6.5) is replaced by

$$\ln(n_2(\underset{\sim}{r}-\underset{\sim}{x})/n) = \ln n - \int [n_2(\underset{\sim}{r}-\underset{\sim}{y})/n - n + \delta(\underset{\sim}{r}-\underset{\sim}{y})]\beta v(\underset{\sim}{y}-\underset{\sim}{x})d\underset{\sim}{y} \qquad (6.7)$$

or choosing $\underset{\sim}{x} = 0$

$$\ln g(\underset{\sim}{r}) = -\beta v(\underset{\sim}{r}) - n\int (g-1)(\underset{\sim}{r}-\underset{\sim}{y})\beta v(\underset{\sim}{y})d\underset{\sim}{y}, \qquad (6.8)$$

the nonlinear Debye-Huckel equation[18] (more recognizable in the Coulomb case if one applies ∇^2 to both sides).

For singular forces, the obvious modification of (6.5) is to mirror (4.4) and hence determine the functional dependence of $ne^{\beta U}$ as $e^{-\beta U}$ is turned on to its final "internal" value. Thus, disregarding the fact that the external potential is hardly slowly varying in this case, we have at once from (4.23)

$$n_2(\underset{\sim}{r}-\underset{\sim}{x})e^{\beta v(\underset{\sim}{r}-\underset{\sim}{x})}/n = n + \int \left(n_2(\underset{\sim}{r}-\underset{\sim}{y}) - n^2\right)\left(e^{-\beta v(\underset{\sim}{y}-\underset{\sim}{x})} - 1\right)d\underset{\sim}{y} + \ldots \qquad (.9)$$

so that to first order

$$g(\underset{\sim}{r})e^{\beta v(\underset{\sim}{r})} - 1 = n\int \left(g(\underset{\sim}{r}-\underset{\sim}{r}') - 1\right)\left(e^{-\beta v(\underset{\sim}{r}')} - 1\right)d\underset{\sim}{r}'. \qquad (6.10)$$

This is the beginning of a Kirkwood-Salsburg[19] type of hierarchy in which, however, the fugacity never enters. It is, like (6.8), a true integral equation for $g(\underset{\sim}{r})$ and cannot

be solved by a mere Fourier transform, but unlike (6.8) it is still a linear integral equation and hence yields (for suitable v) to standard techniques.

Density Expansions

We anticipate more rapid convergence if an expansion with respect to the response - e.g. the density $n(\underset{\sim}{r})$ - rather than the stimulus $U(\underset{\sim}{r})$ is carried out. In this fashion expansion of $ne^{\beta U}$, $\ln n + \beta U$, and $n\nabla\beta U$ give rise in first order, via (6.4), to three well-known approximate integral equations of considerable accuracy. Consider first the quantity $ne^{\beta U}$; we have seen that it remains very smooth across potential discontinuities. But for a finite range potential, $ne^{\beta U}$ reduces to n outside the range of force. Thus if $ne^{\beta U}$ were not only smooth but also analytic, it would be determined by its derivatives outside the range of force, and would thereby be a linear functional of n. In other words, the expansion (5.18) with the identification (6.4) could then literally be cut off at first order:

$$ng(\underset{\sim}{r})e^{\beta v(\underset{\sim}{r})} = n + n^2\int c(\underset{\sim}{r}-\underset{\sim}{r}_1)(g(\underset{\sim}{r}_1) - 1)d\underset{\sim}{r}_1, \quad (6.11)$$

where the origin has been chosen as $\underset{\sim}{x}$. Making use of (5.6), we find at once

$$c(\underset{\sim}{r}) = \left(1-e^{\beta v(\underset{\sim}{r})}\right)g(\underset{\sim}{r}), \quad (6.12)$$

the basic expression of the Percus-Yevick (P.Y.) equation.[20] Eliminating $c(\underset{\sim}{r})$, (5.6) and (6.12) combine to the single integral equation

$$e^{\beta v(\underset{\sim}{r})} g(\underset{\sim}{r}) = 1 + n \int \left(g(\underset{\sim}{r}-\underset{\sim}{r}') - 1 \right) \left(1 - e^{\beta v(\underset{\sim}{r}')} \right) g(\underset{\sim}{r}') d\underset{\sim}{r}', \tag{6.13}$$

which is derived more conventionally in Appendix A.

For a long range potential, the peculiar aptness of (6.11) is not evident, and one may instead concentrate on the positivity of $g(\underset{\sim}{r})$. This suggests expanding $\ln(ne^{\beta U})$ in terms of n. According to (5.15) in conjunction with (6.3), again choosing $\underset{\sim}{x} = 0$,

$$\ln n g(\underset{\sim}{r}) + \beta v(\underset{\sim}{r}) = \ln n + \int c(\underset{\sim}{r}-\underset{\sim}{r}')(ng(\underset{\sim}{r}') - n) d\underset{\sim}{r}' + \ldots \tag{6.14}$$

Thus, truncating at first order and reducing the integral through (5.6),

$$c(\underset{\sim}{r}) = g(\underset{\sim}{r}) - 1 - \ln g(\underset{\sim}{r}) - \beta v(\underset{\sim}{r}), \tag{6.15}$$

which may be recognized as the fundamental equation of the convolution-hypernetted chain (CHNC) approximation.[21] Combining (5.6) and (6.15), there results

$$\ln g(\underset{\sim}{r}) + \beta v(\underset{\sim}{r}) = n\int (g(\underset{\sim}{r}-\underset{\sim}{r}') - 1)(g(\underset{\sim}{r}') - 1 - \ln g(\underset{\sim}{r}') - \beta v(\underset{\sim}{r}'))d\underset{\sim}{r}'. \tag{6.16}$$

In Appendix B, the CHNC equation (and the P.Y. as well) is obtained by the original diagram summation approach.

As the last of this sequence, we consider the expansion of $n\nabla\beta U + \nabla n$, or equivalently insofar as (5.17) is concerned, of $n\nabla\beta U$. Since $\ln(ne^{\beta U})$ and n are, to the extent that (6.14) is valid, linear functionals of each other, we further choose $\ln(ne^{\beta U})$ as independent variable. The motivation of course lies not in these somewhat tenuous equivalences, but in the result to be achieved. We have from (4.18) and (5.11)

$$\frac{\delta(\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r}))}{\delta n(\underset{\sim}{r}')} = c(\underset{\sim}{r},\underset{\sim}{r}'), \quad \frac{\delta(\ln n(\underset{\sim}{r}) + \beta U(\underset{\sim}{r}))}{\delta\beta U(\underset{\sim}{r}')} = -(g-1)(\underset{\sim}{r},\underset{\sim}{r}')n(\underset{\sim}{r}') \tag{6.17}$$

so that in terms of the matrix inverses of c and g-1,

$$\frac{\delta n(\underset{\sim}{r})}{\delta(\ln n(\underset{\sim}{r}') + \beta U(\underset{\sim}{r}'))} = c^{-1}(\underset{\sim}{r},\underset{\sim}{r}'), \quad \frac{\delta\beta U(\underset{\sim}{r})}{\delta(\ln n(\underset{\sim}{r}') + \beta U(\underset{\sim}{r}'))} = -\frac{1}{n(\underset{\sim}{r})}(g-1)^{-1}(\underset{\sim}{r},\underset{\sim}{r}') \tag{6.18}$$

and

$$\frac{\delta(n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r}))}{\delta(\ln n(\underset{\sim}{r}') + \beta U(\underset{\sim}{r}'))} = c^{-1}(\underset{\sim}{r},\underset{\sim}{r}')\nabla\beta U(\underset{\sim}{r}) - n(\underset{\sim}{r})\nabla\frac{1}{n(\underset{\sim}{r})}(g-1)^{-1}(\underset{\sim}{r},\underset{\sim}{r}'). \tag{6.19}$$

Hence expanding $n(\underset{\sim}{r})\nabla\beta U(\underset{\sim}{r})$ about uniform n, via (6.4), yields

$$ng(\underset{\sim}{r})\nabla\beta v(\underset{\sim}{r}) = -\int \nabla_{\underset{\sim}{r}}(g-1)^{-1}(\underset{\sim}{r},\underset{\sim}{r}')[\ln g(\underset{\sim}{r}') + \beta v(\underset{\sim}{r}')]d\underset{\sim}{r}' + \ldots \tag{6.20}$$

or to first order, on applying the matrix g-1,

$$\nabla \ln g(\underset{\sim}{r}) + \nabla\beta v(\underset{\sim}{r}) = -n\int g(\underset{\sim}{r}')\nabla\beta v(\underset{\sim}{r}')(g(\underset{\sim}{r}-\underset{\sim}{r}') - 1)d\underset{\sim}{r}', \tag{6.21}$$

which is the truncated Born-Green-Yvon (BGY) equation.[22] This is examined, in Appendix C, from the more usual viewpoint.

PROPERTIES OF SOLUTIONS

We shall not at this time carry out the solutions of the various approximate integral equations in specific cases. However, some general comments are in order. Since the defining interrelation (5.6) is symmetric in g-1 and -c, it is clear that we cannot expect strictly kinematic differences between the character of g-1 and that of c. In fact, writing (5.6) as

$$\int \left(\frac{1}{n(\underset{\sim}{r'})}\,\delta(\underset{\sim}{r}-\underset{\sim}{r'}) - c(\underset{\sim}{r},\underset{\sim}{r'})\right) n(\underset{\sim}{r'}) \left(\frac{1}{n(\underset{\sim}{r'})}\,\delta(\underset{\sim}{r'}-\underset{\sim}{r''}) + (g-1)(\underset{\sim}{r'},\underset{\sim}{r''})\right) d\underset{\sim}{r'}$$

$$= \frac{1}{n(\underset{\sim}{r})}\,\delta(\underset{\sim}{r}-\underset{\sim}{r''}) \tag{6.22}$$

and taking moments, assuming spherical symmetry, corresponding moments will converge or diverge together, precluding any dramatic qualitative difference in this sense. If for example, in conformity with the thought of Ornstein and Zernike, c is so short-ranged (except in regions of thermodynamic singularities) that all moments exist, then the moments of $g-1+n^{-1}\delta$, satisfying[23]

$$\frac{1}{2s+2}\sum \binom{2s+2}{2j+1} C^{(2s-2j)} G^{(2j)} = \frac{1}{n}\,\delta_{so} \tag{6.23}$$

where

$$C^{(s)} = \int \left(\frac{1}{n}\,\delta(\underset{\sim}{r}) - c(\underset{\sim}{r})\right) r^s d\underset{\sim}{r}, \qquad G^{(s)} = \int \left(\frac{1}{n}\,\delta(\underset{\sim}{r}) + g(\underset{\sim}{r}) - 1\right) r^s d\underset{\sim}{r},$$

will exist as well, suggesting at least an exponential decay for g-1 as $r \to \infty$.

On the other hand, even qualitative knowledge of the interaction lays down important restrictions on the solutions of the approximate equations. For very long range internal forces, fluctuation shielding

should make g-1 of considerably shorter range than c, expected to have a range of the order of that of the force. Indeed the Debye-Huckel and P.Y. approximations (6.6,6.12) both give precisely the asymptotic

$$c \to -\beta v. \tag{6.24}$$

The CHNC expression (6.15), $c \to \frac{1}{2}(g-1)^2 - \beta v$, is easily shown to yield the same result, while the Kirkwood-Salsburg and BGY equations (6.10) and (6.21) must first be transformed to

$$c(\underset{\sim}{r}) + \left(e^{\beta v(\underset{\sim}{r})} - 1\right) g(\underset{\sim}{r}) = n \int c(\underset{\sim}{r}-\underset{\sim}{r}') \left(g(\underset{\sim}{r}') e^{\beta v(\underset{\sim}{r}')} - 1\right)\left(e^{-\beta v(\underset{\sim}{r}')} - 1\right) d\underset{\sim}{r}' \tag{6.10'}$$

$$\nabla \ln g(\underset{\sim}{r}) + \nabla \beta v(\underset{\sim}{r}) = n \int c(\underset{\sim}{r}-\underset{\sim}{r}')[\nabla \ln g(\underset{\sim}{r}') + \nabla \beta v(\underset{\sim}{r}') - g(\underset{\sim}{r}') \nabla \beta v(\underset{\sim}{r}')] d\underset{\sim}{r}' \tag{6.21'}$$

to make this limit apparent.

For strong finite range forces, the short range character of the distributions is a very sensitive function of the potential, while the asymptotic long range behavior rather characterizes the qualitative state of the system, and more importantly depends on the approximation made. For example, (6.6) and (6.12) now aver that the direct correlation function c has exactly the same finite range as the potential, a result which almost certainly fails at several classes of thermodynamic

singularities. In particular, suppose that the compressibility becomes infinite, as in a first order transition. Then according to (5.7), $\int c(\underset{\sim}{r})d\underset{\sim}{r} = 1/n$. But for finite range c, the Fourier transform $c_{\underset{\sim}{k}} = \int e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}} c(\underset{\sim}{r})d\underset{\sim}{r}$ is analytic at the origin of k space, so that

$$c_{\underset{\sim}{k}} = \int c(\underset{\sim}{r})d\underset{\sim}{r} + \tfrac{1}{2}k^2C^{(2)} + \ldots \qquad (6.25)$$

Hence at infinite compressibility, $(g-1)(r) = (2\pi)^{-3}\int e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}} c_{\underset{\sim}{k}}/(1-nc_{\underset{\sim}{k}})d\underset{\sim}{k}$ $= (2\pi)^{-3}(2/n^2C^{(2)})\int (1/k^2 + \ldots)e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}}d\underset{\sim}{k}$ does not even exist in a two-dimensional space. (In one dimension, one would not expect $\partial p/\partial n = 0$ to occur any way.) Conversely, evidence of numerical calculations indicates that (6.10) or (6.21) yield excessively long direct correlation lengths. Quite possibly, modification along the lines of (5.20) is called for in this singular domain.

Mixtures

Finally, two direct but very important generalizations of the procedures of this section should be mentioned. The first is to nonuniform systems, due to an external potential, or perhaps without one in a phase coexistence region. Several possible approaches come to mind; these are discussed and compared in Appendix D. The second extension is to mixtures or multicomponent systems. Formally, the only

change is that each particle requires both position $\underset{\sim}{r}$ and component index α for its specification; correspondingly, one has internal potentials $v_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{r}')$ between a type α particle at $\underset{\sim}{r}$ and one of type β at $\underset{\sim}{r}'$, distribution functions $n_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{r}')$, etc. However, in carrying over the results and techniques of this section, some care must be exercised. To see why this is so, let us consider the nonlinear Debye-Huckel equation (6.7,6.8) for mixtures. It is obtained by observing $\ln n_\alpha(\underset{\sim}{r})$ as the external potential

$$U_\gamma(\underset{\sim}{r}) = v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{x}) \tag{6.26}$$

is turned on for each component γ, thereby fixing a particle of type β at $\underset{\sim}{x}$. We have at once

$$\ln n_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x})/n_\beta(\underset{\sim}{x}) = \ln n_\alpha(\underset{\sim}{r})$$

$$- \left(1/n_\alpha(\underset{\sim}{r})\right) \int \sum_\gamma (n_{\alpha\gamma}(\underset{\sim}{r},\underset{\sim}{y}) - n_\alpha(\underset{\sim}{r})n_\gamma(\underset{\sim}{y}) + n_\alpha(\underset{\sim}{r})\delta_{\alpha\gamma}\delta(\underset{\sim}{r}-\underset{\sim}{y}))\beta v_{\gamma\beta}(\underset{\sim}{y},\underset{\sim}{x})d\underset{\sim}{y}, \tag{6.27}$$

or

$$\ln g_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x}) + \beta v_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x}) + \sum_\gamma \int (g_{\alpha\gamma}(\underset{\sim}{r},\underset{\sim}{y}) - 1)n_\gamma(\underset{\sim}{y})\beta v_{\gamma\beta}(\underset{\sim}{y},\underset{\sim}{x})d\underset{\sim}{y} = 0. \tag{6.28}$$

Equation (6.28) appears a natural extension of (6.8), but it is not satisfactory: if $g_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x})$ is a solution, then $\bar{g}_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x}) \equiv g_{\beta\alpha}(\underset{\sim}{x},\underset{\sim}{r})$ will not be, so that $g_{\alpha\beta}$ cannot be symmetric.

To overcome the difficulty of (6.28), a formulation which is symmetric from the start is called for. Let us therefore instead expand $\ln \Xi$ as two particles are fixed: $U_\gamma(\underset{\sim}{r}) = v_{\gamma\alpha}(\underset{\sim}{r},\underset{\sim}{x}) + v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{y})$. Then from (6.3),

$$\ln \Xi + \ln n_{\alpha\beta}(\underset{\sim}{x},\underset{\sim}{y}) + \beta v_{\alpha\beta}(\underset{\sim}{x},\underset{\sim}{y}) - \ln z_\alpha - \ln z_\beta$$

$$= \ln \Xi - \beta \int \sum_\gamma n_\gamma(\underset{\sim}{r})(v_{\gamma\alpha}(\underset{\sim}{r},\underset{\sim}{x}) + v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{y}))d\underset{\sim}{r}$$

$$+ \tfrac{1}{2}\beta^2 \iint \sum_{\gamma\lambda}' \Big(n_{\gamma\lambda}(\underset{\sim}{r},\underset{\sim}{r}') - n_\gamma(\underset{\sim}{r}) n_\lambda(\underset{\sim}{r}') + n_\gamma(\underset{\sim}{r})\delta_{\gamma\lambda}\delta(\underset{\sim}{r}-\underset{\sim}{r}') \Big)$$

$$\times \Big(v_{\gamma\alpha}(\underset{\sim}{r},\underset{\sim}{x}) + v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{y}) \Big) \Big(v_{\lambda\alpha}(\underset{\sim}{r}',\underset{\sim}{x}) + v_{\lambda\beta}(\underset{\sim}{r}',\underset{\sim}{y}) \Big) d\underset{\sim}{r}\, d\underset{\sim}{r}' + \cdots . \tag{6.29}$$

Now if $\underset{\sim}{x}$ and $\underset{\sim}{y}$ are asymptotically separated with respect to the range of the potential (as well as $n_{\gamma\lambda} - n_\gamma n_\lambda$), (6.29) reduces to

$$\ln n_\alpha(\underset{\sim}{x}) + \ln n_\beta(\underset{\sim}{y}) - \ln z_\alpha - \ln z_\beta$$

$$= -\beta \int \sum_\gamma n_\gamma(\underset{\sim}{r})(v_{\gamma\alpha}(\underset{\sim}{r},\underset{\sim}{x}) + v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{y}))d\underset{\sim}{r}$$

$$+ \tfrac{1}{2}\beta^2 \iint \sum_{\gamma\lambda} (n_{\gamma\lambda}(\underset{\sim}{r},\underset{\sim}{r}') - n_\gamma(\underset{\sim}{r})n_\lambda(\underset{\sim}{r}') + n_\gamma(\underset{\sim}{r})\delta_{\gamma\lambda}\delta(\underset{\sim}{r}-\underset{\sim}{r}'))$$

$$\times(v_{\gamma\alpha}(\underset{\sim}{r},\underset{\sim}{x})v_{\lambda\alpha}(\underset{\sim}{r}',\underset{\sim}{x}) + v_{\gamma\beta}(\underset{\sim}{r},\underset{\sim}{y})v_{\lambda\beta}(\underset{\sim}{r}',\underset{\sim}{y}))d\underset{\sim}{r}\, d\underset{\sim}{r}' + \ldots \tag{6.30}$$

which, for a uniform system, is valid irrespective of the separation. Thus, for a uniform system, we have on subtracting (6.30) from (6.29),

$$\ln g_{\alpha\beta}(\underset{\sim}{x}-\underset{\sim}{y}) + \beta v_{\alpha\beta}(\underset{\sim}{x}-\underset{\sim}{y}) = \tfrac{1}{2}\beta^2 \iint \sum_{\gamma\lambda} \left(g_{\gamma\lambda}(\underset{\sim}{r}-\underset{\sim}{r}') - 1 + \frac{1}{n_\gamma}\delta_{\gamma\lambda}\delta(\underset{\sim}{r}-\underset{\sim}{r}')\right)n_\gamma n_\lambda$$

$$\times\left(v_{\gamma\alpha}(\underset{\sim}{r}-\underset{\sim}{x})v_{\lambda\beta}(\underset{\sim}{r}'-\underset{\sim}{y}) + v_{\gamma\beta}(\underset{\sim}{r}-\underset{\sim}{y})v_{\lambda\alpha}(\underset{\sim}{r}'-\underset{\sim}{x})\right)d\underset{\sim}{r}\, d\underset{\sim}{r}', \tag{6.31}$$

a symmetric version of the nonlinear Debye-Huckel equation.

On the other hand, it is worth pointing out that not all approximations require reformulation to obtain the desired symmetry. For example, in the extended PY approximation,[24] one expands $n_\alpha(\underset{\sim}{r})e^{\beta U_\alpha(\underset{\sim}{r})}$ as U of (6.26) is turned on. Mimicing the procedure (6.11, 6.12), there results

$$g_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x})\left(1 - e^{\beta v_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x})}\right) = c_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x})$$

where

$$g_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x}) - 1 = c_{\alpha\beta}(\underset{\sim}{r},\underset{\sim}{x}) + \int \sum_\gamma (g_{\alpha\gamma}(\underset{\sim}{r},\underset{\sim}{y}) - 1)n_\gamma(\underset{\sim}{y})c_{\gamma\beta}(\underset{\sim}{y},\underset{\sim}{x})d\underset{\sim}{y}, \quad (6.32)$$

and further symmetrization is unnecessary.

VII. ESTIMATION OF THE REMAINDER

The various truncated expansions we have used introduce approximations of unknown magnitude, and often of unknown qualitative content as well. This situation can be improved considerably by writing down explicit remainder terms for the expansions. As prototype, one has the standard single-variable Mac Laurin series with a remainder:[25]

$$H(\gamma) = \sum_0^{\ell} (\gamma^m/m!)H^{(m)}(0) + \frac{1}{\ell!}\int_0^\gamma (\gamma-\lambda)^\ell H^{(\ell+1)}(\lambda)d\lambda, \quad (7.1)$$

obtained by integrating the remainder by parts $\ell+1$ times. If now H is a functional of $h(\underset{\sim}{r})$, into which we introduce a parameter λ such that

$$h(\underset{\sim}{r},\lambda) = (1-\lambda)h_0(\underset{\sim}{r}) + \lambda h(\underset{\sim}{r})$$

or

$$h(\underset{\sim}{r},0) = h_0(\underset{\sim}{r}), \quad \frac{\partial h(\underset{\sim}{r},\lambda)}{\partial\lambda} = h(\underset{\sim}{r}) - h_0(\underset{\sim}{r}), \quad h(\underset{\sim}{r},1) = h(\underset{\sim}{r}), \tag{7.2}$$

application of the chain rule

$$\frac{\partial}{\partial\lambda} = \int \frac{\partial h(\underset{\sim}{r},\lambda)}{\partial\lambda} \frac{\delta}{\delta h(\underset{\sim}{r},\lambda)} \, d\underset{\sim}{r} \tag{7.3}$$

converts (7.1) at once to

$$\begin{aligned} H[h] = H[h_0] &+ \int \frac{\partial h(\underset{\sim}{r},0)}{\partial\lambda} \frac{\delta H(h_0)}{\delta h_0(\underset{\sim}{r})} \, d\underset{\sim}{r} \\ &+ \frac{1}{2!} \iint \frac{\partial h(\underset{\sim}{r}_1,0)}{\partial\lambda} \frac{\partial h(\underset{\sim}{r}_2,0)}{\partial\lambda} \frac{\delta^2 H[h_0]}{\delta h_0(\underset{\sim}{r}_1)\delta h_0(\underset{\sim}{r}_2)} \, d\underset{\sim}{r}_1 d\underset{\sim}{r}_2 \\ &+ \cdots + \frac{1}{\ell!} \int_0^1 \int \cdots \int (1-\lambda)^\ell \frac{\partial h(\underset{\sim}{r}_1,\lambda)}{\partial\lambda} \cdots \frac{\partial h(\underset{\sim}{r}_{\ell+1},\lambda)}{\partial\lambda} \\ &\times \frac{\delta^{\ell+1} H[h_0]}{\delta h_0(\underset{\sim}{r}_1)\cdots\delta h_0(\underset{\sim}{r}_{\ell+1})} \, d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_{\ell+1} d\lambda. \end{aligned} \tag{7.4}$$

Generalization to explicitly λ-dependent $\partial h/\partial\lambda$ is immediate.

Equation (7.4), a natural generalization of (4.20), (4.22), is the desired functional Taylor series with remainder. In some cases, bounds on the remainder are available, and these may be helpful in assessing the associated truncated approximations. In other cases, one can estimate the value of the remainder with sufficient accuracy to yield plausible nonperturbative corrections; this is done principally by interpolation with respect to λ, or by a λ-independent closure relation among distributions. Such closures take the form of generalized superposition approximations, and also appear in many instances simply as conditions that the remainder term be absent, so that the truncation is literally valid.

Bounds on Expansions

Let us first consider the possibility of bounding the remainder term of a typical "turning on" expansion. This possibility is closely tied to the nature of the function

$$\omega(\underset{\sim}{r}) \equiv \frac{\partial h(\underset{\sim}{r},\lambda)}{\partial\lambda} = h(\underset{\sim}{r}) - h_o(\underset{\sim}{r}) \tag{7.5}$$

entering into the expansion of the remainder at any value of λ, and one cannot be cavalier in choosing an approximation whose error is to be bounded. With a nonnegative potential, for example, and an expansion via the potential or Mayer f-function, we may use the fact that

$$\int n_s(\underset{\sim}{r}_1,\cdots,\underset{\sim}{r}_s)\omega(\underset{\sim}{r}_1)\cdots\omega(\underset{\sim}{r}_s)d\underset{\sim}{r}_1\cdots d\underset{\sim}{r}_s \geq 0 \quad \text{if} \quad \omega(\underset{\sim}{r}) \geq 0 \qquad (7.6)$$

to establish bounds on expansions whose successive derivatives generate multiples of the n_s. None of the expansions thus far considered have this property, but those of Kirkwood-Salsburg[19] and Mayer-Montroll[26] do. Let us now derive these expansions. According to (4.10), the sequence of distributions may be generated by regarding $\Xi[U]$, or more generally $\Xi[U]\exp\beta(U(\underset{\sim}{x}_1) + \cdots + U(\underset{\sim}{x}_t))n_t(\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t)$, as a functional of $e^{-\beta U}$. We choose $h(\underset{\sim}{r}) = e^{-\beta U(\underset{\sim}{r})}$, $h_0(\underset{\sim}{r}) = 1$, $\omega(\underset{\sim}{r}) = e^{-\beta U(\underset{\sim}{r})} - 1 \leq 0$, $U(\underset{\sim}{x}) = v(\underset{\sim}{x},\underset{\sim}{r}_1) + \cdots + v(\underset{\sim}{x},\underset{\sim}{r}_s)$. Thus from (6.1),(6.2),(6.3),

$$n_t(\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t|U)\exp(\beta\Sigma U(\underset{\sim}{x}_i))\Xi[U]$$

$$= z^{-s}n_{s+t}(\underset{\sim}{r}_1\cdots\underset{\sim}{r}_s\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t)\exp\beta\Big(\Sigma v(\underset{\sim}{x}_i,\underset{\sim}{r}_j) + \tfrac{1}{2}\Sigma' v(\underset{\sim}{r}_i,\underset{\sim}{r}_j)\Big)\Xi, \qquad (7.7)$$

and we conclude, using (4.10), that[27]

$$z^{-s}n_{t+s}(\underset{\sim}{r}_1\cdots\underset{\sim}{r}_s\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t)\exp\beta\Big(\Sigma v(\underset{\sim}{x}_i,\underset{\sim}{r}_j) + \tfrac{1}{2}\Sigma' v(\underset{\sim}{r}_i,\underset{\sim}{r}_j)\Big)$$

$$= \sum_{m=0}^{\ell}\frac{1}{m!}\int\cdots\int\omega(\underset{\sim}{y}_1)\cdots\omega(\underset{\sim}{y}_m)n_{t+m}(\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t\underset{\sim}{y}_1\cdots\underset{\sim}{y}_m)(d\underset{\sim}{y})^m$$

$$+ \frac{1}{\ell!}\int_0^1\int\cdots\int(1-\lambda)^{\ell}\omega(\underset{\sim}{y}_1)\cdots\omega(\underset{\sim}{y}_{\ell+1})n_{t+\ell+1}(\underset{\sim}{x}_1\cdots\underset{\sim}{x}_t\underset{\sim}{y}_1\cdots\underset{\sim}{y}_{\ell+1}|U(\lambda))$$

$$\exp\beta\Big(\Sigma U(\underset{\sim}{x}_i,\lambda) + \Sigma U(\underset{\sim}{y}_j,\lambda)\Big)\Xi[U(\lambda)]/\Xi(d\underset{\sim}{y})^{\ell+1}d\lambda, \qquad (7.8)$$

where $e^{-\beta U(\underset{\sim}{r},\lambda)} \equiv \lambda e^{-\beta U(\underset{\sim}{r})} + (1-\lambda)$. Equation (7.8) constitutes the truncated Mayer-Montroll sequence when s varies at t = 0, that of Kirkwood-Salsburg when t varies at s = 1. The sign of the remainder is always $(-1)^{\ell+1}$.

The formula (7.8) is appropriate to bounding the distributions themselves and not just relations between them. As a typical case, let us write the Mayer-Montroll sequence for (7.8) as

$$n_s = \sum_0^{\ell} z^s (-1)^m A_{sm} n_m + (-1)^{\ell+1} R_{s\ell} \tag{7.9}$$

where the $R_{s\ell}$ are nonnegative, and the A_{sm} are nonnegative integral operators on nonnegative functions, without z-dependence. Taking $\ell = 0$, we have

$$n_s = z^s A_{so} \cdot 1 - R_{so}; \tag{7.10}$$

then taking $\ell = 1$ and substituting n_1 from (7.10),

$$n_s = z^s A_{so} \cdot 1 - z^{s+1} A_{s1} A_{10} \cdot 1 + z^s A_{s1} R_{10} + R_{s1} - R_{so}. \tag{7.11}$$

Next, $\ell = 2$, with n_1 chosen by (7.11), n_o by (7.10) etc. In this fashion, we iterate out an alternating sequence

$$\begin{aligned} n_s &\geq z^s A_{so} \cdot 1 \\ &\leq z^s A_{so} \cdot 1 - z^{s+1} A_{s1} A_{10} \cdot 1 \\ &\geq z^s A_{so} \cdot 1 - z^{s+1} A_{s1} A_{10} \cdot 1 + z^{s+2} (A_{s1} A_{21} A_{10} \cdot 1 + A_{s2} A_{20} \cdot 1) \end{aligned} \tag{7.12}$$

of successive upper and lower bounds,[27,28,29] each being simply a truncation of the full Mayer fugacity expansion.

For potentials which are not totally repulsive but which have e.g., hard core plus attractive tail, a more careful argument shows that the remainder may again be bounded. However, if we are expanding in a quantity whose variation during the turning on process does not have a predictable sign, the remainder must have a nonnegative (or nonpositive) kernel, by which we mean that

$$\int \cdots \int K(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_s)\omega(\underset{\sim}{r}_1)\cdots\omega(\underset{\sim}{r}_s)d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_s \geq 0 \quad \text{for all } \omega(\underset{\sim}{r}). \tag{7.13}$$

Clearly, s must be even for this to be possible. A prototype kernel of this nature is given by the modified distribution function $\hat{n}_s$,[11]

$$\int \cdots \int \hat{n}_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s)\omega(\underset{\sim}{r}_1)\cdots\omega(\underset{\sim}{r}_s)d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_s \equiv \langle (\Sigma\omega(\underset{\sim}{x}_1))^s \rangle, \tag{7.14}$$

which differs from n_s precisely in that two locations can refer to the same particle. One readily finds $\hat{n}_1(\underset{\sim}{r}) = n_1(\underset{\sim}{r})$, $n_2(\underset{\sim}{r},\underset{\sim}{r}')$ $= n_2(\underset{\sim}{r},\underset{\sim}{r}') + n(\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}')$, $\hat{n}_3(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'') = n_3(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'') + n_2(\underset{\sim}{r},\underset{\sim}{r}')\delta(\underset{\sim}{r}'-\underset{\sim}{r}'')$ $+ n_2(\underset{\sim}{r}',\underset{\sim}{r}'')\delta(\underset{\sim}{r}''-\underset{\sim}{r}) + n_2(\underset{\sim}{r}'',\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}') + n(\underset{\sim}{r})\delta(\underset{\sim}{r}-\underset{\sim}{r}')\delta(\underset{\sim}{r}-\underset{\sim}{r}'')$, etc. Since $\Xi[U] = \Xi\langle e^{-\beta\Sigma U(x_i)} \rangle$, it also follows from (7.14) that

$$\frac{\delta^s \Xi[U]}{\delta - \beta U(\underset{\sim}{r}_1)\cdots\delta - \beta U(\underset{\sim}{r}_s)} = \hat{n}_s(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_s)\Xi[U], \tag{7.15}$$

the analog of (4.10). Thus a number of aspects of (7.8) can immediately be generalized to a hybrid expansion for an arbitrary potential.

The nonnegativity of $\hat{n}_s$ for even s is not as efficient as possible: "smaller" kernels remain nonnegative. For example, not only is $\langle(\Sigma\omega(\underset{\sim}{x}_1))^2\rangle \geq 0$, but also $\langle(\Sigma\omega(\underset{\sim}{x}_1))^2\rangle - \langle\Sigma\omega(\underset{\sim}{x}_1)\rangle^2 \geq 0$. This means that the modified Ursell function

$$\begin{aligned}\hat{\mathcal{J}}_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) &\equiv \hat{n}_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) - n(\underset{\sim}{r}_1)n(\underset{\sim}{r}_2)\\ &= \mathcal{J}_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) + n(\underset{\sim}{r}_1)\delta(\underset{\sim}{r}_1-\underset{\sim}{r}_2)\\ &= n_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) - n(\underset{\sim}{r}_1)n(\underset{\sim}{r}_2) + n(\underset{\sim}{r}_1)\delta(\underset{\sim}{r}_1-\underset{\sim}{r}_2) \qquad (7.16)\end{aligned}$$

is nonnegative. As a direct application, we first note that

$$\hat{\mathcal{J}}_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) = \delta^2 \ln \Xi/\delta\beta U(\underset{\sim}{r}_1)\delta\beta U(\underset{\sim}{r}_2). \qquad (7.17)$$

Hence turning on an s-body potential, as preceding (7.7), and observing $\ln \Xi$, we have $\beta pV + \ln n_s(\underset{\sim}{r}_1\cdots\underset{\sim}{r}_s) - s \ln z + \frac{1}{2}\beta\Sigma' v(\underset{\sim}{r}_i,\underset{\sim}{r}_j)$ $= \beta pV - \int n(\underset{\sim}{x})\beta\Sigma v(\underset{\sim}{x},\underset{\sim}{r}_i)d\underset{\sim}{x} + \frac{1}{2}\int\int\int \hat{\mathcal{J}}_2(\underset{\sim}{x}_1,\underset{\sim}{x}_2,\lambda)[\beta\Sigma v(\underset{\sim}{x}_1\underset{\sim}{r}_i)][\beta\Sigma v(\underset{\sim}{x}_2,\underset{\sim}{r}_i)]$ $\times d\underset{\sim}{x}_1 d\underset{\sim}{x}_2(1-\lambda)d\lambda$, from which

$$z^{-s}n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)e^{\frac{1}{2}\beta\Sigma' v(\underset{\sim}{r}_i,\underset{\sim}{r}_j)} \geq e^{-\beta\int n(\underset{\sim}{x})\Sigma v(\underset{\sim}{x},\underset{\sim}{r}_i)d\underset{\sim}{x}}. \qquad (7.18)$$

$\hat{\mathcal{F}}_2$ of (7.16) is the start of a multidimensional nonnegative sequence of combinations of distribution functions, all of which result from the nonnegativity of the expectation of a square, expressed in general as

$$\left\langle \Big(\sum_{\alpha+\Sigma_j \alpha_j = s} C_{\alpha\alpha_1\alpha_2\ldots} \omega^{\alpha} \langle\omega\rangle^{\alpha_1} \langle\omega^2\rangle^{\alpha_2} \ldots \Big)^2 \right\rangle \geq 0 \quad \text{where} \quad \omega = \Sigma\omega(\underset{\sim}{x}_1). \tag{7.19}$$

Alternatively, from $\langle (\Sigma c_j \omega^j)^2 \rangle = \Sigma c_i \langle \omega^{i+j} \rangle c_j \geq 0$, the matrix $\langle \omega^{i+j} \rangle$ is positive definite. The positivity of its minors is equivalent to (7.19):

$$\begin{vmatrix} \langle\omega^2\rangle & \langle\omega\rangle \\ \langle\omega\rangle & 1 \end{vmatrix} \geq 0, \quad \begin{vmatrix} \langle\omega^4\rangle & \langle\omega^3\rangle & \langle\omega^2\rangle \\ \langle\omega^3\rangle & \langle\omega^2\rangle & \langle\omega\rangle \\ \langle\omega^2\rangle & \langle\omega\rangle & 1 \end{vmatrix} \geq 0, \quad \text{etc.,} \tag{7.20}$$

which indeed can be written as $\langle [\omega - \langle\omega\rangle]^2 \rangle \geq 0$,
$\langle [(\langle\omega^2\rangle - \langle\omega\rangle^2)\omega^2 - (\langle\omega^3\rangle - \langle\omega^2\rangle\langle\omega\rangle)\omega + (\langle\omega^3\rangle\langle\omega\rangle - \langle\omega^2\rangle^2)]^2 \rangle \geq 0$,
etc. The full consequences of (7.19),(7.20) are not known.

Interpolation of Remainder

When rigorous bounds on remainders are not available, an evaluation by interpolation may be a reasonable means of assessment. Since this requires knowledge at several points of the turning-on process, some incidental relation between

the intermediate quantities is a virtual necessity. The simplest instance in which this technique is practicable is that of a one-term expansion of ln Ξ in an external potential which becomes the internal potential based at the origin:

$$\beta pV + \ln n/z = \beta pV + \int_0^1 \frac{\delta \ln \Xi}{\delta_e^{-\beta U_\lambda(\underset{\sim}{r})}} \frac{\partial}{\partial\lambda} e^{-\beta U_\lambda(\underset{\sim}{r})} d\underset{\sim}{r}\, d\lambda$$

$$= \beta pV + \int_0^1 \int n_\lambda(\underset{\sim}{r}) e^{\beta U_\lambda(\underset{\sim}{r})} \frac{\partial}{\partial\lambda} e^{-\beta U_\lambda(\underset{\sim}{r})} d\underset{\sim}{r}\, d\lambda. \tag{7.21}$$

If we choose[30]

$$U_\lambda(\underset{\sim}{r}) = v(\underset{\sim}{r}/\lambda), \tag{7.22}$$

a scale transformation, we are then called upon to compute

$$\frac{\partial \ln \Xi}{\partial\lambda} = -\int n_\lambda(\underset{\sim}{r}) e^{\beta U_\lambda(\underset{\sim}{r})} \underset{\sim}{r}\cdot\nabla e^{-\beta U_\lambda(\underset{\sim}{r})} d\underset{\sim}{r}/\lambda, \tag{7.23}$$

and we shall do this by choosing several values of λ at which $\partial \ln \Xi/\partial\lambda$ is known. Obviously, when λ = 1, $n_\lambda(\underset{\sim}{r}) = ng(\underset{\sim}{r})$, $U_\lambda(\underset{\sim}{r}) = v(\underset{\sim}{r})$, and the virial theorem (2.8) then tells us that

$$\lambda = 1: \quad \frac{\partial \ln \Xi}{\partial \lambda} = 6\left(1 - \frac{\beta p}{n}\right). \tag{7.24}$$

For further information, we must specify the potential somewhat more completely, and for this purpose we assume that v is composed of a hard core φ_o of radius a and a non-singular tail φ from a to a + b. We may then investigate (7.23) as $\lambda \to \infty$; since the volume V must then go to infinity too, we can take the limit of $V \to \infty$ with e.g., $\lambda = (V^{\frac{1}{3}}/a+b)^{\frac{1}{2}}$, guaranteeing that the "test particle" becomes arbitrarily large but remains infinitesimal compared with the size of the system. As $\lambda \to \infty$, (7.23) has two contributions. For the long range φ_λ, changing slowly relative to any correlation length, n_λ may be given its asymptotic value n:

$$\int_{r \geq a\lambda} n_\lambda(\underset{\sim}{r}) e^{\beta\varphi_\lambda(\underset{\sim}{r})}\, \underset{\sim}{r}\cdot\nabla e^{-\beta\varphi_\lambda(\underset{\sim}{r})}\, d\underset{\sim}{r} = -n\beta \int_{r \geq a\lambda} \underset{\sim}{r}\cdot\nabla\varphi_\lambda(\underset{\sim}{r})\, d\underset{\sim}{r}$$

$$= n\beta \int_{r=a\lambda} \varphi_\lambda(\underset{\sim}{r})\, \underset{\sim}{r}\cdot d\underset{\sim}{s} + 3n\beta \int_{r \geq a\lambda} \varphi_\lambda(\underset{\sim}{r})\, d\underset{\sim}{r} = 4\pi n\beta\lambda^3 a^3 \varphi(a)$$

$$+ 3n\beta\lambda^3 \int \varphi(\underset{\sim}{r})\, d\underset{\sim}{r}. \quad \text{For the hard core part,} \int n_\lambda(\underset{\sim}{r}) e^{\beta\varphi_{\lambda o}(\underset{\sim}{r})}\, \underset{\sim}{r}$$

$$\cdot\nabla e^{-\beta\varphi_{\lambda o}(\underset{\sim}{r})}\, d\underset{\sim}{r} = -\beta \int n_\lambda(\underset{\sim}{r})\, \underset{\sim}{r}\cdot\nabla\varphi_{\lambda o}(\underset{\sim}{r})\, d\underset{\sim}{r} = -\beta \left\langle \int \underset{\sim}{r}\cdot\Sigma \underset{\sim}{F}_i(\underset{\sim}{r})\, d\underset{\sim}{r} \right\rangle,$$

where $\underset{\sim}{F}_i(\underset{\sim}{r})$ is the force exerted on the external source by the i^{th} particle, $= 4\pi\beta\lambda^3 a^3 p$, since the hard surface is flat in the limit. Hence

$$\lambda \to \infty: \quad \frac{\partial \ln \Xi}{\partial \lambda} \to -4\pi n\beta\lambda^2 a^3\left[\varphi(a) + \int \varphi(\underset{\sim}{r})d\underset{\sim}{r}/\tfrac{4}{3}\pi a^3 + p/n\right]. \tag{7.25}$$

On the other hand for small λ, there seems to be no problem, since clearly $\partial \ln \Xi/\partial\lambda \to 0$ as $\lambda \to 0$. The trouble however is that $\partial \ln \Xi/\partial\lambda$ has several analytic breaks between $\lambda = 0$ and $\lambda = 1$. For an interpolation employing $\lambda = 1$ and $\lambda = \infty$, one would then want to interpolate only from the closest value of λ to unity where $\partial \ln \Xi/\partial\lambda$ is still known exactly. Let us examine this point in detail. By definition we have

$$\Xi[\lambda]/\Xi[0] = \langle\Pi(1+f_\lambda(\underset{\sim}{r}_i))\rangle$$

$$= 1 + \int n(\underset{\sim}{r})f_\lambda(\underset{\sim}{r})d\underset{\sim}{r} + \tfrac{1}{2}\iint n_2(\underset{\sim}{r},\underset{\sim}{r}')f_\lambda(\underset{\sim}{r})f_\lambda(\underset{\sim}{r}')d\underset{\sim}{r}\, d\underset{\sim}{r}' + \cdots \tag{7.26}$$

where $f_\lambda(\underset{\sim}{r}) \equiv e^{-\beta U_\lambda(\underset{\sim}{r})} - 1$. But $f_\lambda(\underset{\sim}{r}) = 0$ for $r > \lambda(a+b)$ while $n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) = 0$ when any $|\underset{\sim}{r}_i - \underset{\sim}{r}_j| < a$. It follows by obvious geometric considerations that

$$\iint n_2(\underset{\sim}{r}\underset{\sim}{r}')f_\lambda(\underset{\sim}{r})f_\lambda(\underset{\sim}{r}')d\underset{\sim}{r}\, d\underset{\sim}{r}' = 0 \quad \text{for} \quad \lambda \leq \tfrac{1}{2}a/a+b,$$

$$\iiint n_3(\underset{\sim}{r},\underset{\sim}{r}',\underset{\sim}{r}'')f_\lambda(\underset{\sim}{r})f_\lambda(\underset{\sim}{r}')f_\lambda(\underset{\sim}{r}'')d\underset{\sim}{r}\, d\underset{\sim}{r}'\, d\underset{\sim}{r}'' = 0 \quad \text{for} \quad \lambda \leq 3^{-\frac{1}{2}}a/a+b, \quad \text{etc.}$$

Further, by direct integration at contact

$$\iint n_2(\underset{\sim}{r},\underset{\sim}{r}')f_\lambda(\underset{\sim}{r})f_\lambda(\underset{\sim}{r}')d\underset{\sim}{r}\, d\underset{\sim}{r}' = n_2(a)f(a+b)^2\left(\lambda - \tfrac{1}{2}\frac{a}{a+b}\right)^3 \frac{32\pi^2}{3}a^3(a+b)$$

$$+ O\left(\left(\lambda - \tfrac{1}{2}\frac{a}{a+b}\right)^4\right) \quad \text{when} \quad \lambda \geq \tfrac{1}{2}\frac{a}{a+b}, \tag{7.27}$$

with the corresponding n_3 integral starting at still higher order. We see then that

$$\Xi[\lambda]/\Xi[0] = 1 + n\lambda^3 \int f(\underset{\sim}{r})d\underset{\sim}{r} \quad \text{for} \quad \lambda \leq \tfrac{1}{2}\frac{a}{a+b}, \tag{7.28}$$

continuous with its first two derivatives at $\lambda = \frac{1}{2}a/a+b$, with its third as well if $f(a+b) = 0$, but its third derivative known anyway in terms of $n_2(a)$ and hence e.g., in terms of the pressure p for pure hard cores. Although we thus have three or four pieces of information at $\lambda = \frac{1}{2}a/a+b$, let us at this time use only the value

$$\frac{\partial \ln Z}{\partial \lambda} = \frac{3n(\frac{1}{2}a/a+b)^2 \int f(\underset{\sim}{r})d\underset{\sim}{r}}{1 + n(\frac{1}{2}a/a+b)^3 \int f(\underset{\sim}{r})d\underset{\sim}{r}} \qquad (7.29)$$

for $\lambda = \frac{1}{2}\frac{a}{a+b}$ together with the exact

$$\ln \Xi[\tfrac{1}{2}a/a+b] - \ln \Xi[0] = \ln\left[1 + n(\tfrac{1}{2}a/a+b)^3 \int f(\underset{\sim}{r})d\underset{\sim}{r}\right] \qquad (7.30)$$

so that we may restrict our interpolation to the region $\frac{1}{2}a/a+b \leq \lambda \leq \infty$.

The simplest way to interpolate the three-point information (7.24),(7.25),(7.29) is by a quadratic expansion

$$\partial \ln \Xi/\partial\lambda = \alpha + \gamma\lambda - K\lambda^2. \qquad (7.31)$$

Substituting, one easily finds that

$$\alpha = -\frac{K}{2}\frac{a}{a+b} - \frac{6a}{a+2b}\left(1 - \frac{\beta p}{n}\right) - 12\frac{(a+b)^2}{a(a+2b)}\frac{y}{1-y}$$

$$\gamma = \frac{K}{2}\frac{3a+2b}{a+b} + 12\frac{a+b}{a+2b}\left(1 - \frac{\beta p}{n}\right) + 12\frac{(a+b)^2}{a(a+2b)}\frac{y}{1-y} \qquad (7.32)$$

where

$$K = 4\pi\beta a^3 \left[p + n\left(\varphi(a) + \int \varphi(\underset{\sim}{r})d\underset{\sim}{r}/\ \frac{4}{3}\pi a^3\right)\right]$$

$$y = -\frac{n}{8}\left(\frac{a}{a+b}\right)^3 \int f(\underset{\sim}{r})d\underset{\sim}{r},$$

from which it follows that in this approximation

$$\int_{\frac{1}{2}\frac{a}{a+b}}^{1} \frac{\partial \ln \Xi}{\partial\lambda}\, d\lambda = \frac{3}{2}\frac{a+2b}{a+b}\left(1 - \frac{\beta p}{n}\right) - \frac{3}{2}\frac{a+2b}{a}\frac{y}{1-y} + \frac{K}{48}\left(\frac{a+2b}{a+b}\right)^3. \tag{7.33}$$

Combining (7.21),(7.30),(7.33) hence yields the equation of state

$$\ln n/z = \ln(1-y) - \frac{3}{2}\left(\frac{q}{2-q}\right)\frac{y}{1-y} + \frac{3}{2}q\left(1 - \frac{\beta p}{n}\right) + \frac{\pi\beta a^3}{12}q^3 p$$

$$+ \frac{\pi\beta a^3}{12}q^3 n\left[\varphi(a) + \int \varphi(\underset{\sim}{r})d\underset{\sim}{r}/\ \frac{4}{3}\pi a^3\right] \tag{7.34}$$

where

$$q \equiv a+2b/a+b.$$

To eliminate z, one uses the identity

$$\frac{d}{dn} \ln \frac{n}{z} = - \frac{1}{n} \frac{d}{dn} (\beta p - n), \tag{7.35}$$

so that, on n-differentiation,

$$\frac{d}{dn}\left[\left(q - \frac{1}{2}\right)\frac{1}{n} - \frac{\pi a^3}{12} q^3\right)^{1 + \frac{2q}{2q-1}} (\beta p - n)\right]$$

$$= \left(\left(q - \frac{1}{2}\right)\frac{1}{n} - \frac{\pi a^3}{12} q^3\right)^{\frac{2q}{2q-1}} \left\{\frac{\pi a^3}{12} q^3 \left[1 + \beta\varphi(a) + \int \beta\varphi(\underset{\sim}{r}) d\underset{\sim}{r} / \frac{4}{3}\pi a^3\right]\right.$$

$$\left. - \frac{3}{2} \frac{q}{2-q} \frac{1}{(1-y)^2} \frac{y}{n} - \frac{1}{1-y} \frac{y}{n}\right\}. \tag{7.36}$$

In the special case of a pure hard sphere, $\varphi = 0$, $b = 0$, $q = 1$, (7.36) reduces to $d/dy[(\beta p - n)(1-y)^3/y^3] = (n/y)[(1-y)^2/y^2 - 2(1-y)/y^2 - 3/y^2]$ which integrates at once to (Reiss, Frisch, and Lebowitz;[30] see also Section X)

$$\frac{\beta p}{n} = \frac{1+y+y^2}{(1-y)^3}, \qquad y = \frac{\pi a^3}{6} n. \tag{7.37}$$

Comparison of (7.37) with numerical data on hard sphere systems[31] is gratifying in the extreme.

VIII. CLOSURES AND CORRECTIONS

When a remainder in a distribution expansion cannot be bounded or reasonably interpolated, the problem of estimation goes over to that of at least expressing the nature of the assumption made in a truncation and of setting up a correction scheme. Both, and especially the latter, generally involve recognition of an expansion hierarchy of which the approximation in question is prototype (similar to (7.8)), truncated by an explicit approximation.

Kirkwood Approximation

As a first instance of a general class of explicit closures in self-consistent form, let us consider the nonlinear Debye-Huckel approximation of (6.7),(6.8), obtainable by expanding $\ln(ne^{\beta U})$ as a functional of $U(\underset{\sim}{r}) \rightarrow v(\underset{\sim}{r},\underset{\sim}{x})$. This expansion, in common with all of the others, can be extended to a hierarchy in at least three obvious ways (which are not always independent): n can refer to n_s, the derivatives of $\ln(ne^{\beta U})$ rather than it itself can be expanded, or U can fix a set of t particles rather than one. Choosing the first of these methods, we observe[32]

$$\ln(n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)\exp \beta\Sigma U(\underset{\sim}{r}_1)) \quad \text{functional of} \quad U(\underset{\sim}{r};\lambda) = \lambda v(\underset{\sim}{r},\underset{\sim}{x}),$$

(8.1)

leading to a functional series for $\ln[n_{s+1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s,\underset{\sim}{x})/n(\underset{\sim}{x})$ $\exp \beta\Sigma v(\underset{\sim}{r}_i,\underset{\sim}{x})]$. It is also convenient to work directly with the generalized g-functions or intrinsic correlations

$$\ln g_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s) = \sum_{\Omega\subset\{1,\ldots,s\}} (-1)^{s-\Omega} \ln n_\Omega(\ldots\underset{\sim}{r}_j\ldots),$$

(8.2)

found by successive addition and subtraction from (8.1), here Ω refers to a subset of $(1,\ldots,s)$.

The key to the self-consistent character of the approach being described is that the expansion is carried out only to an intermediate stage. Thus stopping at the remainder for the first term in the expansion (8.1), one has

$$\ln \frac{n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s;\ \lambda)}{n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)} + \lambda\Sigma v(\underset{\sim}{r}_i,\underset{\sim}{\lambda})$$

$$= -\int_0^\lambda \iint \left[\frac{n_{s+1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s,\underset{\sim}{y};\ \lambda')}{n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s;\ \lambda')} - n(\underset{\sim}{y};\lambda')\right] \beta v(\underset{\sim}{y},\underset{\sim}{x})d\lambda'\ d\underset{\sim}{y},$$

(8.3)

or interpreting in terms of (8.2),

$$\ln(g_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s;\ \lambda)/g_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s))$$

$$= -\int\int_0^\lambda n(\underset{\sim}{y})\Sigma_\Omega(-1)^{s-\Omega}\prod_{\Omega'\subset\Omega} g_{\Omega'+1}(\ldots\underset{\sim}{r}_j\ldots\underset{\sim}{y};\lambda')\beta v(\underset{\sim}{y},\underset{\sim}{x})d\lambda'\ d\underset{\sim}{y}, \tag{8.4}$$

for $s \geq 2$. Consider now the first case, $s = 1$, of (8.3):

$$\ln\Big(n(\underset{\sim}{r};\lambda)/n(\underset{\sim}{r})\Big) + \lambda v(\underset{\sim}{r},\underset{\sim}{x}) = -\int\int_0^\lambda n(\underset{\sim}{y};\lambda')\Big(g_2(\underset{\sim}{r},\underset{\sim}{y};\ \lambda') - 1\Big)\beta v(\underset{\sim}{y},\underset{\sim}{x})d\lambda'\ d\underset{\sim}{y}. \tag{8.5}$$

To close it, we need an expression for $g_2(\ ;\lambda')$, but this is precisely what is supplied by the $s = 2$ case of (8.4) with remainder dropped:

$$g_2(\underset{\sim}{r},\underset{\sim}{y};\ \lambda') = g_2(\underset{\sim}{r},\underset{\sim}{y}). \tag{8.6}$$

Equations (8.5), (8.6), together with the "boundary condition"

$$n(\underset{\sim}{r};1)/n(\underset{\sim}{r}) = g_2(\underset{\sim}{r},\underset{\sim}{x}) \tag{8.7}$$

then constitute the Kirkwood approximate integral equation. Equation (8.6), it may be noted, reduces on setting $\lambda' = 1$ to

$$g_2(\underset{\sim}{r},\underset{\sim}{y}|\underset{\sim}{x}) = g(\underset{\sim}{r},\underset{\sim}{y}) \tag{8.8}$$

or

$$g_3(\underset{\sim}{x},\underset{\sim}{r},\underset{\sim}{y}) = 1,$$

where $|\underset{\sim}{x})$ indicates a particle fixed at $\underset{\sim}{x}$, and hence (8.6) generalizes the three-body superposition approximation[33] which has already been used for the BGY approximation of Appendix C.

CHNC Extension

The characteristic feature of a closure of the above type is that the quantity to be determined remains a function of the "turning on" parameter, although the latter must ultimately be set equal to one. To avoid the disadvantage of retaining the coupling parameter, another type of superposition termination may be used. It is simply that necessary to annul the remainder term of some stage, and may itself be improved in higher approximation. A possible general procedure is readily described. Suppose that the approximation being considered is based on an expansion of $H(\gamma)$ in $h(\gamma)$, so that [(7.4)] exactly through p terms, is

$$H(\gamma) = \sum_0^p \frac{1}{j!} \int \cdots \int \prod_1^j \Big(h(\underset{\sim}{r}_i,\gamma) - h_o(\underset{\sim}{r}_i) \Big) \frac{\delta^j H(0)}{\delta h_o(\underset{\sim}{r}_1) \cdots \delta h_o(\underset{\sim}{r}_j)} \, d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_j$$

$$+ \frac{1}{p!} \int_0^\gamma \int \cdots \int \prod_1^{p+1} h(\underset{\sim}{r}_i,\lambda) \frac{\delta^j H(\lambda)}{\delta h(\underset{\sim}{r}_1,\lambda) \cdots \delta h(\underset{\sim}{r}_{p+1},\lambda)} \, d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_{p+1} d\lambda . \tag{8.9}$$

If (8.9) is functionally differentiated successively with respect to $h(\gamma)$ (or (8.9) used for the derivatives, with decreasing p) and γ set equal to one, we obtain the chain

$$\frac{\delta^l H(1)}{\delta h(\underset{\sim}{r}_1) \cdots \delta h(\underset{\sim}{r}_l)} = \sum_0^{p-l} \frac{1}{j!} \int \cdots \int \prod_{l+1}^{l+j} [h(\underset{\sim}{r}_i) - h_o(\underset{\sim}{r}_i)]$$

$$\times \frac{\delta^{l+j} H(0)}{\delta h_o(\underset{\sim}{r}_1) \cdots \delta h_o(\underset{\sim}{r}_{l+j})} \, d\underset{\sim}{r}_{l+1} \cdots d\underset{\sim}{r}_{l+j}$$

$$+ \frac{1}{p-l!} \int_0^1 (1-\lambda)^{p-l} \int \cdots \int \prod_{l+1}^{p+1} h(\underset{\sim}{r}_i,\lambda)$$

$$\times \frac{\delta^{p+1} H(\lambda)}{\delta h(\underset{\sim}{r}_1,\lambda) \cdots \delta h(\underset{\sim}{r}_{p+1},\lambda)} \, d\underset{\sim}{r}_{l+1} \cdots d\underset{\sim}{r}_{p+1} d\lambda \tag{8.10}$$

for $\ell = 0,1,\ldots,p$. Thus the "single approximation"

$$\delta^{p+1}H/\delta h(\underset{\sim}{r}_1)\ldots\delta h(\underset{\sim}{r}_{p+1}) = 0 \qquad (8.11)$$

for any h (including $h(\lambda)$) truncates every member of the chain and allows in principle solving for $\delta^p H/\delta h(\underset{\sim}{r}_1)\ldots\delta h(\underset{\sim}{r}_p)$, $\delta^{p-1}H/\delta h(\underset{\sim}{r}_1)\ldots\delta h(\underset{\sim}{r}_{p-1}),\ldots$ until H itself is found with p + 1-order error.

Let us examine the CHNC equation from this viewpoint.[34] It resulted from applying the external potential $U(\underset{\sim}{r}) = v(\underset{\sim}{r},\underset{\sim}{x})$ to $\ln(n(\underset{\sim}{r})e^{\beta U(\underset{\sim}{r})})$ initially at U = 0, (generalization to $U(\underset{\sim}{r}) = \Sigma v(\underset{\sim}{r},\underset{\sim}{x}_1)$ is trivial, but would produce the type of hierarchy we do not consider until the next section) and expanding in $n(\underset{\sim}{r})$. The required derivatives are then the higher order c_s defined by (5.12). As prototype, we have for p = 1 in (8.10),(8.11)

$$c_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3) = 0$$

implies

$$\ln g_2(\underset{\sim}{r}_1,\underset{\sim}{x}) + \beta v(\underset{\sim}{r}_1,\underset{\sim}{x}) = \int c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)n(\underset{\sim}{r}_2)\left(g_2(\underset{\sim}{r}_2,\underset{\sim}{x}) - 1\right)d\underset{\sim}{r}_2, \qquad (8.12)$$

the CHNC approximation. Choosing p = 2,

$$c_4(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3,\underset{\sim}{r}_4) = 0$$

implies

$$\ln g_2(\underset{\sim}{r}_1,\underset{\sim}{x}) + \beta v(\underset{\sim}{r}_1,\underset{\sim}{x}) = \int c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)n(\underset{\sim}{r}_2)\Big(g_2(\underset{\sim}{r}_2,\underset{\sim}{x}) - 1\Big)d\underset{\sim}{r}_2$$

$$+ \tfrac{1}{2}\iint c_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3)n(\underset{\sim}{r}_2)\Big(g_2(\underset{\sim}{r}_2,x) - 1\Big)n(\underset{\sim}{r}_3)$$

$$\times\Big(g(\underset{\sim}{r}_3,\underset{\sim}{x}) - 1\Big)d\underset{\sim}{r}_2 d\underset{\sim}{r}_3,$$

(8.13)

$$c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2|U) = c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) + \int c_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3)n(\underset{\sim}{r}_3)\Big(g_2(\underset{\sim}{r}_3,\underset{\sim}{x}) - 1\Big)d\underset{\sim}{r}_3$$

where

$$U(\underset{\sim}{r}) = v(\underset{\sim}{r},\underset{\sim}{x}),$$

and may be continued in obvious fashion.

It is of course necessary to evaluate the quantities $c_s(|U)$ which enter prominently in the sequence (8.12),(8.13),... . The process of (6.2),(6.3),(6.3′) does not go through cleanly in this case because the change in the independent variable $n(\underset{\sim}{r}|U) = n(\underset{\sim}{r}) + \Big(n_2(\underset{\sim}{r},\underset{\sim}{x}) - n(\underset{\sim}{r})n(\underset{\sim}{x})\Big)/n(\underset{\sim}{x})$ is no longer independent of the initial choice of $n(\underset{\sim}{r})$. In order to represent $c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s|U)$ explicitly but concisely, it is best therefore to express $c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)$ in terms of the n_s or combinations

thereof, and then turn on the external U. For this purpose, we shall introduce the full modified Ursell sequence $\hat{\mathfrak{F}}_s$ patterned after (7.17). The $\hat{\mathfrak{F}}_s$ are defined by

$$\hat{\mathfrak{F}}_s(\underset{\sim}{r}_1,\ldots\underset{\sim}{r}_s) = \delta^s \ln Z/\delta - \beta U(\underset{\sim}{r}_1)\ldots\delta - \beta U(\underset{\sim}{r}_s) \tag{8.14}$$

and hence satisfy

$$\delta\hat{\mathfrak{F}}_{s-1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)/\delta - \beta U(\underset{\sim}{r}_s) = \hat{\mathfrak{F}}_s(\underset{\sim}{r}_1\ldots\underset{\sim}{r}_s); \tag{8.15}$$

they bear the same relation to the modified $\hat{n}_s$ of (7.14) and (7.15) that the $\mathfrak{F}_s$ do to the n_s. Indeed they are also related quite directly to expectations, for by the discussion preceding (7.15) or from (8.10) itself, it follows that

$$\Sigma \frac{1}{s!} \int\cdots\int \hat{\mathfrak{F}}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)\omega(\underset{\sim}{r}_1)\ldots\omega(\underset{\sim}{r}_s)d\underset{\sim}{r}_1\ldots d\underset{\sim}{r}_s = \ln\langle \exp \Sigma\omega(\underset{\sim}{r}_i)\rangle. \tag{8.16}$$

The distribution $\hat{\mathfrak{F}}_2(\underset{\sim}{r},\underset{\sim}{r}') = \delta n(\underset{\sim}{r})/\delta - \beta U(\underset{\sim}{r}')$ always plays a dominant role in operations involving density differentiation. By means of its matrix inverse

$$\hat{\mathfrak{F}}_2^{-1}(\underset{\sim}{r},\underset{\sim}{r}') = \delta - \beta U(\underset{\sim}{r})/\delta n(\underset{\sim}{r}'), \tag{8.17}$$

we can in fact set up an iterative procedure for the sequence c_s:

$$c_{s+1}(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_{s+1}) = \delta c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)/\delta n(\underset{\sim}{r}_{s+1})$$

$$= \int \hat{\mathfrak{F}}_2^{-1}(\underset{\sim}{r}_{s+1},\underset{\sim}{y}) \frac{\delta}{\delta - \beta U(\underset{\sim}{y})} c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)d\underset{\sim}{y}. \tag{8.18}$$

It is only necessary to start the sequence off by observing, from (5.12) and (5.2), that

$$\begin{aligned} c_1(\underset{\sim}{r}_1) &= \ln n(\underset{\sim}{r}_1) + \beta U(\underset{\sim}{r}_1) \\ c_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2) &= \delta(\underset{\sim}{r}_1-\underset{\sim}{r}_2)/n(\underset{\sim}{r}_1) - \mathfrak{F}_2^{-1}(\underset{\sim}{r}_1,\underset{\sim}{r}_2), \end{aligned} \tag{8.19}$$

and continue it according to (8.18) by using

$$\frac{\delta}{\delta - \beta U(\underset{\sim}{r}'')} \hat{\mathfrak{F}}_2^{-1}(\underset{\sim}{r},\underset{\sim}{r}') = - \int \hat{\mathfrak{F}}_2^{-1}(\underset{\sim}{r},\underset{\sim}{y})\hat{\mathfrak{F}}_3(\underset{\sim}{y},\underset{\sim}{y}',\underset{\sim}{r}'')\hat{\mathfrak{F}}_2^{-1}(\underset{\sim}{y}',\underset{\sim}{r}')d\underset{\sim}{y}\, d\underset{\sim}{y}' \tag{8.20}$$

in conjunction with (8.15). For example, the next member of (8.19) is seen to be

$$c_3(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_3) = -\frac{\delta(\underset{\sim}{r}_1-\underset{\sim}{r}_2)\delta(\underset{\sim}{r}_1-\underset{\sim}{r}_3)}{n(\underset{\sim}{r}_1)^2} + \int \hat{\mathcal{F}}_2^{-1}(\underset{\sim}{r}_3,\underset{\sim}{x}_3)\hat{\mathcal{F}}_2^{-1}(\underset{\sim}{r}_2,\underset{\sim}{x}_2)$$

$$\times\hat{\mathcal{F}}_2^{-1}(\underset{\sim}{r}_1,\underset{\sim}{x}_1)\hat{\mathcal{F}}_3(\underset{\sim}{x}_1,\underset{\sim}{x}_2,\underset{\sim}{x}_3)d\underset{\sim}{x}_1d\underset{\sim}{x}_2d\underset{\sim}{x}_3. \tag{8.21}$$

Now one can fix the requisite number of particles in $c_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s|U)$ with little difficulty by applying this equation to the component $\hat{\mathcal{F}}_s$. The latter in turn may either be done by treating the n_s similarly, or alternatively from (6.3) and (8.14), by

$$\hat{\mathcal{F}}_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s|U) = \prod_1^s \left(e^{-\beta U(\underset{\sim}{r}_1)}\frac{\delta}{\delta_e - \beta U(\underset{\sim}{r}_1)}\right)\ln\left[n_t(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_t)e^{\beta\Sigma U(\underset{\sim}{x}_1)}\Xi[U]\right]\Bigg|_{U=0} \tag{8.22}$$

where

$$U(\underset{\sim}{r}) = \sum_1^t v(\underset{\sim}{r},\underset{\sim}{x}_1).$$

The notation $c_s(|U)$ in (8.13) is thus computationally meaningful.

It may be pointed out that the equation $c_3 = 0$ which terminates (8.12) is a true superposition principle, as are the further members of the series. This follows at once from (8.21) by applying three $\hat{\mathfrak{F}}_2$ matrix multiplications to cancel the three $\hat{\mathfrak{F}}_2^{-1}$, yielding the condition

$$\hat{\mathfrak{F}}_3(\chi_1,\chi_2,\chi_3) = \int \hat{\mathfrak{F}}_2(\underset{\sim}{r},\chi_1)\hat{\mathfrak{F}}_2(\underset{\sim}{r},\chi_2)\hat{\mathfrak{F}}_2(\underset{\sim}{r},\chi_3)d\underset{\sim}{r}/n(\underset{\sim}{r})^2. \tag{8.23}$$

Corrections to PY Equation

When a specific physical rationale is available for the expansions employed, completion to a hierarchy is not at all arbitrary. Thus at least for short range forces, the PY approximation is based upon the smooth transition of $ne^{\beta U}$ from its value of n outside the range of force, as one particle is being fixed. To pick up higher distributions required for corrections, one must clearly turn on several particles instead, but otherwise adopt the same strategy. In other words, we consider

$$n(\underset{\sim}{r})e^{\beta U(\underset{\sim}{r})} \quad \text{functional of} \quad n(\underset{\sim}{r}) \tag{8.24}$$

as

$$U(\underset{\sim}{r}) \to \sum_{1}^{s} v(\underset{\sim}{r},\underset{\sim}{x}_1).$$

Making use of (5.12) together with the fact that

$$\frac{\delta^t}{\delta n(\underset{\sim}{r}_1)\ldots\delta n(\underset{\sim}{r}_t)} e^F = e^F \prod_{1}^{t}\left(\frac{\delta}{\delta n(\underset{\sim}{r}_i)} + \frac{\delta F}{\delta n(\underset{\sim}{r}_i)}\right)\cdot 1, \qquad (8.25)$$

(where all operators are applied to the function 1) the corresponding expansion with remainder may be written as

$$g_{1,s}(\underset{\sim}{r},\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_s)e^{\beta\sum_{1}^{s} v(\underset{\sim}{r},\underset{\sim}{x}_1)} = 1 + \sum_{1}^{\ell}\frac{1}{j!}\int\ldots\int n(\underset{\sim}{y}_1)\ldots n(\underset{\sim}{y}_j)$$

$$\times\left(g_{1,s}(\underset{\sim}{y}_1,\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_s) - 1\right)\ldots\left(g_{1,s}(\underset{\sim}{y}_j,\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_s) - 1\right)\left(\frac{\delta}{\delta n(\underset{\sim}{y}_1)} + c(\underset{\sim}{r},\underset{\sim}{y}_1)\right)$$

$$\ldots\left(\frac{\delta}{\delta n(\underset{\sim}{y}_j)} + c(\underset{\sim}{r},\underset{\sim}{y}_j)\right)\cdot 1(d\underset{\sim}{y})^j + \frac{1}{\ell!}\int\ldots\iint_{0}^{1} n(\underset{\sim}{r};\lambda)e^{\beta U(\underset{\sim}{r};\lambda)} n(\underset{\sim}{y}_1)\ldots n(\underset{\sim}{y}_{\ell+1})$$

$$\times\left(g_{1,s}(\underset{\sim}{y}_1,\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_s) - 1\right)\ldots\left(g_{1,s}(\underset{\sim}{y}_{\ell+1},\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_s) - 1\right)$$

$$\times\left(\frac{\delta}{\delta n(\underset{\sim}{y}_1;\lambda)} + c(\underset{\sim}{r},\underset{\sim}{y}_1;\ \lambda)\right)\ldots\left(\frac{\delta}{\delta n(\underset{\sim}{y}_{\ell+1};\lambda)} + c(\underset{\sim}{r},\underset{\sim}{y}_{\ell+1};\ \lambda)\right)\cdot 1(1-\lambda)^{\ell}d\lambda(d\underset{\sim}{y})^{\ell+1}$$

(8.26)

where

$$g_{t,s}(\underset{\sim}{y}_1,\cdots,\underset{\sim}{y}_t,\underset{\sim}{x}_1,\cdots,\underset{\sim}{x}_s) \equiv \frac{n_{t+s}(\underset{\sim}{y}_1,\cdots,\underset{\sim}{y}_t,\underset{\sim}{x}_1,\cdots,\underset{\sim}{x}_s)}{n_t(\underset{\sim}{y}_1,\cdots,\underset{\sim}{y}_t)n_s(\underset{\sim}{x}_1,\cdots,\underset{\sim}{x}_s)} .$$

In the present case, the expansion coefficients are not symmetric in all of their variables, so that we can anticipate that the condition imposed by identical vanishing of the integrand of a remainder term will be essentially self-inconsistent. Instead, we can only demand the vanishing of the term following the last one we retain. Thus the PY equation of (6.11)-(6.13) results from $s = 1$, $\ell = 1$ of (8.26) if the first correction ($\ell = 2$) vanishes:

$$\int\cdots\int n(\underset{\sim}{y}_1)n(\underset{\sim}{y}_2)\left(g(\underset{\sim}{y}_1,\underset{\sim}{x}) - 1\right)\left(g(\underset{\sim}{y}_2,\underset{\sim}{x}) - 1\right) \times \left(c_3(\underset{\sim}{r},\underset{\sim}{y}_1,\underset{\sim}{y}_2) + c_2(\underset{\sim}{r},\underset{\sim}{y}_1)c_2(\underset{\sim}{r},\underset{\sim}{y}_2)\right)d\underset{\sim}{y}_1 d\underset{\sim}{y}_2 = 0 \tag{8.27}$$

which at least is only a one-variable assumption. More importantly, it appears explicitly in the next approximation, $s = 1$, $\ell = 2$ plus $s = 2$, $\ell = 1$:

$$g(\underset{\sim}{r},\underset{\sim}{x})e^{\beta v(\underset{\sim}{r},\underset{\sim}{x})}$$

$$= 1 + \Big(g(\underset{\sim}{r},\underset{\sim}{x}) - 1 - c(\underset{\sim}{r},\underset{\sim}{x})\Big) + \frac{1}{2}\int\cdots\int n(\underset{\sim}{y}_1)n(\underset{\sim}{y}_2)\Big(g(\underset{\sim}{y}_1,\underset{\sim}{x}) - 1\Big)\Big(g(\underset{\sim}{y}_2,\underset{\sim}{x}) - 1\Big)$$

$$\times\Big(c_3(\underset{\sim}{r},\underset{\sim}{y}_1,\underset{\sim}{y}_2) + c_2(\underset{\sim}{r},\underset{\sim}{y}_1)c_2(\underset{\sim}{r},\underset{\sim}{y}_2)\Big)d\underset{\sim}{y}_1 d\underset{\sim}{y}_2$$

(8.28a)

$$g_{1,2}(\underset{\sim}{r},\underset{\sim}{x}_1,\underset{\sim}{x}_2)e^{\beta v(\underset{\sim}{r},\underset{\sim}{x}_1) + \beta v(\underset{\sim}{r},\underset{\sim}{x}_2)} = 1 + \int n(\underset{\sim}{y})\Big(g_{1,2}(\underset{\sim}{y},\underset{\sim}{x}_1,\underset{\sim}{x}_2) - 1\Big)c(\underset{\sim}{r},\underset{\sim}{y})d\underset{\sim}{y}.$$

(8.28b)

At the p^{th} step in the correction series, one correspondingly makes use of the p-tuplet $s = 1,\ t = p;\ s = 2,\ t = p - 1;\ldots$ $s = p,\ t = 1$.

Although the solution of the (8.28a,b) is far from trivial, a brief qualitative observation as to the nature of the correction is in order. Using (8.21) and the relation

$$\int \Big(g(\underset{\sim}{r}_1,\underset{\sim}{r}_2) - 1\Big)n(\underset{\sim}{r}_2)\hat{\mathcal{J}}_2^{-1}(\underset{\sim}{r}_2,\underset{\sim}{r}_3)d\underset{\sim}{r}_2 = c(\underset{\sim}{r}_1,\underset{\sim}{r}_3)$$

(8.29)

as well as (5.6), (8.28a) can be rewritten as

$$g(\underset{\sim}{r},\underset{\sim}{x})e^{\beta v(\underset{\sim}{r},\underset{\sim}{x})} = g(\underset{\sim}{r},\underset{\sim}{x}) - c(\underset{\sim}{r},\underset{\sim}{x}) - c(\underset{\sim}{r},\underset{\sim}{x})\left(g(\underset{\sim}{r},\underset{\sim}{x}) - 1 - \tfrac{1}{2}c(\underset{\sim}{r},\underset{\sim}{x})\right)$$

$$+ \tfrac{1}{2}\int\cdots\int \hat{\mathfrak{J}}_2^{-1}(\underset{\sim}{r},\underset{\sim}{r}')\hat{\mathfrak{J}}_3(\underset{\sim}{r}',\underset{\sim}{y}_1,\underset{\sim}{y}_2)c(\underset{\sim}{y}_1,\underset{\sim}{x})c(\underset{\sim}{y}_2,\underset{\sim}{x})d\underset{\sim}{y}_1 d\underset{\sim}{y}_2 d\underset{\sim}{r}', \tag{8.30}$$

which is the reasonable form for simultaneous solution with (8.28b). For a quick estimate, however, let us reduce (8.30) by the CHNC condition $c_3 = 0$. It thus becomes

$$g(\underset{\sim}{r},\underset{\sim}{x})\left(e^{\beta v(\underset{\sim}{r},\underset{\sim}{x})} - 1\right) + c(\underset{\sim}{r},\underset{\sim}{x}) = \tfrac{1}{2}\left(g(\underset{\sim}{r},\underset{\sim}{x}) - 1 - c(\underset{\sim}{r},\underset{\sim}{x})\right)^2, \tag{8.31}$$

which avers that c, instead of vanishing outside the range of force, has a positive value. Consider the effect of this change on e.g., a hard sphere fluid. Since $\int(g-1) > 0$, the increase of the tail of c presumably provokes a second-order change or perhaps even a first-order increase in $g(\underset{\sim}{r}) - 1 - c(\underset{\sim}{r}) = n\int(g(\underset{\sim}{r}') - 1)c(\underset{\sim}{r}-\underset{\sim}{r}')d\underset{\sim}{r}'$. Hence g will increase over the full range. This means, according to (2.8), written as

$$\frac{\beta p}{n} = 1 + \tfrac{2}{3}\pi a^3 n g(a) \qquad (8.32)$$

for hard spheres, and (2.9), that the compressibility-derived pressure will come down and the virial pressure up when the correction is applied - in accord with numerical calculations.[35]

IX. THE LATTICE PY EQUATION

It has been observed on numerous occasions that the distributions and thermodynamic parameters computed from the PY approximation are almost uniformly superior to those achieved by other current simple approximation schemes. This is particularly fortunate, since the dominating characteristic of the PY equation - the finite range of the direct correlation function for finite range forces - permits explicit solutions to be obtained in several interesting cases. In these concluding sections, we shall examine a general instance of this type - the lattice gas with finite range forces - as well as the specific example of a hard core fluid.

Lattice Gas[36]

Suppose that the position of a particle in space is described not by a continuous vector $\underset{\sim}{x}$, but as a point on a finitely spaced mesh or lattice of lattice vectors

$$\underset{\sim}{\alpha} = \Sigma \alpha_h \underset{\sim}{e}_h \qquad (9.1)$$

where the α_h are integers and the $\underset{\sim}{e}_h$ lattice unit vectors and that a lattice site can be occupied by at most one particle. Thus the interparticle potential at zero spacing is always $+\infty$, and the density of particles defined in the obvious way can never exceed unity. The statistical mechanics of such a lattice gas of N particles is determined by its appropriate canonical configuration space distribution and partition function

$$\rho_N(\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_N) = \frac{v_o^N}{N!} \exp - \beta\Big(\frac{1}{2} \sum_{1\neq j} v(\underset{\sim}{\alpha}_i,\underset{\sim}{\alpha}_j) + U(\underset{\sim}{\alpha}_i)\Big)/Q_N$$

$$Q_N = \frac{v_o^N}{N!} \sum_{\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_N} \exp - \beta\Big(\frac{1}{2} \sum_{1\neq j} v(\underset{\sim}{\alpha}_i,\underset{\sim}{\alpha}_j) + U(\underset{\sim}{\alpha}_i)\Big), \tag{9.2}$$

where v_o is the volume of the lattice unit cell. One can obtain the normal continuum analog simply by the scaling $\underset{\sim}{e}_h \to \lambda \underset{\sim}{e}_h$ where $\lambda \to 0$.

Conversion of (9.2) to a grand ensemble proceeds automatically; defining the grand partition function

$$\Xi(z) = \Sigma z^N Q_N, \tag{9.3}$$

it is readily verified that

$$n_s(\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_s)e^{\beta\Sigma U(\underset{\sim}{\alpha}_i)}\Xi = \partial^s\Xi/\partial_e^{-\beta U(\underset{\sim}{\alpha}_1)}\ldots\partial^{-\beta U(\underset{\sim}{\alpha}_s)}, \quad (9.4)$$

where the distribution n_s is defined by the expectation

$$n_s\left(\underset{\sim}{\alpha}_1',\ldots,\underset{\sim}{\alpha}_s'\right) = \Big\langle \sum_{\underset{\sim}{\alpha}_1\neq\ldots\neq\underset{\sim}{\alpha}_s} \delta_{\underset{\sim}{\alpha}_1'\underset{\sim}{\alpha}_1}\ldots\delta_{\underset{\sim}{\alpha}_s'\underset{\sim}{\alpha}_s} \Big\rangle \quad (9.5)$$

using Kronecker δ-functions, and the N-particle contribution to the expectation is weighted by $z^N Q_N/\Xi$. Similarly, one has

$$\mathfrak{F}_s(\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_s)e^{\beta\Sigma U(\underset{\sim}{\alpha}_i)} = \partial^s \ln \Xi/\partial_e^{-\beta U(\underset{\sim}{\alpha}_1)}\ldots\partial_e^{-\beta U(\underset{\sim}{\alpha}_s)}$$

$$\hat{n}_s(\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_s)\Xi = \partial^s\Xi/\partial - \beta U(\underset{\sim}{\alpha}_1)\ldots\partial - \beta U(\underset{\sim}{\alpha}_s) \quad (9.6)$$

$$\hat{\mathfrak{F}}_s(\underset{\sim}{\alpha}_1,\ldots,\underset{\sim}{\alpha}_s) = \partial^s \ln \Xi/\partial - \beta U(\underset{\sim}{\alpha}_1)\ldots\partial - \beta U(\underset{\sim}{\alpha}_s),$$

the relations among the $n_s, \mathfrak{F}_s, \hat{n}_s, \hat{\mathfrak{F}}_s$ being as in the continuum case, with the replacement of Dirac by Kronecker δ-functions, e.g.,

$$\hat{\mathfrak{F}}_2(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2) = n_2(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2) - n(\underset{\sim}{\alpha}_1)n(\underset{\sim}{\alpha}_2) + n(\underset{\sim}{\alpha}_1)\delta_{\underset{\sim}{\alpha}_1 \underset{\sim}{\alpha}_2}. \quad (9.7)$$

The specification of a configuration by the position of each particle, $\{\underset{\sim}{\alpha}_i\} = (\underset{\sim}{\alpha}_1, \ldots, \underset{\sim}{\alpha}_N)$, is only one of several which are useful in lattice gases. For example, if $\nu_{\underset{\sim}{\alpha}} = 0,1$ denotes the number of particles at position $\underset{\sim}{\alpha}$, we have the equally valuable occupation number representation $\{\nu_{\underset{\sim}{\alpha}}\}$. Now (9.3), via (9.2), becomes

$$\Xi = \sum_{\{\nu_{\underset{\sim}{\alpha}}\}} \exp - \left[\frac{1}{2} \beta \sum_{\alpha_1 \neq \alpha_2} v(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2) \nu_{\underset{\sim}{\alpha}_1} \nu_{\underset{\sim}{\alpha}_2} + \Sigma(\beta U(\underset{\sim}{\alpha}) - \ln v_o z) \nu_{\underset{\sim}{\alpha}} \right] \quad (9.8)$$

and of course (9.4),(9.6) remain valid. In particular, one has the elementary

$$\hat{n}_s(\underset{\sim}{\alpha}_1, \ldots, \underset{\sim}{\alpha}_s) = \langle \nu_{\underset{\sim}{\alpha}_1} \nu_{\underset{\sim}{\alpha}_2} \cdots \nu_{\underset{\sim}{\alpha}_2} \rangle, \quad (9.9)$$

with the kernel of Ξ in (9.8) as unnormalized probability. The $\{\nu_{\underset{\sim}{\alpha}}\}$ representation is appropriate, e.g., for a binary alloy with the two species represented by 0 or 1. Equally suitable for this case, and particularly so for a spin $\frac{1}{2}$ system, is the spin representation in which

$$\mu_{\underset{\sim}{\alpha}} = 2\nu_{\underset{\sim}{\alpha}} - 1 \tag{9.10}$$

takes on the values ± 1. Dropping a μ-independent factor, (9.8) then transforms to

$$\Xi' = \sum_{\{\mu_{\underset{\sim}{\alpha}}\}} \exp - \left[\frac{1}{8}\beta \sum_{\alpha_1 \neq \alpha_2} v(\underset{\sim}{\alpha}_1, \underset{\sim}{\alpha}_2)\mu_{\underset{\sim}{\alpha}_1}\mu_{\underset{\sim}{\alpha}_2} + \frac{1}{2}\sum_{\underset{\sim}{\alpha}} \left(\frac{1}{2}\beta \sum_{\underset{\sim}{\alpha}' \neq \underset{\sim}{\alpha}} v(\underset{\sim}{\alpha}, \underset{\sim}{\alpha}') + \beta U(\underset{\sim}{\alpha}) - \ln v_o z \right)\mu_{\underset{\sim}{\alpha}} \right]. \tag{9.11}$$

In the special case of the Ising model, the composite external plus chemical potential $\frac{1}{2}\beta\Sigma v + \beta U - \ln v_o z$ vanishes, and the resulting symmetry in μ and $-\mu$ establishes that $\langle \mu_{\underset{\sim}{\alpha}} \rangle = 0$; alternatively the lattice gas density $n = \langle \nu_{\underset{\sim}{\alpha}} \rangle = \frac{1}{2}$. The chemical potential required to produce $n \neq \frac{1}{2}$ uniformly can then be interpreted as due to a uniform external magnetic field.

To derive the PY equation for a lattice gas, we proceed precisely as in the continuum case. A direct correlation is defined by

$$\partial(-\beta U(\underset{\sim}{\alpha}_1))/\partial n(\underset{\sim}{\alpha}_2) = \delta_{\underset{\sim}{\alpha}_1\underset{\sim}{\alpha}_2}/n(\underset{\sim}{\alpha}_2) - c(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) \qquad (9.12)$$

which according to (9.7) satisfies

$$g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) - 1 - c(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) = \sum_{\underset{\sim}{\alpha}_3}\Big(g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_3) - 1\Big)n(\underset{\sim}{\alpha}_3)c(\underset{\sim}{\alpha}_3,\underset{\sim}{\alpha}_2) \qquad (9.13)$$

where

$$g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) \equiv n_2(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)/n(\underset{\sim}{\alpha}_1)n(\underset{\sim}{\alpha}_2).$$

If the smooth function $n(\underset{\sim}{\alpha}_1)e^{\beta U(\underset{\sim}{\alpha}_1)}$ is then expanded as a function of its extrapolated tail $n(\underset{\sim}{\alpha}_1)$ as $U(\underset{\sim}{\alpha}_1)$ rises from 0 to $v(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)$, we obtain (see (8.28),(8.30))

$$g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)e^{\beta v(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)} = 1 + [g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) - 1 - c(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)]$$

$$+\Bigg[-c\Big(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2\Big)\Big(g(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2) - 1 - \tfrac{1}{2}\,c(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_2)\Big) + \tfrac{1}{2}\sum\hat{\mathfrak{F}}_2^{1}\Big(\underset{\sim}{\alpha}_1,\underset{\sim}{\alpha}_1'\Big)$$

$$\times c\Big(\underset{\sim}{\alpha}_2',\underset{\sim}{\alpha}_2\Big)c\Big(\underset{\sim}{\alpha}_3',\underset{\sim}{\alpha}_2\Big)\hat{\mathfrak{F}}_3\Big(\underset{\sim}{\alpha}_1',\underset{\sim}{\alpha}_2',\underset{\sim}{\alpha}_3'\Big)\Bigg] + \cdots .$$

Truncating at first order yields the PY approximation, which for a uniform system may be written as

$$f(\underset{\sim}{\alpha})g(\underset{\sim}{\alpha}) = (1 + f(\underset{\sim}{\alpha}))c(\underset{\sim}{\alpha})$$

where

$$f(\underset{\sim}{\alpha}) = e^{-\beta v(\underset{\sim}{\alpha})} - 1. \tag{9.15}$$

PY Solution[37]

Now from (9.15), if v vanishes after a finite range, $c(\underset{\sim}{\alpha})$ is required only for a finite number of points. This is the power of the method, for it means that only algebraic equations will have to be solved. Suppose for example that only nearest neighbor forces, with $f = e^{-\beta v} - 1$, are present in a simple D-dimensional cubic lattice of unit spacing; we must therefore solve the problem

$$g(0) = 0, \quad c(0) \neq 0, \quad fg(1) = (1+f)c(1); \quad c(\underset{\sim}{\alpha}) = 0 \quad \text{for} \quad |\underset{\sim}{\alpha}| > 1. \tag{9.16}$$

To do this, we apply a discrete Fourier transform to (9.13):

$$(g-1)_{\underset{\sim}{\theta}} = \frac{c_{\underset{\sim}{\theta}}}{1-nc_{\underset{\sim}{\theta}}} = -\frac{1}{n} + \frac{1}{n}\frac{1}{1-nc_{\underset{\sim}{\theta}}} \tag{9.17}$$

where

$$c_{\underset{\sim}{\theta}} \equiv \sum_{\underset{\sim}{\alpha}} e^{i\underset{\sim}{\alpha}\cdot\underset{\sim}{\theta}} c(\underset{\sim}{\alpha}).$$

From (9.16),

$$c_{\underset{\sim}{\theta}} = c(0) + 2c(1)\sum_{1}^{D} \cos\theta_j, \tag{9.18}$$

so that, reverse transforming g - 1 of (9.17) at $|\underset{\sim}{\alpha}| = 0,1$ and applying (9.16)

$$-1 = -\frac{1}{n} + \frac{1}{n}\oint\cdots\oint \left(1 - nc(0) - 2nc(1)\sum_{1}^{D}\cos\theta_j\right)^{-1} \prod d\theta_j/2\pi \tag{9.19a}$$

$$\frac{1+f}{f}c(1) - 1 = \frac{1}{n}\oint\cdots\oint e^{i\theta_1}\left(1 - nc(0) - 2nc(1)\sum_{1}^{D}\cos\theta_j\right)^{-1} \prod d\theta_j/2\pi, \tag{9.19b}$$

to be solved for c(0) and c(1), after which e.g., $\partial\beta p/\partial n = 1 - nc(0) - 2nDc(1)$ may be computed.

Let us simplify (9.19) somewhat. Multiplying (9.19a) by 1 - nc(0), (9.19b) by -2Dnc(1) and adding yields the direct relation

$$- (1-n)c(0) + 2Dc(1)\left(n - \frac{n}{f}(1+f)c(1)\right) = 1. \qquad (9.20)$$

On the other hand, transforming the denominator in (9.19a) by $1/\text{Den.} = \int_0^\infty e^{-s\,\text{Den.}}\,ds$ and carrying out the θ integrations,

$$1 - n = \int_0^\infty e^{-s(1-nc(0))}[I_0(2nsc(1))]^D ds. \qquad (9.21)$$

Numerical and asymptotic solutions of (9.20),(9.21) are feasible. For definiteness, we shall only examine the special case of the Ising model, $n = \frac{1}{2}$, in one dimension, D = 1. We then find from (9.20) and (9.21),

$$\text{PY:}\quad 1 + \tfrac{1}{2}c(0) + c(1)\left(-1 + \frac{1+f}{f}c(1)\right) = 0$$

$$c(0) = -2\left[\left(4 + c(1)^2\right)^{1/2} - 1\right]. \qquad (9.22)$$

This is particularly believable, because in the Ising model without external field, according to (9.11), all expectations are unchanged if each $\nu_{\underset{\sim}{\alpha}}$ is replaced by $1 - \nu_{\underset{\sim}{\alpha}}$ (particle interchanges with "hole"). Thus $\langle e^{\Sigma \nu_{\underset{\sim}{\alpha}} \omega_{\underset{\sim}{\alpha}}} \rangle = e^{\Sigma \omega_{\underset{\sim}{\alpha}}} \langle e^{-\Sigma \nu_{\underset{\sim}{\alpha}} \omega_{\underset{\sim}{\alpha}}} \rangle$, or taking the logarithm and applying (8.16), $\Sigma \hat{\mathfrak{F}}_s(\alpha_{\underset{\sim}{1}}, \ldots, \alpha_{\underset{\sim}{s}}) \omega_{\alpha_{\underset{\sim}{1}}} \ldots \omega_{\alpha_{\underset{\sim}{s}}} / s! = \Sigma \omega_{\underset{\sim}{\alpha}} + \Sigma \hat{\mathfrak{F}}_s(\alpha_{\underset{\sim}{1}}, \ldots, \alpha_{\underset{\sim}{s}})(-\omega_{\alpha_{\underset{\sim}{1}}}) \ldots (-\omega_{\alpha_{\underset{\sim}{s}}})/s!$, so that[38]

$$\hat{\mathfrak{F}}_{2m+1}(\alpha_{\underset{\sim}{1}}, \ldots, \alpha_{\underset{\sim}{2m+1}}) = 0. \tag{9.23}$$

The second-order correction in (9.14) hence does not affect the condition that $c = 0$ outside the range of force.

The second-order correction does affect the relation between c and g inside the range of force, so that a comparison with exact results is not out of place. Using (9.11) for the Ising model with V sites and periodic boundary conditions, we have

$$g(s) = \mathrm{Tr}[(1+\mu)E^s(1+\mu)E^{V-s}]/\mathrm{Tr}\ E^V \tag{9.24}$$

where

$$\mu = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad E = \begin{pmatrix} \varepsilon & \varepsilon^{-1} \\ \varepsilon^{-1} & \varepsilon \end{pmatrix}, \quad \varepsilon = e^{-\beta v/4}.$$

But $\mathrm{Tr}(\mu E^V) = 0$, $\mu E^s \mu = (\mu E \mu)^s = (2\epsilon - E)^s$ and the maximum eigenvalue of E is $\epsilon + \epsilon^{-1}$. It follows at once that as $V \to \infty$,

$$\text{exact:} \quad g(s) = 1 + \left(\frac{\epsilon - \epsilon^{-1}}{\epsilon + \epsilon^{-1}}\right)^s, \qquad s > 0$$
$$= 1 + \left(\frac{(1+f)^{\frac{1}{2}} - 1}{(1+f)^{\frac{1}{2}} + 1}\right)^s. \tag{9.25}$$

A brief calculation using (9.17) in the form $c_\theta = 2(g-1)_\theta/(2 + (g-1)_\theta)$ shows that c(s) is indeed of finite range for this one-dimensional Ising model, but that

$$c(0) = -2\left(\frac{2+f}{(1+f)^{\frac{1}{2}}} - 1\right)$$
$$c(1) = f/(1+f)^{\frac{1}{2}}. \tag{9.26}$$

The second relation of (9.22) is thus exact, as it must be, having been derived on the single assumption of finite range c. The first relation of (9.22) bears the stigma of the approximation, but this improves substantially if, with no additional effort, the full (9.14) is utilized.

X. THE HARD SPHERE PY EQUATION

Let us return to continuum fluids. The PY equation asserts a relation between the quantities

$$\tau = ge^{\beta v} \tag{10.1}$$

and g - c, as does the CHNC equation. For the case of pure hard cores, the character of τ is best seen by writing it as

$$\tau(\underset{\sim}{r}_1,\underset{\sim}{r}_2) = (z/n)^2 \Xi[U]/\Xi \tag{10.2}$$

from (6.3), where

$$U(\underset{\sim}{r}) = v(\underset{\sim}{r},\underset{\sim}{r}_1) + v(\underset{\sim}{r},\underset{\sim}{r}_2).$$

In this form, τ is determined by the grand partition function with an excluded volume of two (overlapping if $|\underset{\sim}{r}_1-\underset{\sim}{r}_2| < a$, the hard core diameter) spheres. It is thus continuous as a function of $\underset{\sim}{r}_1,\underset{\sim}{r}_2$, with several derivatives, but suffers a "primary" analytic break whenever the pair of holes can touch a new number of particles, as well as "secondary" breaks corresponding to topological changes in other shells. But g - c, written as

$$g(\underset{\sim}{r}_1,\underset{\sim}{r}_2) - c(\underset{\sim}{r}_1,\underset{\sim}{r}_2) = 1 + \int (g-1)(\underset{\sim}{r}_1,\underset{\sim}{x})n(\underset{\sim}{x})(g-1)(\underset{\sim}{x},\underset{\sim}{r}_2)d\underset{\sim}{x}$$

$$+ \iint (g-1)(\underset{\sim}{r}_1,\underset{\sim}{x}_1)n(\underset{\sim}{x}_1)(g-1)(\underset{\sim}{x}_1,\underset{\sim}{x}_2)n(\underset{\sim}{x}_2)(g-1)(\underset{\sim}{x}_2,\underset{\sim}{r}_2)d\underset{\sim}{x}_1 d\underset{\sim}{x}_2 + \ldots, \tag{10.3}$$

also suffers a primary break due to the spherical discontinuity of g - 1 whenever a new chain can just fit between $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_2$ in the first integral - equivalent to the condition on τ - and appropriate secondary breaks in the higher integrals.

Thus a direct functional relation between τ and g - c is suggested:

$$F(\tau) = g - c$$
$$g = (1+f)\tau \tag{10.4}$$

where $f(\underset{\sim}{r}) = e^{-\beta v_{core}(\underset{\sim}{r})} - 1$ is the Mayer f-bond:

$$f(\underset{\sim}{r}) = -1 \quad \text{for} \quad r \leq a$$
$$= 0 \quad \text{for} \quad r > a. \tag{10.5}$$

Now the solution of (10.4) can be expressed diagrammatically[39] using f-bonds. For small r, these bonds permit only diagrams of small diameter. Hence, if it develops that the difference

between, say τ of (10.4) and the exact τ can be written so as to have one or more links $1 + f$ - insisting on a large diameter - τ would be represented with small error inside the core. This requires checking that exact and approximate τ coincide when f becomes the function -1, and perhaps some functional derivatives as well. If (10.4) is combined with the definition of τ to read

$$F(\tau(\underset{\sim}{r})) = 1 + n\int F(\tau(\underset{\sim}{r}'))d\underset{\sim}{r}' - n\int\Big(1 + f(\underset{\sim}{r}-\underset{\sim}{r}')\Big)(\underset{\sim}{r}-\underset{\sim}{r}')\Big(1 + F(\tau(\underset{\sim}{r}'))\Big)d\underset{\sim}{r}'$$

$$+ n\int\Big(1 + f(\underset{\sim}{r}-\underset{\sim}{r}')\Big)\Big(1 + f(\underset{\sim}{r}')\Big)\tau(\underset{\sim}{r}-\underset{\sim}{r}')\tau(\underset{\sim}{r}')d\underset{\sim}{r}', \tag{10.6}$$

then by taking successive f-derivatives and evaluating at $f = -1$, one finds for system volume V, at $f = -1$,

$$F(\tau(r)) = \frac{1}{1-nV}, \tag{10.7}$$

$$F'(\tau(r))\,\frac{\delta\tau(\underset{\sim}{r})}{\delta f(\underset{\sim}{r}')} = -\,n\tau(\underset{\sim}{r})\,\frac{2-nV}{(1-nV)^2},\ldots$$

On the other hand, for the exact system in a grand canonical ensemble, we know that (see (2.13))

$$n = z + z^2\int (1 + f(\underset{\sim}{r}))d\underset{\sim}{r} + \ldots/1 + zV + \tfrac{1}{2}z^2V\int (1 + f(\underset{\sim}{r}))d\underset{\sim}{r} + \ldots \tag{10.8}$$

which yields in precisely the same way at f = -1,

$$z = \frac{n}{1-nV}, \qquad \frac{\delta z}{\delta f(\underset{\sim}{r})} = -\frac{z^2}{2}\frac{2-nV}{1-nV},\ldots . \tag{10.9}$$

Since

$$\tau(\underset{\sim}{r}) = (z^2+\ldots)/n^2\Big(1 + zV + \tfrac{1}{2}z^2V\int (1 + f(\underset{\sim}{r}'))d\underset{\sim}{r}' + \ldots\Big), \tag{10.10}$$

it follows that at f = -1,

$$\tau(\underset{\sim}{r}) = \frac{1}{1-nV}, \qquad \frac{\delta\tau(\underset{\sim}{r})}{\delta f(\underset{\sim}{r}')} = -n\,\frac{2 - \frac{3}{2}\,nV}{(1-nV)^3},\ldots . \tag{10.11}$$

We conclude that there will be at least one 1 + f bond in the error term of τ only if $F(\tau) = \tau$, the PY equation, which should therefore be particularly suitable inside the core. There will not be two 1 + f's in general in the error in τ, but since

$$\frac{\partial n}{\partial \beta p} = 1 + n \int [\tau(\underset{\sim}{r})(1 + f(\underset{\sim}{r})) - 1] d\underset{\sim}{r}, \qquad (10.12)$$

there will be in the error for the virial coefficients of the equation of state.

Low Density Limit

In attempting to solve the PY hard core equation from scratch, it is advisable to work from limiting cases; since temperature is irrelevant for hard cores, low and high density are the limits available. At low density, -c, which continues g into the hard core interior, would be expected to be smooth and hence representable by a low order polynomial; in fact, a direct virial expansion (power series in n) to the first few orders uncovers only a polynomial for c in its non-vanishing domain, of degree one or three for the case of one or three dimensional hard cores. Assuming the adequacy of a low order polynomial, its coefficients can be obtained with a minimum of work.

We first observe that although both g and c have finite discontinuities at the hard core surface, nonetheless from the representation

$$g(\underset{\sim}{r}) - c(\underset{\sim}{r}) = 1 - n\int c(\underset{\sim}{r}')d\underset{\sim}{r}' + n\int g(|\underset{\sim}{r}-\underset{\sim}{r}'|)c(\underset{\sim}{r}')d\underset{\sim}{r}' \tag{10.13}$$

g - c will be continuous at r = a with its first D derivatives, for dimension D. Next, going over to the hard core PY equation in the form

$$\begin{aligned} g(r) &= 0 \quad r \leq a \\ c(r) &= 0 \quad r > a, \end{aligned} \tag{10.14}$$

(10.13) is evaluated near the origin r = 0 by expanding g and c in the integral about r' = a, since only this region contributes. We find for one and three dimensions respectively (using for the latter case the general convolution relation

$$\int \mu(|\underset{\sim}{r}-\underset{\sim}{r}'|)\nu(\underset{\sim}{r}')d\underset{\sim}{r}' = \frac{2\pi}{r}\iint_{\text{strip}} r''\mu(\underset{\sim}{r}'')r'\nu(r')dr'\,dr'' \tag{10.15}$$

integrated over the strip $|r'-r''| \leq r \leq |r'+r''|$, $r' \geq 0$, $r'' \geq 0$)

$$\text{1d:}\quad -c(r) = 1 - n\int_{-a}^{a} c(r')dr' + nrg(a)c(a) + \tfrac{1}{2}nr^2(g'(a)c(a) - g(a)c'(a)) + \ldots \tag{10.16}$$

$$\begin{aligned}\text{3d:}\quad -c(r) = 1 - n\int c(r')dr' &+ 2\pi n\Big\{\tfrac{1}{2}r[ag(a)ac(a)] + \tfrac{1}{6}r^2[(ag(a))'ac(a) - ag(a)(ac(a))'] \\ &+ \tfrac{1}{24}r^3[(ag(a))''ac(a) - (ag(a))'(ac(a))' + ag(a)(ac(a))''] \\ &+ \tfrac{1}{120}r^4[(ag(a))'''ac(a) - 2(ag(a))''(ac(a))' \\ &+ 2(ag(a))'(ac(a))'' - ag(a)(ac(a))'''] + \ldots\Big\}\end{aligned} \tag{10.17}$$

Hence differentiating at the origin and using the equality of derivatives of g and -c at r = a,

$$\text{1d:}\quad -c(0) = 1 - n\int_{-a}^{a} c(r')dr', \quad -c'(0) = -nc(a)^2, \quad -c''(0) = 0 \tag{10.18}$$

$$\text{3d:}\quad -c(0) = 1 - n\int c(r')dr', \quad -c'(0) = -\pi n a^2 c(a)^2, \quad -c''(0) = 0, \tag{10.19}$$

$$-c'''(0) = \tfrac{1}{12}\, n\pi\left[a^2(c'(a)^2 - 2c(a)c''(a)) + 2ac(a)c'(a) + c(a)^2\right],$$

$$-c''''(0) = 0.$$

Excluding a spurious solution $c \propto r$ in each case, these are easily solved to yield the polynomials[40]

$$\text{1d:}\quad c(r) = -\frac{1}{(1-na)^2} + \frac{n}{(1-na)^2}\, r$$

$$\text{3d:}\quad c(r) = -\frac{(1+2\eta)^2}{(1-\eta)^4} + \frac{6\eta(1+\frac{1}{2}\eta)^2}{(1-\eta)^4}\,\frac{r}{a} - \frac{\eta(1+2\eta)^2}{2(1-\eta)^4}\left(\frac{r}{a}\right)^3 \tag{10.20}$$

$$\eta = \tfrac{1}{6}\,\pi n a^3$$

which it may be verified, cause g to vanish identically inside the core. Thus (10.20) solves the PY equation exactly. There is a bonus in the one dimensional case where the exact c can

be shown[41] to have precisely the range a, so that c(r) of (10.20) is not an approximation at all. Finally from $\partial\beta p/\partial n = 1 - n\int c(r)d\underset{\sim}{r} = -c(0)$, there follow the PY equations of state

$$\begin{aligned} &1d: \quad \beta p = n/1 - na \\ &3d: \quad \beta p = n(1+\eta+\eta^2)/(1-\eta)^3, \end{aligned} \tag{10.21}$$

identical with the RFL equations of state.[30]

High Density, One Dimension

There is no reason to believe that the solution of the hard core PY equation is given uniquely by its low density form. Indeed (10.20),(10.21) allow super-close packing in three dimensions, $n = [4\pi/3(a/2)^3]^{-1}$ compared with a physical maximum of $n = [4\sqrt{2}(a/2)^3]^{-1}$. This might be in the nature of the approximation, but might also indicate the presence of another solution of lower free energy, valid in the high density region. The PY equation in this region has not yet been exactly solved for two or three dimensions, but the major considerations involved are clearly evident in the fairly trivial one dimensional case, which we therefore examine in some detail.

In one dimension near close packing, $n \sim 1/a$, we know that the correct distribution must be a set of δ-functions:

$$g(r) = \sum_{m\neq 0} \frac{1}{n}\,\delta(r-ma). \tag{10.22}$$

Hence the Fourier transform $[\delta(r) + n(g(r) - 1)]_k = \Sigma e^{imka}$ for $k \neq 0$ becomes

$$[\delta(r) + n(g(r) - 1)]_k = \frac{2\pi}{a} \sum_j \delta(k-K_j) \qquad (10.23)$$

where

$$K_j = (2\pi/a)j$$

Thus the reciprocal $1 - nc_k$ of (10.23) at close packing must become infinite at all k except for the K_j, at each of which it approaches a zero - a double zero since $1 - nc_k$ is nonnegative - whose reciprocal area is $2\pi/a$. This is meaningful, since we know that $1 - n\int c(r)dr = -c(0) = \partial\beta p/\partial n$ should go to infinity at close packing; here then $1 - nc_k$ is a function with isolated near-zeros at the K_j for $j \neq 0$ and a very large factor multiplying it.

To secure the δ-function normalization, we must examine the structure of $1 - nc_k$ more carefully. Suppose that γ is a parameter of largeness associated with the inverse compressibility $n\partial\beta p/\partial n$, and we expand $1 - nc_k$ about the point K_j in a series in $\Delta = k - K_j$. We can assume that the leading term goes as $\gamma^s\Delta^2$, and hence without loss of generality that to second order in γ and Δ

$$1 - nc_k = \gamma^s a_2\Delta^2 + \gamma^{s-1}(b_o+b_1\Delta) + \gamma^{s-2}(c_o+c_1\Delta+c_2\Delta^2), \qquad (10.24)$$

where the a_i, b_i, c_i are suitable constants.

The normalization $\int dk/1 - nc_k = 2\pi/a$ then leads immediately to $4(a_2\gamma^s+c_2\gamma^{s-2})(b_o\gamma^{s-1}+c_o\gamma^{s-2}) - (b_1\gamma^{s-1}+c_1\gamma^{s-2})^2 = a^2$, which as an identity in γ requires $s = 2$ and simplifies (10.24) to the form

$$1 - nc_k = (\gamma\alpha\Delta+\delta)^2 + (a^2/4\delta^2)\Delta^2. \qquad (10.25)$$

Equation (10.25), which is quite general, becomes very restrictive when taken with the PY assumption that c vanishes outside r = a. That assumption implies that

$$1 - nc_k = A_o(k) + A_+(k)e^{ika} + A_-(k)e^{-ika} \qquad (10.26)$$

where the A(k) are regular, except at k = 0, where $1 - nc_k$ is regular. In fact one can show directly by integration that

$$c_k = 2\mathcal{R}\left[e^{ika}c\left(a + \frac{1}{i}\frac{d}{dk}\right)\frac{1}{ik} - c\left(\frac{1}{i}\frac{d}{dk}\right)\frac{1}{ik}\right]. \qquad (10.27)$$

Now according to (10.25),

$$2(nc_k-1)nc_k'' = n^2\left(c_k'\right)^2 + a^2 \qquad (10.28)$$

for large γ at $k = K_j$, $j \neq 0$. With (10.27), this leads directly to the form

$$1 - nc_k = \left(\gamma\frac{\sin ka/2}{ka/2} + \cos ka/2\right)^2 + \sin^2 ka/2. \qquad (10.29)$$

To relate γ to the density n, we see that the maximum of g_k now occurs at $\Delta = -K_J/\gamma$ or $k = K_J(1 - 1/\gamma)$, so that $n = (1/a)(1 - 1/\gamma)$ or

$$\gamma = 1/1 - na. \tag{10.30}$$

The equation of state then arises as usual from $\partial\beta p/\partial n = 1 - nc_o = (1+\gamma)^2$.

It so happens that (10.29),(10.30) is exact, and in fact identical with the low density solution (a one-dimension phase transition should not occur). In general however, one would have to extend (10.29) to lower γ, making explicit use of the vanishing of g inside the hard core. In fact, one even knows why g vanishes inside the hard core: this is because the close packing δ-function at $r = ma$ spreads at lower density to a regular function $g^{(m)}$ - the m^{th} neighbor distribution - which starts at $r = ma$ and goes to ∞ with area $1/n$. But it is easy to see that the Fourier transform for $m > 0$ must look like

$$g_k^{(m)} = G_m(k)e^{imka} = g_{-k}^{(m)} \tag{10.31}$$

where $G_m(k)$ is regular in the upper half complex k-plane. Thus in the expansion at $k \neq 0$

$$\Big(\delta(r) + n(g-1)\Big)_k = 1 + \sum_1^{\infty} G_m(k)e^{imka} + \sum_{-1}^{-\infty} G_m(-k)e^{imka} \tag{10.32}$$

the coefficient of the 0-frequency exponential must be precisely unity. It is trivial to establish that

$$\text{coef } e^{ik0}(A_0+A_+e^{ika}+A_-e^{-ika})^{-1} = \left(A_0^2 - 4A_+A_-\right)^{-1/2} \qquad (10.33)$$

is unity for (10.29), and the higher Fourier terms also have the correct analytic properties.

High Density, Dimension D > 1

The close packing configuration in two or three dimensional space depends upon the crystal type which is attained, e.g., cubic or hexagonal close packed in three dimensions, and only free energy considerations can tell which will dominate. At any rate, suppose the close packed lattice, anchored at the origin, to be given by the points $\underset{\sim}{R}_m$, so that

$$g(\underset{\sim}{r}) = \sum_{\underset{\sim}{m}\neq 0} \frac{1}{n}\,\delta(\underset{\sim}{r}-\underset{\sim}{R}_m). \qquad (10.34)$$

There is and will be the explicit assumption that only a vanishing energy per particle is required to orient the system so that it maintains the lattice symmetry. Now the Fourier transform for $\underset{\sim}{k} \neq 0$ becomes

$$[\delta(\underset{\sim}{r}) + n(g(\underset{\sim}{r}) - 1)]_{\underset{\sim}{k}} = \frac{(2\pi)^D}{v_0}\,\Sigma\delta(\underset{\sim}{k}-\underset{\sim}{K}_j) \qquad (10.35)$$

where the $\underset{\sim}{K}_{\underset{\sim}{l}}$ are the Bragg vectors of the lattice, and v_o the volume per particle. The first condition on $1 - nc_{\underset{\sim}{k}}$ at close packing is then that it vanish at each $\underset{\sim}{K}_{\underset{\sim}{l}}$ except $\underset{\sim}{l} = 0$, and since $c_{\underset{\sim}{k}}$ overwhelms 1, this is a condition on c itself. But suppose $\bar{c}$ is the lattice translated sum of c:

$$\bar{c}(r) = \sum_{\underset{\sim}{m}} c(\underset{\sim}{r}-\underset{\sim}{R}_{\underset{\sim}{m}}); \qquad (10.36)$$

$\bar{c}_{\underset{\sim}{k}}$ exists only at $\underset{\sim}{k} = \underset{\sim}{K}_{\underset{\sim}{l}}$ where it is proportional to $c_{\underset{\sim}{k}}$. Hence the leading condition is that

$$\bar{c}(\underset{\sim}{r}) = \text{constant}.$$

Proceeding to δ-function normalization, since the distributions for a regular lattice will certainly maintain lattice symmetry normal to the radius vector, and since the maximum of $g_{\underset{\sim}{k}}$ will move out along the radius vector, (10.25) readily generalizes to the form

$$1 - nc_{\underset{\sim}{k}} = \gamma^{2(D-1)}\left[\left(\gamma\alpha_o\Delta_o+\delta\right)^2 + b_o\Delta_o^2 + (\gamma\alpha+\beta)^2\Delta^2\right] \qquad (10.37)$$

where Δ_o refers to radial direction, Δ to normal. The relation of the constants is easily obtained, but irrespective of this, one result is strongly suggested; since it appears that now $\partial\beta p/\partial n \sim \gamma^{2D}$ while again the close packing density deviation $n - n_{cp} \sim 1/\gamma$, we surmise that unless the leading term is higher order in γ,

$$\partial\beta p/\partial n \sim (n-n_{cp})^{-2D}. \qquad (10.38)$$

For further information, the detailed nature of the PY approximation must be taken into account. The finite range of c means, together with (10.37), that at close packing, $c(\underset{\sim}{r})$ will consist of a constant inside the second Brillouin zone together with a portion antisymmetric about the first zone. At lower density, the nearest neighbor singularities remain and determine the Fourier form. On the other hand, the point singularities of g will expand outwardly to type $\underset{\sim}{m}$ neighbor distributions with characteristic exponential Fourier form. If each particle moves in its kinematically possible region, this distribution has a scalloped boundary, but if the approximation demands collective movements, the boundary becomes planar. In any event, the absence of the first zone g then determines the full distribution, a matter which is now under investigation.

A final remark is in order. Once the correct spherical c and g are obtained, sphericalization induced physically by energy degeneracy will not alter the thermodynamic parameters since their defining integrals are linear and include all orientations. However, the direct correlation function c derived from the sphericalized g suffers a qualitative change: it is no longer of finite range, even in the PY approximation.

Acknowledgments

The writer would like to thank Drs. J. L. Lebowitz, G. Stell, M. S. Wertheim, and G. J. Yevick for numerous valuable discussions, and to acknowledge the hospitality of the Aspen Physics Institute during the writing of this paper.

APPENDIX A

THE PY EQUATION

Percus and Yevick have derived[7] an approximate integral equation for the pair distribution function in a fluid by using self-consistently the assumption of independent propagation of longitudinal sound waves, i.e., waves of density fluctuation. From the viewpoint of the present paper, their argument may be compressed as follows.

On a microscopic basis, a density fluctuation cannot be small, for it is due to the arrival or departure of a particle. Let us however consider a fluid or continuum picture in which all fluctuations are regarded as weak. We may then ask for that potential $v^*(\underset{\sim}{r})$ whose effect at the origin on a linear basis is to produce the same forced density fluctuation $ng(\underset{\sim}{r}) - n$ which the true interparticle potential $v(\underset{\sim}{r})$ due to the arrival of a particle would produce. Since

$$\frac{\delta\beta U(\underset{\sim}{r})}{\delta n(\underset{\sim}{r}')} = c(\underset{\sim}{r}-\underset{\sim}{r}') - \frac{\delta(\underset{\sim}{r}-\underset{\sim}{r}')}{n}, \qquad (A.1)$$

we have $\beta v^*(\underset{\sim}{r}) = \int (c(\underset{\sim}{r}-\underset{\sim}{r}') - \delta(\underset{\sim}{r}-\underset{\sim}{r}')/n)n(g(\underset{\sim}{r}') - 1)d\underset{\sim}{r}'$, and hence conclude that

$$\beta v^*(\underset{\sim}{r}) = - c(\underset{\sim}{r}). \qquad (A.2)$$

Next we construct a model in which microscopic fluctuations can be arbitrarily small. Suppose we have a special fluid expressed in Lagrangian coordinates $\underset{\sim}{x}(\underset{\sim}{x}_o,t)$, the position at time t of a volume element which started at $\underset{\sim}{x}_o$ at time 0. The fluid is to be described by the equation of motion

$$\ddot{\underset{\sim}{x}} + (1/m)\int \nabla v\left(\underset{\sim}{x} - \underset{\sim}{x}(\underset{\sim}{x}_o',t)\right)\rho_o\left(\underset{\sim}{x}_o'\right)d\underset{\sim}{x}_o' + (1/m)\nabla U(\underset{\sim}{x}) = 0, \tag{A.3}$$

with mass conservation given by the definition

$$\rho(\underset{\sim}{x}) = \rho_o(\underset{\sim}{x}_o)/|\partial\underset{\sim}{x}/\partial\underset{\sim}{x}_o|. \tag{A.4}$$

Now if the initial density $\rho_o(\underset{\sim}{x}_o')$ is a sum of δ-functions $\sum_i \delta\left(\underset{\sim}{x}_o' - \underset{\sim}{x}_{io}\right)$ and $\underset{\sim}{x}$ in (A.3) is initially any $\underset{\sim}{x}_{jo}$, (A.3) together with initial velocities precisely describes the motion of a particle system with internal potential v and external U; the density $\rho(\underset{\sim}{x})$ remains a set of δ-functions. To generalize to a continuum or fluid, we need only drop the special form of ρ_o. It is then convenient to transform to a Eulerian frame, using the variables $(\underset{\sim}{x},t)$ rather than $(\underset{\sim}{x}_o,t)$. Defining

$$\underset{\sim}{u} = \dot{\underset{\sim}{x}}, \tag{A.5}$$

it is clear that

$$\frac{\partial}{\partial t}\Big|_{\underset{\sim}{x}_o} = \frac{\partial}{\partial t}\Big|_{\underset{\sim}{x}} + \underset{\sim}{u}\cdot\nabla|_t, \qquad \nabla_o|_t = \nabla|_t. \tag{A.6}$$

Thus (A.3) yields[41]

$$\frac{\partial \underset{\sim}{u}}{\partial t} + \underset{\sim}{u}\cdot\nabla\underset{\sim}{u} + \frac{1}{m}\int \nabla v(\underset{\sim}{x}-\underset{\sim}{x}')\rho(\underset{\sim}{x}')d\underset{\sim}{x}' + \frac{1}{m}\nabla U(\underset{\sim}{x}) = 0, \tag{A.7}$$

and from the time derivative of (A.4), $0 = \partial\rho(\underset{\sim}{x})|\partial\underset{\sim}{x}/\partial\underset{\sim}{x}_o|/\partial t\Big|_{\underset{\sim}{x}_o}$ $= |\partial\underset{\sim}{x}/\partial\underset{\sim}{x}_o|(\partial\rho/\partial t|_x + \underset{\sim}{u}\cdot\nabla\rho) + \rho\partial|\partial\underset{\sim}{x}/\partial\underset{\sim}{x}_o|/\partial t\Big|_{\underset{\sim}{x}_o}$, we have

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\underset{\sim}{u}\rho = 0. \tag{A.8}$$

If U = 0, (A.7) and (A.8) as microscopic equations are satisfied by $\rho = \bar{\rho}$, a constant, and $\underset{\sim}{u} = 0$. If U is small and $\rho - \bar{\rho} \equiv \delta\rho$, (A.7) and (A.8) become to first order

$$\frac{\partial \underset{\sim}{u}}{\partial t} + \frac{1}{m}\int \nabla v(\underset{\sim}{x}-\underset{\sim}{x}')\delta\rho(\underset{\sim}{x}')d\underset{\sim}{x}' + \frac{1}{m}\nabla U(\underset{\sim}{x}) = 0$$

$$\frac{\partial}{\partial t}\delta\rho + \bar{\rho}\nabla\cdot\underset{\sim}{u} = 0, \tag{A.9}$$

so that

$$\frac{\partial^2}{\partial t^2}\delta\rho = \frac{\bar{\rho}}{m}\int \nabla^2 v(\underset{\sim}{x}-\underset{\sim}{x}')\delta\rho(\underset{\sim}{x}')d\underset{\sim}{x}' + \frac{\bar{\rho}}{m}\nabla^2 U(\underset{\sim}{x}). \tag{A.10}$$

The steady-state response to an external $U(\underset{\sim}{x}) = V^{-1}\sum U_{\underset{\sim}{k}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{x}}$ is thus given by Fourier transformation as

$$U_{\underset{\sim}{k}} = -v_{\underset{\sim}{k}}\delta\rho_{\underset{\sim}{k}}. \tag{A.11}$$

When higher order terms are taken into account, at most higher harmonics of $U_{\underset{\sim}{k}}$ are required in (A.9) to produce a pure sinusoidal density change at wave vector $\underset{\sim}{k}$. Now the $\underset{\sim}{k}^{th}$ component of the radial distribution may be obtained as the $\underset{\sim}{k}^{th}$ component of the density at particle 1 due to a small sinusoidal impressed force at particle 2:

$$g_{\underset{\sim}{k}} = \lim_{\alpha\to 0} V\langle e^{i\underset{\sim}{k}\cdot\underset{\sim}{r}_1} e^{2\beta\gamma\cos \underset{\sim}{k}\cdot\underset{\sim}{r}_2}\rangle / \beta\gamma\langle e^{2\beta\gamma\cos \underset{\sim}{k}\cdot\underset{\sim}{r}_2}\rangle$$

$$= V\langle e^{i\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\rangle. \tag{A.12}$$

Accordingly, our discussion of (A.7, A.8) suggests the approximation that in computing (A.12), $v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)$ be replaced by the $\underset{\sim}{k}$-harmonic components of its continuum equivalent v^*:

$$g_{\underset{\sim}{k}} = V\langle e^{i\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)} \exp -\beta\left[\frac{1}{V}\sum_{s} v^*_{s\underset{\sim}{k}} e^{-is\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)} - v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)\right]\rangle. \tag{A.13}$$

Equating (A.12) and (A.13), we have

$$V\left\langle e^{i\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\left[e^{\beta v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\exp\left\{-\frac{\beta}{V}\sum v^*_{\underset{\sim}{r}\underset{\sim}{k}}e^{-is\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\right\}-1\right]\right\rangle = 0. \tag{A.14}$$

But as the system volume $V \to \infty$, $\left\langle e^{\beta v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\right\rangle \to 1$ (since $g \to 1$ asymptotically), whereas $\left\langle e^{it\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}e^{\beta v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\right\rangle$ $= (1/V)\int g(\underset{\sim}{r})e^{\beta v(\underset{\sim}{r})}e^{it\underset{\sim}{k}\cdot\underset{\sim}{r}}d\underset{\sim}{r} \to 0$. Thus if (A.14) converges uniformly, only the terms

$$V\left\langle e^{i\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\left[e^{\beta v(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\left\{1-\frac{\beta}{V}v^*_{\underset{\sim}{k}}e^{-i\underset{\sim}{k}\cdot(\underset{\sim}{r}_1-\underset{\sim}{r}_2)}\right\}\right]-1\right\rangle = 0 \tag{A.15}$$

remain, and these may be rewritten as

$$\left(e^{\beta v(\underset{\sim}{r})}g(\underset{\sim}{r})\right)_{\underset{\sim}{k}} - \beta v^*_{\underset{\sim}{k}} - g_{\underset{\sim}{k}} = 0, \tag{A.16}$$

or extending to $\underset{\sim}{k} = 0$ by continuity,

$$\left(e^{\beta v(\underset{\sim}{r})}-1\right)g(\underset{\sim}{r}) = \beta v^*(\underset{\sim}{r}) \tag{A.17}$$

which is the PY approximation.

APPENDIX B

CHNC AND PY EQUATIONS[21,7]

Several approximate integral equations for fluids may be based upon partial summation of the Mayer diagrammatic representation of the grand partition function and associated distributions. For a uniform system in fugacity expansion, we recall that[39]

$$\Xi = \sum \frac{z^N}{N!} \int \cdots \int e^{-\frac{1}{2}\beta \Sigma' v(\underset{\sim}{r}_i - \underset{\sim}{r}_j)} d\underset{\sim}{r}_1 \cdots d\underset{\sim}{r}_N \qquad (B.1)$$

$$= \sum \{\text{All distinct numbered diagrams with links } f_{ij} \equiv e^{-\beta v(\underset{\sim}{r}_i - \underset{\sim}{r}_j)} - 1, \text{ with z and integration at each vertex, and weighted by 1/N!, where N is the number of vertices}\}. \qquad (B.2)$$

$$= \exp \sum \{\text{All distinct connected numbered diagrams with links } f_{ij}, \text{ with z and integration at each vertex, and weighted by 1/N!}\}. \qquad (B.3)$$

It then follows that

$$n_2(\underset{\sim}{r}_1 - \underset{\sim}{r}_2) = \delta \ln \Xi / \delta(-\beta v(\underset{\sim}{r}_1 - \underset{\sim}{r}_2))$$

$$= \sum \{\text{All distinct numbered diagrams with links } f_{ij}, \text{ connected to vertex 1 and/or 2, with z at each vertex and integration at all but 1 and 2, weighted by 1/N - 2!}\}, \qquad (B.4)$$

$$n = \delta \ln \Xi/\delta(-\beta U(\underset{\sim}{r}_1))\Big|_{U=0}$$

$$= \sum \{\text{All distinct connected numbered diagrams with links } f_{ij}\text{, with z at each vertex and integration at all but 1, weighted by } 1/N\text{-}1!\}, \tag{B.5}$$

and

$$n_2(r_1,r_2) - n^2 = \sum \{\text{All distinct connected numbered diagrams with links } f_{ij}\text{, with z at each vertex and integration at all but 1 and 2, weighted by } 1/N\text{-}2!\}. \tag{B.6}$$

Clusters hanging by a single point in (B.6) are precisely representations of n of (B.5), so that (B.6) becomes

$$n_2(r_1,r_2) - n^2 = \sum \{\text{All distinct connected numbered diagrams which remain connected to 1 and/or 2 on removal of a single vertex, with links } f_{ij}\text{, with n at each vertex and integration at all but 1 and 2, weighted by } 1/N\text{-}2!\}. \tag{B.7}$$

If the symmetry factor of a diagram is defined as the number of permutations which change its form, and double connectivity allows a virtual link between 1 and 2, (B.7) also reads

$$g(\underset{\sim}{r}_1-\underset{\sim}{r}_2) - 1 = \left(n_2(\underset{\sim}{r}_1-\underset{\sim}{r}_2) - n^2\right)/n^2$$

$$= \sum \Big\{ \text{All distinct doubly connected unnumbered diagrams with links } f_{ij}\text{, with n and integration at each vertex but 1 and 2, weighted by 1/symmetry factor} \Big\}. \quad \text{(B.8)}$$

We now make one further diagram compression, introducing the link

$$F_{12} \equiv F(r_1-r_2) = \sum \Big\{ \text{All distinct doubly connected unnumbered series-parallel diagrams with links } f_{ij}\text{, with n and integration at each vertex but 1 and 2, weighted by 1/symmetry factor} \Big\}. \quad \text{(B.9)}$$

Since no series-parallel combinations remain, (B.8) becomes

$$g(\underset{\sim}{r}_1-\underset{\sim}{r}_2) - 1 = \sum \Big\{ \text{All distinct doubly connected unnumbered diagrams with links } F_{ij}\text{, with } \geq 3 \text{ links at every vertex but 1 and 2, with n and integration at each vertex but 1 and 2, weighted by 1/symmetry factor} \Big\}. \quad \text{(B.10)}$$

The CHNC approximation now consists of retaining only the first term F_{12} of $g_{12} - 1$. To express this analytically, we construct the sequence ($x = \underset{\sim}{r}_1 - \underset{\sim}{r}_2$)[43]

$$\varphi^0(\underset{\sim}{x}) = -\beta v(\underset{\sim}{x})$$

$$f^0(\underset{\sim}{x}) = e^{\varphi^0(\underset{\sim}{x})} - 1$$

$$\cdots$$

$$\varphi^s(\underset{\sim}{x}) = \sum \left\{ \begin{array}{l} \text{All distinct series combinations of links} \\ f^{s-1}(\underset{\sim}{x}), \text{ with appropriate weights} \end{array} \right\}$$

$$f^s(\underset{\sim}{x}) = \sum \left\{ \begin{array}{l} \text{All distinct parallel combinations of links} \\ \varphi^s(\underset{\sim}{x}), \text{ with appropriate weights} \end{array} \right\}, \qquad \text{(B.11)}$$

of which $F(\underset{\sim}{x})$ is the common limit, and use the fact that for parallel and series combinations of diagrams,

$$\begin{aligned} a(\underset{\sim}{x}) \| b(\underset{\sim}{x}) &= a(\underset{\sim}{x}) b(\underset{\sim}{x}) \\ a(\underset{\sim}{x}) \oplus b(\underset{\sim}{x}) &= n \int a(\underset{\sim}{x} - \underset{\sim}{y}) b(\underset{\sim}{y}) d\underset{\sim}{y} \end{aligned} \qquad \text{(B.12)}$$

or

$$(a \oplus b)_{\underset{\sim}{k}} = n a_{\underset{\sim}{k}} b_{\underset{\sim}{k}}.$$

Now φ^s consists of series combinations of f^{s-1}-type links in which two consecutive f^{s-1} links cannot both be from φ^{s-1}, since two φ^{s-1} links in series are already contained in φ^{s-1}. Let $\varphi^{s,\ell}$ denote the allowable chains of length ℓ, and consider $\varphi^{s,\ell+2}$. If the last link on say the right end is not in φ^{s-1}, it can be preceded by $\varphi^{s,\ell+2}$; if the last link is in φ^{s-1}, it must be preceded by the combination $\varphi^{s,\ell}(f^{s-1}-\varphi^{s-1})$. Hence

$$\underset{\sim}{\varphi}_k^{s,\ell+2} = n\underset{\sim}{\varphi}_k^{s,\ell+1}\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right) + n^2\underset{\sim}{\varphi}_k^{s,\ell}\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)\underset{\sim}{\varphi}_k^{s-1}. \tag{B.13}$$

Summing from $\ell = 1$ to ∞, and using the boundary conditions $\underset{\sim}{\varphi}_k^{s,1} = \underset{\sim}{f}_k^{s-1}$, $\underset{\sim}{\varphi}_k^{s,2} = n\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)\left(\underset{\sim}{f}_k^{s-1} + \underset{\sim}{\varphi}_k^{s-1}\right)$, we have

$$\underset{\sim}{\varphi}_k^{s} - \underset{\sim}{f}_k^{s-1} - n\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)\left(\underset{\sim}{f}_k^{s-1} + \underset{\sim}{\varphi}_k^{s-1}\right)$$

$$= n\left(\underset{\sim}{\varphi}_k^{s} - \underset{\sim}{f}_k^{s-1}\right)\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right) + n^2\underset{\sim}{\varphi}_k^{s}\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)\underset{\sim}{\varphi}_k^{s-1}, \text{ so}$$

that for $s > 1$

$$\underset{\sim}{\varphi}_k^{s} = \frac{\underset{\sim}{\varphi}_k^{s-1} + \left(1 + n\underset{\sim}{\varphi}_k^{s-1}\right)\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)}{1 - n\left(1 + n\underset{\sim}{\varphi}_k^{s-1}\right)\left(\underset{\sim}{f}_k^{s-1} - \underset{\sim}{\varphi}_k^{s-1}\right)}. \tag{B.14}$$

On the other hand, f^s consists of parallel combinations of φ^s in which (a) no two φ^s can be in f^{s-1}, (b) parallel paths cannot be numerically ordered. Therefore,

$$f^s(\underset{\sim}{x}) = \sum_1^\infty (\varphi^s(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x}))^\ell/\ell! + f^{s-1}(\underset{\sim}{x}) \sum_0^\infty (\varphi^s(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x}))^\ell/\ell!:$$

$$f^s(\underset{\sim}{x}) = f^{s-1}(\underset{\sim}{x}) + \left(1 + f^{s-1}(\underset{\sim}{x})\right)\left(e^{\varphi^s(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x})} - 1\right). \tag{B.15}$$

If (B.14) and (B.15) are rewritten as

$$\frac{\varphi_{\underset{\sim}{k}}^s}{1 + n\varphi_{\underset{\sim}{k}}^s} - \frac{\varphi_{\underset{\sim}{k}}^{s-1}}{1 + n\varphi_{\underset{\sim}{k}}^{s-1}} = f_{\underset{\sim}{k}}^{s-1} - \varphi_{\underset{\sim}{k}}^{s-1} \tag{B.16}$$

$$\ln(1 + f^s(\underset{\sim}{x})) - \ln(1 + f^{s-1}(\underset{\sim}{x})) = \varphi^s(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x}), \tag{B.17}$$

a "constant of the motion" is available, since by addition $\left(\varphi_{\underset{\sim}{k}}^s/1 + n\varphi_{\underset{\sim}{k}}^s\right)(\underset{\sim}{x}) - \left(\varphi_{\underset{\sim}{k}}^{s-1}/1 + n\varphi_{\underset{\sim}{k}}^{s-1}\right)(\underset{\sim}{x}) + \ln(1 + f^s(\underset{\sim}{x}))$ $- \ln(1 + f^{s-1}(\underset{\sim}{x})) + \varphi^{s-1}(\underset{\sim}{x}) - \varphi^s(\underset{\sim}{x}) = 0$, whence

$$\ln(1 + f^s(\underset{\sim}{x})) + \left(\frac{\varphi_{\underset{\sim}{k}}^s}{1 + n\varphi_{\underset{\sim}{k}}^s}\right)(\underset{\sim}{x}) - \varphi^s(\underset{\sim}{x}) = K(\underset{\sim}{x}) \tag{B.18}$$

Combining with (B.17): $K(\underset{\sim}{x}) = \ln(1 + f^{s-1}(\underset{\sim}{x})) - f^{s-1}(\underset{\sim}{x})$ $+ \left(\varphi^{s}_{\underset{\sim}{k}}/1 + n\varphi^{s}_{\underset{\sim}{k}}\right)(\underset{\sim}{x})$ and using $\varphi^{1}_{\underset{\sim}{k}} = f^{o}_{\underset{\sim}{k}}/1 - nf^{o}_{\underset{\sim}{k}}$, we conclude that

$$K(\underset{\sim}{x}) = \varphi^{o}(\underset{\sim}{x}) = -\beta v(\underset{\sim}{x}). \qquad (B.19)$$

Thus if the sequences φ^{s} and f^{s} converge to a common $G = g - 1$, (B.18) and (B.19) reduce to

$$\beta v(\underset{\sim}{x}) + \ln(1 + G(\underset{\sim}{x})) - G(\underset{\sim}{x}) = -C(\underset{\sim}{x})$$

where

$$C_{\underset{\sim}{k}} = G_{\underset{\sim}{k}}/1 + nG_{\underset{\sim}{k}}, \qquad (B.20)$$

the usual CHNC approximation.

From the diagrammatic viewpoint, a poorer approximation in that it includes fewer terms is obtained from the PY equation, which may be described as consisting of all plane convex diagrams which are doubly connected (after 1 and 2 are connected). Thus the n^{th} order parallel diagrams are composed of n^{th} order series diagrams plus single f-links joining any n^{th} order series diagrams which are not themselves $n\text{-}1^{st}$ order parallel. Equations (B.14) and (B.15) now go over to

$$\varphi_{\underset{\sim}{k}}^{s} = \frac{\varphi_{\underset{\sim}{k}}^{s-1} + \left(1 + n\varphi_{\underset{\sim}{k}}^{s-1}\right)\left(f_{\underset{\sim}{k}}^{s-1} - \varphi_{\underset{\sim}{k}}^{s-1}\right)}{1 - n\left(1 + n\varphi_{\underset{\sim}{k}}^{s-1}\right)\left(f_{\underset{\sim}{k}}^{s-1} - \varphi_{\underset{\sim}{k}}^{s-1}\right)} \tag{B.21}$$

$$f^{s}(\underset{\sim}{x}) = \varphi^{s}(\underset{\sim}{x}) + f(\underset{\sim}{x})(\varphi^{s}(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x})) \tag{B.22}$$

which may be rewritten

$$\frac{\varphi_{\underset{\sim}{k}}^{s}}{1 + n\varphi_{\underset{\sim}{k}}^{s}} - \frac{\varphi_{\underset{\sim}{k}}^{s-1}}{1 + n\varphi_{\underset{\sim}{k}}^{s-1}} = f_{\underset{\sim}{k}}^{s-1} - \varphi_{\underset{\sim}{k}}^{s-1} \tag{B.23}$$

$$f^{s}(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x}) = (1 + f(\underset{\sim}{x}))(\varphi^{s}(\underset{\sim}{x}) - f^{s-1}(\underset{\sim}{x})). \tag{B.24}$$

The combination implies

$$\left(\frac{\varphi_{\underset{\sim}{k}}^{s}}{1 + n\varphi_{\underset{\sim}{k}}^{s}}\right)(\underset{\sim}{x}) + \frac{f^{s}(\underset{\sim}{x})}{1 + f(\underset{\sim}{x})} = \varphi^{s}(\underset{\sim}{x}) + K(\underset{\sim}{x}), \tag{B.25}$$

and from the $s = 1$ case, we have $K(x) = \frac{f(\underset{\sim}{x})}{1 + f(\underset{\sim}{x})}$. Hence if the sequences f^{s}, φ^{s} converge to G, there results

$$\left(e^{\beta v(\underset{\sim}{x})} - 1\right)\left(1 + G(\underset{\sim}{x})\right) = -\, C(\underset{\sim}{x}) \tag{B.26}$$

$$C_{\underset{\sim}{k}} = G_{\underset{\sim}{k}}/1 + nG_{\underset{\sim}{k}},$$

It is to be noted that the CHNC and PY equations are linearizations of each other: $e^{\beta v(\underset{\sim}{x})}(1 + G(\underset{\sim}{x}) = e^{G(\underset{\sim}{x})-C(\underset{\sim}{x})}$ $\rightarrow e^{\beta v(x)}(1 + G(\underset{\sim}{x})) = 1 + G(\underset{\sim}{x}) - C(\underset{\sim}{x})$, while

$\beta v(\underset{\sim}{x}) + \ln(1 + G(\underset{\sim}{x})) = \ln(1 + G(\underset{\sim}{x}) - C(\underset{\sim}{x})) \rightarrow \beta v(\underset{\sim}{x}) + \ln(1 + G(\underset{\sim}{x}))$ $= G(\underset{\sim}{x}) - C(\underset{\sim}{x})$, so that there is no preference on this basis alone. One implication is that they both have correct 3^{rd} order coefficients in the equation of state.

APPENDIX C

THE BGY EQUATION[22]

When a distribution is functionally expanded in terms of a potential (or another distribution) an infinite series of successively higher distributions ordinarily results. There is a general class of expansions however in which an early truncation occurs. Suppose that F is a functional of W, F[W] and that W depends upon some parameter λ, W_λ. Then $\partial F[W_\lambda]/\partial\lambda$ will be a linear functional of $\partial W_\lambda/\partial\lambda$, leading at once to the full "expansion" in $\partial W_\lambda/\partial\lambda$

$$\frac{\partial}{\partial\lambda} F[W_\lambda] = \int \frac{\partial W_\lambda(\underset{\sim}{r})}{\partial\lambda} \frac{\delta F[W_\lambda]}{\delta W_\lambda(\underset{\sim}{r})} \, d\underset{\sim}{r}. \tag{C.1}$$

In particular, λ may simply be some spatial position, $\underset{\sim}{r}'$, in the independent function:

$$\nabla' F[W(\underset{\sim}{r}', \,)] = \int \nabla' W(\underset{\sim}{r}', \underset{\sim}{r}) \frac{\delta F[W(\underset{\sim}{r}', \,)]}{\delta W(\underset{\sim}{r}', \underset{\sim}{r})} \, d\underset{\sim}{r}. \tag{C.2}$$

As a special and important case, let us choose W as the external potential which fixes particles at $\underset{\sim}{r}'_1, \ldots, \underset{\sim}{r}'_t$,

$$U(\underset{\sim}{r}) = \sum_1^t v\left(\underset{\sim}{r}, \underset{\sim}{r}'_j\right), \tag{C.3}$$

and λ as the coordinate $\underset{\sim}{r}_1'$, while F is the s-body distribution $n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s)$. Since the relation

$$\frac{\delta n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U)}{\delta - \beta U(\underset{\sim}{r})} = n_{s+1}(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s \underset{\sim}{r}|U) - n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U) n(\underset{\sim}{r}|U) + n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U) \sum_1^s \delta(\underset{\sim}{r} - \underset{\sim}{r}_i) \tag{C.4}$$

may be established directly from the definition (2.13) (and indeed implies (7.15)), (C.2) becomes

$$\nabla_1' n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U) = \int \beta \nabla_1' v(\underset{\sim}{r}, \underset{\sim}{r}_1') [n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U) n(\underset{\sim}{r}|U) - n_{s+1}(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s \underset{\sim}{r}|U) - n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U) \Sigma \delta(\underset{\sim}{r} - \underset{\sim}{r}_i)] d\underset{\sim}{r}. \tag{C.5}$$

It is convenient to write (C.5) in terms of a generalized dimensionless distribution

$$g_{s+1}(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s; \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t') \equiv \frac{n_{s+t}(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t')}{n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s) n_t(\underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t')} = n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s|U)/n_s(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s). \tag{C.6}$$

After trivial manipulation, (C.5) then appears as

$$\nabla_1' \left[\ln g_{s+t}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t'\right) + \sum_1^s \beta v\left(\underset{\sim}{r}_i, \underset{\sim}{r}_1'\right) \right]$$

$$= \int \beta \nabla_1' v\left(\underset{\sim}{r}, \underset{\sim}{r}_1'\right) \left[g_{1+t}\left(\underset{\sim}{r};\ \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t'\right) - g_{s+t+1}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t';\ \underset{\sim}{r}\right) \right] n(\underset{\sim}{r}) d\underset{\sim}{r}. \tag{C.7}$$

The hierarchy (C.7), as s and t are varied, constitutes the full set of Born-Green-Yvon equations.

To close or truncate the sequence (C.7), one may proceed as follows. At zero density, it is clear from (C.6) that in the presence of both internal potential v and external U,

$$g_{s+t}\left(\underset{\sim}{r}_1, \ldots, \underset{\sim}{r}_s;\ \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t'\right) \to \exp\left(-\beta \sum_{i,j} v\left(\underset{\sim}{r}_i, \underset{\sim}{r}_j'\right)\right) \tag{C.8}$$

Thus if U is adjusted to fix an additional particle at $\underset{\sim}{r}$, g_{s+t} is unaltered. Further, Eq. (C.7) when valid remains valid on imposition of this added U. This suggests the superposition approximation that

$$g_{s+t}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t' \,\middle|\, \underset{\sim}{r}\right) = g_{s+t}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}_1' \cdots \underset{\sim}{r}_t'\right) \tag{C.9}$$

at any density, or employing (C.6), that

$$g_{s+t+1}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s \underset{\sim}{r}'_1 \cdots \underset{\sim}{r}'_t;\ \underset{\sim}{r}\right) = g_{s+1}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}\right) g_{t+1}\left(\underset{\sim}{r}'_1 \cdots \underset{\sim}{r}'_t;\ \underset{\sim}{r}\right). \quad (C.10)$$

Hence (C.7) reduces to

$$\nabla'_1 \left[\ln g_{s+t}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}'_1 \cdots \underset{\sim}{r}'_t\right) + \sum_1^s \beta v\left(\underset{\sim}{r}_i, \underset{\sim}{r}'_1\right) \right]$$

$$+ \int \beta \nabla'_1 v\left(\underset{\sim}{r}, \underset{\sim}{r}'_1\right) g_{1+t}\left(\underset{\sim}{r};\ \underset{\sim}{r}'_1 \cdots \underset{\sim}{r}'_t\right) \left[g_{s+1}\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s;\ \underset{\sim}{r}\right) - 1 \right] n\left(\underset{\sim}{r}\right) d\underset{\sim}{r} = 0, \quad (C.11)$$

the superposition-terminated BGY approximation. (C.11) is a self-consistent relation for s = 1 and/or t = 1, an explicit solution for s,t > 1.

The set (C.11) is not truly extensive. By virtue of (C.10), it is easily seen by iteration that

$$n_s\left(\underset{\sim}{r}_1 \cdots \underset{\sim}{r}_s\right) = \prod n\left(\underset{\sim}{r}_i\right) \prod_{i>j} g\left(\underset{\sim}{r}_i, \underset{\sim}{r}_j\right), \quad (C.12)$$

the more customary form of the suberposition approximation,[33] and it then follows that the s = 1, t = 1 case of (C.11),

$$\nabla'[\ln g(\underset{\sim}{r},\underset{\sim}{r}') + \beta v(\underset{\sim}{r},\underset{\sim}{r}')] + \int \beta \nabla' v(\underset{\sim}{x},\underset{\sim}{r}') g(\underset{\sim}{x},\underset{\sim}{r}')(g(\underset{\sim}{x},\underset{\sim}{r}) - 1) n(\underset{\sim}{x}) d\underset{\sim}{x} = 0, \tag{C.13}$$

implies the whole set. We generally refer to (C.13) as the BGY approximation.

APPENDIX D

RESPONSE TO EXTERNAL FORCES

The "turning on" argument can be used in a variety of ways to assess the effect of an external potential on the distribution functions of a system interacting via pair forces, depending upon the nature of the initial and final states. Suppose that we start with the inhomogeneous system with all forces present, and then fix one particle by applying the internal potential tethered to $\underset{\sim}{x}$. Now the fundamental relations (6.1) and (6.2) are unaffected by any external potential which is also present. Hence dropping the condition of initial uniformity in the process (6.4), all of the arguments of Section 6 remain valid. In particular, we may combine the appropriate generalizations of Eqs. (6.6,6.12,6.15),

$$\text{D.H.} \qquad c(\underset{\sim}{x},\underset{\sim}{r}) = -\beta v(\underset{\sim}{x},\underset{\sim}{r}) \tag{D.1}$$

$$\text{P.Y.} \qquad c(\underset{\sim}{x},\underset{\sim}{r}) = \left(1-e^{\beta v(\underset{\sim}{x},\underset{\sim}{r})}\right) g(\underset{\sim}{x},\underset{\sim}{r}) \tag{D.2}$$

$$\text{C.H.N.C.} \qquad c(\underset{\sim}{x},\underset{\sim}{r}) = g(\underset{\sim}{x},\underset{\sim}{r}) - 1 - \ln g(\underset{\sim}{x},\underset{\sim}{r}) - \beta v(\underset{\sim}{x},\underset{\sim}{r}), \tag{D.3}$$

with the general definition (5.6) relating c and g to obtain the corresponding first order approximate integral equations with external potential.

The quantity $v(\underset{\sim}{x},\underset{\sim}{r})$ occurring in (D.1-3) is the internal potential alone. Thus a necessary part of the solution is a subsidiary calculation of the imposed external potential $u(\underset{\sim}{r})$ which does not explicitly appear. In principle, this can be achieved by the same process which gives rise to the approximation in question. For example, the Debye-Huckel result emanates from the expansion of n relative to u, or vice versa. But

$$\frac{\delta\beta u(\underset{\sim}{r})}{\delta n(\underset{\sim}{x})} = c(\underset{\sim}{x},\underset{\sim}{r}) - \frac{\delta(\underset{\sim}{x},\underset{\sim}{r})}{n(\underset{\sim}{x})} \tag{D.4}$$

becomes, on applying (D.1),

$$\frac{\delta\beta u(\underset{\sim}{r})}{\delta n(\underset{\sim}{x})} = -\beta v(\underset{\sim}{x},\underset{\sim}{r}) - \frac{\delta(\underset{\sim}{x},\underset{\sim}{r})}{n(\underset{\sim}{x})}, \tag{D.5}$$

which integrates directly to

$$\beta u(\underset{\sim}{r}) = -\int \beta v(\underset{\sim}{x},\underset{\sim}{r})(n(\underset{\sim}{x}) - n)d\underset{\sim}{x} - \ln(n(\underset{\sim}{r})/n), \tag{D.6}$$

the desired result.

If one tries to follow the method of (D.4) and (D.6) in treating the PY and CHNC approximations, the equivalent of (D.5) turns out not to be integrable. A more systematic method is to expand that function which first on fixing a particle reduces to the dependent function

of the approximation in question, or at least to within a constant. For this purpose, we recall that in the persistent presence of an external potential u(r), Eq. (6.3) has precisely the form desired,

$$z^{-s} n_s(\underset{\sim}{r}_1,\ldots,\underset{\sim}{r}_s)\exp\beta\Big(\sum_{i<j\le s} v(\underset{\sim}{r}_i,\underset{\sim}{r}_j) + \sum u(\underset{\sim}{r}_i)\Big)\Xi = \Xi[U]. \quad (D.7)$$

Hence in the PY and CHNC cases, instead of expanding $ne^{\beta U}$ and $\ln n + \beta U$ in terms of n, one expands Ξ and $\ln \Xi$; this yields directly on turning on the internal v

$$\text{PY:}\quad z^{-1} n(\underset{\sim}{x}) e^{\beta u(\underset{\sim}{x})} = 1 + \int n(\underset{\sim}{r})(g(\underset{\sim}{r},\underset{\sim}{x}) - 1)d\underset{\sim}{r}$$

$$- \iint n(\underset{\sim}{r}')c(\underset{\sim}{r}',\underset{\sim}{r})n(\underset{\sim}{r})(g(\underset{\sim}{r},\underset{\sim}{x}) - 1)d\underset{\sim}{r}, \quad (D.8)$$

or simply

$$\text{PY:}\quad n(\underset{\sim}{x}) e^{\beta u(\underset{\sim}{x})} = z\Big(1 + \int n(\underset{\sim}{r})c(\underset{\sim}{r},\underset{\sim}{x})d\underset{\sim}{r}\Big), \quad (D.9)$$

and similarly

$$\text{CHNC:}\quad \ln n(\underset{\sim}{x}) + \beta u(\underset{\sim}{x}) = \ln z - \int n(\underset{\sim}{r})c(\underset{\sim}{r},\underset{\sim}{x})d\underset{\sim}{r}. \quad (D.10)$$

The presence of z of course is due to the fact that $u(\underset{\sim}{r})$ and $v(\underset{\sim}{r},\underset{\sim}{r}')$ alone do not determine the system. However, we can equally well normalize to the density n in a region in which $u(\underset{\sim}{x})$ vanishes; employing (5.7) to eliminate z, then Eq. (D.9) becomes

$$(n(\underset{\sim}{x})/n)e^{\beta u(\underset{\sim}{x})} = \left(1 + \int n(\underset{\sim}{r})c(\underset{\sim}{r},\underset{\sim}{x})d\underset{\sim}{r}\right)/(2 - \partial\beta p/\partial n). \qquad (D.11)$$

Typical among the multitude of possibilities of inserting an external potential is that of starting with the uniform system and then turning on both the over-all external potential and the internal type of external potential which fixes a particle. Carrying this out to first order for the CHNC variation yields directly $\beta v(\underset{\sim}{r},\underset{\sim}{x}) + \beta u(\underset{\sim}{r}) + \ln(n_2(\underset{\sim}{r},\underset{\sim}{x})/n(\underset{\sim}{x})) = \ln n_o + \int c_o(\underset{\sim}{r},\underset{\sim}{r}')(n_2(\underset{\sim}{r}',\underset{\sim}{x})/n(\underset{\sim}{x}) - n_o)d\underset{\sim}{r}'$ where subscript zero denotes the uniform state, or

$$\beta v(\underset{\sim}{r},\underset{\sim}{x}) + \beta u(\underset{\sim}{r}) + \ln g(\underset{\sim}{r},\underset{\sim}{x}) + \ln(n(\underset{\sim}{r})/n_o)$$

$$= \int c_o(\underset{\sim}{r}-\underset{\sim}{r}')n(\underset{\sim}{r}')g(\underset{\sim}{r}',\underset{\sim}{x})d\underset{\sim}{r}' - \int c_o(\underset{\sim}{r}-\underset{\sim}{r}')n_o d\underset{\sim}{r}'. \qquad (D.12)$$

Separating $\underset{\sim}{r}$ and $\underset{\sim}{x}$ asymptotically, then subtracting from (D.12),

$$\beta u(\underset{\sim}{r}) + \ln(n(\underset{\sim}{r})/n_o) = \int c_o(\underset{\sim}{r}-\underset{\sim}{r}')(n(\underset{\sim}{r}') - n)d\underset{\sim}{r}'$$

$$\beta v(\underset{\sim}{r},\underset{\sim}{x}) + \ln g(\underset{\sim}{r},\underset{\sim}{x}) = \int c_o(\underset{\sim}{r}-\underset{\sim}{r}')n(\underset{\sim}{r}')(g(\underset{\sim}{r}',\underset{\sim}{x}) - 1)d\underset{\sim}{r}'. \qquad (D.13)$$

The second equation of (D.13) is a repeat of the relation (D.3), but possibly less accurate, since c is replaced by c_o in the integral. The first equation of (D.13), which determines the external potential, is precisely (5.15) in which only the external potential is turned on. With no loss of generality, we can then use only the elementary process of Section 5.

Let us look then in somewhat greater detail at the direct application of an external potential to a uniform system, as in Section 5. If the potential is weak and asymptotically zero, it is reasonable to expand about the asymptotic uniform system. The PY-type expansion, $ne^{\beta u}$ a functional of n, yields

$$n(\underset{\sim}{r})e^{\beta u(\underset{\sim}{r})} = n\left(1 + \int c_o(\underset{\sim}{r}-\underset{\sim}{r}')(n(\underset{\sim}{r}') - n)d\underset{\sim}{r}'\right) \qquad (D.14)$$

which can more profitably be written as

$$\int \left[\left(e^{\beta u(\underset{\sim}{r})} - 1\right)\delta(\underset{\sim}{r}-\underset{\sim}{r}') + \left(\delta(\underset{\sim}{r}-\underset{\sim}{r}') - nc_o(\underset{\sim}{r}-\underset{\sim}{r}')\right)\right](n(\underset{\sim}{r}') - n)d\underset{\sim}{r}' = -\left(e^{\beta u(\underset{\sim}{r})} - 1\right)n, \qquad (D.15)$$

or employing a Taylor expansion of $n(\underset{\sim}{r}')$ about $n(\underset{\sim}{r})$, as

$$\left[\left(e^{\beta u(\underset{\sim}{r})}-1\right)+\left(\frac{\partial\beta p}{\partial n}+\frac{1}{6}c^{(2)}\nabla^2+\frac{1}{120}c^{(4)}(\nabla^2)^2+\cdots\right)\right](n(\underset{\sim}{r})-n)=-\left(e^{\beta u(\underset{\sim}{r})}-1\right)n \tag{D.16}$$

in the notation of (6.23). In the potential-bearing region, $n(\underset{\sim}{r}) \sim (n\partial\beta p/\partial n)/\left(e^{\beta u(\underset{\sim}{r})} - 1 + \partial\beta p/\partial n\right)$, while outside this region, $n(\underset{\sim}{r}) - n$ is a linear combination of oscillating Yukawa factors $r^{-1}e^{-\alpha r}\cos(\gamma r+\delta)$, with the pure oscillating case $\alpha = 0$, $\gamma \neq 0$ being prohibited by the nonnegativity of the integral operator $\delta(\underset{\sim}{r}-\underset{\sim}{r}') - nc_o(\underset{\sim}{r}-\underset{\sim}{r}')$ (a fact which follows directly from (3.3)). It is interesting to note that in the approximation (D.15), the operator applied to $n(\underset{\sim}{r}') - n$ is positive definite for repulsive $u \geq 0$, so that an applied repulsive potential never produces a singular density reaction even in a phase coexistence region.

The linearity of (D.15) in $n(\underset{\sim}{r}') - n$ is peculiar to the choice of dependent and independent variable. It is to be noted however that the corresponding CHNC expansion has a certain self-consistent advantage. If $n(\underset{\sim}{r};\ \lambda) = n + \lambda(n(\underset{\sim}{r}) - n)$ and we expand $\ln n + \beta\alpha$ in terms of n, with remainder as in Section 7,

$$\ln n(\underset{\sim}{r}) + \beta u(\underset{\sim}{r}) = \ln n + \int\int_0^1 c(\underset{\sim}{r},\underset{\sim}{r}';\ \lambda)(n(\underset{\sim}{r}') - n)d\lambda\, d\underset{\sim}{r}', \tag{D.17}$$

then noting from Section 8 that the CHNC approximation says precisely that $c(\underset{\sim}{r},\underset{\sim}{r}')$ is unchanged under an external potential, (D.17) reduces to

$$\beta u(\underset{\sim}{r}) = \ln(n/n(\underset{\sim}{r})) + \int c_o(\underset{\sim}{r},\underset{\sim}{r}')(n(\underset{\sim}{r}') - n)d\underset{\sim}{r}'. \qquad (D.18)$$

REFERENCES

1. J. Moser, Symp. Nonlinear Diff. Eq., Colorado Springs, (Aug. 1961).
2. T. L. Hill, Statistical Mechanics, (McGraw-Hill, New York, 1956), 190.
3. L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amst. 17, 793 (1914).
4. Reference 2, p. 233.
5. See e.g. J. L. Lebowitz and J. K. Percus, Phys. Rev. 122, 1675 (1961).
6. J. E. Mayer, in Many-Body Problem, ed. J. K. Percus, (Interscience, New York, 1963), 383.
7. J. K. Percus and G. J. Yevick, Bull. Am. Phys. Soc. 5, 275 (1960).
8. C. Garrod and J. K. Percus, J. Math. Phys., to be published.
9. J. K. Percus, Bull. Am. Phys. Soc. 7, 267 (1962).
10. G. Joos, Theoretical Physics, (Hafner, New York), 181.
11. J. L. Lebowitz and J. K. Percus, J. Math. Phys. 4, 116 (1963).
12. See e.g. Reference 5, p. 1689.
13. J. K. Percus, Phys. Rev. Letters, 8, 462 (1962).
14. V. Volterra, Theory of Functionals, (Dover, New York 1959), 26.

References (Cont'd.) - 2

15. J. Yvon, Suppl. Nuovo cimento 9, 144 (1958).
16. J. K. Percus and G. J. Yevick, Nuovo cimento 5, 1057 (1957).
17. L. Landau and I. Lifshitz, Statistical Physics, (Addison-Wesley, Reading 1958), 231.
18. R. H. Fowler and E. A. Guggenheim, Statistical Thermodynamics (Cambridge Press, Cambridge 1939) 390.
19. J. G. Kirkwood and Z. Salsburg, Discussions Faraday Soc. 15, 28 (1953).
20. J. K. Percus and G. J. Yevick, Phys. Rev. 110, 1 (1958).
21. Van Leuwen, Groeneveld, and de Boer, Physica 25, 792 (1959); E. Meeron, J. Math. Phys. 1, 192 (1960).
22. M. Born and H. S. Green, Proc. Roy. Soc. (London) A188, 10 (1946); J. Yvon, Actualities Scientifiques et Industrielles (Hermann and Cie, Paris 1935), No. 203.
23. L. Goldstein, Phys. Rev. 85, 35 (1952).
24. J. L. Lebowitz, Phys. Rev. 133, A895 (1964).
25. See e.g. H. Jeffreys and B. Swirles, Mathematical Physics (Cambridge Press, Cambridge 1956), 51.
26. J. E. Mayer and E. Montroll, J. Chem. Phys. 9, 2 (1941).
27. J. L. Lebowitz and J. K. Percus, J. Math. Phys. 4, 1495 (1963).
28. E. Lieb, J. Math. Phys. 4, 671 (1963).
29. O. Penrose, J. Math. Phys. 4, 1488 (1963).

References (Cont'd.) - 3

30. Reiss, Frisch, and Lebowitz, J. Chem. Phys. 31, 369 (1959). E. Helfand and E. H. Stillinger, J. Chem. Phys. 37, 2646 (1962).

31. B. J. Alder and T. E. Wainwright, J. Chem. Phys. 33, 1439 (1960); Wood, Parker and Jacobsohn, Suppl. Nuovo cimento 9, 493 (1958).

32. See e.g. Hirschfelder, Curtiss and Bird, Molecular Theory of Gases and Liquids (Wiley, New York 1954), 330.

33. J. G Kirkwood and E. M. Boggs, J. Chem. Phys. 10, 394 (1942).

34. L. Verlet, Physica, to be published.

35. A. A. Broyles, J. Chem. Phys. 35, 493 (1961); Bull. Am. Phys. Soc. 7, 67 (1962).

36. See e.g. Reference 2, Chapter 7.

37. J. K. Percus and G. Stell, to be published.

38. F. H. Stillinger, Phys. Rev. 126, 1239 (1962).

39. J. E. Mayer and M. G. Mayer, Statistical Mechanics, (Wiley, New York 1940) Chapter 13. G. Stell, this volume.

40. M. S. Wertheim, Phys. Rev. Lett. 10, 321 (1963). E. Thiele, J. Chem. Phys. 38, 1959 (1963).

41. Lebowitz, Percus, and Zucker, to be published.

42. J. K. Percus in Reference 6, p. 257.

43. J. K. Percus, unpublished paper under ONR contract, 1955.

CLUSTER EXPANSIONS FOR CLASSICAL SYSTEMS IN EQUILIBRIUM*

G. Stell
Courant Institute of Mathematical Sciences
New York University, New York

SECTION 1

INTRODUCTION

In this paper we shall consider one method of deriving the cluster expansions relevant to a classical nonionic statistical mechanical system at equilibrium.

A large number of alternative derivations and summaries of existing derivations have appeared since the introduction of cluster expansions by Ursel,[1] Yvon,[2] and Mayer and his collaborators.[3] We shall not attempt to rederive and rediscuss all of the results that can be found in these various treatments.[4-8] Instead we shall derive a set of expansions, including all these that are commonly used in discussing nonionic systems, by means of a simple unified approach, various aspects of which have been developed independently by a number of authors whose work we shall review.

The basis of our method consists of the application of two operations. The first of these is variational or functional differentiation. There seems to be no single phrase in use to describe the second operation, which offers a convenient means of resumming series. For want of a better term we shall call it topological reduction. It enables one

---

*This work was supported by the National Science Foundation under Grant GP-2003

to reduce the sums of graphs that appear in fugacity expansions to the sums of more highly connected graphs that appear in number-density expansions. For a system in which the potential consists of a sum of one- and two-body terms, it then enables one to further reduce these sums to sums of still more highly connected graphs that involve only the one- and two-particle distribution functions, ρ_1 and ρ_2. When the use of functional differentiation to define distribution functions is combined with topological reduction, the resulting formalism offers a very effective means of dealing with cluster expansions. It applies to the graphs associated with a non-uniform system, in which ρ_1 is not a constant, as easily as it does to the graphs associated with a uniform system. It can also be immediately generalized to apply to mixtures and to the case in which the potential energy includes n-body terms for $n \geq 3$.[9] Finally, it enables us to obtain the prescription for the general term of cluster expansions without the use of combinatorial arguments.

The first use of functional differentiation to define the distribution functions of a classical statistical mechanical system seems to have been made in 1945 by N. N. Bogoliubov[10] who introduced a generating functional as a purely formal device suggested to him by the previous use of generating functions to consider stochastic processes. Later Bogoliubov realized that his generating functional is closely related to objects with thermodynamical meaning.[11] In fact his functional is essentially the Helmholtz free energy of a system, viewed as a functional of the external

field to which the system is subjected, and his test function is essentially the external field itself. Thus the basis of his formalism turned out to be quite similar to that of the theory of Green's functions developed by Schwinger.[12] Furthermore his definition of the n-particle distribution function as an nth functional derivative, which at first appeared somewhat artificial, proved to have a fundamental thermodynamical meaning. It introduces the ρ_n for $n = 1,2,3,\ldots$ as a sequence of functions that measure the successively higher derivatives of the change in the free energy of a system with respect to the change in the external field. In light of the fact that the change in the volume of a system is just a special case of the much more general class of possible changes in the external field, the distribution functions thus emerge as meaningful objects even if their probabilistic definition is never introduced. In a sense, they are quite natural generalizations of the partial derivatives with which thermodynamics abounds.

Generating functionals and functional derivatives have also been used extensively in treating quantum mechanical systems and systems not at equilibrium. However, even a summary of the ways in which functionals appear in these treatments would constitute a section in itself,[13] so we shall restrict ourselves to references that deal with classical systems at equilibrium. Bazarov,[14] Zumino,[15] Lewis,[16]

Lewis and Keller,[17] and the author[18] have used Bogoliubov's generating functional as a starting point for various studies, and Percus[19] and Lebowitz and Percus[20] have used functional differentiation extensively, especially in connection with functional expansions that are based upon the identification of the pair potential $\varphi(1,\alpha)$ and the external field due to the distinguished particle 1 acting upon any other particle α. Yvon,[21] Hiroike,[22] Morita and Hiroike,[23,24] and deDominicis[25] have also used functional derivatives in considering classical systems. It is the work of Morita and Hiroike and deDominicis to which we shall refer in the main body of this paper.

The technique we have called topological reduction can be characterized in the following way. A sum S of graphs with vertices representing some function F_0, bonds representing F_1, and, in general, d-dimensional faces representing F_d, is shown to be equal to a single graph with F_d replaced by some G_d for a particular d. The equality is established by the reduction of all the graphs in S to the same graph under some operation that involves removing faces of dimension d, and the function G_d is the function represented by the sum of all the distinct faces that are removed. This kind of reduction depends only on the topological structure of the graphs and the associated relationship between F_d and G_d and not at all on the nature of F_d and G_d as functions.

Thus when it is used to get expansions of the ρ_n as power series in the number density from the fugacity expansions it does not require that the connected cluster integrals factor into products of irreducible clusters. The factorization will only occur when the system is translation invariant, i.e., when ρ_1 is the constant, ρ. The original derivations of Mayer and his collaborators,[3] as well as a number of variations of their derivations and alternatives to them,[26] do require translation invariance[27] and hence apply only to a field-free infinite system, whether canonical or grand canonical ensembles are considered.

Systematic use of topological reduction in rederiving the Mayer results was made by E. E. Salpeter[6] in 1958. Shortly after Salpeter's work appeared, Meeron,[28] M. S. Green,[29] Morita and Hiroike,[23] and Van Leeuwen, Groeneveld, and DeBoer[30] independently applied the same ideas of topological reduction to replace the ordinary cluster expansion of ψ_2, the potential of mean force of a fluid, by a (g_2-1)-bond, ρ-vertex expansion, where $g_2 = e^{-\beta\psi_2} = \rho_2/\rho^2$. Thus they obtained an expression for g_2, that can be viewed as an exact but intractable integral equation for g_2, involving an infinite number of integrals. An approximation that was suggested by all of these workers is obtained by setting one class of integrals equal to zero. This approximation leads to an integral equation that was also independently suggested

by Verlet[31] and by Rushbrooke.[32] (The equation has come to be known as the hypernetted chain equation.) At the same time topological reduction was also used by M. S. Green[33] to rederive the virial expansion of the equation of state.

Subsequently Morita systematized and generalized his approach and combined it with a functional definition of the distribution functions, equivalent to Bogoliubov's, that evolved out of work Hiroike had previously initiated.[22] With the comprehensive formalism that resulted, Morita and Hiroike[24] were able to obtain a large set of expansions for thermodynamical quantities and for ρ_1 and ρ_2 in a nonuniform grand canonical system. These included the expressions that Morita has previously obtained for a uniform system as well as an interesting expression for the entropy[34] that is given entirely in terms of ρ_1 and ρ_2. Morita and Hiroike also obtained a number of variational principles, one of which was also found by M. S. Green.[29] DeDominicis[25] independently considered a classical system by using the same methods that he, Balian, and Bloch had developed for treating quantum systems.[35] Applied to the classical case, the methods proved to yield identical results to those of Morita and Hiroike.

The technique of topological reduction can also be used to formally solve the closure problem of finding higher order distribution functions completely in terms of ρ_1 and

ρ_2,[36] and more generally, of investigating the relationships between ρ_n and the ρ_s for $s < n$. Abe[37] and Verlet[31] independently found the first few terms of an expansion of ρ_3 involving only ρ_1 and ρ_2, and pointed out the possibility of obtaining the whole expansion in principle by iteration. The author[38] obtained the general term of this expansion as well as similar expansions for ρ_n, $n > 3$, by using topological reduction. Meeron[39] has also considered closely related expressions.

In Sections 3 through 10 we shall derive all of the cluster expansions we have just discussed. In Section 11 we shall analyze in some detail two approximations, the hypernetted chain [hyperchain, convolution, series-parallel] approximation and the Percus-Yevick[40] approximation, that are closely related to the developments reviewed here. Finally, in Section 12, we consider the graphical interpretation of the observation that the pair potential $\varphi(1,\alpha)$ can be identified with the external field due to a distinguished particle 1 acting upon any other particle α. As we have mentioned, this observation is the basis of a number of interesting functional expansions obtained by Percus[19] and Lebowitz and Percus,[20] and it can be used as a starting point for the effective-density approach to the theory of fluids considered by Helfand and Stillinger[41] and by Lebowitz and Praestgaard.[42]

SECTION 2

THE DEFINITIONS OF THE FUNCTIONS TO BE CONSIDERED

The canonical partition function for a system of N particles is given by

$$Z_N = \frac{1}{N!}\int d(1)\ldots d(N)\exp\{-\beta V(1,\ldots,N)\}. \qquad (2\text{-}1)$$

The function $V(1,\ldots,N)$ is the potential energy of the system and we shall assume it to be of the form

$$V(1,\ldots,N) = \sum_{1\le i\le N}\varphi_1(r_1) + \sum_{1\le i<j\le N}\varphi_2(|r_1-r_j|). \qquad (2\text{-}2)$$

In Eq. (2-1), $\beta = 1/kT$ where k is Boltzmann's constant and T is the absolute temperature. For simplicity we write the arguments of functions as 1,2,... and differentials as d(1),d(2),... rather than $\underset{\sim}{r}_1,\underset{\sim}{r}_2,\ldots$, and $d\underset{\sim}{r}_1, d\underset{\sim}{r}_2,\ldots$, respectively.

In this article we shall consider a grand ensemble. Letting z be the fugacity, the grand partition function can be written as

$$\Xi = \sum_{s\ge 0} z^s Z_s. \qquad (2\text{-}3)$$

In (2-2), $\varphi_1(\underset{\sim}{r}_i)$ is the potential energy of a particle at $\underset{\sim}{r}_i$ due to external forces and φ_2 is the interaction potential of two particles. One of the useful aspects of the formal treatment of cluster integrals we are about to examine is that it is no more difficult to apply when φ_1 is arbitrary than when φ_1 is assumed to be identically zero. It can also be readily used even if terms of the form φ_s $(1,2,\ldots,s)$ for any $s \leq N$ are added to the right-hand side of (2-2). However, adding such terms necessitates generalizing the definitions of a number of graphical concepts in a way that somewhat extends the length of their description, and since such terms do not play an important role in classical equilibrium statistical mechanics we shall not consider them in this paper.

The canonical s-particle [s-point, s-body] distribution function is given by

$$n_s^{(N)}(1,\ldots,s) = \frac{N!}{(N-s)!} \int d(s+1)\ldots d(N) D_N(1,\ldots,N), \qquad (2\text{-}4)$$

where

$$D_N = \frac{1}{N! Z_N} \exp\{-\beta V(1,\ldots,N)\}. \qquad (2\text{-}5)$$

The grand canonical s-particle distribution functions can be written as

$$\rho_s(1,\ldots,s) = \Xi^{-1} \sum_{N \geq s} n_s^{(N)}(1,\ldots,s) z^N Z_N. \qquad (2\text{-}6)$$

It can be directly verified from the definition of functional differentiation that $\rho_s(1,\ldots,s)$ can be defined alternatively in terms of the sth functional or variational derivative[43] of Ξ. Introducing for convenience the function $z_1(1) = z \exp\{-\beta\varphi_1(1)\}$, we have

$$\rho_s(1,\ldots,s) = \frac{\prod_{1\le i\le s} z_1(1)}{\Xi} \frac{\delta^s \Xi}{\prod_{1\le i\le s} \delta z_1(1)}. \qquad (2\text{-}7)$$

Another set of s-particle functions that arise in a natural fashion in the study of statistical mechanical systems is the set $\{u_s(1,\ldots,s)\}$ of cluster functions or Ursell functions. The u_s are usually defined by their relationship with the $\rho_s(1,\ldots,s)$:

$$\rho_s(1,\ldots,s) = \Sigma\pi u_\alpha(i_1,\ldots,i_{n_\alpha}). \qquad (2\text{-}8)$$

The sum of products in (8) is carried out over all partitions of the set $\{1,\ldots,s\}$. That is, we consider an arbitrary partitioning of $\{1,\ldots,s\}$ into N subsets, where the αth subset has the n_α elements $i_1, i_2, \ldots, i_{n_\alpha}$. We associate with this partitioning the term $\prod_{1<\alpha\le N} u_\alpha(i_1,\ldots,i_{n_\alpha})$ and then sum over all possible partitions.

The $u_s(1,\ldots,s)$ can equally well be defined by a functional derivative of $\ln \Xi$.[43]

$$u_s(1,\ldots,s) = \prod_{1\le i\le s} z_1(i) \frac{\delta^s \ln \Xi}{\prod_{1\le i\le s} \delta z_1(i)}. \qquad (2\text{-}9)$$

The equivalence of (2-8) and (2-9) given (2-7) is readily verified by noting the way in which the nth functional derivative of a function can be expressed as a sum of products of the sth functional derivatives of its logarithm for $s \le n$. A comparison of (2-7) and (2-9) then yields (2-8). It is worth noting in particular the relationship between the first two distribution and cluster functions

$$\rho_1(1) = u_1(1); \quad \rho_2(1,2) - \rho_1(1)\rho_1(2) = u_2(1,2). \qquad (2\text{-}10)$$

Instead of using the ρ_s and the u_s, we shall often find it more convenient to work with related functions defined by the equations

$$g_s(1,\ldots,s) = \frac{\rho_s(1,\ldots,s)}{\rho_1(1)\rho_1(2)\ldots\rho_1(s)} \quad \text{for} \quad s \ge 2, \qquad (2\text{-}11)$$

$$h_s(1,\ldots,s) = \frac{u_s(1,\ldots,s)}{\rho_1(1)\rho_1(2)\ldots\rho_1(s)} \quad \text{for} \quad s \geq 2, \tag{2-12}$$

$$h_1(1) = \ln\{u_1(1)/z_1(1)\}. \tag{2-13}$$

From (2-10) we have

$$h_2(1,2) = g_2(1,2) - 1. \tag{2-14}$$

The function $h_2(1,2)$ is often called the two-particle correlation function. Alas, so is g_2. When g_2 is just a function of $r_{12} = |\underset{\sim}{r}_1 - \underset{\sim}{r}_2|$, as it is in a uniform fluid, it is often called the radial distribution function. As will become clear from the expansions we shall develop, when z approaches zero, $\rho_1(1)$ approaches $z_1(1) = ze^{-\beta\varphi_1(1)/kT}$ and $g_s(1,\ldots,s)$ approaches $\exp\left\{\sum_{1\leq i<j\leq s} \beta\varphi_2(i,j)/kT\right\}$. [We shall denote the function, $\exp\{-\beta\varphi_2(i,j)\}$, which appears often, simply as $e(i,j)$.] The function $h_2(i,j)$ approaches the Mayer f-function, $f(i,j) = e(i,j) - 1$.

Another function that plays an important role in the theory of fluids is the Ornstein-Zernike direct correlation function,[44] which can be defined in terms of $h_2(1,2)$ by the equation

$$h_2(1,2) = c(1,2) + \int c(1,3)h_2(3,2)\rho_1(3)d(3). \tag{2-15}$$

As we shall see, this definition is equivalent to the functional definition[45]

$$c(1,2) = \frac{\delta \ln\{\delta \ln \Xi/\delta z_1(1)\}}{\delta \rho_1(2)}. \qquad (2\text{-}16)$$

SECTION 3

GRAPHICAL DEFINITIONS AND LEMMAS

In the following definitions, some alternative terms that have appeared in the literature are given in brackets. We have made no attempt to be exhaustive in listing such alternatives.

A linear graph is a collection of circles [vertices, points] between some pairs of which there are bonds [lines]. Two circles are adjacent [contiguous] if there is a bond joining them directly. A path is a sequence of adjacent circles joined by bonds. Two paths from circle γ_1 to circle γ_2 are independent if they have no intermediate circles in common. A graph is connected if there exists a path between any pair of circles in the graph. Our graphs will consist of two kinds of circles, black [free, field, unlabeled] circles and white [rooted, reference, labeled] circles. If there are at least n independent paths between any pair of circles in a graph we shall say that the graph is at-least-n-tuply connected. At-least-singly connected = connected.

In general a graph will consist of several unconnected parts or components. More precisely, a component is a maximal connected subgraph. By a maximal subgraph possessing any property we shall mean one that is contained in no

other graph possessing that property. Similarly by a minimal subgraph with a given property we shall mean one that contains no other subgraph with that property. Given a graph, there is no guarantee that a unique maximal or minimal subgraph possessing a property exists.

In a simple graph, between any pair of adjacent circles there is only one path that consists of a single bond. Most of the graphs we consider will be simple. A nonsimple graph is sometimes called composite. (This use of the term simple is different from that found in references 6, 25, and 30.)

We shall associate integrals with graphs. The simplest integrals we shall consider - for example, the integral in (2-1) - have the form

$$\int d(n+1)\ldots d(N) \left[\prod_{S} B(i,j)\right]\left[\prod_{1\le i\le n} \gamma_1(i)\right]\left[\prod_{n+1\le i\le N} \gamma_2(i)\right]. \tag{3-1}$$

Here S is some subset of all pairs $\{i,j\}$ such that $1 \le i < j \le N$ and $B(i,j) = B(j,i)$. In (2-1), S is the whole set of such pairs and integration is performed over all variables, so that $n = 0$. A considerably wider class of integrals than those given by (3-1) can be made to correspond to graphs. For our purposes it is sufficient

to consider an only slightly more general integral, for which the condition that $B(i,j) = B(j,i)$ is still satisfied:

$$\int d(n+1)\ldots d(N)\left[\prod_S\prod_a\left\{B_a(i,j)^{m_a(i,j)}\right\}\right]\left[\prod_{1\le i\le n}\gamma_1(i)\right]\left[\prod_{n+1\le i\le N}\gamma_2(i)\right]. \tag{3-2}$$

For example we can have

$$\int d(3)d(4)B_1(1,3)^4B_1(1,2)B_2(1,2)^2B_1(2,3)\gamma_1(2)\gamma_2(3)\gamma_2(4)\gamma_1(1). \tag{3-3}$$

Each such integral we let correspond to a graph times the factor

$$\nu = \sigma \prod_{1\le i<j\le N}\prod_a\{q_a(i,j)!\} \tag{3-4}$$

The graph is obtained by associating to each $\gamma_b(i)$ a γ_b-circle and to each $B_a(i,j)$ an B_a-bond that joins the γ_b-circle corresponding to $\gamma_b(i)$ to the γ_b-circle corresponding to $\gamma_b(j)$. The γ_2-circles, which are associated with i's for which d(i)'s appear in (3-2) are colored black and are not labeled; this is consistent with their corresponding to dummy variables over which integration has been performed.[46] Each γ_1-circle is associated with an i over which no integration is performed. It is colored white and is labeled by i. In expression (3-4)

σ is the symmetry number of the graph. This is the order of the group of permutations of the black circles that leaves the connections of the graph invariant.[47] The number $q_a(i,j)$ is associated with each pair $\{i,j\}$ of circles in the graph. It is the number of single B_α-bonds joining that pair of circles and in a simple graph it is zero or one. An example of the correspondence between integrals and graphs is given by Fig. 1, which shows the graph corresponding to (3-3). The factor given by (3-4) is 48 for this graph.

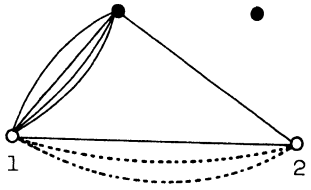


Figure 1. The graph associated with the expression (3-3). Dotted lines correspond to B_1-bonds and solid lines to B_2-bonds. White circles are γ_1-circles and black circles are γ_2-circles. The graph times 48 represents the integral in (3-3). The number 48 is obtained from (3-4).

We shall use the concept of the product of graphs. If Γ_1 is a connected graph with the n white 1-circles[48] labeled by $i_1,\ldots,i_n$ while Γ_2 is a connected graph with m white 1-circles labeled by $j_1,\ldots,j_m$ we consider the white 1-circles labeled by $k_1,\ldots,k_s$, where $k_1,\ldots,k_s$ are the labels common to both the set $\{j_1,\ldots,j_m\}$ and the set $\{i_1,\ldots,i_n\}$. The product of Γ_1 and Γ_2, which we shall

denote by $\Gamma_1 * \Gamma_2$, is defined as the graph Γ_3 containing the white 1-circles $k_1,\ldots,k_s$ such that the removal of the circles $k_1,\ldots,k_s$ will separate it into two pieces, one of which is Γ_2 minus its white circles, the other of which is Γ_1 minus its white circles. We describe this situation by saying that Γ_3 consists of Γ_1 and Γ_2 connected in parallel at the circles $k_1,k_2,\ldots,k_s$.[49] If either Γ_1 or Γ_2 contains no white circles, Γ_3 is a disconnected graph with Γ_1 and Γ_2 as its two components. The extension of the definition of the product to an arbitrary Γ_1 and Γ_2 is straightforward. It is important to note that if I_1 is an integral corresponding to the graph Γ_1 times the factor ν_1 (see Eq. (3-4)) and I_2 is an integral corresponding to $\nu_2\Gamma_2$ then the product of I_1 and I_2 does not necessarily correspond to $\nu_1\nu_2(\Gamma_1 * \Gamma_2)$ since ν_3, the factor corresponding to $\Gamma_1\Gamma_2$, will not necessarily be equal to $\nu_1\nu_2$. By product of I_1 and I_2 here we mean the algebraic object obtained by formally multiplying two integrals considered as algebraic objects according to the usual rules, without regard to the analytic nature of the functions for which integrals stand. Of course, if I_1 and I_2 are finite numbers or sufficiently well-behaved functions of the vectors $\underset{\sim}{r}_1, \underset{\sim}{r}_2$, etc., then it also makes sense to compare the values of I_1, I_2, and $I_1 I_2$. In this paper we shall not always explicitly state whether we are considering a graph as a geometrical object, an algebraic

object, the function for which it stands, or a particular value of that function for particular values of its arguments. We hope our meaning will be clear from the context in which our statements are made. Had we attempted a more precise treatment the result would have been either considerably more formal or considerably more cluttered.

Despite the annoying fact that sometimes $I_1 I_2/\nu_1\nu_2 \neq \Gamma_1 * \Gamma_2$, we have the following lemma,[50] which we shall use a great deal.

Let G be a set of distinct connected graphs, $\Gamma_1, \Gamma_2, \ldots, \Gamma_j, \ldots$ each consisting of some or no black circles, some or no white 1-circles, and some or no bonds. We assume that no products of members of G are in G. Then if F is the set consisting of all graphs in G and all products of graphs in G,

$$\text{The function represented by the sum of all graphs in F} = \exp\{\text{the function represented by the sum of all graphs in G}\} - 1. \tag{3-5}$$

The proof is immediate. A typical graph in F has the form $\varphi_i = \Gamma_1^{*p_1} * \Gamma_2^{*p_2} * \ldots * \Gamma_n^{*p_n}$ where $\Gamma^{*p} = \Gamma * \ldots * \Gamma$. Hence $\varphi_i = \dfrac{I_1^{p_1} I_2^{p_2} \ldots I_n^{p_n}}{\nu_{\varphi_i}}$, where ν_{φ_i} is the factor of φ_i, and I_i is the integral represented by $\nu_{\Gamma_i}\Gamma_i$. The factor of φ_i is

clearly equal to $p_1!p_2!\ldots p_n! \prod_{i=1}^{n} \left[{p_i \atop \nu_{\Gamma_i}} \right]$ so $\varphi_1 = \dfrac{I_1^{p_1} I_2^{p_2} \ldots I_n^{p_n}}{p_1!p_2!\ldots p_n! \prod_{i=1}^{n} \left[{p_i \atop \nu_{\Gamma_i}} \right]}$.

Hence summing up over all the φ_1 yields

$$\exp\left\{\frac{I_1}{\nu_{\Gamma_1}} + \frac{I_2}{\nu_{\Gamma_1}} + \ldots + \frac{I_i}{\nu_{\Gamma_i}} + \ldots\right\} - 1 = \exp\{\Gamma_1+\Gamma_2+\ldots+\Gamma_i+\ldots\} - 1,$$

which is the desired result.

Linear graphs provide a natural representation for a large class of functionals of functions of one and two variables, and hence it is to be expected that there are very simple and useful relationships between functional differentiation of such functions and certain elementary graphical operations. The relationships we need are embodied in two statements. The proofs are not trivial, but are nevertheless straightforward exercises in functional differentiation and we shall not reproduce them.[51]

Let Γ be a connected graph consisting of bonds and black circles such that some (or none) of the black circles are γ-circles and some (or none) of the bonds are B-bonds. Then

$$\frac{\delta^n \Gamma}{\prod_{1\le i\le n} \delta^n \gamma(i)} = \text{the sum of all the distinct graphs that are obtained from } \Gamma \text{ by changing n black } \gamma\text{-circles into white 1-circles labeled by 1,2,...,n, respectively.}$$

(3-6)

If n exceeds the number of black γ-circles, then the derivative will of course be zero.

We have also

$$\frac{\delta \Gamma}{\delta B(i,j)} = \frac{1}{2}\left[\text{the sum of all the distinct graphs that are obtained from } \Gamma \text{ by removing a B-bond from } \Gamma \text{ and changing the two circles at either end of B from black circles to white ones labeled i and j, respectively.}\right]$$

(3-7)

The extensions of these statements to the case of repeated differentiation with respect to a whole set of different $\gamma_\alpha(i)$ and $B_\alpha(i,j)$ can be given easily, but the statements above will suffice for our purposes.

SECTION 4

FUGACITY EXPANSIONS

We are now in a position to express the series we wish to consider in graphical terms. To start, we can write (2-3) as

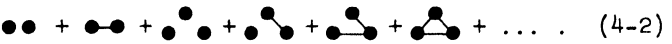
Ξ = 1 + the sum of all distinct simple graphs consisting of black z_1-circles and some or no e-bonds such that there is one e-bond between each pair of z_1-circles,[52]



(4-1)

This is not a very useful expansion; the magnitude of a graph with n circles is proportional to Ω^n of the system in in a container of volume Ω. To obtain a useful series we first rewrite Eq. (4-1) in terms of the Mayer f-function, $f(i,j) = e(i,j) - 1$. We have then

Ξ = 1 + the sum of all distinct simple graphs consisting of black z_1-circles and some or no f-bonds,



(4-2)

We can now express ln Ξ as a sum of graphs with z_1-circles and f-bonds by using (3-5) which can be applied to Eq. (4-2) to yield

$\ln \Xi$ = the sum of all distinct connected simple graphs composed of black z_1-circles and some or no f-bonds,



(4-3)

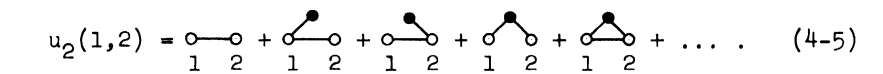
This is the fugacity expansion of the equation of state.

If we apply (2-9) and (3-6) to Eq. (4-3) we find

$u_n(1,\ldots,n)$ = the sum of all distinct connected simple graphs composed of n white z_1-circles labeled by $1,\ldots,n$, respectively, some or no black z_1-circles, and some or no f-bonds.

(4-4)

For example



(4-5)

From (4-4) it follows that $u_n(1,\ldots,n)/z^n$ at $z = 0$ is the sum of all distinct connected simple graphs composed of n white $e^{-\beta\varphi_1}$-circles labeled by $1,\ldots,n$, respectively, and f-bonds. Hence it further follows that

$$\frac{u_n(1,\ldots,n)}{z^n} = \left\{\frac{u_n(1,\ldots,n)}{z^n}\right\}_{z=0} + z\int \left\{\frac{u_{n+1}(1,\ldots,n+1)}{z^{n+1}}\right\}_{z=0} d(n+1)$$

$$+ \ldots + z^s \int \left\{\frac{u_{n+s}(1,\ldots,n+s)}{z^{n+s}}\right\}_{z=0} d(n+1)\ldots d(n+s) + \ldots .$$

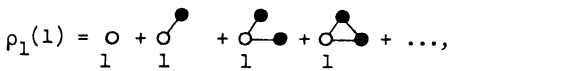
(4-6)

Using (2-8), we obtain from (4-4)

$\rho_n(1,\ldots,n) =$ the sum of all distinct simple graphs consisting of white z_1-circles labeled by $1,\ldots,n$, respectively, some or no black z_1-circles, and some or no f-bonds, such that there is at least one path from each black circle to a white circle.[53]

(4-7)

For instance



(4-8)

A somewhat more useful characterization of $\rho_n(1,\ldots,n)$ is given by the expression

$\rho_n(1,\ldots,n) =$ the sum of all distinct connected simple graphs consisting of n white z_1-circles labeled by $1,\ldots,n$, respectively, some or no black z_1-circles, some or no f-bonds, and an e-bond between every pair of white circles.

(4-9)

One can see that (4-9) is equivalent to (4-7) by replacing each e-bond in (4-9) by f + 1. If the sum of terms so obtained is expressed as a sum of graphs, all of whose bonds are f-bonds, the result is clearly (4-7). Equation (4-9) shows that for $n \geq 2$, as $z \to 0$, $\left\{ \prod_{1\leq i\leq n} z_1(i) \right\}^{-1} \rho_n(1,\ldots,n) \to \prod_{1\leq i<j\leq n} e(i,j)$. Since $e(i,j) = \exp\{-\beta\varphi_2(i,j)\}$, one can also see from (4-9) that $\rho_n = 0$ for $r_{ij} < a$, if the potential of interaction is infinite for $r_{ij} < a$. This property of ρ_n follows immediately from its meaning as a probability distribution.

SECTION 5

THE REDUCTION FROM GRAPHS WITH z_1-CIRCLES TO GRAPHS WITH ρ_1-CIRCLES

We have succeeded in expressing ln Ξ, u_n and ρ_n as functionals of f and z_1. We now consider the problem of expressing them as functionals of f and ρ_1. For many purposes ρ_1 is a more useful independent function than z_1 since its use yields expansions that are power series in ρ, the expected number density of the system, which is a parameter that is more easily associated with physical measurements than is the fugacity z.

In order to obtain expansions in ρ_1 we introduce the concept of an articulation circle, defined as follows. A circle, α, is an articulation circle if, upon its removal, the component of which it is a part, separates into two or more pieces in such a way that at least one piece contains no white circles.[49] The multiplicity of α is the number of pieces into which the component separates. A graph that is free of articulation circles we shall call 1-irreducible.

For convenience, we shall first obtain the expansion of h_n, defined by Eqs. (2-12) for $n \geq 2$ and (2-13) for $n = 1$. We note that

h_n = the sum of all distinct connected simple graphs consisting of white 1-circles labeled by 1,...,n, respectively, some or no black z_1-circles, and at least one f-bond, such that the graphs are free of white articulation circles.

(5-1)

For $n \geq 2$ this equation can be verified immediately by forming the product

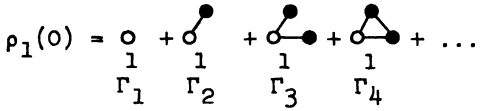
$$u_n(1,\ldots,n) = h_n(1,\ldots,n)\rho_1(1)\rho_1(2)\ldots\rho_1(n)$$

and substituting into the right-hand side of this expression the expansions of h_n and ρ_1 given by Eqs. (5-1) and (4-7) respectively. Carrying out the indicated multiplication yields (4-4). For n = 1, (5-1) follows from the application of Eq. (3-5) to (2-13), with u_1 given by (4-4).

A typical graph, G, in the f-bond, z_1-circle expansion of h_n given by (5-1) contains what we shall call a maximal 1-irreducible rooted subgraph, defined as the largest subgraph of G that contains all the white circles but contains no articulation circles. We shall denote this graph by G_m. It is clear that a unique G_m always exists for graphs in (5-1).

We note that graph G can be obtained by starting with some G_m and replacing each black circle of G_m by some graph, Γ_α, in the f-bond z_1-circle expansion of ρ_1 given by (4-7). The replacement can always be done so that the

graph Γ_α is attached to G_m by the white circle of Γ_α, which is first stripped of its label and blackened. It should be noted that we might require the use of the same graph in the expansion of ρ_1 for a number of replacements in order to get G from G_m. On the other hand, for a given G it will often be unnecessary to replace all the black circles of G_m to get G. However it is always possible to conceive of a replacement being made at <u>every</u> black circle (since one of the graphs in the expansion of ρ_1 is a single circle), and it turns out to be most convenient to think of getting G from G_m in these terms. An example of these graphical manipulations is illustrated by the graphs in Fig. 2. The G shown in that figure



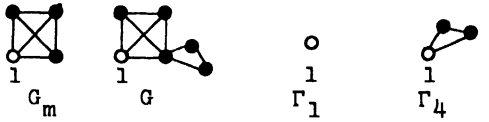


Figure 2. The graph labeled G is a typical graph in the series whose sum is $h_1(r)$ (see Eq. (5-1)). It can be obtained by replacing the black circles of G_m by the graphs Γ_1 and Γ_4 as indicated in the text. The graphs Γ_1 and Γ_4 are two graphs in the f-bond, z_1-circle expansion of ρ_1.

can be obtained from the G_m shown there by replacing two of the black circles of the G_m by Γ_1 and the other black circle of G_m by Γ_4 in the manner we have indicated. We shall hereafter speak of replacing the black circles of G by graphs $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$ in this manner as "hanging the graphs $\Gamma_1,\Gamma_2,\Gamma_3,\ldots$, on G." We now show that

> the sum of all graphs appearing in the expansion of h_n given by (5-1) that have the same G_m is equal in value to the graph, $\tilde{G}_m$, obtained by replacing all the black z_1-circles of G_m by ρ_1-circles (all the bonds are to remain f-bonds).
>
> (5-2)

To see this we first denote a typical graph that appears in expansion of ρ_1 given by (4-7) as Γ_α so that we can write this sum as

$$\rho_1(1) = \sum_\alpha \Gamma_\alpha(1). \qquad (5\text{-}3)$$

We consider the sum of integrals that result when we explicitly write the value of $\tilde{G}_m$ as $1/\sigma_m$ times an integral (where σ_m is the symmetry number of $\tilde{G}_m$) and substitute the sum given by (5-3) for $\rho_1(i)$ wherever the latter appears in this integral. In doing this we write each Γ_α itself as $1/\sigma_\alpha$ times an integral, where σ_α is the symmetry number of Γ_α. The result is a product of sums of integrals, which we can expand

into a single sum by doing all of the indicated multiplication. If the graph G_m has m black circles, then a typical term t in this new sum will be an integral I times $1/\sigma_m\sigma_{\alpha_1}\sigma_{\alpha_2}\ldots\sigma_{\alpha_m}$ where the integrand is just a product of $f(i,j)$'s and $z_1(i)$'s. This term t will correspond to a graph times some number, but the number will not necessarily be one, since I corresponds to a graph times a symmetry number and this symmetry number will not necessarily be the product $\sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}$. However in general there will always be other terms in the sum that are indistinguishable from this one in the sense that they are all products of the same number, $1/\sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}$, times integrals that differ from one another and I only by a different labeling of the dummy variables of integration. Suppose that there are a total of s such terms, including t. They can be summed up and written as a single term, $\sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}$, times an integral whose integrand is again just a product of f's and z_1's.

We claim that $\sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}/s$ is the symmetry number σ_I of the graph corresponding to I. To verify this we characterize σ_I in terms of a second labeling. We can do this by noting that if we consider the graph obtained from G_m by whitening its black circles and labeling them n+1,...,n+m, (the n white circles are already labeled 1 through n) then s is the number of distinguishable ways of labeling the circles n+1,...,n+m with the labels $\alpha_1,\alpha_2,\ldots,\alpha_m$. (The

labels α_i and α_j may be the same, for $i \neq j$, i.e., Γ_{α_i} may be the same graph as Γ_{α_j}. If all the Γ_{α_i}'s are different, $s = m!$ If all the Γ_{α_i}'s are the same, $s = 1$.) Now σ_I can be expressed as $\tilde{\sigma}\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}$ where $\tilde{\sigma}$ is the number of permutations among the $\alpha_1,\alpha_2,\ldots,\alpha_m$ (or equivalently, among the labels $n+1,\ldots,n+m$) that leave the doubly labeled graph invariant. But we have also the equation

$$s\tilde{\sigma} = \sigma_m. \tag{5-4}$$

This is simply an expression of the fact that the group of permutations among the labels $n+1,\ldots,n+m$ that leave the doubly labeled graph invariant is a subgroup of the larger group of permutations that leave the singly labeled graphs invariant, and the ratio of the orders of the two groups is just the number of distinguishable ways that the double labeling can be done.[55] Therefore, using (5-4) we can write $\sigma_I = \sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}$ as $\sigma_m\sigma_{\alpha_1}\ldots\sigma_{\alpha_m}/s$. Since $\sigma_I I$ corresponds to the graph obtained by hanging $\Gamma_{\alpha_1},\Gamma_{\alpha_2},\ldots,\Gamma_{\alpha_m}$ onto G_m we have shown that the value of the graph obtained by replacing all the black z_1-circles of G_m by ρ_1-circles equals the value of the sum of all the distinct graphs obtainable by hanging onto G_m the graphs in the f-bond, z_1-circle expansion of ρ_1. But this sum is clearly identical to the sum of all graphs appearing in (5-1) that have the same G_m that G has. Thus we have proved (5-2). From this it immediately follows that

$h_n(1,\ldots,n) =$ the sum of all distinct connected simple graphs consisting of white 1-circles labeled by 1,2,...,n, respectively, some or no black ρ_1-circles, and at least one f-bond, such that the graphs are free of articulation circles, i.e., are 1-irreducible.

(5-5)

We shall refer so often to the special case of $n = 1$ that we write it separately,

$h_1(1) = \ln \rho_1(1) - \ln z_1(1) =$ the sum of all distinct connected simple graphs consisting of a white 1-circle labeled 1, one or more black ρ_1-circles, and f-bonds, such that the graphs are free of articulation circles.

(5-6)

SECTION 6

THE DISTRIBUTION FUNCTIONS AS FUNCTIONALS OF ρ_1 AND f

From the expansion of h_n we can obtain the f-bond, ρ_1-circle expansion of g_n by using (2-8), since the relationship between the g_n and h_n is the same as that between the ρ_n and u_n. Therefore Eq. (5-5) immediately implies that

$$g_n(1,\ldots,n) = \text{the sum of all distinct simple graphs consisting of white 1-circles labeled } 1,2,\ldots,n, \text{ respectively, some or no black } \rho_1\text{-circles, and some or no f-bonds, such that there is a path from each black circle to white circle and the graphs are free of articulation circles, i.e., the graphs are root-connected and 1-irreducible.} \tag{6-1}$$

It is useful to re-express this sum in terms of graphs with e-bonds between white circles rather than f-bonds. The same step that takes one from (4-7) to (4-9) yields the desired expression from (6-1):

$$g_n(1,\ldots,n) = \text{the sum of all distinct, at-least-doubly-connected, simple graphs consisting of white 1-circles labeled } 1,2,\ldots,n, \text{ respectively, some or no black } \rho_1\text{-circles, some or no f-bonds, and an e-bond between each pair of white circles.} \tag{6-2}$$

Expansions of u_n and ρ_n are obtained from (5-5), (6-1), and (6-2) simply by letting the white circles in these equations be ρ_1-circles rather than 1-circles.

We note that (5-6) can be viewed as an integral equation for ρ_1 if z_1 and f are considered known functions. Using Eq. (3-5) we can rewrite this equation as

$z_1 =$ the sum of all distinct connected simple graphs composed of one white $\{(-1)^m \rho_1\}$-circle labeled 1, some or no black ρ_1-circles, and some or no f-bonds, such that the white circle is the only articulation circle. Here m = the number of bonds attached at the white circle, or the order of the circle.

(6-3)

Thus we have found z_1 in terms of ρ_1 and f. In the special case that ρ_1 defines a container of volume Ω but is otherwise zero, (6-3) reduces to the familiar expansion of z as a power series in ρ. In that case (6-3) is simply the inversion of the expansion that gives ρ as a power series in z (Eq. (4-9) with n = 1).

From Eq. (2-9) and the definition of h_1 (Eq. (2-12)) it follows that $h_1 = \ln\{\delta \ln \Xi/\delta z_1\}$ so the definition of c (Eq. (2-16)) can be rewritten as

$$c(1,2) = \frac{\delta h_1(1)}{\delta \rho_1(2)}. \qquad (6\text{-}4)$$

The effect of functionally differentiating with respect to ρ_1 each of the terms in (5-6) is found by applying Eq. (3-6). In order to express the result of the differentiation we observe that since each graph in (5-6) contains only one white circle, its characterization as articulation-circle free implies that it cannot be disconnected by the removal

of any one circle. Let us call a circle a connecting circle if the component of which it is a part is disconnected upon its removal.[56] Since the graphs in (5-6) are free of connecting circles, they are still free of connecting circles after being differentiated with respect to $\rho_1(1)$. Thus we have

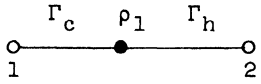
$$c(1,2) = \begin{array}{l}\text{the sum of all distinct simple graphs consisting}\\ \text{of two white 1-circles labeled by 1 and 2,}\\ \text{respectively, some or no black } \rho_1\text{-circles, and}\\ \text{one or more f-bonds, such that the graphs are}\\ \text{free of connecting circles.}\end{array} \tag{6-5}$$

Alternatively, we can characterize c(1,2) in terms of the somewhat different notion of a nodal-circle. A circle γ is a nodal circle [bridge point, cutting circle, node, s-circle] if there are at least two white circles in a component, neither of them γ, such that all the paths between that pair of white circles pass through γ. It is clear that a characterization of c(1,2) equivalent to (6-5) is given by the statement

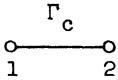
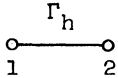
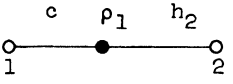
$$c(1,2) = \begin{array}{l}\text{the sum of all distinct simple graphs consisting}\\ \text{of two white 1-circles labeled by 1 and 2,}\\ \text{respectively, some or no black } \rho_1\text{-circles, and}\\ \text{f-bonds, such that the graphs are free of nodal}\\ \text{circles and of articulation circles.}\end{array} \tag{6-6}$$

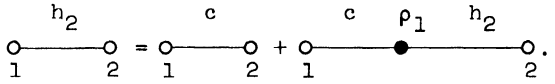
This characterization is the one that naturally arises when Eq. (2-16) is used as the definition of c(1,2) since if we split the sum of graphs in the f-bond, ρ_1-circle expansion

of h_2 into the subsum Σ_c of those that are free of nodal circles and the subsum, Σ_t of those that have at least one such circle, we see that any graph in Σ_t can be written as



(6-7)

where  is a graph in Σ_t and  is a graph in the f-graph in the f-bond, ρ_1-circle expansion of $h_2(1,2)$. As a result Σ_t itself equals the function  and we have



(6-8)

This is Eq. (2-15), which can be shown to be equivalent to (2-16) even more directly from the identity

$$\int d(2)\, \frac{\delta z_1(1)}{\delta \rho_1(2)} \frac{\delta \rho_1(2)}{\delta z_1(3)} = \delta(1,3). \tag{6-9}$$

The δ here is the Dirac δ-function. To see the equivalence of (2-15) and (2-16) we first rewrite (2-16) as

$$\frac{\delta z_1(1)}{\delta \rho_1(2)} = \frac{z_1(1)}{\rho_1(1)^2}\, \delta(1,2) - c(1,2)\, \frac{z_1(1)}{\rho_1(1)}. \tag{6-10}$$

But keeping in mind that $u_1 = \rho_1$ we also have from (2-9) and (2-12)

$$\frac{\delta\rho_1(2)}{\delta z_1(3)} = \frac{\delta}{\delta z_1(3)}\left\{z_1(2)\,\frac{\delta \ln \Xi}{\delta z_1(2)}\right\}$$

$$= \delta(2,3)\,\frac{\rho_1(2)}{z_1(2)} + \rho_1(2)\,\frac{\rho_1(3)}{z_1(3)}\,h_2(2,3). \quad (6\text{-}11)$$

Substituting (6-10) and (6-11) into (6-9) yields (2-15).

In the remainder of this section we shall consider a number of expressions that are of particular value in exhibiting certain asymptotic relationships that exist among the n-point distribution functions when the separation between subsets of the n points becomes great. The Kirkwood superposition approximation is the best known and perhaps most useful of these relationships. As its name implies, it has been used as an approximation to describe the relative behavior of the 2- and 3-point distribution functions at all separations and there is evidence that it does so surprisingly well, even where the three points are close together.[57] It can be written as

$$g_3(1,2,3) = g_2(1,2)g_2(1,3)g_2(2,3). \quad (6\text{-}12)$$

We shall introduce here potentials of average force, which are convenient functions to use for our purposes. For the rest of this section, we assume always that the subscript n is greater than 1. The potential of average force, $\psi_n(1,\ldots,n)$,

associated with $g_n(1,\ldots,n)$ is usually defined as $-\beta^{-1}\log g_n(1,\ldots,n)$. From this quantity we shall subtract the actual potential, V_n, multiply by β, and work with the function,

$$W_n(1,\ldots,n) = \ln g_n(1,\ldots,n) - \beta V_n(1,\ldots,n). \quad (6\text{-}13)$$

The function $W_n(1,\ldots,n)$ can be conveniently written as a sum of a set of graphs with two white circles plus the sum of a set with three white circles, etc. This double sum can be written as

$$W_n(1,\ldots,n) = \sum_{2\le s\le n} \sum_{\{i_1,\ldots,i_s\}\subset \mathfrak{N}} \omega_s(i_1,\ldots,i_s), \quad (6\text{-}14)$$

Here the $\mathfrak{N}$ under the Σ stands for the set of all s-tuples of the form $\{i_1,\ldots,i_s\}$ such that $i_i < i_j$ and $1 \le i_1 \le n$. Using Eq. (3-5) to obtain the logarithm of the expansion of $g(1,\ldots,n)/e^{-\beta V_n(1,\ldots,n)}$ given by (6-2), we find that $\omega_n(1,\ldots,n)$ can be characterized by the following expression

$\omega_n(1,\ldots,n)$ = the sum of all distinct connected simple graphs consisting of n white 1-circles labeled by $1,\ldots,n$ respectively, at least one black ρ_1-circle, and f-bonds, such that each graph is free of articulation circles. Furthermore each pair of black circles in a graph is connected by at least one path that does not pass through any white circle, while there is no path consisting of a single bond between any pair of white circles.

(6-15)

It follows easily by induction from (6-14) that

$$\omega_n(1,\ldots,n) = \sum_{2\le s\le n} \sum_{\{i_1,\ldots,i_s\}\subset n} (-1)^{s+n} W_s(i_1,\ldots,i_s). \tag{6-16}$$

Taking the exponent of (6-16), we can rewrite the resulting right-hand side in terms of g_s for $s \le n$ by using (6-13). The left hand side, exp ω_n, is easily obtained by using (6-15) and (3-5). For $n \ge 3$, we find

$$g_n(1,\ldots,n) / \prod_{s=2}^{n-1} \left[\prod_{\{i_1,\ldots,i_s\}\subset n} g_s(i_1,\ldots,i_s) \right]^{(-1)^{s+n+1}}$$

= 1 + the sum of all distinct simple graphs consisting of n white 1-circles labeled by 1,...,n, respectively, at least one black ρ_1-circle, and f-bonds, such that every graph is free of articulation circles. Furthermore, every black circle is connected to every white circle by at least one path that does not pass through any intermediate white circles, while there is no path consisting of a single bond between any pair of white circles. (6-17)

From (6-14), we can also write W_n in terms of $\omega_2 = W_2$ and ω_s for $3 \le s \le n$, where the ω's are given by (6-15). This means that by again using (3-5) and (6-15) we can express g_n for $n \ge 2$ entirely in terms of g_2, ρ_1, and f. We obtain the prescription:

$$g_n(1,\ldots,n)/\prod_{1\le i\le j\le n} g_2(i,j)$$

= 1 + the sum of all distinct simple graphs consisting of n white 1-circle labeled by 1,...,n, respectively, at least one black ρ_1-circle, and f-bonds, such that every graph is free of articulation circles. Furthermore, every black circle is connected to each of at least 3 white circles by a path that does not pass through any intermediate white circles, while there is no path consisting of a single bond between any pair of white circles.

(6-18)

The functions $g_n \to 1$ for $r_{ij} \to \infty$ where $1 \le i < j \le n$. As all the distances between the n points increase, the left-hand sides of (6-17) and (6-18) go to unity quite rapidly, the first expression approaching its limit faster than the second. If any single point, $\underset{\sim}{r}_\alpha$, is moved away from the others, (6-17) becomes unity in the limit $|\underset{\sim}{r}_\alpha| = \infty$ and its left-hand side is close to unity for large $r_{1\alpha}$ regardless of the relative positions of the other n-1 points. This is not true of (6-18) for $n > 3$.

Since the g's go to unity as their arguments are separated, the question is whether the left-hand sides of (6-17) and (6-18) approach unity fast enough compared to the g's themselves to provide useful expressions. There is some direct evidence[67] that this is the case for hard spheres; it is almost certainly the case for a wide variety of other

potentials, too. Order of magnitude calculations in terms of various parameters can be used to support the plausibility of this conclusion, although the results of any one such calculation are rather weak. For example, if we consider an interparticle potential $\varphi_2(i,j)$ that is identically zero only for $r_{ij} > R$, we can compare the ρ expansions of the numerator and denominator of the left-hand side of (6-17) for a configuration of n points, each pair of points separated by a distance of $(2i+1)R$, where i is some integer. Then both expansions start with terms of order ρ^{2i+1} and agree through terms of order less than ρ^{ni+1}. Similar results are easily obtained for other separations of points and for Eq. (6-18) as well as Eq. (6-17).

We note that in the expansions that appear in (6-6), (6-15), (6-17), and (6-18), the ρ_1-circles can be replaced by z_1-circles if the condition that the graphs be free of articulation circles is deleted.

Setting the right-hand side of (6-17) equal to unity yields a convenient generalization of the Kirkwood superposition approximation. Fisher and Kopeliovick[58] used this approximation for n = 4 in the Born-Green-Bogoliubov-Yvon hierarchy to obtain a pair of coupled equations for g_2 and g_3. However, these equations are rather complicated. Weissberg and Prager,

Ree, and Rowlinson independently found analytic expressions for the first term in the sum in (6-18) when $n = 3$ and $\varphi_2(1,2)$ is a hard sphere potential.[72] Each of the three expressions has a somewhat different form.

SECTION 7

THE FUNCTION ln Ξ AS A FUNCTIONAL OF ρ_1 AND f

We have expressed a number of relevant functions of one or more variables, i.e., the h_n, W_n, etc., in terms of $f(1,2)$ and $\rho_1(1)$. There remains the task of obtaining ln Ξ in these terms.

Perhaps the simplest way of doing this is due to Morita and Hiroike.[24] Their argument begins with Eq. (2-7). For s = 1, this equation can be written as

$$\delta \ln \Xi = \int d(1)\rho_1(1)\delta \ln z_1(1). \tag{7-1}$$

This in turn can be rewritten as

$$\delta \ln \Xi = \delta \int d(1)\rho_1(1)\ln z_1(1) - \int d(1)\ln z_1(1)\delta\rho_1(1). \tag{7-2}$$

Using (5-6), we have from this

$$\delta \ln \Xi = \delta \int d(1)\rho_1(1)\ln z_1(1) - \int d(1)\delta\rho_1(1)[\ln \rho_1(1) - \{\text{the sum on the right-hand side of (5-6)}\}]. \tag{7-3}$$

From (3-6), this means that aside from an additive function that is independent of $\rho_1(1)$,

$$\ln \Xi = \int d(1)\rho_1(1)\ln z_1(1) - \int d(1)\rho_1(r)\{\ln \rho_1(r) - 1\} + H[\rho_1,f],$$

where

$$H[\rho_1,f] = \text{the sum of all distinct connected simple graphs composed of black } \rho_1\text{-circles and one or more f-bonds such that the graphs are free of articulation circles.}$$

(7-4)

$H[\rho_1,f]/\beta$ is sometimes called the Helmhotz free energy of interaction or excess Helmholtz free energy.

The function that is independent of $\rho_1(1)$ must be zero, since when $\rho_1(i) = 0$, all the terms in (7-4) vanish. Thus, since $\ln \rho_1 - \ln z_1 = h_1$, we have

$$\ln \Xi = \int d(1)\rho_1(1) + H[\rho_1,f] - \int d(1)\rho_1(1)h_1(1). \quad (7\text{-}5)$$

Using (7-4) and (5-6) to characterize $H[\rho_1,f]$ and $h_1(1)$ respectively yields the desired expression for $\ln \Xi$ as a functional of f and ρ_1. When φ_1 defines the walls of a container of volume Ω but is otherwise zero, and $\Omega \to \infty$, then (7-5) becomes the familiar equation of state in terms of Mayer's irreducible cluster integrals, β_k. In fact, if the sum of graphs with k black vertices in (5-6) is denoted as $\rho^k B_k(1)$, then $B_k(1) \to B_k$ in this limit.

Somewhat different derivations of (7-5) have been given by Stillinger and Buff,[59] M. S. Green,[33] and deDominicis.[25] We shall consider the derivation of deDominicis in Section 10. Green's derivation is based on the same topological relationships as that of deDominicis, and although it was given only for the case of a uniform system it is general enough to be readily applied to a nonuniform system as well.

If one restricts oneself to a uniform system there is a great variety of simple alternative ways of obtaining (7-5). For example, we can start with the closed form of the equation of state in terms of z,

$$\frac{dp}{kT\,dz} = \frac{\rho}{z}, \tag{7-6}$$

or

$$\frac{p}{kT} = \int_0^\rho \eta \frac{d(\ln z)}{d\eta}\, d\eta.$$

Using this and (5-6) we obtain

$$\frac{p}{kT} = \rho - \sum_{k\geq 1} \frac{k}{k+1} \beta_k \rho^{k+1}. \tag{7-7}$$

Alternatively we can start with the equation, from fluctuation theory,

$$\frac{dp}{kT\,d\rho} = 1 - \rho \int c(1,2)d(1). \qquad (7\text{-}8)$$

Again a single integration with respect to ρ yields (7-7) if the sum in (6-5) is used to represent $c(1,2)$.

SECTION 8

BOND-REDUCTION OF GRAPHS THE HIGHER ORDER DISTRIBUTION FUNCTIONS IN TERMS OF ρ_1 AND h_2

We shall now express the distribution functions and grand partition function as functionals of ρ_1 and h_2. These expansions are of interest from several standpoints. First the expansion of $\ln \Xi$ can be used to establish a number of variational properties of a grand ensemble. For example, it implies that for prescribed z_1 and f, $\ln \Xi$ is stationary with respect to independent variations in ρ_1 and h_2. Furthermore the expansion of g_2 yields an interesting expression for $\varphi(1,2)$ as a functional of ρ_1 and h_2. Finally the expansions of g_n for $n \geq 3$ are free from any explicit reference to the potential or fugacity and so give a formal answer to the closure problem of completely characterizing the g_n for $n \geq 3$ in terms of ρ_1 and g_2.

We shall first look at the distribution functions. In the development of this section we closely mimic the development of Section 5, since the operations that lead from f-bond, ρ_1-circle expansions to h_2-bond, ρ_1-circle expansions can almost be put in a one-to-one correspondence with those that lead from the f-bond, z_1-circle expansions to the f-bond, ρ_1-circle expansions.

We start by introducing the concept of an articulation pair of circles [bifocal points]. The circles γ_1 and γ_2 constitute an articulation pair if upon their removal a component separates into two or more parts, one of which contains at least one circle but no white circle, another of which contains at least one circle. We shall call a graph that is free of articulation circles and free of articulation pairs a 2-irreducible graph.

Let G be a typical graph in the expansions (6-15), (6-17), or (6-18), all for $n \geq 3$. We consider the largest subgraph of G that is free of articulation circles and contains all the white circles of G. We shall call this the maximal 2-irreducible rooted subgraph of G and denote it by G_m. It is clear that every graph in the f-bond, ρ_1-circle expansion under consideration possesses a unique G_m.

The graph G can be obtained by taking G_m and, roughly speaking, replacing each bond of G_m by some graph in the f-bond, ρ_1-circle expansion of h_2 given by Eq. (5-5). More precisely, we note that a typical bond B in G is touching two circles, γ_1 and γ_2, one on each of its ends. We consider the process of replacing B, γ_1, and γ_2 by a graph Γ_α in the f-bond, ρ_1-circle expansion of h_2 so that γ_1 is replaced by one of the white circles of Γ_α and γ_2 is replaced by the other. The white circles of Γ_α must be delabeled and if γ_1 or γ_2 is black the circle replacing it is blackened. We shall hereafter refer to

this whole operation, which involves appropriately delabeling and recoloring γ_1 and γ_2, simply as "the replacement of a bond in G by a graph."

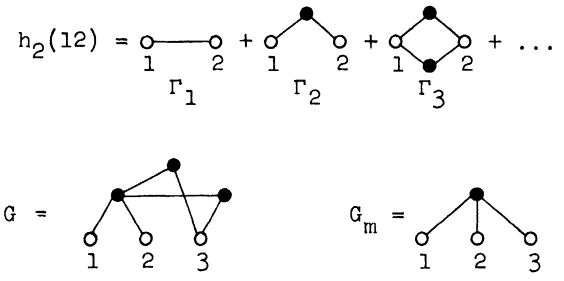


Figure 3. The graph G is obtained from G_m by the replacement of two of its bonds by Γ_1 and one of its bonds by Γ_3, as described in the text.

To get from G_m to G we successively make such a replacement for every bond of G_m and the order in which we replace the bonds is clearly immaterial. Some of the replacements, however, need only be the trivial ones that consist of replacing a single bond between two circles by a single bond between 2 circles, which is indeed one of the graphs in the expansion of h_2 given by Eq. (5-5).

As an example, we show a G in Fig. 3 that is obtained from G_m, also shown, by the replacement of two of its bonds by Γ_1 and one of its bonds by Γ_3, which are two typical graphs in the expansion of h_2.

We claim now that

> the sum of all graphs appearing in either (6-15), (6-17) or (6-18) that have the same G_m is equal in value to the graph obtained by letting all the bonds of G_m be h_2-bonds. All the black circles are to remain ρ_1-circles and the white circles, 1-circles.

(8-1)

It is probably clear that we can use the same sort of argument to show this as that used to prove (5-2). It turns out, in fact, that we can use an argument that is identical in structure provided that we set up an appropriate new correspondence between the permutation groups considered in the proof and the graphical objects to which the groups refer. Thus in our discussion we must replace the phrase black circle by bond, etc. and the phrase hanging $\Gamma_1,\Gamma_2,\ldots,$ on the G by replacing the bonds of G by $\Gamma_1,\Gamma_2,\ldots,$. Most such changes are obvious. The only nontrivial one lies in the nature of the double labeling used. This time each label in the second set of labels is attached to a bond of G_m rather than a circle and it has a direction as well as an index. Thus a typical label will be $\alpha(\rightarrow)$ or $\alpha(\leftarrow)$. If $\alpha(\rightarrow)$ stands for the graph $\Gamma_\alpha(i,j)$, then $\alpha(\leftarrow)$ stands for $\Gamma_\alpha(j,i)$ and $\alpha(\rightarrow)$ is a different label from $\alpha(\leftarrow)$ or not depending on whether or not the graph $\Gamma_\alpha(i,j)$ is different from $\Gamma_\alpha(j,i)$.

The statement (8-1) immediately implies that for $n \geq 3$, we can recharacterize the functions that are given by the expansions (6-15), (6-17) and (6-18) by changing f to h_2

and adding the condition that each graph be free of articulation pairs. This result for Eq. (6-18) implies that we can modify (6-2) by changing e to g_2 as well as f to h_2 and changing the phrase at-least-doubly-connected to at-least-triply-connected. The resulting restatement of (6-2) provides what is perhaps the most striking way of characterizing g_n in terms of ρ_1 and h_2.

It should also be noted that for $n \geq 3$ we can recharacterize the functions given by (5-5) and (6-1) by changing the f-bonds that appear in those expansions to h_2-bonds and adding the condition that each graph be free of articulation pairs of circles. The result for (6-1) follows from our results for Eq. (6-18) if the identity $g_2 - 1 = h_2$ is used, in the same way that (6-1) as it stands follows from (6-2) when the identity $e - 1 = f$ is used. The result for (5-5) then follows from the use of Eq. (2-8), which relates g_n and h_n through (2-11) and (2-12).

SECTION 9

THE FUNCTION g_2 IN TERMS OF ρ_1 AND h_2

Since a typical graph in the expansion for $\omega_2(1,2)[= W_2(1,2)]$ given by (6-15) need not contain a unique maximal 2-irreducible rooted subgraph G_m, we cannot use the preceding argument directly when $n = 2$ to obtain h_2-bond, ρ_1-circle expansions. However, it is clear that any graph that is free of nodal circles in the expansion of W_2 given by (6-13) will contain a unique G_m. Therefore, let us first subtract from the f-bond, ρ_1-circle expansion of $W_2(1,2)$ all those graphs containing nodal circles. The remaining sum, call it Σ_b, defines a function $b(i,j)$ that can easily be expressed in terms of h_2-bonds. Applying to Σ_b the same argument that we used to get from the f-bond to the h_2-bond expansion of ω_n for $n \geq 3$ we find immediately that

$$b(1,2) = \begin{array}{l}\text{the sum of all distinct connected simple graphs consisting of two white 1-circles labeled 1 and 2, respectively, black } \rho_1\text{-circles, and } h_2\text{-bonds, such that the graphs are free of articulation circles, nodal circles, and articulation pairs of circles. Furthermore there is no bond between the two white circles.}\end{array} \qquad (9\text{-}1)$$

These graphs have been called basic, prototype, and irreducible in a wider sense.

We now note that the subset of graphs that we subtracted from $W_2(1,2)$ to get $b(1,2)$ is the same set as the subset of graphs in the sum used to define $t(1,2) = h_2(1,2) - c(1,2)$ in Section 6, since both subsets are characterized by the presence of a nodal circle. (It is true that they are subsets of different sets, but it follows from the characterization of the f-bond expansions of h_2 and W_2 that their difference cannot include graphs containing nodal circles.) Therefore we have

$$W_2(1,2) = t(1,2) + b(1,2) \tag{9-2}$$

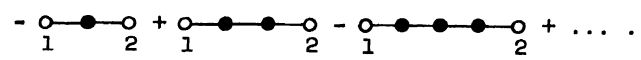
or

$$W_2(1,2) = h_2(1,2) - c(1,2) + b(1,2). \tag{9-3}$$

Hence if we can characterize $t(1,2)$ or, equivalently, $c(1,2)$ as a functional of ρ_1 and h_2 we shall have a complete description of $W_2(1,2)$ as a functional of ρ_1 and h_2. Happily, $c(1,2)$ is one of the simplest of all functions to express in terms of ρ_1-circles and h_2-bonds. This follows directly from Eq. (2-15), which can be solved for $c(1,2)$ by iteration to give

$c(1,2) =$ the sum of all distinct connected simple graphs consisting of two white 1-circles labeled 1 and 2, respectively, some or no black $(-\rho_1)$-circles, and at least one h_2-bond, such that every graph is a <u>chain</u> of black $(-\rho_1)$-circles and h_2-bonds on either end of which is a white 1-circle. (9-4)

Pictorially, in terms of h_2-bonds and black ρ_1-circles, we have

$c(1,2) =$  $+ \ldots$.

It is worthwhile pointing out that these equations enable us to formally answer the question: Given the 1- and 2-point distribution functions, what interparticle potential gave rise to them? Since from the definition of W_2,

$$-\beta\varphi_2 = \ln[h_2+1] - W_2, \qquad (9\text{-}5)$$

the answer is provided by Eq. (9-3) used with (9-1) and (9-4). Equation (9-1), (9-4) and (9-5) can be combined into a single expression that can be viewed as an exact integral equation for h_2. When the approximation $b = 0$ is made, it becomes the hypernetted chain equation.

SECTION 10

THE FUNCTIONS h_1 AND $\ln \Xi$ IN TERMS OF ρ_1 AND h_2

To obtain the h_2-bond, ρ_1-circle expansions of h_1 and $\ln \Xi$ we shall use the method of deDominicis.[25] A somewhat different approach to finding these expansions was taken by Morita and Hiroike,[24] who based their method on an elaboration of the same techniques that they used to obtain Eq. (7-5), which gives $\ln \Xi$ as a functional of ρ_1 and f_1. We derived that equation using their approach; in passing, we shall rederive it in this section using the approach of Green[33] and deDominicis.

Suppose we are given a function, $L[z_1,f]$ expressed as the sum of all distinct graphs that satisfy some condition where each graph has z_1-circles and f-bonds. In obtaining the relationships used by deDominicis we must consider the general problem of generating various functions that can be expressed as sums of the same graphs that describe $L[z_1,f]$, but not necessarily sums in which each graph appears exactly once.

Our interest in this problem is motivated by the observation that in both the f-bond, ρ_1-circle expansion and h-bond, ρ_1-circle expansion of $\ln \Xi$, a typical graph will appear more than once.[60] Thus we cannot hope to find expansions that are as simple as those for the ω_n when $n \geq 2$, but we can try to express $\ln \Xi$ as a linear combination of functions of the type we wish to generate. The $L[z_1,f]$ with which we deal will be the z_1-circle f-bond expansion of $\ln \Xi$, given by the sum of all distinct unlabeled connected simple graphs.

There is one obvious way of generating from $L[z_1,f]$ a function of the type we want. We can functionally differentiate with respect to $z_1(1)$, multiply the result by $z_1(1)$ and integrate with respect to $d\underset{\sim}{r}_1$. It is clear from Eq. (3-6) that the resulting function,

$$\int \frac{\delta L}{\delta z_1(1)} z_1(1)d(1) = \int \rho_1(1)d(1), \qquad (10\text{-}1)$$

can be expressed as a sum of graphs such that each graph that appears in $L[z_1,f]$ once will appear $N(\gamma)$ times where $N(\gamma)$ is the number of circles contained in the graph.

We can also differentiate twice, or many times, since

$$\int d(1)\ldots d(s) \frac{\delta^s L[z_1,f]}{\prod\limits_{1\le i\le s} z_1(i)} \prod_{1\le i\le s} [z_1(i)] = \int d(1)\ldots d(s)\rho_s(1,\ldots,s) \qquad (10\text{-}2)$$

is also a linear combination of the graphs in $L[z_1,f]$. In particular, from Eq. (3-6) it can be seen that in

$$\int d(1)d(2)z_1(2) \frac{\delta^2 L[z_1,f]}{\delta z_1(1)\delta z_1(2)} = \int d(1)d(2)\rho(1)\rho_1(2)h_2(1,2) \qquad (10\text{-}3)$$

each graph in $L[z_1,f]$ appears $N(\pi)$ times, where $N(\pi)$ is the number of ordered pairs of circles contained in the graphs. Differentiating with respect to $f(i,j)$ instead of $z_1(i)$ yields a similar result; the function

$$2\int d(1)d(2)\,\frac{\delta L[z_1,f]}{\delta f(1,2)}\,\{f(1,2)+1\} \qquad (10\text{-}4)$$

is also equal to $\int \rho_1(1)\rho_1(2)\rho_1(1)\rho_1(2)h_2(1,2)$ and so is expressible in terms of the graphs in $L[z_1,f]$ as is

$$\int d(1)\ldots d(2s)\,\frac{\delta^s L[z_1,f]}{\prod_{\substack{1\le n\le s\\ i=2n-1;\ j=2n}}\delta f(i,j)}\prod_{\substack{1\le n\le s\\ i=2n-1;\ j=2n}}[f(i,j)+1]. \qquad (10\text{-}5)$$

There is still another way of getting the sort of function in which we are interested. We can find a sum of graphs $S[\gamma,B]$ with γ-circles and B-bonds, such that the z_1-circle, f-bond expansion of the function defined by $S[\gamma,B]$ consists of a linear combination of the graphs in $L[z_1,f]$. One such sum is the sum of all graphs in $L[z_1,f]$ that are free of articulation circles. We denote the sum as $H[z_1,f]$. The sum $H[\rho_1,f]$ defined in Eq. (7-4) is expressible as a sum of all the graphs appearing in $L[z_1,f]$ and each graph in $L[z_1,f]$ appears $N(H)$ times in $H[\rho_1,f]$, where $N(H)$ is the number of maximal 1-irreducible subgraphs in the graph. We note that

$$\frac{\delta H[\rho_1,f]}{\delta\rho_1(1)} = h_1(1). \qquad (10\text{-}6)$$

Another example of such a sum is provided by $W[\rho_1,h_2]$, consisting of the sum of all connected black ρ_1-circle, h_2-bond graphs that are 2-irreducible, (i.e., free of articulation circles and articulation pairs) minus the sum of all polygons[61] consisting of h_2-bonds and four or more black $(-\rho_1)$-circles. This perhaps curious sounding sum arises in a natural way in this context because it serves as a generating functional for the ρ_1-circles, h_2-bond expansion of W_2. To be more specific,

$$\frac{\delta W[\rho_1,h_2]}{\delta h_2(1,2)} = \tfrac{1}{2}\rho_1(1)\rho_1(2)W_2(1,2)$$

$$= \tfrac{1}{2}\rho_1(1)\rho_1(2)[b(1,2) + t(1,2)]. \qquad (10\text{-}7)$$

The number of times each graph in $L[z_1,f]$ appears in the f-bond, z_1-circle expansion of the function defined by $W[\rho_1,h_2]$ has been obtained by deDominicis. To characterize this number, call it $N(W)$, we assign the weight $\binom{2}{c} - \binom{1}{c} + \binom{0}{c} = \frac{c^2-3c+2}{2}$ to each minimal polygon containing c circles that we find in the graph and a weight of unity to each maximal 2-irreducible subgraph. We then sum over the weights of all such subgraphs of both types. The resulting number is $N(W)$.

Finally we can consider various combinations of these ways of generating functions of the sort we are discussing. For instance, since $H[\rho_1,f]$ can be described in terms of the same graphs that define $L[\rho_1,f]$, we can also so describe

$$\tilde{H}[\rho_1,f] = \int \frac{\delta H[\rho_1,f]}{\delta\rho_1(1)}\,\rho(1)d(1). \qquad (10\text{-}8)$$

It can be verified easily that the number of times a graph in $L[z_1,f]$ appears in the f-bond, z_1-circle expansion of $\tilde{H}[\rho,f]$ is equal to $N(\mu\gamma)$, the total number of articulation circles contained in the graph where each circle is weighted by its multiplicity, which is the number of pieces the graph falls into if the circle is removed. A circle that is not an articulation circle can be thought of as having multiplicity one. Similarly

$$2\int d(1)d(2)\,\frac{\delta W[\rho_1,h_2]}{\delta h_2(1,2)}\,[h_2(1,2)+1]$$

$$= \int d(1)d(2)\rho_1(1)\rho_1(2)W_2(1,2)[h_2(1,2)+1]$$ can be described as a sum of graphs in $L[z_1,f]$. The number of times a graph in the f-bond, z_1-vertex expansion of this function appears is $N(\mu p)$, the total number of ordered articulation pairs, each pair weighted by its multiplicity, where the multiplicity is the number of nontrivial pieces (i.e., pieces containing at least one circle) that separate from one another when the pair is removed. A pair of circles that is not an articulation pair can be assigned multiplicity one.

We have been able to identify the number of times graphs appear in the expansions of various functions with numbers that can be assigned to the graphs on the basis of their topolological structure. It has been found that there exist certain extremely simple relationships among these numbers and therein lies their great utility for our purposes.

The relevance of one such relationship to the equation of state of an imperfect gas has been pointed out by several authors.[25,33] We can write it as

$$N(\gamma) - N(\mu\gamma) + N(H) = 1. \qquad (10\text{-}9)$$

On the basis of our above discussion we can immediately translate (10-9) into a statement relating $L[z_1,f]$, $\int \rho_1(1)d(1)$, $H[z_1,f]$ and $\tilde{H}[\rho,f]$. This statement turns out to be Eq. (7-5).

Similarly, for any connected unlabeled graph, deDominicis[25] has pointed out that

$$N(H) = \tfrac{1}{2}[N(\pi) - N(\mu\pi)] + N(W). \qquad (10\text{-}10)$$

From our discussion we see that this implies that

$$H[\rho_1,f] = \frac{1}{2}\int d(1)d(2)\rho_1(1)\rho_1(2)[h_2(1,2) - W_2(1,2)\{1 + h_2(1,2)\}] + W[\rho_1,h_2].$$

Functionally differentiating with respect to ρ_1 while holding h_2 fixed and using (10-6), we obtain

$$h_1(1) = \int d(2)\rho_1(2)\{h_2(1,2) - [1 + h_2(1,2)]W_2(1,2)\} + \frac{\delta W[\rho_1,h_2]}{\delta\rho_1(1)}. \tag{10-12}$$

Since we already know the ρ_1-circle, h_2-bond expansion of $W_2(1,2) = t(1,2) + b(1,2)$ from the work of Section 8, (10-12) gives the ρ_1-circle, h_2-bond expansion of h_1.

Combining (10-11) and (7-5), we have

$$\ln \Xi = \int d(1)\rho_1(1) - \int d(1)\rho_1(1)h_1(1)$$

$$+ \tfrac{1}{2}\int d(1)d(2)\rho_1(1)\rho_1(2)\{h_2(1,2) - [1 + h_2(1,2)]W_2(1,2)\} + W[\rho_1,h_2]. \tag{10-13}$$

With the aid of (10-12) we can readily express all the functions in (10-13) as sums of graphs with ρ_1-circles and h_2-bonds. Thus we have succeeded in eliminating z_1 and f from our expressions for h_1 and ln Ξ, which are now given completely in terms of ρ_1 and h_2.

These expressions lend themselves to simple proofs that the ensemble possesses a number of interesting variational properties. For example, let us consider the right-hand side

of (10-13) as a functional of the independent functions ρ_1 and h_2, where z_1 and f are fixed, prescribed functions. Since z_1 and f_2 are fixed, $h_1 = \ln\{\rho_1/z_1\}$ and $W_2 = \ln\{h_2 + 1/(f_2+1)\}$ vary independently as ρ_1 and h_2 do. Therefore, from (10-13) we see that the values of ρ_1 and h_2 that leave the functional stationary are given by

$$\ln\{\rho_1/z_1\} = \int d(2)\rho_1(2)\{h_2(1,2) - [1 + h_2(1,2)]\}W_2(1,2)\} + \frac{\delta W[\rho_1,h]}{\delta\rho_1(1)},$$

$$\ln\left\{\frac{h_2(1,2) + 1}{f(1,2) + 1}\right\} = \frac{2\delta W[\rho_1,h_2]}{\delta h_2(1,2)}. \qquad (10\text{-}14)$$

The stationary value of the functional itself is hence just ln Ξ.

Finally we note that the entropy of the ensemble can be written as

$$S = \ln \Xi - (\ln z)\int d(1)\rho_1(1) + \beta\left\{\int d(1)\rho_1(1)\varphi_1(1) + \tfrac{1}{2}\int d(1)d(2)\rho_2(1,2)\varphi_2(1,2)\right\}$$

$$= \ln \Xi + \int d(1)\rho_1(1)\ln z_1(1) + \tfrac{1}{2}\int d(1)d(2)\rho_1(1)\rho_1(2)[h_2(1,2) + 1]\beta\varphi_2(1,2). \qquad (10\text{-}15)$$

Therefore S can be expressed completely in terms of ρ_1 and h_2 since ln Ξ, $z_1(1)$, and $\beta\varphi_2(1,2)$ can be so expressed. The result can be written as

$$S = \int d(1)\rho_1(1)[1 - \ln \rho_1(1)] + \tfrac{1}{2}\int d(1)d(2)\rho_1(1)\rho_1(2)[h_2(1,2) + 1]$$

$$\times\{\ln[1 + h_2(1,2)] + h_2(1,2)\} + W[\rho_1,h_2].$$

SECTION 11

THE GRAPHS ASSOCIATED WITH THE PERCUS-YEVICK AND HYPERNETTED CHAIN EQUATIONS

At the end of Section 9 we noted that the hypernetted chain approximation is obtained by setting $b = 0$, where b is the sum of graphs in ω_2 that are free of articulation circles, nodal circles, and articulation pairs of circles, i.e., the basic graphs.

Letting $b = 0$ is certainly an attractive procedure to consider, since b contains all the integrals that are intractably difficult to evaluate. It is clear that it will be an adequate approximation for small ρ, simply because $b = O(\rho^2)$. It is also clear that it yields a relationship between ω_2 and t that becomes asymptotically exact for large r_{12} in much the same sense that the Kirkwood superposition approximation becomes exact for large values of the separation between any pair of the triplet of points for which $g_3(1,2,3)$ is considered.[62] Thus, despite the lack of intuitive probabilistic meaning associated with t, one can justify setting b equal to zero on grounds similar to those that have been used to justify the Kirkwood superposition approximation, which we shall abbreviate as KSA for the rest of this article. However, it is not at all clear from these considerations whether the equation for g_2 obtained by setting $b = 0$ will yield a good or a bad approximation for moderate to large ρ.

Such an approximation can be tested by comparing it and the equations of state that result from its use with exact and near-exact results obtained by other methods. These consist of actual experimental results, Monte Carlo and molecular dynamical calculations done with fast computing machines, and analytic and machine evaluations of the cluster integrals in the first few terms of the f-bond, ρ_1-circle expansions of $p\beta/\rho$ and g_2. Rather extensive comparisons between the results of the HNC approximation and these exact and near-exact results have been made,[63] and they indicate that although the HNC approximation is not an outstandingly good approximation for any particular potential or density, it is an adequate one for a variety of potentials over a large range of number densities.

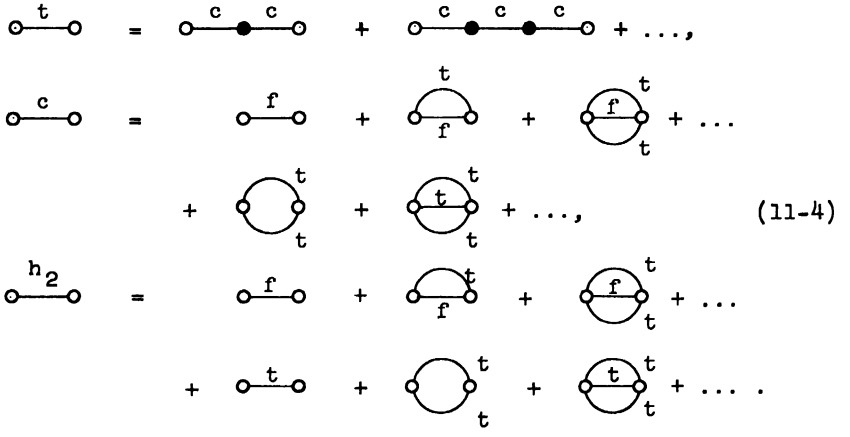
We shall now examine the structure of this approximation in some detail. The condition that b = 0 in Eq. (9-3), along with Eq. (9-5) and (6-8), can be rewritten as

$$t(1,2) = \int d(3)c(1,3)\rho(3)c(3,2) + \int d(3)c(1,3)\rho(3)t(3,2), \tag{11-1}$$

$$c(1,2) = f(1,2)\exp t(1,2) + \{\exp t(1,2)\} - 1 - t(1,2), \tag{11-2}$$

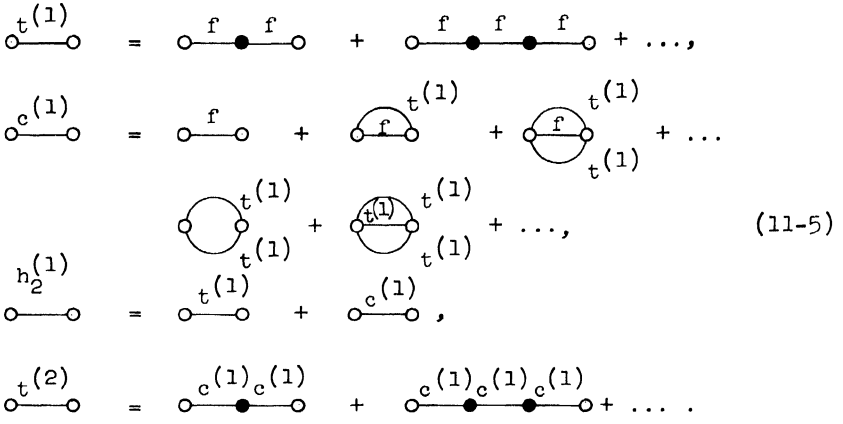
$$h_2(1,2) = t(1,2) + c(1,2). \qquad (11\text{-}3)$$

If (11-3) is considered as an integral equation for t(1,2) its formal solution can be written as an infinite sum of chain graphs with c-bonds. The right-hand side of (11-2) can also be expressed as an infinite sum of graphs. We can therefore express (11-1)-(11-3) pictorially as shown in (11-4). We drop the labels 1 and 2 and always let the black circles stand for ρ_1 and the white circles stand for unity.



(11-4)

We can now iterate by letting $t^{(0)} = 0$. Then $c^{(0)} = f$, $h_2^{(0)} = f$ and



(11-5)

If it exists, $\lim h_2^{(n)}$ will be the sum of all those graphs that can be obtained by such "series and parallel" summation of f-bonds.

Equations (11-1) and (11-2) can be written as a single nonlinear integral equations for t. An equivalent equation for g_2 can also be written. It is

$$g_2(1,2) = \exp\Big[-\beta\varphi_2(1,2) + \int\{g_2(1,3) - 1 - \ln g_2(1,3) - \beta\varphi_2(1,3)\}\{g(2,3) - 1\}\rho(3)d(3)\Big]. \tag{11-6}$$

An approximation that is considerably superior to the HNC approximation for most potentials and densities for which comparison with exact results have been made is[63] the Percus-Yevick approximation, which can be obtained by setting

$$\ln W_2 - [t+1] = 0. \qquad (11\text{-}7)$$

We shall denote the left-hand side of (11-7) as d(1,2). For most potentials of interest d(1,2) seems to be smaller than b(1,2).[64] Furthermore the PY approximation is also easier to handle analytically than the HNC approximation. For example, the PY approximation to the direct correlation function can be obtained in closed form for a hard-sphere system,[65] and the cluster expansion of the solution of the PY equation is simpler than the expansion for the solution of the HNC equation.

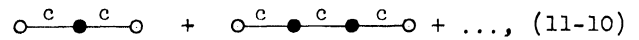
To compare the cluster expansions associated with the two equations we give first the exact expansions for d. The f-bond expansion is

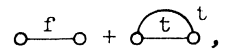
d(1,2) = the sum of all distinct connected simple graphs consisting of two white 1-circles labeled by 1 and 2, respectively, black ρ_1-circles, and f-bonds, such that no graph contains an f(1,2)-bond or a nodal circle.

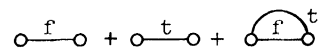
The h_2-bond expansion of d(1,2) is given by

$d(1,2) =$ the sum of all distinct connected simple graphs consisting of the two white 1-circles labeled by 1 and 2 respectively, black ρ_1-circles, and h_2-bonds, such that every graph is either a basic graph or a product of two or more subgraphs connected in parallel at the white circles. Each such subgraph is either a basic graph or $(-1)^s$ times a chain composed of $s \geq 2$ h_2-bonds. (11-9)

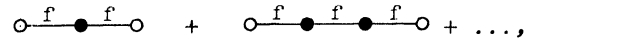
If we assume that d = 0, we obtain, instead of (11-1), (11-2) and (11-3),

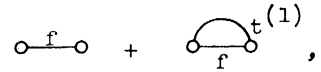
$t =$  (11-10)

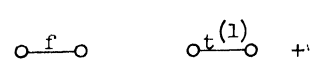
$c =$  (11-11)

$h_2 =$  (11-12)

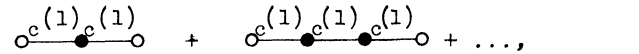
We can consider the iteration, $t^{(0)} = 0$, $c^{(0)} = f$, $h_2^{(0)} = f$ and, instead of (11-5), we have

$t^{(1)} =$ 

$c^{(1)} =$ 

$h_2^{(1)} =$ 

(11-13)

$t^{(2)} =$ 

etc.

For $\lim_{n\to\infty} h_2^{(n)} = h_2$ we have

$h_2(1,2) =$ a line plus all distinct planar graphs such that each is obtainable by following the prescription: take a convex polygon, color a pair of contiguous vertices white, and label them 1 and 2, respectively. Then delete or don't delete the line between 1 and 2 and add some or no nonintersecting straight lines between any vertex. Lines are f-bonds, black vertices are ρ_1-circles, and the labeled white vertices are 1-circles.

(11-14)

Equations (11-10), (11-11), and (11-12) can be rewritten as a single equation for either t or g_2. The equation for g_2 is

$$g_2(1,2) = e^{-\beta\varphi_2(1,2)}\left[1 + \int\left\{f(1,3)g(1,3)e^{+\beta\varphi_2(1,3)}\right\}\left\{g(2,3) - 1\right\}\rho_1(3)d(3)\right]. \tag{11-15}$$

It should be noted that in the case of potentials with hard cores of diameter σ, prescribing c for $r_{12} > \sigma$ fully determines c(1,2) for $r_{12} < \sigma$ if g(1,2) is required to be zero for $r_{12} < \sigma$. This can be seen from the expressions

$$\text{PY} \qquad c = f[t+1] \tag{11-16}$$

and

$$\text{HNC} \qquad c = e^{-\beta\varphi+t} - [t+1]. \tag{11-17}$$

For $r_{12} < \sigma$, Eqs. (11-16) and (11-17) simply state that $c = c$ if we require that $g_2 = 0$ for $r_{12} < \sigma$. As a result, for potentials with hard cores the assumption that $b = 0$ or $d = 0$ should be replaced by the weaker assumption that eb or ed are zero, i.e., that b and d are undetermined for $r < \sigma$. The use of the stronger assumptions will make no difference in the values obtained for g_2, but their use will lead to poor values of $W_2(1,2)$ for $r_{12} < \sigma$,[66] whereas assuming that b or d are undetermined for $r_{12} < \sigma$ will result in the impossibility of finding $W_2(1,2)$ for $r_{12} < \sigma$.

Equations (11-16), and hence the PY approximation, is exact for a uniform system of hard spheres in one dimension (hard rods).

SECTION 12

THE GRAPHICAL ANALYSIS OF EFFECTIVE DENSITY TECHNIQUES

A number of useful relationships among probability distribution functions can be derived by viewing the pair potential between a given particle of a system and the rest of the particles as an external field acting on these particles. The given particle, call it α, need not be distinguished from its fellows in any way except conceptually. On the other hand it is sometimes useful to assume that it is distinguished in a real way, interacting with any other particle 1 through a parameterized potential $v(1,j)$ between any pair of other particles. If this assumption is made, it is usually convenient to further assume that $v_\lambda(1,\alpha)$ can vary continuously from $v_0(1,\alpha) = 0$ to $v_1(1,\alpha) = v(1,\alpha)$ as α is "turned on" into the system, i.e., as λ is changed from 0 to 1. It is also often useful to consider introducing a whole set of distinguished particles, $\{\alpha,\beta,\ldots,\nu\}$.

Let us write $\rho_1(1;\alpha)$ for the conditional one-particle density function that measures the probability of finding particle 1 somewhere in a system if the location of particle α is known. We have, from (2-11) and the definition of conditional probability,

$$\rho_1(1;\alpha) = \frac{\rho_1(1,\alpha)}{\rho_1(\alpha)} = \rho_1(1)g_2(1,\alpha). \qquad (12\text{-}1)$$

Similarly for the general conditional probability density we have

$$\rho_n(1,2,\ldots,n;\,\alpha,\beta,\ldots,\nu) = \frac{\rho_{n+\nu}(1,2,\ldots,n,\alpha,\beta,\ldots,\nu)}{\rho_\nu(\alpha,\beta,\ldots,\nu)}. \tag{12-2}$$

If the potential is assumed to be a sum of one- and two-body terms it is a simple matter to modify the prescriptions for such functions as $h_1(1) = \ln\{\rho_1(1)/z_1(1)\}$ and $\rho_n(1,\ldots,n)$ in terms of graphs with f-bonds and ρ_1- or z_1-circles in order to get the corresponding prescriptions for $h_1(1;\alpha)$ and $\rho_n(1,\ldots,n;\,\alpha)$. To each graph one adds a white 1-circle labeled by α and then draws a g_2-bond from each ρ_1-circle to this 1-circle and an e-bond from each z_1-circle to this 1-circle. The generalization to a whole set $\{\alpha,\beta,\ldots,\nu\}$ of distinguished particles is obvious.

The case of $\rho_2(1,2;\,\alpha)$ is of particular interest. Equation (11-2) yields in this case

$$\rho_2(1,2;\,\alpha) = \frac{\rho_3(1,2,\alpha)}{\rho_1(\alpha)} = \rho_1(1)\rho_1(2)g_3(1,2,\alpha). \tag{12-3}$$

Combining this with $\rho_n(1,\ldots,n) = g_n(1,\ldots,n)\rho_1(1)\ldots\rho_1(n)$ we have

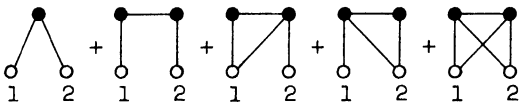
$$g_3(1,2,\alpha) = g_2(1,\alpha)g_2(2,\alpha)g_2(1,2;\,\alpha). \tag{12-4}$$

This can also be written in terms of the potentials of average force defined by Eq. (6-13),

$$W_3(1,2,\alpha) = W_2(1,2) + W_2(2,\alpha) + W_2(1,2;\alpha).^{67} \qquad (12\text{-}5)$$

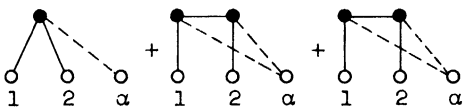
We note that the function $g_2(1,2)$ is a functional of $\rho_1(1)$ and $f(1,j)$ as well as a function of $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_2$.[68] From its definition and the fact that the potential is a sum of singlet and doublet terms, $g_2(1,2;\alpha)$ is the same functional of $f(1,j)$ and $\rho_1(1)g_2(1,\alpha)$ that $g_2(1,2)$ is of $f(1,j)$ and $\rho(1)$. It is of course a function of $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_2$, just as $g_2(1,2)$ is, but it is also a function of $\underset{\sim}{r}_\alpha$, through its dependence on $\rho_1(1)g_2(1,\alpha)$.

The nature of these functional relations is nicely shown by the graphical prescription associated with conditional distributions. The f-bond, ρ_1-circle expansion of $W_2(1,2)$ (with black ρ_1-circles and white 1-circles) is given by

$W_2(1,2) =$  $+ \ldots .$

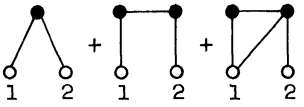
(12-6)

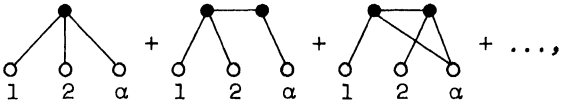
Thus we have from our prescription

$W_2(1,2;\alpha) =$  $+ \ldots$

(12-7)

where the dashed line are g_2-bonds. The g_2-bond can itself be expressed fn terms of f-bonds and ρ_1-circles itself to yield

$W_2(1,2;\alpha) =$  + ...

+  + ...,

$$W_2(1,2;\alpha) = W_2(1,2) + \omega_3(1,2,\alpha). \qquad (12\text{-}8)$$

Equations (12-8) and (12-5) together imply that

$$W_3(1,2,3) = W_2(1,2) + W_2(1,3) + W_2(2,3) + \omega_3(1,2,3). \qquad (12\text{-}9)$$

This is as it should be, since (12-9) is just (6-16) for $n = 3$.

We note that for a one-dimensional system with only nearest-neighbor interactions, $W(1,2;\alpha)$ is zero for r_{12} greater than the interaction range, when α is between 1 and 2. This well-known result, which can be expressed as $g_3(1,2,\alpha) = g_2(1,\alpha)g_2(\alpha,2)$, can be proved easily by means of our graphical representation of $W(1,2;\alpha)$ when the interaction potential is a sum of

one- and two-body terms. To prove it, it is only necessary to note that every black circle in a graph of $W(1,2;\alpha)$ is attached to a g_2-bond ending at α. In one dimension, with α between 1 and 2, the integrand associated with such a graph will always be zero except when 1 and 2 are within an interaction distance of one another because one of the g_2-bonds will always be evaluated at a point for which it is zero.

We have considered (12-2) only for $n = 2$ and $\nu = 1$ but a similar analysis can be made for arbitrary n and ν. The special case of $h(1;\alpha)$ is also of interest, since this quantity is simply β times the work necessary to bring a particle from infinity to point $\underset{\sim}{r}_1$, in the presence of a fixed particle at $\underset{\sim}{r}_\alpha$. Since $\rho_1(1;\alpha)/z_1(1;\alpha) = \{\rho(1)g_2(1,\alpha)\}/\{z_1(1)e_2(1,\alpha)\}$, we have

$$h_1(1;\alpha) = \ln \rho_1(1;\alpha) - \ln z_1(1;\alpha) = \ln\{\rho_1(1)/z_1(1)\} + W_2(1,\alpha)$$

$$= h_1(1) + W_2(1,\alpha). \quad (12\text{-}10)$$

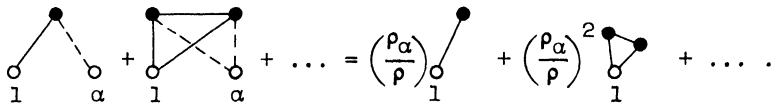
So far, all the considerations in this section have been exact, but one reason for looking at conditional probabilities is that they suggest approximations that are somewhat different from those suggested by the results in the previous sections. Under certain conditions these approximations, which all involve the concept of an "effective density", promise to be more useful than truncations of those cluster

expansions considered in previous sections. In discussing these approximations we shall use the notation $h_1(1:\rho]$ and $g_2(1,2:\ \rho]$ to denote $h_1(1)$ and $g_2(1,2)$ respectively. We do this to emphasize that h_1 and g_2 are to be evaluated at number density ρ. We shall also use an extension of this notation, writing, for example, $g_2(1,2;\ \gamma(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_\alpha)]$ to mean $g_2(1,2)$ evaluated at $\rho = \gamma(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_\alpha)$ where the $\underset{\sim}{r}_1$ and $\underset{\sim}{r}_2$ are the same arguments as those indicated by the 1 and 2 in $g_2(1,2:\ \gamma(\underset{\sim}{r}_1,\underset{\sim}{r}_2,\underset{\sim}{r}_\alpha)]$. Also we recall that $g_2(1,2)$ stands for $g_2(\underset{\sim}{r}_1,\underset{\sim}{r}_2)$; it will be convenient to reinstate the less abbreviated notation in writing $g_2\left(\frac{\underset{\sim}{r}_1+\underset{\sim}{r}_2}{2},\ \underset{\sim}{r}_\alpha\right)$.

The first approximation we shall consider involves Eq. (12-10). We introduce the effective density ρ_α^* (it will turn out to be a function of $r_{1\alpha}$) by setting

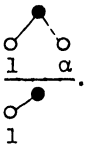
$$h_1(1;\alpha) = h_1\left(1:\rho_\alpha^*\right], \tag{12-11}$$

or



(12-12)

Here black circles are ρ_1-circles, white circles are 1-circles, the dashed lines are g_2-bonds and the solid lines are f-bonds. This will be true for all the graphs pictured throughout the rest of this section. Dropping all integrals but the first on each side of (12-12) we have

$$\rho_\alpha^* = \rho$$  (12-13)

Equation (12-13) has already been considered by Helfand and Stillinger[41] as a possible definition of an effective number density, and Eq. (12-11) is the starting point of an effective density technique recently developed by Lebowitz and Praestgaard.[42] The approximation defined by Eqs. (12-11) and (12-13) clearly becomes exact in the limit $r_{1\alpha} \to \infty$. It is also reasonable for small $r_{1\alpha}$, and in the special case of a hard-sphere potential, for which $\exp W_n(1,2,\ldots,n)\Big|_{1=2=\ldots=n} = (z/\rho)^{n-1}$, it becomes exact as $r_{1\alpha} \to 0$.

These limiting results are still true, and are of special interest, when the system is in two phases.[69] For such a system we have a relation, first suggested by Mayer and Montroll[3] and recently proved by Fisher under very weak assumptions:[70]

$$g_n = \left(\frac{\rho_a}{\rho}\right)^n A g_n^{(a)} + \left(\frac{\rho_b}{\rho}\right)^n B g_n^{(b)},$$

where (12-14)

$$A = \frac{\rho-\rho_b}{\rho_a-\rho_b}, \qquad B = \frac{\rho_a-\rho}{\rho_a-\rho_b}.$$

Here ρ_a refers to the limiting density of the heavier phase; ρ_b to the limiting density of the lighter phase, and $g_n^{(a)}$ and $g_n^{(b)}$ refer to the g's at these densities. It is easy to verify that Eq. (12-11) and (12-13) are consistent with Eq. (12-4) when $r_{1\alpha} = \infty$ and (for hard spheres) when $r_{1\alpha} = 0$. Such asymptotic consistency with (12-14) seems to be characteristic of effective density approximations.

We turn now to the expressions relating the two- and three-particle functions. For a system in which $g_2(1,2)$ approaches the asymptotic value of unity rapidly and in an oscillatory fashion as r_{12} increases we can see that replacing g_2 by unity in the graphs of Eq. (12-7) is likely to result in a satisfactory approximation if $g_2(1,2) \to 1$ when $r_{12} \to \infty$, except when $r_{1\alpha}$ and $r_{2\alpha}$ are very small. In each graph the integration tends to average out the effect of the oscillations of g_2, and the resulting integral is relatively insensitive to the fact that g_2 differs pointwise from its average over a few oscillations. But for most densities and temperatures this average will be appreciably different from unity only over a small region of the domain of integration.

The replacement of g_2 by unity yields the already familiar Kirkwood superposition approximation, (Eq. (6-12)) which we shall abbreviate as KSA.

Equation (12-7) also leads immediately to a pair of modifications of the KSA that have previously been suggested by others on the basis of rather different considerations.

The form of (12-7) suggests that if $g_2(1,2)$ were a slowly varying function of r_{12} it would be appropriate to let

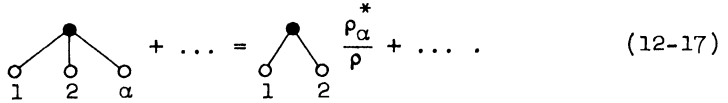
$$g_2(1,2;\alpha) = g_2\left(1,2: \quad \rho g_2\left(\frac{\underset{\sim}{r}_1+\underset{\sim}{r}_2}{2} - \underset{\sim}{r}_\alpha\right)\right]. \qquad (12\text{-}15)$$

Using this approximation in (12-4) yields a g_3 proposed by Fixman[71] in order to obtain a better theory for a one-phase system near the critical point. For such a system it seems likely that $g_2(1,2)$ approaches unity slowly enough for (12-15) to yield a considerably better approximation than the KSA.

An expression obtained by Helfand and Stillinger[4] can also be recovered from (12-4) by comparing the exact expansion of $W(1,2;\alpha)$ with the expansion of $W\left(1,2: \quad \rho_\alpha^*\right]$ where ρ_α^* is again an effective density to be determined. Setting

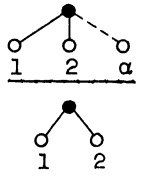
$$W_2(1,2;\alpha) = W_2\left(1,2: \quad \rho_\alpha^*\right] \qquad (12\text{-}16)$$

we have



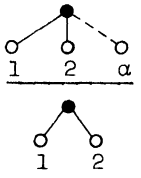
(12-17)

Dropping all but the first integral on each side leaves

$$\rho_\alpha^* = \rho$$ 

(12-18)

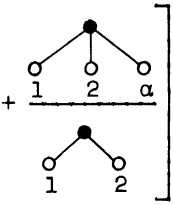
This expression was derived by Helfand and Stillinger from another standpoint. Here ρ_α^* turns out to be a quite complicated function of r_{12}, $r_{1\alpha}$, and $r_{2\alpha}$. (In such expressions, when 1 and 2 are far apart for finite ranged potentials, the resulting 0/0 should be interpreted by taking the limit as r_{12} becomes large.) If one assumes that the g_2 in

$\rho_\alpha^* = \rho$ 

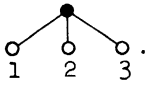
is slowly varying one recovers Fixman's approximation,

$$\rho_\alpha^* = \rho g\left(\left|\frac{\underset{\sim}{r}_1+\underset{\sim}{r}_2}{2} - \underset{\sim}{r}_\alpha\right|\right). \qquad (12\text{-}19)$$

One can also simplify (12-18) in a somewhat different direction by expanding the right-hand side of (12-18) in powers of ρ and retaining only the first term to get

$$\rho_\alpha^* = \rho \left[1 + \right.$$  $$\left. \right].$$ (12-20)

For a hard-sphere potential, the right-hand side of (12-20) can actually be expressed analytically, as a result of the recent work of Rowlinson[72] in evaluating



If $g_2(1,2;\alpha)$ is set equal to $g_2(1,2;\rho^*]$ in (12-14), where ρ^* is given by either (12-18), (12-19) or (12-20), the resulting g_3 is not symmetric with respect to the interchange of 1, 2, and α. This lack of symmetry is an annoying feature that is likely to occur whenever an approximate effective density is introduced into (12-4). It is acceptable if it appears in connection with some expansion or equation in which $\underset{\sim}{r}_\alpha$ is already a distinguished variable and where the final expressions sought are functions only of r. This will be the case when $g_3(1,2,\alpha)$ appears in a hierarchy of equations in which g_2 is the unknown of interest and (12-4) is used simply to get rid of g_3. If g_3 is desired for its own sake, however, it is necessary to symmetrize the approximate g_3 or W_3. In

the absence of any special information one can simply add to the expression being considered the two expressions that result from cyclically permuting the 1, 2, and α; and then divide by three. A somewhat less crude alternative is to introduce the angles Φ_1, Φ_2, and Φ_3 where Φ_1 is the angle formed by the vectors $\underset{\sim}{r}_2 - \underset{\sim}{r}_1$ and $\underset{\sim}{r}_3 - \underset{\sim}{r}_1$ at $\underset{\sim}{r}_1$, Φ_2 is the angle formed by the vectors $\underset{\sim}{r}_3 - \underset{\sim}{r}_2$ and $\underset{\sim}{r}_1 \sim \underset{\sim}{r}_2$ at $\underset{\sim}{r}_2$, etc. Setting $\theta_1 = \Phi_1/2\pi$ we can let

$$W_3(1,2,3) = \theta_2\left[W(1,2) + W(2,3) + W\left(1,3:\ \rho_2^*\right)\right]$$

$$+ \theta_3\left[W(1,3) + W(2,3) + W\left(1,2:\ \rho_3^*\right)\right]$$

$$+ \theta_1\left[W(1,3) + W(1,2) + W\left(2,3:\ \rho_1^*\right)\right]. \qquad (12\text{-}21)$$

When (12-18) is used, this symmetrization has the advantage of coinciding with the exact limiting result for a hard-sphere potential as $r_{12} \to 0$:

$$W_3(1,2,3) \to W_2(1,1) + W_2(1,3).$$

Equations (12-18) and (12-21) will yield an exact fourth virial coefficient if used with a hierarchy of equations such as the B-B-G-K-Y hierarchy, and they are consistent with (12-14) when $\underset{\sim}{r}_1$, $\underset{\sim}{r}_2$, and $\underset{\sim}{r}_3$ are far apart. It is not clear, however, whether they constitute a general improvement over the KSA.

More nearly exact (and complicated) expressions than (12-13) and (12-18) can be obtained if one retains more integrals or starts with larger n and ν in (12-2). However, progress seems less likely to come from such extensions than from the felicitous choice of expressions in which the affective densities are to be used.

ACKNOWLEDGMENTS

The author is grateful to Professors J. L. Lebowitz, M. Fisher, O. Penrose, and J. K. Percus for encouragement and helpful discussions.

FOOTNOTES

1. Ursell, H. D., Proc. Cambridge Phil. Soc. 23, 685 (1927).
2. Yvon, J., Actualities Scientifiques et Industrielles, No. 203 (Hermann et Cie, Paris, 1935).
3. Mayer, J. E., J. Chem. Phys. 5, 67 (1937); Mayer, J. E., and P. F. Ackermann, ibid 5, 74 (1937); Mayer, J. E., and S. F. Harrison, ibid 6, 87 (1938); Mayer, J. E., and E. Montroll, ibid 9, 2 (1941). See also Mayer, J. E., and M. G. Mayer, Statistical Mechanics (John Wiley and Sons, Inc., New York, 1940).
4. Particularly comprehensive reviews of cluster theory are given in references 5 through 8. For alternative derivations, see, for example, reference 26. We wish to emphasize that there is a huge literature on the subject of cluster expansions that we shall not touch upon. We consider the problem of characterizing the general term of a cluster series. There is also the problem, which we shall ignore, of rigorously establishing the existence and uniqueness of such series for a given potential of interaction (and a fast growing literature celebrating the partial solution of this problem). We shall also ignore the problem of calculating the functions given by the cluster integrals for a given potential. Finally, the references we cite in

the remainder of this article are only those that are related in some specific way to the development in the article. Hence even among the various purely formal derivations of cluster series that have been given, there are some that we shall not mention at all.

5. DeBoer, J., Rep. Progr. Phys. 12, 305 (1949).

6. Salpeter, E. E., Ann. Phys. 5, 183 (1958).

7. Uhlenbeck, G. E., and G. W. Ford, in Part B of Studies in Statistical Mechanics, Vol. I (North-Holland Publishing Company, Amsterdam, 1962).

8. Friedman, H. L., Ionic Solution Theory Based on Cluster Expansion Methods (John Wiley and Sons, New York, 1962).

9. We shall not consider these generalizations in this paper. For the case of mixtures, see reference 24 and for the case of $n \geq 3$, see reference 8.

10. Bogoliubov, N. N., Problems of Dynamical Theory in Statistical Physics, Gostekhizdat, Moscow 1946. (In Russian.) An English translation can be found as Part A in Studies in Statistical Mechanics, Vol. I (North-Holland Publishing Company, Amsterdam, 1962).
Earlier, Fock had used generating functional in studying quantum systems. (Fock, V. A., Zeit. f. Phys. 49, 339 (1928).

11. Bogoliubov, N. N., Vestnik Moskov. Univ. 10, 115-24 (1955).
In English as AEC - I.I. 3487, and also as the introductory article of The Method of Functionals in the Quantum Theory of Fields, reference 13.

12. Schwinger, J., Proc. Nat. Acad. Sci. 37, 452, 455 (1951).

13. Partial summaries can be found in Novoshilov and Tulub The Method of Functionals in the Quantum Theory Fields (Gordon and Breach, New York, 1961) and in Green, M. S., Some Applications of the Generating Functional of the Molecular Distribution Function, which appears in Lectures in Theoretical Physics, Vol. III (Interscience, New York, 1960). The use of functional derivatives in quantum statistical mechanics that is closest in form and purpose to the use we consider in this paper is found in T. D. Lee, and C. N. Yang, Phys. Rev. 117, 22 (1960) and in reference 25.

14. Bazarov, I., JETP (SSSR) 32, 1065 (1957).
In English in Soviet Phys. - JETP 5, 872 (1957).
Bazarov makes use of functional differentiation with respect to functions of two variables, thus generalizing somewhat Bogoliubov's approach. In the first paper cited in reference 20 functions of n variables are used in obtaining the ρ_n.

15. Zumino, B., Phys. Fluids 2, 20 (1959).

16. Lewis, R. J., J. Math. Phys. 2, 222 (1961).

17. Lewis, R. J., and J. B. Keller, Phys. Rev. 121, 1022 (1961).

18. Dissertation, New York University 1961 (unpublished).

19. Percus, J. K., Phys. Rev. Letters 8, 462 (1962).

20. Lebowitz, J., and J. K. Percus, Phys. Rev. 122, 1675 (1961); 124, 1673 (1961); J. Math. Phys. 4, 116, 248 (1963).

21. Yvon, J., Suppl. Nuovo Cimento 9, 144 (1958).

22. Hiroike, K., J. Phys. Soc. Japan 12, 864 (1957); 15, 771 (1960).

23. Morita, T., and K. Hiroike, Prog. Theor. Phys. 23, 385, 1003 (1960).

24. Morita, T., and K. Hiroike, Prog. Theor. Phys. 25, 531 (1961).

25. DeDominicis, C., J. Math. Phys. 3, 983 (1962); 4, 255 (1963).

26. Born, M., and K. Fuchs, Proc. Roy. Soc. (London) A166, 391 (1938) and Husimi, K., J. Chem. Phys. 18, 682 (1950).

27. However, it had long been taken for granted that the cluster expansions that describe a nonuniform system have exactly the same topological form as those used to describe a uniform system. This identity of form is already implied by the early work of Yvon (reference 2)

who described an iterative procedure that can be used to get the cluster integrals in the more general case of a nonuniform system.

28. Meeron, E., J. Math. Phys. 1, 192 (1960), Hughes Aircraft Co. Report (1959) (unpublished).

29. M. S. Green, J. Chem. Phys. 33, 1403 (1960).

30. Van Leeuwen, J. M. J., J. Groeneveld, and J. DeBoer, Physica 25, 792 (1959).

31. Verlet, L., Nuovo Cimento 18, 77 (1960).

32. Rushbrooke, G. S., Physica 26, 259 (1960).

33. Green, M. S., J. Math. Phys. 1, 391 (1960).

34. Earlier, Nettleton and Green [J. Chem. Phys. 29, 1365 (1958)] had also obtained an expression for the entropy of a system in terms of distribution functions. Their expression, which involves all the distribution functions, is extremely complex, but more general than that of Morita and Hiroike in that it is applicable to a system in which the potential contains n-body terms for all n.

35. Balian, R., and C. deDominicis, Nuclear Phys. 16, 502 (1960); Compt. rend. 250, 3285, 4111 (1960). Balian, C., Block, C., and C. deDominicis, Compt. rend. 250, 2850 (1960); Nuclear Phys. 25, 529 (1961); 27, 294 (1961).

36. This can only be done if the potential is additive. In general, if the potential includes s-body terms, ρ_n can be expressed as a functional of ρ_1 through ρ_s for $s < n$.

37. Abe, R., Prog. Theor. Phys. 22, 213 (1959).

38. Stell, G., Physica 29, 517 (1963).

39. Meeron, E., Phys. Fluids 1, 130 (1958); Meeron, E., and E. R. Rodemick, Phys. Fluids 1, 246 (1958).

40. Percus, J. K., and G. L. Yevick, Phys. Rev. 110, 1 (1958).

41. Helfand, E., and F. H. Stillinger, Jr., J. Chem. Phys. 37, 2649 (1962).

42. Lebowitz, J. L., and E. Praestgaard (to be published in J. Chem. Phys., 1964).

43. See e.g., reference 24 for a definition of functional differentiation. A detailed treatment of the differential calculus of functionals can be found in V. Volterra's Theory of Functionals (Dover Publications, Inc., New York, 1959). When functionally differentiating with respect to $z_1(1)$ we assume that the quantity being differentiated is fully determined by z_1 and f and that f is held fixed in the differentiation. (Thus for a system undergoing a phase transition, Eq. (2-9) cannot be used since $\ln \Xi$ is not a unique differentiable functional of z_1 and f.) When differentiating with respect to ρ_1 we assume that the quantity being differentiated is fully determined by ρ_1 and f and that f is held fixed in the differentiation.

36. This can only be done if the potential is additive. In general, if the potential includes s-body terms, ρ_n can be expressed as a functional of ρ_1 through ρ_s for $s < n$.

37. Abe, R., Prog. Theor. Phys. 22, 213 (1959).

38. Stell, G., Physica 29, 517 (1963).

39. Meeron, E., Phys. Fluids 1, 130 (1958); Meeron, E., and E. R. Rodenick, Phys. Fluids 1, 246 (1958).

40. Percus, J. K., and G. L. Yevick, Phys. Rev. 110, 1 (1958).

41. Helfand, E., and F. H. Stillinger, Jr., J. Chem. Phys. 37, 2649 (1962).

42. Lebowitz, J. L., and E. Praestgaard J. Chem. Phys. 40 2951 (1964).

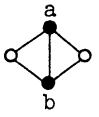
43. See e.g., reference 24 for a definition of functional differentiation. A detailed treatment of the differential calculus of functionals can be found in V. Volterra's Theory of Functionals (Dover Publications, Inc., New York, 1959). When functionally differentiating with respect to $z_1(1)$ we assume that the quantity being differentiated is fully determined by z_1 and f and that f is held fixed in the differentiation. (Thus for a system undergoing a phase transition, Eq. (2-9) cannot be used since $\ln \Xi$ is not a unique differentiable functional of z_1 and f.) When differentiating with respect to ρ_1 we assume that the quantity being differentiated is fully determined by ρ_1 and f and that f is held fixed in the differentiation.

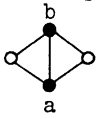
44. Ornstein, L. S., and F. Zernike, Proc. Acad. Sci. Amst. 17, 793 (1914).

45. This seems to have been first pointed out by Yvon, J., Suppl. Nuovo Cimento 9, 144 (1958).

46. Sometimes graphs are used to represent just the integrands appearing in integrals such as (3.2) and all the circles are labeled. These graphs are then called specific graphs, whereas the graphs associated with the integral are called generic. We shall speak only of the latter. Of course, if the integrand of (3.2) were to appear by itself, it would be a legitimate function to represent by a graph according to our rules of correspondence. All the circles of the graph would be white and labeled.

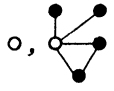
47. To find σ, one can temporarily label the black circles of a graph Γ in any way so that they are distinguishable from one another and then consider all permutations of these labels. Some of the permutations take Γ into a graph distinguishable from Γ; some take Γ into itself. The total number of permutations that do the latter is equal to σ. This definition of σ necessitates having a precise enough definition of Γ to be able to recognize when two graphs are distinguishable, i.e., not "the same". The intuitive sort of definitions we have given in this section may seem a little hazy on this score but are adequate if we note that a graph is defined by the set of all connections between circles

and bonds. This is consistent with visualizing all graphs as being realized in Euclidian 3-space rather than 2-space. Thus 

and 

are the same graph, defined by the fact that 1 and 2 are each connected to a and b; a is connected to 1, 2, and b; and b is connected to 1, 2, and a. We can get from one graph to the other by taking it out of the plane of this page and into 3-space and then flopping it over.

48. A 1-circle is a circle corresponding to the function identically equal to unity.

49. When we speak of the removal of a circle from a graph, we really mean removal. Thus, upon the removal of  becomes  . Note that all the connections among bonds at the white circle are assumed broken and that the white circle itself vanishes from the scene and hence is not counted as one of the remaining pieces.

50. Variants of this result have been rederived a great many times in the literature.

51. For proofs, see reference 23 or the first item of reference 25.

52. Graphs such that there is a bond between each pair of circles are sometimes called <u>complete</u>.

53. Graphs such that there is at least one path from each black circle to a white circle are sometimes called root-connected.

54. In the case of connected graphs with one or no white circles, this definition is identical to Mayer's definition of irreducible (see reference 3). For such connected graphs, excepting a graph consisting of 2 circles and a single bond, 1-irreducibility is identical to the condition that a graph be at least doubly connected, which is the condition used by Mayer in considering h_1. In the case of connected graphs with two or more white circles, a graph is free of articulation circles if and only if there can be found a path between any black circle and each of two white circles such that the two paths share only the black circle. This characterization was the one originally used by Mayer and Montroll in considering h_n for $n \geq 2$.

55. This follows immediately from Lagrange's theorem relating the order of a group, the order of any subgroup and the number of cosets generated by the subgroup. Equation (5-4) is the statement of that theorem as applied to the permutation groups under consideration. The theorem itself can be found in any elementary abstract algebra text.

56. In a graph with no white circles a circle is a connecting circle if and only if it is an articulation circle, but this is not true in general. What we call a connecting circle, Uhlenbeck and Ford [reference 7] call an articulation point. Our definition of articulation circle, however, is consistent with the definitions used in [24], [25], and [30].

57. Alder, B. J., has made a careful molecular dynamical study comparing g_3 and g_2 for hard spheres (to be published). I am indebted to Dr. Alder for information concerning these results.

58. Fisher, I. Z. and B. L. Kopeliovich, Soviet Phys. - Doklady, 5, 761 (1960/61). From Doklady Akad. Nauk. SSSR, 133, 81 (1960).

59. Stillinger, F. H., Jr., and F. P. Buff, J. Chem. Phys. 57, 1 (1962).

60. The fact that each graph does not appear exactly once in these series is already clear from their first few terms which can be obtained easily be straightforward iterative procedures.

61. A polygon or ring is a simple graph with an equal number of circles and bonds such that there are two bonds incident upon every circle.

62. In particular, the same kind of order of magnitude comparisons can be made using parameters such as ρ. Thus for a potential $\varphi_2(1,2)$ that is nonzero only for $r_{12} \leq R$ and an r_{12} that is equal to $1R$, $W_2(1,2)$ and $t(1,2)$ start with terms of order ρ^1 and are identical through terms of order less than ρ^{21}.

63. Comparisons among the results of the HNC and PY approximations and near-exact machine results have been made by several groups. See for example A. A. Broyles, J. Chem. Phys. 35, 493 (1961) [PY; Lennard-Jones]; A. A. Broyles, S. U. Chang, and H. L. Sahlen, J. Chem. Phys. 37 2462 (1962) [HNC, PY; Lennard-Jones]; A. A. Broyles, H. L. Sahlin, D. D. Carley, Phys. Rev. Letters 10, 319 (1963) [HNC, PY; Coulomb], D. D. Carley (to be published) [HNC, PY; Coulomb], L. Verlet and D. Levesque, Physica 28, 1124 (1962) [HNC, PY; Lennard-Jones]; M. Klein and M. S. Green, J. Chem. Phys. 39, 1367 (1963) [HNC; Lennard-Jones]; M. Klein, J. Chem. Phys. 39, 1388 [HNC; hard spheres]. In addition to these results, further results have been obtained by P. Hutchinson, Dissertation, Newcastle upon Tyne, 1964 and by J. deBoer and his co-workers, (to appear in Physica).

Comparisons between the integral equation results and exact number-density expansions for various potentials have also been made. See G. S. Rushbrooke, J. Chem. Phys.

38, 1262 (1963) [HNC, PY; hard spheres]; W. G. Hoover and J. C. Poirier, J. Chem. Phys. 38, 327 (1963) [HNC, PY; hard squares, hard cubes]; E. Helfand and R. L. Kornegay (to be published) [HNC, PY; Gaussian f-function]. All PY virial coefficients are known exactly for hard spheres; see M. Wertheim's article in this volume.

In the above references, we have indicated only if the HNC and PY approximations are considered; several other approximations are also compared in some of these papers.

64. By examining functional expansion that leads to the PY and HNC approximations when truncated, Percus was able to conclude that this is to be expected for any potential that changes rapidly wherever it changes at all, e.g., the hard-sphere potential. For details see the article by Percus in this volume.

65. See the results of M. Wertheim in this volume.

66. Hoover, W. C., and J. C. Poirier, J. Chem. Phys. 37, 1041 (1962).

67. Despite their highly asymmetric appearance, the right-hand sides of (12-4) and (12-5) are invariant under interchange of the points 1, 2, and α.

68. In the case of a uniform system this functional dependence becomes somewhat degenerate, since $\rho_1(1)$ is simply ρ, the expected number density.

69. Equations like (11-3) and (11-4) are true simply by definition, whether canonical or grand canonical distribution functions are considered, and hence these equations are applicable to a system in an arbitrary number of phases.

70. Fisher, M. (to be published).

71. Fixman, M., J. Chem. Phys. 33, 1357 (1960).

72. Weissberg, H. L. and S. Prager, Phys. Fluids 5, 1390 (1962), F. H. Ree, unpublished; Rowlinson, J. S., Mol. Phys. 6, 517 (1963). The result of Weissberg and Prager was obtained in connection with a somewhat different problem than the one considered here, and hence is couched in somewhat different language.

Reprinted from *Printed in U.S.A.*
JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 5 MAY 1964

Analytic Solution of the Percus–Yevick Equation*

M. S. WERTHEIM†
Courant Institute of Mathematical Sciences, New York University, New York, New York
(Received 1 November 1963)

The properties of the Percus–Yevick approximate integral equation for the pair distribution function in classical statistical mechanics are examined for the class of pair potentials consisting of a hard core plus a short-range tail. For one-dimensional systems, some elementary theorems of complex variable applied to the Laplace-transformed equations enable one to express the direct correlation function in a very simple form, one which becomes explicit and trivial in the absence of a short-range tail. In the presence of the tail, the direct correlation function satisfies a (coupled) integral equation over a finite domain. The impossibility of a phase transition in one dimension is strongly indicated. Analysis of the case of three dimensions proceeds similarly, but is complicated by the appearance of essential parameters other than the density and compressibility. The character of the direct correlation function is qualitatively unchanged. Principal differences in three dimensions are that a phase transition is no longer prohibited, and the pair distribution function cannot be reasonably expressed as a sum of nth-neighbor contributions.

I. INTRODUCTION

SEVERAL approximate integral equations have been proposed for the pair distribution function in classical statistical mechanics. On the basis of numerical comparisons[1] made so far, the Percus–Yevick (PY) equation[2] is the most successful of these. Furthermore the PY equation has been solved exactly for hard spheres,[3,4] representing the first nontrivial system solved for any of these integral equations. In the solution for hard spheres a factorization of the PY equation occurred, which in the original method of solution appeared as a lucky accident. This suggested a more detailed study of the PY equation, exploiting the fact that it is a quadratic integral equation.

The method we have used consists of the application of some elementary theory of the functions of a complex variable to the Laplace-transformed PY equation. We consider the case of hard spheres with a potential tail which becomes identically zero beyond a distance a which is less than the hard sphere diameter R. The potential tail is assumed to be finite or have only integrable singularities, but is otherwise arbitrary. For this system we find a factorization of the PY equation, which may be used to eliminate the pair distribution function, leaving an equation for the direct correlation function alone. In the case of zero tail the solution for hard spheres is obtained in a simple fashion. For the case of nonzero tail the equation for the direct correlation function $C(x)$ is expressed in a form believed suitable for numerical computation. The method used clearly shows the similarity as well as some important differences between the one-dimensional and the three-dimensional case. In the following we carry out both cases explicitly. Since the one-dimensional problem can be solved exactly, any information on the solution of the PY equation in one dimension serves as a welcome check on the adequacy of the PY approximation.

II. ONE DIMENSION

A. Analytic Character of the Integral Equation

In one dimension, the PY equation for a system of particles interacting pairwise with a potential $V(x)$ can be written[2]

$$\tau(x) = 1 - \rho \int_{-\infty}^{\infty} \tau(x')f(x')\,dx' + \rho \int_{-\infty}^{\infty} \tau(x')f(x')\tau(x - x')\,\mathrm{e}(x - x')\,dx', \qquad (1)$$

where $\mathrm{e}(x) \equiv \exp[-\beta V(x)]$, $f(x) = \mathrm{e}(x) - 1$, ρ is the density, and $\beta = (kT)^{-1}$. The pair distribution function $g(x)$ and the direct correlation function $C(x)$ of Ornstein and Zernike[5] are in the PY approximation related to $\tau(x)$ by

$$g(x) = \tau(x)\mathrm{e}(x), \qquad C(x) = \tau(x)f(x).$$

In our case $V(x) = V_H + V_T$, where the hard-

* The work presented in this paper is supported by the AEC Computing and Applied Mathematics Center, Courant Institute of Mathematical Sciences, New York University, under Contract AT(30-1)-1480 with the U. S. Atomic Energy Commission.
† On leave from Los Alamos Scientific Laboratories.
[1] A. A. Broyles, J. Chem. Phys. **33**, 456 (1960); **34**, 359 (1961); **34**, 1068 (1961); **35**, 493 (1961). A. A. Broyles, S. U. Chung, and H. L. Sahlin, J. Chem. Phys. **37**, 2462 (1962).
[2] J. K. Percus and G. J. Yevick, Phys. Rev. **110**, 1 (1958).
[3] M. S. Wertheim, Phys. Rev. Letters **10**, 321 (1963).
[4] E. Thiele, J. Chem. Phys. **38**, 1959 (1963).
[5] L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amsterdam **17**, 793 (1914).

rod potential V_H is given by

$$V_H = 0 \quad (|x| > l), \qquad V_H = \infty \quad (|x| < l),$$

and the tail potential V_T satisfies $V_T = 0$ for $|x| > l + a$.

We take the one-side Laplace transform (LT) of (1) for a potential of this form and obtain

$$F(s) + G(s) = (1 + \rho K)/s - \rho[F(s) + F(-s)]G(s) - \rho Y(s) + \rho Y(-s), \qquad (2)$$

where

$$F(s) \equiv -\int_0^{l+a} C(x) \exp(-sx)\, dx,$$

$$G(s) \equiv \int_l^{\infty} g(x) \exp(-sx)\, dx,$$

$$Y(s) \equiv \int_0^{a} y(x) \exp(-sx)\, dx,$$

$$y(x) \equiv -\int_{l+x}^{l+a} C(x')g(x - x')\, dx' \quad (0 < x < a),$$

$$y(x) \equiv 0 \quad (x < 0 \text{ and } x > a),$$

$$K \equiv -2\int_0^{l+a} C(x')\, dx'.$$

Solving for $G(s)$, and letting $1 + \rho K \equiv Q^2$, this becomes

$$G(s) = \frac{Q^2 s^{-1} - F(s) - \rho Y(s) + \rho Y(-s)}{1 + \rho F(s) + \rho F(-s)}. \qquad (3)$$

We assume that $C(x)$ and $g(x)$ are everywhere finite. Therefore $F(s)$ and $Y(s)$, being LT's over a finite interval, are entire functions of s. $G(s)$ is an LT over an infinite interval and is therefore regular in the right half-plane (RHP). In the left-half plane (LHP) $G(s)$ is defined by analytic continuation of the function $G(s)$ in the RHP.

From (3) it follows that

$$G(-s) = \frac{-Q^2 s^{-1} - F(-s) - \rho Y(-s) + \rho Y(s)}{1 + \rho F(-s) + \rho F(s)} \qquad (4)$$

and that the quantity

$$G(s)[Q^2 s^{-1} + F(-s) - \rho Y(s) + \rho Y(-s)]$$

is an even function of s. According to (3), $G(s)$ has a simple pole with principal part s^{-1} at $s=0$. Therefore the quantity $s^2 G(s)[Q^2 s^{-1} + F(-s) - \rho Y(s) + \rho Y(-s)]$ is regular at the origin and in the RHP, since both factors are regular in the RHP. Since it is also an even function of s, it is also regular in the LHP, and hence it is an entire function of s. Noting that the growth are $O[s^{-1} \exp(-l - a)s]$ and $O[s^{-1} \exp(-as)]$, respectively, we find by (3) that $G(s)$ goes to $-\rho^{-1}$ as we go to infinity in the LHP, and that the growth rate of $s^2 G(s)[Q^2 s^{-1} + F(-s) - \rho Y(s) + \rho Y(-s)]$ is $O[s \exp(-as)]$.

Now consider the function

$$H(s) = s^2 G(s)[Q^2 s^{-1} + F(-s) - \rho Y(s) + \rho Y(-s)] - s^2[Y(s) + Y(-s)].$$

Clearly it is also an even entire function of s. Furthermore, by (3) H is given by

$$H(s) = s^2[1 + \rho F(s) + \rho F(-s)]^{-1}\{Q^4 s^{-2} + Q^2 s^{-1} \times [F(-s) - F(s) - 2\rho Y(s) + 2\rho Y(-s)] - F(s)F(-s) - 2\rho F(s)Y(-s) - 2\rho F(-s)Y(s) - Y(s) - Y(-s) + \rho^2[Y(s) - Y(-s)]^2\}. \qquad (5)$$

In the LHP, the term with the largest growth rate is $F(s)$ which is $O[s^{-1} \exp(-l - a)s]$. Using $F(-s) = -\tau(0)s^{-1} + O(s^2)$, $Y(-s) = -y(0)s^{-1} + O(s^{-2})$, and $\tau(0) = Q^2 - 2\rho y(0)$ from (1) as we go to infinity in the LHP, we find $H(s) = O(1)$ as we go to infinity along any radius in the LHP. Since $H(s)$ is even, $H(s) = O(1)$ holds also in the RHP. But by Liouville's theorem a bounded entire function is a constant. We evaluate the constant by taking $H(0)$ in (5) and find $H = Q^2$. Therefore

$$G(s)[Q^2/s + F(-s) - \rho Y(s) + \rho Y(-s)] = Q^2/s^2 + Y(s) + Y(-s). \qquad (6)$$

At this point we note that this factorization holds only for a nearest-neighbor-type tail. For $a > l$ the dominant term in the numerator of (5) is $Y^2(s)$, which is $O[s^{-2} \exp(-2as)]$ as we go to infinity in the LHP, and $H(s)$ is $O[s^{-1} \exp(-a + l)s]$ rather than $O(1)$.

Combining (6) with (2), we obtain

$$F(s) + G(s) = Q^2\left(\frac{1}{s} - \frac{\rho}{s^2}\right) - 2\rho Y(s) + \rho^2 Y(-s)G(s) - \rho^2 Y(s)G(s) - \rho F(s)G(s) + \frac{\rho Q^2 G(s)}{s}. \qquad (7)$$

Let us transform (7) back to coordinate space. For hard rods with no tail, $Y = 0$. Moreover $G(s)$, $\rho F(s)G(s)$ and $\rho Q^2 s^{-1} G(s)$ are LT's of functions which are zero for $x < l$. Hence we obtain for $x < l$:

$$-C(x) = \tfrac{1}{2}Q^2(1 - \rho x), \qquad (8)$$

and, since by (2)

$$\tau(0) = 1 + 2\rho \int_0^l C(x)\, dx = Q^2 = 1 + \rho^2 \int_0^{2l} C(x)\, dx \int_0^{|x-l|}$$

and

$$Q^2 = (1 - \rho l)^{-2}. \tag{9}$$

In this fashion the PY equation is solved for hard rods in an extremely simple way.

B. Elimination of $g(x)$ and Reduction to Finite-Range Equations

Now consider the case of nonzero tail. The right-hand side of (7) contains four types of terms. The term $Q^2(s^{-1} - \rho s^{-2})$ is the LT of $Q^2(1 - \rho x)$, the term $-2\rho Y(s)$ is the LT of $-2\rho y(x)$, which is nonzero only for $0 < x < a$. The term $-\rho^2 Y(-s)G(s)$ is the LT of a function which is zero for $x < l - a$. All the remaining terms are LT's of functions which are zero for $x < l$.

For future convenience we now define

$$-C(x) = -C_0(x) - 2\rho y(x) + p(x - l) + q(x - l),$$

where $-C_0(x) = Q^2(1 - \rho x)$ for $(0 < x < l)$ and $-C_0(x) = 0$ for $x > l$; $p(x) = 0$ for $x < -a$ and $x > 0$; $q(x) = 0$ for $x < 0$ and $x > a$. According to (7), p is given by

$$p(x - l) = \rho^2 \int_{-a}^{x-l} y(-x')g(x - x')\,dx'. \tag{10}$$

We omit the equation for $q(x - l)$, since it still involves $g(x)$ for $x > l + a$ in the inverse LT of the term $\rho^2 Y(-s)G(s)$.

In order to eliminate $g(x)$ for $x > l + a$, and obtain an equation for $C(x)$ alone, it is necessary to return to Eqs. (2) and (6). We define

$$X(s) = [F(s) + 2\rho Y(s) - Q^2 s^{-1} + \rho Q^2 s^{-2}],$$
$$Z(s) = Q^2 s^{-2} + Y(s) + Y(-s).$$

Equations (2) and (6) combined now become

$$[X(s) - \rho Z(s)][X(-s) - \rho Z(s)]$$
$$= -Z(s)[1 + \rho X(s) + \rho X(-s) - 2\rho^2 Z(s)] \tag{11}$$

or

$$X(s)X(-s) = -Z(s) + \rho^2 Z^2(s). \tag{12}$$

Now let

$$X(s) = e^{-sl}\mu(s)\nu(s), \tag{13}$$

$$Z(s) = -\mu(s)\mu(-s). \tag{14}$$

The result is

$$\nu(s)\nu(-s) = \rho^2 \mu(s)\mu(-s) + 1 \tag{15}$$

and

$$G(s) = \frac{-\mu(s)}{e^{sl}\nu(-s) + \rho\mu(s)}. \tag{16}$$

The functions $s^2X(s)$ and $s^2Z(s)$ are entire and can therefore be written as canonical Weierstrass products. In the factorization $X(s) = \mu(s)\nu(s) \exp(-sl)$, all the roots of $X(s)$ which are also roots of $Z(s)$ appear in $\mu(s)$, while the common roots of $X(s)$ and $\rho Z(s) - 1$ appear in $\nu(s)$.

So far, $\mu(s)$ and $\nu(s)$ are determined only up to an exponential factor, since simultaneously multiplying $\mu(s)$ by $\exp \phi(s)$ and $\nu(s)$ by $\exp[-\phi(s)]$—with $\phi(s)$ an odd function of s—has no effect on any of the preceding equations. The factor can certainly be chosen to be an entire function. Then $\mu(s)$ and $\nu(s)$ are regular except for a simple pole at $s = 0$. The poles $\rho Q^2 s^{-2}$ in $X(s)$ and $Q^2 s^{-2}$ in $Z(s)$ imply poles Qs^{-1} in $\mu(s)$ and $\rho Q s^{-1}$ in $\nu(s)$. We define

$$\mu(s) = Q/s + M(s), \quad \nu(s) = \rho Q/s - 1 + N(s), \tag{17}$$

where $M(s)$ and $N(s)$ are entire functions. The explicit terms taken out of $\mu(s)$ and $\nu(s)$ are just the hard-rod values of $\mu(s)$ and $\nu(s)$, with the actual Q replacing the hard-rod value of Q.

Substituting (17) into (13), (14), (15), and (16), we obtain

$$M(s)N(s) + \frac{Q}{s}[N(s) + \rho M(s)] - M(s)$$
$$= \frac{Q}{s}[1 - Q(1 - \rho l)] + P(s) + \Pi(s), \tag{18}$$

where

$$P(s) = \int_{-a}^{0} p(x) \exp(-sx)\,dx$$

and

$$\Pi(s) = \int_{0}^{a} q(x) \exp(-sx)\,dx,$$

$$N(s)N(-s) + \frac{\rho Q}{s}[N(-s) - N(s)] - N(-s) - N(s)$$
$$= \rho^2\left\{M(s)M(-s) + \frac{Q}{s}[M(-s) - M(s)]\right\}$$
$$= -\rho^2[Y(s) + Y(-s)], \tag{19}$$

and

$$G(s)\{e^{sl}[-\rho Q/s - 1 + N(-s)]$$
$$+ \rho[Q/s + M(s)]\} = -Q/s - M(s). \tag{20}$$

The right-hand side of (18) is the LT of a function which has the constant value $Q[1 - Q(1 - \rho l)]$ for $x > a$. The inverse LT of the left-hand side of (18) will certainly be of this form if $M(s)$ and $N(s)$ are the LT's of functions $m(x)$ and $n(x)$, one of which is identically zero except in the interval

$0 < x < a$, while the other vanishes identically outside the interval $-a < x < 0$. Similarly the right-hand side of (19) is the LT of a function which is zero for $x > a$, and the left-hand side of (19) will automatically have this behavior if $m(x)$ and $n(x)$ are confined to $-a < x < 0$ and $0 < x < a$.

The correct choice between these two alternatives can be made by testing the consistency of each choice with the two relations between $g(x)$, $q(x)$, $p(x)$, and $y(x)$ already obtained, namely the definition of $y(x) = \int_{x+l}^{a+l} q(x')g(x' - x)\,dx'$ and (10). The correct choice turns out to be $m(x)$ confined to $0 < x < a$, $n(x)$ confined to $-a < x < 0$.

Taking the inverse LT of (18) we obtain

$$p(-x) = Q\int_x^a n(-x')\,dx' + \int_x^a m(x' - x)n(-x')\,dx', \tag{21}$$

$$q(x) = -m(x) - \rho Q\int_x^a m(x')\,dx' + \int_x^a m(x')n(x - x')\,dx', \tag{22}$$

$$1 - Q(1 - \rho l) = \rho\int_0^a m(x')\,dx' + \int_0^a n(-x')\,dx'. \tag{23}$$

The inverse LT of (19) yields

$$\int_x^a n(-x')n(x - x')\,dx' + \rho Q\int_x^a n(-x')\,dx' - n(-x) = \rho^2\left[\int_x^a m(x')m(x' - x)\,dx' + Q\int_x^a m(x')\,dx'\right] = -\rho^2 y(x), \tag{24}$$

and from (20) we obtain, defining $\gamma(x) = g(x + l)$, for the interval $0 < x < a$

$$-\gamma(x) - \rho Q\int_0^x \gamma(x')\,dx' + \int_0^x \gamma(x')n(x' - x)\,dx' = -Q - m(x). \tag{25}$$

Equations (22), (24), and (25) form a closed set of three coupled quadratic integral equations for the quantities $m(x)$, $n(x)$ and $q(x)$—$\gamma(x)$ being related to $q(x)$ by $-\gamma(x)/f_T(x + l) = q(x)e_T(x + l)$—with the auxiliary equation (23) relating the value of the parameter Q to $m(x)$ and $n(x)$. Thus the pair distribution function has been eliminated completely, and the interval of integration in the equation for the direct correlation function has been reduced to the range of the tail potential alone. It is a characteristic feature of this reduced form that the square root of the compressibility appears as a parameter, whereas the compressibility itself appears in the original PY equation.

C. Further Considerations

It is still necessary to demonstrate the consistency of (21) through (25) with the definition of y,

$$y(x) = \int_x^a q(x')\gamma(x' - x)\,dx', \tag{26}$$

and (10), which may be rewritten as

$$p(-x) = -\rho^2\int_x^a y(x')\gamma(x' - x)\,dx'. \tag{27}$$

Writing (22) as an equation for $q(x')$, multiplying by $\gamma(x' - x)$ and integrating from x to a, then using (25) to eliminate $\gamma(x)$, we obtain precisely $y(x)$ in terms of $m(x)$ as given by (24). Similarly we may use (24) in the form which gives $y(x')$ as a function of $n(x')$. Multiplying by $\gamma(x' - x)$, integrating from x to a, and using (25) to eliminate $\gamma(x)$, we obtain exactly (21). Had we assumed $m(x)$ confined to $-a < x < 0$ and $n(x)$ confined to $0 < x < a$, then we would have found a contradiction at this point.

One may further verify that, defining a function $\xi(x) = Q\theta(x, l) - \rho^{-1}n(-x) - m(x - l)$, where $\theta(x, l)$ is the step function which is 1 for $0 < x < l$, and zero otherwise, the direct correlation function is given

$$-C(x) = \xi(-x) + \xi(x) + \rho\int_{-\infty}^{\infty} \xi(x')\xi(x' - x)\,dx'. \tag{28}$$

Finally we consider some important consequences of Eq. (15). We may expand both sides of (15) in a Laurent series and equate coefficients for each power of s. In the case $Q = 0$, the leading equation is

$$[N(0) - 1]^2 = \rho^2 M^2(0) + 1. \tag{29}$$

But (24) becomes, for the case $Q = 0$,

$$N(0) - 1 = \rho M(0). \tag{30}$$

The two equations are obvioulsy inconsistent for finite values of $M(0)$ and $N(0)$. Therefore Q can

never attain the value zero for finite $M(0)$ and $N(0)$. It follows that there can be neither a straight line segment in the P–V curve, nor a van der Waals loop. We conclude that the system does not exhibit a first-order phase transition. This seems to be a strong point in favor of the PY equation, since the exact treatment of this system also shows that no phase transition can occur, as first shown by Gürsey.[6]

There remains some question whether Q can become zero through $m(x)$ and $n(x)$ increasing without limit as ρ approaches some density ρ_0. By (15) this would imply that $m(x)$ approaches $n(-x)$ and $g(x)$ approaches $\delta(x-l)$ for $1 \leq x < 2$. We also have $q(x)f(x+l) = -e(x+l)g(x+l)$. On the other hand, $q(x)$ is related to $m(x)$ and $n(x)$ via (22). To make the two relations for $q(x)$ consistent, enormous cancellation of large quantities on the right-hand side of (22) would be required, and it seems unlikely that this is possible.

Next we show how to write $g(x)$ as a sum of nth nearest neighbor distribution functions.

In taking the inverse LT of (16), we can expand the denominator in powers of $\rho\mu(s)/\exp(sl)\nu(-s)$ if the inequality $\rho|\mu(s)| < |\nu(-s)\exp sl|$ is satisfied on the contour. The path of integration consists of the y axis from $-i\infty$ to $-i\epsilon$, a small semicircle of radius ϵ in the RHP, and the y axis from $i\epsilon$ to $i\infty$. On the y axis (15) is equivalent to

$$|\nu(iy)|^2 = \rho^2|\mu(iy)|^2 + 1 \tag{31}$$

and the inequality is obviously satisfied. Letting $s = \epsilon\exp i\theta$ on the small semicircle, we have

$$\begin{aligned} |s\nu(-s)e^{sl}|^2 &= \rho^2Q^2 \\ &\quad + 2\rho Q[1 - N(0) + Q\rho l]\epsilon\cos\theta + O(\epsilon^2), \\ \rho^2|\mu(s)|^2 &= \rho^2Q^2 + 2\rho^2QM(0)\epsilon\cos\theta + O(\epsilon^2), \end{aligned} \tag{32}$$

and, using (23),

$$|s\nu(-s)e^{sl}|^2 - \rho^2|\mu(s)|^2 = 2\rho Q\epsilon\cos\theta + O(\epsilon^2). \tag{33}$$

Since $Q > 0$, and $\cos\theta > 0$, the inequality is satisfied on the small semicircle also. Therefore we can expand $G(s) = \sum_{n=1}^{\infty} G_n(s)$, where $G_n(s) = [\mu(s)/\nu(-s)\exp sl]^n$.

Because of the factor $\exp sl$, the inequality is also satisfied on a large semicircle in the RHP. Thus we have a closed contour on which $|\nu(-s)\exp sl| > |\mu(s)|$. Our original hypothesis that $G(s)$ has no pole in the RHP implies that $\nu(-s)\exp sl + \rho\mu(s)$ has no zero in the RHP, since in (16) a zero of the denominator cannot be canceled by a zero of $\mu(s)$, $\mu(s)$ and $\nu(-s)$ by (15) having no common zero. It follows that we can apply Rouché's theorem, and we find that $\nu(-s)$ has no root in the RHP.

In calculating

$$g_n(x) = \frac{1}{2\pi}\int_{-i\infty}^{+i\infty} G_n(s)e^{sx}\,dx \tag{34}$$

we close in the RHP for $x < nl$, obtaining zero. For $x > nl$ we close the contour in the LHP, obtaining contributions from the residues of the poles at the zeros of $\nu(-s)$ in the LHP. We also have the relation

$$\int_0^{\infty} g_n(x)\,dx = G_n(0) = \rho^{-1}, \tag{35}$$

which tells us that each $g_n(x)$ is normalized to exactly one particle. It is therefore natural to interpret $g_n(x)$ as the nth nearest neighbor distribution function.

III. THREE DIMENSIONS

A. Preliminary Reduction

In the following we omit mathematical details whenever the argument closely parallels the one-dimensional case, concentrating instead on the important differences between the two cases.

In three dimensions the PY equation can be written[2]:

$$\begin{aligned} \tau(\mathbf{r}) &= 1 - \rho\int \tau(\mathbf{r})f(\mathbf{r})\,d\mathbf{r} \\ &\quad + \rho\int \tau(\mathbf{r}')f(\mathbf{r}')\tau(\mathbf{r}-\mathbf{r}')\,e(\mathbf{r}-\mathbf{r}')\,d\mathbf{r}'. \end{aligned} \tag{36}$$

We assume a spherically symmetric solution, and introduce the coordinates $r' = |\mathbf{r}'|$, $|\mathbf{r}-\mathbf{r}'|$, and a trivial polar angle ϕ. We make the equations dimensionless by defining $x = (r/R)$, $\eta = (\pi R^3\rho/6)$, $a' = (a/R)$, where R is hard-sphere radius, $R + a$ the distance beyond which $V_T(r) = 0$, and take the one side LT of the equation satisfied by $\sigma(x) \equiv x\tau(x)$.

Defining

$$\begin{aligned} F(t) &\equiv -\int_0^{1+a'} xC(x)\exp(-tx)\,dx, \\ G(t) &\equiv \int_1^{\infty} xg(x)\exp(-tx)\,dx, \\ K &\equiv -F'(0) = -\int_0^{a'} x^2C(x)\,dx, \\ Y(t) &\equiv \int_0^{a'} y(x)\exp(-tx)\,dx, \\ y(x) &\equiv -\int_{1+x}^{1+a'} x'C(x')(x-x')g(x-x')\,dx', \end{aligned} \tag{37}$$

[6] F. Gürsey, Proc. Cambridge Phil. Soc. **46**, 182 (1950).

and noting that $1 + 24\eta K$ is the inverse compressibility $\beta(\partial p/\partial\rho)_\beta$, we obtain

$$t[F(t) + G(t)] = (1 + 24\eta K)/t + 12\eta \times \{[F(t) - F(-t)]G(t) - Y(t) + Y(-t)\}. \quad (33)$$

Solving for $G(t)$, this becomes

$$G(t) = \frac{(1 + 24\eta K)/t^2 - F(t) + (12\eta/t)[Y(-t) - Y(t)]}{1 + (12\eta/t)[F(-t) - F(t)]}. \quad (39)$$

Arguing as in the one-dimensional case, $F(t)$ and $Y(t)$ are entire functions, and $G(t)$ is regular in the RHP and has a double pole at $t = 0$. Furthermore,

$$G(-t) = \frac{(1 + 24\eta K)/t^2 - F(-t) + (12\eta/t)[Y(-t) - Y(t)]}{1 + (12\eta/t)[F(-t) - F(t)]} \quad (40)$$

and hence the function

$$t^4G(t)\{(1 + 24\eta K)t^{-2} - F(-t) + 12\eta t^{-1}[Y(-t) - Y(t)]\}$$

is an even entire function of t. We define

$$H(t) = t^4G(t)\{(1 + 24\eta K)t^{-2} - F(-t) + 12\eta t^{-1}[Y(-t) - Y(t)]\} + Y(t) + Y(-t).$$

By (39), $H(t)$ is given by

$$\begin{aligned} H(t) = {} & t^4\{1 + 24\eta t^{-1}[F(-t) - F(t)]\}^{-1} \\ & \times \{(1 + 24\eta K)^2t^{-4} + (1 + 24\eta K)t^{-2} \\ & \times [-F(t) - F(-t) + 24\eta t^{-1}Y(-t) \\ & - 24\eta t^{-1}Y(t)] + F(t)F(-t) \\ & - 24\eta t^{-1}F(t)Y(-t) + 24\eta t^{-1}F(-t)Y(t) \\ & + (12\eta t^{-1})^2[Y(t) - Y(-t)]^2\}. \end{aligned} \quad (41)$$

Examining the growth rates at large t in a fashion completely analogous to the one-dimensional case we find $H(t) = O(t^2)$ at large t. Therefore, by a well known extension of Liouville's theorem, $H(t)$ is a quadratic polynomial $H(t) = \lambda_1 + \lambda_2 t^2$.

Here the first important difference between one and three dimensions appears. In the original PY equation there appears one parameter, the value of which is given as an integral of the unknown function appearing in the integral equation, namely the inverse compressibility. In one dimension the factored equation also contains just one parameter, which is the square root of the inverse compressibility. In three dimensions, however, two parameters appear in the factored equation. As will be shown later, one of them is again the square root of the inverse compressibility.

Evaluating the constants by expanding $H(t)$ in powers of t in (41), we find

$$\lambda_1 = (1 + 24\eta K),$$

$$\lambda_2 = -2F(0) + 4\eta F'''(0) - 48Y'(0), \quad (42)$$

and

$$G(t)\left\{\frac{1 + 24\eta K}{t^2} - F(-t) + \frac{12\eta}{t}[Y(-t) - Y(t)]\right\} = \frac{\lambda_1}{t^4} + \frac{\lambda_2}{t^2} - Y(t) - Y(-t). \quad (43)$$

Combining (43) and (33) one obtains

$$\begin{aligned} tF(t) = {} & \frac{\lambda_1}{t} + \frac{12\eta\lambda_2}{t^2} + \frac{12\eta\lambda_1}{t^4} - 24\eta Y(t) \\ & - \frac{(12\eta)^2Y(-t)G(t)}{t} + \frac{(12\eta)^2Y(t)G(t)}{t} \\ & + 12\eta F(t)G(t) - \frac{12\eta\lambda_1G(t)}{t^2} - tG(t). \end{aligned} \quad (44)$$

We first solve the case of zero tail. Then

$$G(t)[12\eta F(t) - 12\lambda_1t^{-2} - t]$$

is the LT of a function which is zero for $x < 1$. For $x < 1$ we obtain

$$-C(x) = \lambda_1 + 6\eta\lambda_2x + \tfrac{1}{2}\eta\lambda_1x^3. \quad (45)$$

Combining (45) with the relations (42), we obtain two simultaneous linear equations for λ_1 and λ_2:

$$\begin{aligned} \lambda_1 &= 1 + 24\eta(\tfrac{1}{3}\lambda_1 + \tfrac{3}{2}\eta\lambda_2 + 1\tfrac{1}{2}\eta\lambda_1), \\ \lambda_2 &= -2(\tfrac{1}{2}\lambda_1 + 2\eta\lambda_2 + 1\tfrac{1}{10}\eta\lambda_1) \\ &\quad - 4\eta(\tfrac{1}{5}\lambda_1 + \eta\lambda_2 + 1\tfrac{1}{6}\eta\lambda_1), \end{aligned} \quad (46)$$

the solution of which is

$$\lambda_1 = (1 + 2\eta)^2/(1 - \eta)^4, \quad \lambda_2 = -(1 + \tfrac{1}{2}\eta)^2/(1 - \eta)^4. \quad (47)$$

From λ_1 we obtain the equation of state

$$\beta P\rho^{-1} = (1 + \eta + \eta^2)/(1 - \eta)^3. \quad (48)$$

The factorization method is a much simpler and physically more transparent way of solving the PY equation for hard spheres than the method originally used.[3,4]

B. Finite-Range Equations

Returning to (44) for the case of a nonvanishing tail, we see that $-[xC(x)]'$ can be written

$$-[xC(x)]' = -[xC_0(x)]' - 24\eta y(x) + p(x-1) - q(x-1), \quad (49)$$

where $-[xC_0(x)]' = \lambda_1 + 2\eta\lambda_2 + 2\eta\lambda_1 x^3$ for $0 < x < 1$, and $-[xC_0(x)]' = 0$ for $x > 1$, $y(x)$ is confined to the interval $0<x<a'$, $p(x)$ to $-a'<x<0$, and $q(x)$ to $0 < x < a'$.

Defining

$$X(t) = F(t) - \lambda_1 t^{-2} - 12\eta\lambda_2 t^{-3} - 12\eta\lambda_1 t^{-5} + 24\eta t^{-1} Y(t),$$

$$Z(t) = \lambda_1 t^{-4} + \lambda_2 t^{-2} - Y(t) - Y(-t),$$

we obtain from (39) and (43)

$$\left[X(t) + \frac{12\eta}{t} Z(t)\right]\left[X(-t) - \frac{12\eta}{t} Z(t)\right] = Z(t)\left[1 - \frac{12\eta}{t} X(t) + \frac{12\eta}{t} X(-t) - \frac{2(12\eta)^2}{t^2} Z(t)\right], \quad (50)$$

or

$$X(t)X(-t) = Z(t) - \frac{(12\eta)^2}{t^2} Z^2. \quad (51)$$

Letting $X(t) = e^{-t}\mu(t)\nu(t)$, and $Z(t) = \mu(t)\mu(-t)$, then

$$\nu(t)\nu(-t) = 1 - (12\eta/t)^2\mu(t)\mu(-t) \quad (52)$$

and

$$G(t) = \frac{-\mu(t)}{e^t\nu(-t) - (12\eta/t)\mu(t)}. \quad (53)$$

The singularities of $X(t)$ and $Z(t)$ tell us immediately that the leading powers in the Laurent series are t^{-2} for $\mu(t)$ and t^{-3} for $\nu(t)$. We let

$$\mu(t) = A_2 t^{-2} + A_1 t^{-1} + M(t),$$
$$\nu(t) = B_3 t^{-3} + B_2 t^{-2} + B_1 t^{-1} + 1 + N(t), \quad (54)$$

where $M(t)$ and $N(t)$ are entire functions. We now compare the two equations for $G(t)$, (39) and (53). By (39) $G(t)$ is given in the neighborhood of $t = 0$ by $G(t) = t^{-2} + O(1)$. In (39) the numerator yields $-\mu(t) = -A_2 t^{-2} - A_1 t^{-1} + O(1)$, hence consistency with (53) requires that near zero the denominator satisfy $e^t\nu(-t) - 12\eta t^{-1}\mu(t) = -A_2 - A_1 t + O(t^2)$. Equating the principal part of $e^t\nu(-t) - 12\eta t^{-1}\mu(t)$ to zero yields

$$B_3 = -12\eta A_2, \qquad B_2 = 12\eta(A_1 - A_2),$$
$$B_1 = 12\eta[A_1 - \tfrac{1}{2}A_2 - M(0)].$$

The equations for the coefficients of t^1 and t^0 yield, after expressing the B's in terms of A_2 and A_1:

$$(1 + 2\eta)A_2 - 6\eta A_1 = -1 - N(0) - 12\eta M(0) + 12\eta M'(0),$$
$$\tfrac{3}{2}\eta A_2 + (1 - 4\eta)A_1 = -1 - N(0) + N'(0) - 6\eta M(0) - 12\eta M''(0). \quad (55)$$

Having obtained the parameters A_2 and A_1 in terms of the functions $M(t)$ and $N(t)$, we obtain λ_1 and λ_2 in terms of A_1 and A_2 by equating coefficients in the Laurent series of $\mu(t)\nu(t)e^{-t} = X(t)$. The coefficients of t^{-5} and t^{-3} tell us that

$$A_2^2 = \lambda_1, \qquad A_1^2 - 2A_2 M(0) = -\lambda_2. \quad (56)$$

By an argument closely paralleling the one-dimensional case, $N(t)$ is the LT of a function $n(x)$ confined to $-a' < x < 0$ and $M(t)$ is the LT of $m(x)$ which is confined to $0<x<a'$. We note that $t^{-1}[M(t)-M(0)]$ is also the LT of a function confined to $0 < x < a'$ and equal to $\hat{m}(x) = \int_x^a m(x')\,dx'$ in that interval. Hence $e^{-t}\nu(t) + 12\eta t^{-1}\mu(-t)$ is the LT of $n(1 + x) + \hat{m}(-x) + \delta(x-1) + 12\eta[M(0) - A_1 x + \frac{1}{2}A_2 x^2]\theta(x, 1)$. By (50), expressed in terms of μ and ν, we have

$$1 + \frac{12\eta}{t}[F(-t) - F(t)] = \left[e^{-t}\nu(t) + \frac{12\eta}{t}\mu(-t)\right] \times \left[e^t\nu(-t) - \frac{12\eta}{t}\mu(t)\right]. \quad (57)$$

Recalling the definition of $F(t)$, the inverse LT of (57) can be written

$$\int_{-\infty}^{x} x'C(x')\,dx' = \xi(x) + \xi(-x) + 12\eta\int_{-\infty}^{\infty}\xi(x')\xi(x' - x)\,dx', \quad (58)$$

where

$$\xi(x) = (12\eta)^{-1}n(-x) + \hat{m}(x-1) + [M(0) - A_1(1-x) + \tfrac{1}{2}A_2(1-x)^2]\theta(x, 1).$$

Defining $\gamma(x-1) = xg(x)$ for $0 < x < a'$, we now write down the three coupled integral equations for $m(x)$, $n(x)$, and $\hat{q}(x)$, where $\hat{q}(x) = \int_0^a q(x')\,dx'$. From $\mu(t)\nu(t)e^{-t} = X(t)$ we obtain for the interval $0<x<a$

$$\hat{q}(x) = m(x) + \int_x^{a'} m(x')n(x - x')\,dx' - 12\eta\int_x^{a'} m(x')[A_1 - \tfrac{1}{2}A_2 - M(0) + (A_1 - A_2)(x - x') - \tfrac{1}{2}A_2(x - x')^2]\,gx'. \quad (59)$$

From (52) for $0 < x < a'$ we obtain

$$n(-x) + \int_x^{a'} n(-x')n(x - x')\,dx' - 12\eta \int_x^{a'} n(-x')[A_1 - \tfrac{1}{2}A_2 - M(0) + (A_1 - A_2)(x - x') - \tfrac{1}{2}A_2(x - x')^2]\,dx' = 144\eta^2\Big[\int_x^{a'} \hat{m}(x')\hat{m}(x' - x)\,dx' + \int_x^{a'} \hat{m}(x') \times [M(0) - A_1(x - x') + \tfrac{1}{2}A_2(x - x')^2]\,dx'\Big]. \quad (60)$$

From (53) for $0 < x < a'$ we obtain

$$\gamma(x) + \int_0^x \gamma(x - x')\{n(-x') + 12\eta[\tfrac{1}{2}A_2x'^2 + (A_1 - A_2)x' - A_1 + \tfrac{1}{2}A_2 + M(0)]\}\,dx' = -[\tfrac{1}{2}A_2x + A_1 + m(x)]. \quad (61)$$

Here $\gamma(x)$ and $\hat{q}(x)$ are related by $-\gamma(x)f(1 + x) = \hat{q}(x)e(1 + x)$.

The equations corresponding to (59) and (60) for $x > a'$ just duplicate (55).

We note at this point that vanishing of the inverse compressibility does not lead to a contradiction analogous to the one encountered in one dimension. This does not prove that the PY equation in three dimensions can show a phase transition, or even that a solution for zero inverse compressibility does in fact exist, but at least this possibility is not excluded in the same obvious way that occurred in one dimension.

C. Structure of the Radial Distribution Function

Finally we consider the question of zeros of $\nu(-t)$ in the RHP, and the expansion of $g(x)$ obtained by expanding the denominator of (53) in powers of $\mu(t)/\nu(-t)$ exp t.

We consider first the case of positive compressibility. Taking the same contour considered for the one-dimensional case, we find that the inequality $|12\eta t^{-1}\mu(t)| < |\nu(-t)|$ is satisfied on the y axis and on the large semicircle, but not on the small semicircle. We compute the variation of the argument of $\nu(t)$ around the contour in two steps.

By our original hypothesis, $D(t) = \nu(-t)e^t - 12\eta t^{-1}\mu(t)$ has no zero in the RHP. Hence the variations of the argument $\phi(D)$ of $D(t)$ around the contour considered here is zero. Near $t=0$, $D(t) = -A + O(t)$. This means that the change of $\phi(D)$ on the small semicircle is zero, and by subtraction, the change of $\phi(D)$ going from $-i\epsilon$ to $i\epsilon$ counterclockwise is also zero. On this part of the contour $|12\eta t^{-1}\mu(t)| < |\nu(-t)|$, so that $|\phi(D) - \phi[\nu(-t)]| < \frac{1}{2}\pi$ everywhere, and it suffices to calculate $\phi(D) - \phi[\nu(-t)]$ at the end points to obtain the total change of $\phi[\nu(-t)]$. At $y = \epsilon$, $D(t) = -A_2 + O(\epsilon)$, $\nu(-t) = -12\eta A_2 i\epsilon^{-3} + O(\epsilon^{-2})$; at $y = -\epsilon$, $D(t) = -A_2 + O(\epsilon)$, $\nu(-t) = 12\eta A_2 i\epsilon^{-3} + O(\epsilon^{-2})$. Hence $\phi[\nu(-t)]$ changes by $-\pi$ on this part of the contour. The point $t = 0$ is a triple pole of $\nu(-t)$, and therefore $\phi[\nu(-t)]$ increases by 3π on the small semicircle. Thus the total change in $\phi[\nu(-t)]$ around the contour is 2π, and $\nu(-t)$ has exactly one root in the RHP. In the only solved case,[3,4] hard spheres with no tail, $\nu(-t)$ has two complex-conjugate roots in the LHP and one real root in the RHP.

For zero inverse compressibility, $D(t) = -A_1t + O(t^2)$ and $\nu(-t) = 12\eta A_1t^{-2}$ near $t = 0$. $\phi(D)$ increases by π going from $y = -\epsilon$ to $y = \epsilon$ counterclockwise. At $y = \epsilon$, $D(t) = -A_1i\epsilon + O(\epsilon^2)$; $\nu(-t) = -A_1\epsilon^{-2}$; at $y = -\epsilon$, $D(t) = A_1i\epsilon + O(\epsilon^2)$, $\nu(-t) = -A_1\epsilon^{-2} + O(\epsilon^{-1})$. Hence $\phi[\nu(-t)]$ increases by 0 on this part of the contour. On the small semicircle $\phi[\nu(-t)]$ increases by 2π, since $t = 0$ is a double pole. Thus the total variation of $\phi[\nu(-t)]$ around the contour is 2π. Therefore a solution with zero inverse compressibility also has one root of $\nu(-t)$ in the RHP.

In order to expand $G(t)$ in powers of $\mu(t)/\nu(-t)$ exp t, it is necessary to find a contour on which $|12\eta t^{-1}\mu(t)| < |\nu(-t)e^t|$ is satisfied. From our proof that $\nu(-t)$ has one root in the RHP it follows that there can be no such contour passing from $-i\infty$ to $i\infty$ to the left of the positive root of $\nu(-t)$, since the existence of such a contour would imply that $D(t)$ also has one root in the RHP. On the other hand it is clear that on a sufficiently large semicircle $|t| = P$ in the RHP $|12\eta t^{-1}\mu(t)| < |\nu(-t)|$ is satisfied. If one replaces the segment of the y axis between $y = -P$ and $y = P$ by $|t| = P$, one can expand on the contour. Then $g(x) = \sum_{n=1}^{\infty} g_n(x)$ and

$$xg_n(x) = \frac{1}{12\eta}\sum_i R_i\left\{e^{t(x-n)}\left[\frac{\mu(t)}{12\eta t\nu(-t)}\right]^n\right\} \quad (x > n),$$
$$xg_n(x) = 0 \quad (x < n), \quad (62)$$

where the R_i are the residues of the quantity in braces at the roots of $\nu(-t)$. Since there is always one root in the RHP, each $g_n(x)$ includes a term of exponentially increasing character. Therefore the $g_n(x)$ are never normalizable and never represent nth nearest-neighbor distributions.

In addition to being suitable for numerical computation, the coupled equations are suitable

for obtaining two expansions: a virial expansion in powers of the density, and a perturbation expansion in powers of $f_T(x)$. In the latter case the advantage of eliminating $g(x)$ is particularly striking; while even the first order is difficult to obtain from the original PY equation, the coupled equations in every order of $f_T(x)$ lead to simple linear integral equations which can be solved explicitly by reduction to integrals involving $f_T(x)$ and known functions.

ACKNOWLEDGMENT

The author wishes to thank Professor J. K. Percus for several helpful discussions throughout the course of this investigation.

FURTHER COMMENTS ON THE PRECEDING PAPER

by

M. Wertheim

(a) Comment on Eq. (7)

In both (2) and (6) the $F(-s)G(s)$ term is the only one that has a nonvanishing inverse LT for $x < \ell - a$. It is therefore appropriate to take the linear combination which eliminates this term.

(b) Comments on Eq. (9)

Several comments are in order.

In the original method of solving the PY equation for hard spheres,[3,4] an additional spurious solution with physically unacceptable $g(r)$ appeared in both one and three dimensions. In the present calculation the spurious solutions no longer appear, since the assumption that $G(s)$ is regular in the RHP eliminates them.

The correct sign of Q is found by reinserting (7) in (2) and (6). This yields $Q = + (1-\rho\ell)^{-1}$.

The PY equation is exact for hard rods, as first shown by graph theoretical methods.[7]

(c) Comments on Eq. (28)

Expressing the bracketed factor on the RHS of (11) in terms of F(s), all other factors in (11) in terms of $\mu(s)$ and $\nu(s)$ yields

$$1 + \rho F(s) + \rho F(-s) = [e^{-s\ell}\nu(-s) - \rho\mu(s)][e^{s\ell}\nu(s) - \rho\mu(-s)]$$

The inverse LT of this is (23).

(d) Comments on the Absence of a Phase Transition [paragraphs following Eqs. (30) and (31)].

Two comments may be made. As far as the question of a phase transition is concerned, the restriction to a short range tail potential is unnecessary, as subsequently shown by the following more general argument due to J. K. Percus[8] (and verified independently by J. Lebowitz):

Letting C_k and u_k be the Fourier transforms of $C(r)$ and $u(r) \equiv g(r) - 1$, the definition of $C(r)$ is expressed by $u_k = C_k(1-\rho C_k)^{-1}$. Integrating over k, one obtains

$$u(0) = \int \frac{C_k}{1-\rho C_k}\, dk$$

On the left-hand side, the hard sphere part of the potential implies $g(0) = 0$, or $u(0) = -1$. On the right-hand side zero inverse compressibility implies that $1 - \rho C_o = 0$.

Hence near $\bar{k} = 0$ $1 - \rho C_k = O(k^2)$. The integral diverges near $k = 0$ in the cases of one or two dimensions, leading to a contradiction. In three dimensions $d\bar{k} = 4\pi k^2 dk$, so that the integral converges and a phase transition is not ruled out. The validity of the preceding argument depends on $C(\bar{r})$ remaining an integrable function. In the special case treated in the paper this is represented by $M(0)$ and $N(0)$ remaining finite.

The second comment refers only to the case of short range $V_T(x)$ treated in the paper. Here we can rigorously exclude zero inverse compressibility, if $V_T(x) \geq 0$. Correcting a misprint in the paper, we have $g(x)e(x+\ell) = -\gamma(x)\varphi(x+\ell)$. Using this in (26) shows that for positive $V_T(x)$, $y(0) > 0$. Hence by (24)

$$\int_0^a m^2(x)dx + QM(0) < 0$$

Furthermore, by Schwarz's inequality

$$a\int m^2(x)dx \geq M^2(0).$$

Combining the two inequalities,

$$Qa > |M(0)|$$

and the possibility of Q going to zero and $|M(0)|$ going to infinity is ruled out. In the preceding it was not assumed that $\gamma(x) \geq 0$. No parallel argument is known for the more interesting case: $V_T(x) < 0$.

(e) Comment on Eq. (35)

The relation $G_n(s) = G_1^n(s)$ between the n^{th} nearest neighbor distribution and the nearest neighbor distribution is the same for the PY approximation and for the exact solution of the problem. The exact nearest neighbor distribution for a one dimensional system interacting with arbitrary nearest neighbor forces is[9]

$$\rho g_1(x) = \frac{e^{-\beta V(x)-\beta P x}}{\int_0^\infty e^{-\beta V(x')-\beta P x'} dx'}$$

where $V(x)$ is the pair potential and P is the pressure. In the case of hard rods plus tail the corresponding PY result is of course not explicit, and a detailed assessment of the validity of the PY approximation requires the numerical solution of Eqs. (22)-(25).

(f) Comment on Eq. (48)

The PY equation of state of hard spheres (48) is identical with the result obtained by Reiss, Frisch, and Lebowitz[10] from the scaled particle model. It is also in good

agreement with molecular dynamics[11] and Monte Carlo[12] calculations of the equation of state of hard spheres at densities below the phase transition indicated by the latter.

(g) Comment on Eq. (56)

In the case of vanishing tail the values of A_1 and A_2 are $A_2 = -(1+2y)/(1-y)^2$ and $A_1 = -(1+\frac{1}{2}y)/(1-y)^2$.

(h) Comment on End of Section III.B

It should be noted, however, that there is no possibility of a negative compressibility, since it is always the square of a real number. Thus, if there is a phase transition in three dimensions, it will not appear as a Van der Waals loop. Instead one may conjecture that a phase transition appears in the following manner: as ρ increases $\beta \frac{dp}{d\rho}$ reaches the value 0 at some ρ_1, for $\rho_1 < \rho < \rho_2$ no solution of the PY equation exists, and at ρ_2 another solution begins, with $\beta \frac{dp}{d\rho}\Big|_{\rho_2} = 0$. Such a behavior has been found by Percus and Stell[13] for a lattice gas. See also comment (α).

(i) Comment on End of Section III.C

In concluding this section, it seems appropriate to mention an important unsolved problem: under what conditions will the g(r) given by the PY equation be positive for all r? In the case of hard rods, the PY equation is exact, and the predicted g(r) is, of course, properly positive. In

the case of hard spheres, however, the first minimum of $g(r)$ reaches zero at a density near $\eta = 0.61$ and becomes increasingly negative with increasing density.[14] The solution is therefore no longer physically acceptable for $\eta > 0.61$, despite the fact that the equation of states shows no discontinuity whatever at this density, and continues smoothly up to $\eta = 1$.

A well-known condition for positivity of a function $g(r)$ in terms of the properties of its LT $G(s)$ is the following: $g(r)$ is nonnegative if and only if $(-1)^n \frac{d^n G(s)}{ds^n} > 0$ for all integral $n \geq 0$, and all real positive s. The occurrence of all derivatives of $G(s)$ makes it unlikely that this condition can be applied in practice in cases where $G(s)$ is not known explicitly.

Extension to Mixtures of Hard Spheres

The method developed in the preceding sections can also be applied successfully to mixtures of hard spheres. This calculation has actually been carried out by Lebowitz.[15] In summarizing his results, we deviate somewhat from his presentation in order to make the formal similarity to a one component system of hard spheres plus tail even more obvious.

The generalization of the PY equation to an m component mixture can be written[15]

$$\underline{\underline{g}}(\underset{\sim}{r}) - \underline{\underline{C}}(\underset{\sim}{r}) = \underline{\underline{J}}\left[\underline{\underline{I}} - \int \underline{\underline{\rho}}\underline{\underline{C}}(\underset{\sim}{r})d\underset{\sim}{r}\right] + \int \underline{\underline{g}}(\underset{\sim}{r}-\underset{\sim}{r}')\underline{\underline{\rho}}\underline{\underline{C}}(\underset{\sim}{r}')d\underset{\sim}{r}'$$

$$g_{ij}(\underset{\sim}{r}) = \tau_{ij}(\underset{\sim}{r})e(\underset{\sim}{r}_{ij}) \qquad (63)$$

$$C_{ij}(\underset{\sim}{r}) = \tau_{ij}(\underset{\sim}{r})f(\underset{\sim}{r}_{ij})$$

where $\underline{\underline{g}}$ is an m-by-m matrix of pair distribution functions and $\underline{\underline{C}}$ the corresponding matrix of direct correlation functions. $\underline{\underline{I}}$ is a unit matrix, $\underline{\underline{J}}$ is a matrix which has all elements J_{ij} equal to one, and $\underline{\underline{\rho}}$ is the diagonal matrix of number densities. Furthermore the matrices $\underline{\underline{G}}$ and $\underline{\underline{C}}$ are symmetric:

$$\underline{\underline{g}}(\underset{\sim}{r}) = \underline{\underline{g}}(\underset{\sim}{r}); \quad \tilde{\underline{\underline{C}}}(\underset{\sim}{r}) = \underline{\underline{C}}(\underset{\sim}{r})$$

where the tilde indicates the transposed matrix.

A. One Dimension

As in the case of a single component, we start with the one dimensional case. For convenience we arrange the components in order of increasing rod length ℓ_i. We define the matrix U(x) by

$$\underline{\underline{U}}(x) = -\theta(x)\int \underline{\underline{g}}^{(-)}(x-x')\underline{\underline{\rho}}\underline{\underline{C}}(x')dx'$$

where $\theta(x)$ is the step function which is equal to one for positive x, zero for negative x, and

$$g^{(-)}(x) = \begin{cases} 0 & (x > 0) \\ g(x) & (x < 0) \end{cases}$$

From the geometry it is clear that U_{ij} is identically zero for $i \leq j$, and that is confined to the interval $0 < x < \lambda_{ij} = \frac{1}{2}|t_j - t_i|$ for $i > j$.

Defining

$$\underline{\underline{F}}(x) = -\int_0^\infty \underline{\underline{C}}(x)\exp(-sx)dx$$

$$\underline{\underline{G}}(x) = \int_0^\infty \underline{\underline{g}}(x)\exp(-sx)dx$$

$$\underline{\underline{X}}(s) = \int_0^\infty \underline{\underline{U}}(x)\exp(-sx)dx$$

$$\underline{\underline{K}} = -2\int_0^\infty \underline{\underline{C}}(x)dx,$$

we can write the generalized PY equation

$$\underline{\underline{F}}(s) + \underline{\underline{G}}(s) = \frac{1}{s}\underline{\underline{J}}[\underline{\underline{I}}+\underline{\underline{\rho}}\underline{\underline{K}}] - \underline{\underline{G}}(s)\underline{\underline{\rho}}[\underline{\underline{E}}(s) + \underline{\underline{F}}(-s)] - \underline{\underline{X}}(s) + \underline{\underline{X}}(-s) \tag{64}$$

or

$$\underline{\underline{G}}(s)[\underline{\underline{I}} + \underline{\underline{\rho}}\underline{\underline{F}}(s) + \underline{\underline{\rho}}\underline{\underline{F}}(-s)] = \frac{1}{s}\underline{\underline{J}}\underline{\underline{P}} - \underline{\underline{F}}(s) - \underline{\underline{X}}(s) + \underline{\underline{X}}(-s) \tag{65}$$

where

$$\underline{\underline{P}} \equiv \underline{\underline{I}} + \underline{\underline{\rho}}\underline{\underline{K}}$$

We introduce the symmetric matrices $\underline{\underline{\mathcal{F}}} \equiv \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{F}}\underline{\underline{\rho}}^{\frac{1}{2}}$, and $\underline{\underline{\mathcal{G}}} = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{G}}\underline{\underline{\rho}}^{\frac{1}{2}}$ and rewrite (65) as

$$\underline{\underline{\mathcal{G}}}[\underline{\underline{I}} + \underline{\underline{\mathcal{F}}}(s) + \underline{\underline{\mathcal{F}}}(-s)] = \frac{1}{s}\underline{\underline{\mathcal{P}}} - \underline{\underline{\mathcal{J}}}(s) - \underline{\underline{\mathcal{H}}}(s) + \underline{\underline{\mathcal{H}}}(-s) \tag{66}$$

where $\underline{\underline{\mathcal{H}}} = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{Y}}\underline{\underline{\rho}}^{\frac{1}{2}}$ and $\underline{\underline{\mathcal{P}}} = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{P}}\underline{\underline{\rho}}^{\frac{1}{2}}$.

Consider the matrix

$$\underline{\underline{\mathcal{G}}}(s)\left[\frac{1}{s}\underline{\underline{\mathcal{P}}} + \underline{\underline{\mathcal{F}}}(-s) - \underline{\underline{\tilde{\mathcal{H}}}}(s) + \underline{\underline{\tilde{\mathcal{H}}}}(-s)\right] = \underline{\underline{L}}(s) \tag{67}$$

which can be written in the form $\underline{\underline{A}}(s)\underline{\underline{D}}^{-1}(s)\underline{\underline{\tilde{A}}}(-s)$, where $\underline{\underline{A}}(s) = \frac{1}{s}\underline{\underline{\mathcal{P}}} - \underline{\underline{\mathcal{J}}} - \underline{\underline{\mathcal{H}}} + \underline{\underline{\mathcal{H}}}(-s)$, and $\underline{\underline{D}} = \underline{\underline{I}} + \underline{\underline{\mathcal{F}}} + \underline{\underline{\mathcal{F}}}(-)$ satisfies $\underline{\underline{D}}(s) = \underline{\underline{D}}(-s) = \underline{\underline{\tilde{D}}}(s)$. Hence $\underline{\underline{L}}(s)$ satisfies $\underline{\underline{\tilde{L}}}(s) = \underline{\underline{L}}(-s)$.

Now $s\underline{\underline{A}}(s)$ is entire - by this we mean that each A_{ij} is an entire function - and $s\underline{\underline{G}}(s)$ must be regular in the RHP. The condition $\tilde{\underline{\underline{L}}}(s) = \underline{\underline{L}}(-s)$ then implies that each element L_{ij} is entire. We next consider the matrix

$$\underline{\underline{H}}(s) = s^2\left\{\underline{\underline{C}}(s)\left[\frac{1}{s}\tilde{\underline{\underline{\rho}}} + \underline{\underline{J}}(-) - \tilde{\underline{\underline{N}}} + \tilde{\underline{\underline{N}}}(-)\right] - \tilde{\underline{\underline{N}}} - \underline{\underline{N}}(-)\right\} \quad (68)$$

which also satisfies $\tilde{\underline{\underline{H}}}(s) = \underline{\underline{H}}(-s)$.

We consider the growth rate on a ray in the LHP. From (64) we obtain for the limiting form of $\underline{\underline{C}}$: $\underline{\underline{C}}(s) = -\underline{\underline{I}}$. Thus the limiting form of $\underline{\underline{H}}(s)$ is

$$H(s) = s\underline{\underline{\rho}}^{\frac{1}{2}}[-\tilde{\underline{\underline{P}}}\underline{\underline{J}} + \underline{\underline{\tau}}(0) + \underline{\underline{U}}(0) + \tilde{\underline{\underline{U}}}(0)]\underline{\underline{\rho}}^{\frac{1}{2}} + O(1) \quad (69)$$

Applying Liouville's theorem, we conclude that $\underline{\underline{H}}(s) = \underline{\underline{A}} + \underline{\underline{B}}s$. $\underline{\underline{B}}$ is of course antisymmetric and therefore identically zero in the one component case. We proceed to evaluate $\underline{\underline{A}}$ and $\underline{\underline{B}}$ by expanding $\underline{\underline{H}}(s)$ around $s = 0$, using (65). The result is

$$\underline{\underline{A}} = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{P}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}} \qquad \underline{\underline{B}} = \tfrac{1}{2}\underline{\underline{\rho}}^{\frac{1}{2}}[\underline{\underline{P}}\underline{\underline{J}} - \underline{\underline{J}}\underline{\underline{P}}]\underline{\underline{\rho}}^{\frac{1}{2}} \quad (70)$$

Combining (64) with (70) to eliminate $\underline{\underline{C}}\underline{\underline{\tilde{J}}}(-)$ leads to

$$\underline{\underline{\tilde{J}}} = \frac{1}{2s}\underline{\underline{\rho}}^{\frac{1}{2}}(\underline{\underline{J}}\underline{\underline{P}}+\underline{\underline{\tilde{P}}}\underline{\underline{J}})\underline{\underline{\rho}}^{\frac{1}{2}} - \frac{1}{s^2}\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{\tilde{P}}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}} - \underline{\underline{H}} - \underline{\underline{\tilde{H}}} - \underline{\underline{C}}\left[\underline{\underline{I}} + \underline{\underline{\tilde{J}}} + \frac{1}{s}\underline{\underline{\rho}} - \underline{\underline{\tilde{H}}}(s) + \underline{\underline{\tilde{H}}}(-s)\right. \tag{71}$$

After taking the inverse LT, we have

$$-\underline{\underline{C}}(x) = \tfrac{1}{2}(\underline{\underline{J}}\underline{\underline{P}}+\underline{\underline{\tilde{P}}}\underline{\underline{J}}) - \underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{\tilde{P}}}\underline{\underline{J}}x - \underline{\underline{U}}(x) - \underline{\underline{\tilde{U}}}(x) \tag{72}$$

In the preceding $\underline{\underline{U}}(x)$ has played a role similar to that of $y(x)$ in the case of a one component system of hard spheres plus tail. There is, however, one welcome difference: $\underline{\underline{U}}(x)$ is known explicitly. We obtain directly from (63), that

$$\tau_{ij}(x) = (\underline{\underline{J}}\underline{\underline{P}})_{ij} \quad \text{for} \quad 0 < x < \lambda_{ij} \equiv \tfrac{1}{2}|\ell_i - \ell_j|, \quad i < j \tag{73}$$

Combining this with (72) shows us that U_{ij} is linear in x, that we have in fact

$$\begin{aligned} \tau_{ij} &= \tfrac{1}{2}(\underline{\underline{J}}\underline{\underline{P}}+\underline{\underline{\tilde{P}}}\underline{\underline{J}})_{ij} - dx \qquad \lambda_{ij} < x < \ell_{ij} \\ \tau_{ij} &= \tfrac{1}{2}(\underline{\underline{J}}\underline{\underline{P}}+\underline{\underline{\tilde{P}}}\underline{\underline{J}})_{ij} - d\lambda_{ij} \qquad 0 < x < \lambda_{ij} \end{aligned} \tag{74}$$

and that

$$(\tilde{\underline{\underline{P}}}\underline{\underline{J}}-\underline{\underline{J}}\underline{\underline{P}})_{ij} = d(R_j-R_i) \quad (i < j) \tag{75}$$

Here d is the value of every element of the matrix $\underline{\underline{J}}\rho\tilde{\underline{\underline{P}}}\underline{\underline{J}}$.

We could now solve simultaneous linear equations to obtain the constants in the matrix $\underline{\underline{P}}$. Fortunately, the form allows us to guess the solution. It is then easy to verify that it satisfies all conditions. By analogy with the one component system we look for a solution of the form

$$-\underline{\underline{C}}(x) = \underline{\underline{\xi}}(x) + \tilde{\underline{\underline{\xi}}}(-x) + \int_{-\infty}^{\infty} \underline{\underline{\xi}}(x')\rho\tilde{\underline{\underline{\xi}}}(x'-x)dx'$$

To obtain a piecewise linear $\underline{\underline{C}}$, with the breaks in the $\underline{\underline{C}}'(x)$ the right distance apart, it is clear that we need a matrix of step functions, all of the same height, with widths R_i. Actually

$$\xi_{ij} = Q\theta[x - \tfrac{1}{2}(R_j-R_i),R_i]$$

We now verify that (74) is satisfied. We have from (75)

$$\underline{\underline{J}}\underline{\underline{P}} = [\underline{\underline{J}}+Q\underline{\underline{J}}\rho\underline{\underline{R}}\underline{\underline{J}}+Q\underline{\underline{J}}\rho\underline{\underline{J}}\underline{\underline{R}}+Q\underline{\underline{J}}\rho\underline{\underline{R}}\underline{\underline{J}}\rho\underline{\underline{J}}\underline{\underline{R}}] \tag{76}$$

which reduces to

$$(\underline{\underline{J}}\underline{\underline{P}})_{ij} = [1 + Q\ \mathrm{Tr}(\rho\underline{\underline{R}})][1 + QR_j\mathrm{Tr}\ \underline{\underline{\rho}}] \tag{77}$$

Also from (75) we obtain

$$\tau_{ij}(0) = Q + Q^2 R_< \mathrm{Tr}\ \underline{\underline{\rho}} \tag{78}$$

where $R_< = \mathrm{Min}(R_i, R_j)$ and Tr indicates the trace. Equating the two expressions, we obtain two roots, the correct $Q = [1 - \mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}})]^{-1}$, as well as an extraneous root $Q = -[\mathrm{Tr}\ \rho]^{-1}$.

Next we test relation (73). It turns out that

$$(\underline{\underline{\tilde{P}}}\underline{\underline{J}} - \underline{\underline{J}}\underline{\underline{P}})_{ij} = Q\ \mathrm{Tr}\ \underline{\underline{\rho}}[1 + Q\ \mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}})](R_j - R_i) \tag{79}$$

and

$$d(R_j - R_i) = \mathrm{Tr}\ \underline{\underline{\rho}}[1 + Q\ \mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}})]^2(R_j - R_i) \tag{80}$$

Thus there are again two roots, the extraneous root $Q = -\mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}})$, and the correct root $Q = [1 - \mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}})]^{-1}$. This step completes the proof that the guessed solution is correct.

If we compare the explicit $\underline{\underline{C}}$ with $\underline{\underline{\xi}}$, then the greater simplicity of $\underline{\underline{\xi}}$ is even more striking for mixtures, then for the one component system.

As in the case of a single component, the solution of the PY equation is exact for a mixture of hard rods.

B. Three dimensions

Following the preceding detailed calculation in one dimension, we present the three-dimensional case with little comment. The Laplace transform of the generalized PY equation in three dimensions is

$$\underline{\underline{G}}(s)\left\{\underline{\underline{I}} + \frac{1}{s}\rho[\underline{\underline{F}}(-s) + \underline{\underline{F}}(s)]\right\} = \frac{1}{s^2}\underline{\underline{J}}[\underline{\underline{I}}+\rho\underline{\underline{K}}] = \underline{\underline{F}}(s) + \frac{1}{s}\underline{\underline{X}}(s) - \frac{1}{s}\underline{\underline{X}}(-s) \tag{81}$$

where

$$\underline{\underline{F}}(s) = -\int_0^\infty \underline{\underline{C}}(r)r\exp(-sr)dr$$

$$\underline{\underline{G}}(s) = \int_0^\infty \underline{\underline{g}}(r)r\exp(-sr)dr$$

$$\underline{\underline{K}} = -2\int_0^\infty \underline{\underline{C}}(r)r^2dr = -2\underline{\underline{F}}'(0)$$

$$\underline{\underline{X}}(s) = \int_0^\infty \underline{\underline{U}}(x)\exp(-sx)\,dx$$

$$\underline{\underline{U}}(r) = \theta(r)\int_0^\infty \underline{\underline{g}}^-(r-r')\underline{\underline{C}}(r')(r-r')r'\,dr'$$

where $\theta(r)$ is the step function which is zero for $r < 0$ and 1 for $r > 0$ and $g^-(r)$ satisfies

$$g^-(r) = \begin{cases} g(-r) & r < 0 \\ 0 & r > 0 \end{cases}$$

and $\underline{\rho}$ is 2π times the diagonal matrix of number densities. Introducing $\underline{\mathcal{J}} = \underline{\rho}^{\frac{1}{2}}\underline{\rho}\underline{F}^{\frac{1}{2}}$, $\underline{\mathcal{G}} = \underline{\rho}^{\frac{1}{2}}\underline{G}\underline{\rho}^{\frac{1}{2}}$, $\underline{\kappa} = \underline{\rho}^{\frac{1}{2}}\underline{X}\underline{\rho}^{\frac{1}{2}}$, and $\underline{\mathcal{P}} = \underline{\rho}^{\frac{1}{2}}\underline{J}[\underline{I}+\underline{\rho}\underline{K}]\underline{\rho}^{\frac{1}{2}}$ this becomes

$$\underline{\mathcal{G}}(s)\left[\underline{I} - \frac{1}{s}\underline{\mathcal{J}}(s) + \frac{1}{s}\underline{\mathcal{J}}(-s)\right] = \frac{1}{s^2}\underline{\rho} - \underline{\mathcal{J}}(s) + \frac{1}{s}\underline{\kappa}(s) - \frac{1}{s}\underline{\kappa}(-s) \tag{82}$$

Arguing as before, we obtain the additional relation

$$\underline{\mathcal{G}}(s)\left[\frac{1}{s^2}\underline{\rho} - \underline{\mathcal{J}}(-s) - \frac{1}{s}\underline{\mathcal{K}}(-s) + \frac{1}{s}\underline{\mathcal{K}}(-s)\right] = \sum_{n=1}^{4} s^{-n}\underline{A}_n - \underline{\mathcal{K}}(s) + \underline{\kappa}(-s) \tag{83}$$

In evaluating the constant matrices $\underline{A}_n$ we use (82) multiplied by $\left[\underline{\mathcal{J}} - \frac{1}{s}\underline{\mathcal{J}}(s) + \frac{1}{s}\underline{\mathcal{J}}(-s)\right]^{-1}$. The assumption of the existence of the inverse matrix can be justified from the final solution. Using the fact that $(\underline{L}+\underline{\varepsilon})^{-1} = \underline{L}^{-1} + \underline{L}^{-1}\underline{\varepsilon}\underline{L}^{-1} + O(\underline{\varepsilon}^2)$, we obtain

$$\underline{\underline{A}}_4 = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}(\underline{\underline{\rho}}+\underline{\underline{\rho}}\underline{\underline{K}}\underline{\underline{\rho}})\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}$$

$$\underline{\underline{A}}_3 = 0$$

$$\underline{\underline{A}}_2 = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}[-\underline{\underline{\mathfrak{F}}}(0) + 2\underline{\underline{\mathfrak{X}}}'(0)] + [-\underline{\underline{\mathfrak{F}}}(0) + 2\underline{\underline{\mathfrak{X}}}'(0)]\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}} + \tfrac{1}{3}\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{\mathfrak{F}}}'''(0)\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}} \tag{84}$$

$$\underline{\underline{A}}_1 = \underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{\mathfrak{F}}}'(0) - \underline{\underline{\mathfrak{F}}}'(0)\underline{\underline{\rho}}^{\frac{1}{2}}\underline{\underline{J}}\underline{\underline{\rho}}^{\frac{1}{2}}$$

Combining (82), (83), and (84) and giving back to coordinate space we obtain for $0 < r < \frac{1}{2}(R_i+R_j)$

$$r\underline{\underline{C}}'(r) = (\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{F}}'(0) + \underline{\underline{F}}'(0)\underline{\underline{\rho}}\underline{\underline{J}})$$

$$+ [\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{F}}(0) - \underline{\underline{F}}(0)\underline{\underline{\rho}}\underline{\underline{J}} + 2\underline{\underline{\rho}}\underline{\underline{J}}\underline{\underline{X}}(0) + 2\underline{\underline{X}}(0)\underline{\underline{\rho}}\underline{\underline{J}} + \tfrac{1}{3}\underline{\underline{\rho}}\underline{\underline{J}}\underline{\underline{F}}'''(0)\underline{\underline{\rho}}\underline{\underline{J}}]r$$

$$+ \tfrac{1}{6}[\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{J}} + 2\underline{\underline{J}}\underline{\underline{\rho}}\underline{\underline{F}}(0)\underline{\underline{\rho}}\underline{\underline{J}}]r^3 + \underline{\underline{U}}(r) + \underline{\underline{U}}(-r) \tag{85}$$

By analogy with the single component case, we expect that $\underline{\underline{C}}(r)$ can be represented in the form

$$\underline{\underline{\xi}}(r) + \underline{\underline{\xi}}(-r) + \int_{-\infty}^{\infty} \underline{\underline{\xi}}(r')\underline{\underline{\rho}}\underline{\underline{\xi}}(r'-r)dr' = \int_{r}^{\infty} \underline{\underline{C}}(r')r'\,dr' \tag{86}$$

Instead of solving the preceding (85), we again prefer to construct $\underline{\underline{\xi}}(r)$ immediately. We expect the following analytic form

$$\xi_{ij} = [\alpha_{ij}(R_{ij}-r) + \beta_{ij}(R-R)^2]\theta[r - \tfrac{1}{2}(R_j-R_i),R_i] \qquad (87)$$

where $\underline{\underline{\alpha}}$ and $\underline{\underline{\beta}}$ are matrices to be determined. Since $\int_r^\infty r'\underline{\underline{C}}(r')dr'$ is continuous and has a continuous first derivative, the possibility of having all polynomials the same is immediately eliminated. Moreover, on the left hand side of (86) the jump in the first derivative of the ij element of the convolution term at $\frac{1}{2}(R_j-R_i)$ must be cancelled by the jump in the first derivative of $\xi_{ij}(r) + \xi_{ji}(-r)$. Thus the structure of α and β is certainly not trivially simple. We proceed by inserting (87) in (86) and imposing the symmetry condition $\underline{\underline{C}}(r) = \underline{\underline{\tilde{C}}}(r) = \underline{\underline{\tilde{C}}}(-r)$.

After discarding some redundant ones, the conditions obtained are the following:

$$\underline{\underline{\alpha}} - \underline{\underline{\tilde{\alpha}}} = 0 \qquad (88)$$

$$(\underline{\underline{\alpha}}+\underline{\underline{R}}\underline{\underline{\beta}})\underline{\underline{\rho}}\underline{\underline{\beta}} - \underline{\underline{\beta}}\underline{\underline{\rho}}(\underline{\underline{\tilde{\alpha}}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}) = 0 \qquad (89)$$

$$(\underline{\underline{R}}\underline{\underline{\alpha}}+\underline{\underline{R}}^2\underline{\underline{\beta}})\underline{\underline{\rho}}\underline{\underline{\tilde{\alpha}}} - \underline{\underline{\alpha}}\underline{\underline{\rho}}(\underline{\underline{\tilde{\alpha}}}\underline{\underline{R}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}^2) - 2\underline{\underline{\beta}} + 2\underline{\underline{\tilde{\beta}}} = 0 \qquad (90)$$

$$(\underline{\underline{R}}\underline{\underline{\alpha}}+\underline{\underline{R}}^2\underline{\underline{\beta}}) - (\underline{\underline{\tilde{\alpha}}}\underline{\underline{R}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}^2) = 0 \qquad (91)$$

$$(\underline{\alpha}+2\underline{R}\underline{\beta}) + (\underline{\tilde{\alpha}}+2\underline{\tilde{\beta}}\underline{R}) = -(\underline{R}\underline{\alpha}+\underline{R}^2\underline{\beta})\underline{\rho}(\underline{\tilde{\alpha}}\underline{R}+\underline{\tilde{\beta}}\underline{R}^2) \qquad (92)$$

$$-2(\underline{R}\underline{\alpha}+\underline{R}^2\underline{\beta})\underline{\rho}\underline{\tilde{\beta}} - 2\underline{\beta}\underline{\rho}(\underline{\tilde{\alpha}}\underline{R}+\underline{\tilde{\beta}}\underline{R}^2)$$

$$+ (\underline{\alpha}+2\underline{R}\underline{\beta})\underline{\rho}(\underline{\tilde{\alpha}}+2\underline{\tilde{\beta}}\underline{R}) - \underline{\alpha}\underline{\rho}\underline{\tilde{\alpha}} = 0 \qquad (93)$$

We note that the conditions of continuity (91) and of continuous derivative (92) are automatically included.

The simultaneous solution of these equations appears difficult. Fortunately the problem can again be simplified by using other information and by judicious guessing. The equality of all coeffic ents of r^3 in (85) tells us that $\underline{\beta}$ is of the form $\underline{\beta} = \underline{J}\underline{B}$, where B may be chosen as diagonal with no loss of generality. We expect that $\underline{\alpha}$ contains the ubiquitous $\underline{J}$, too, and guess a symmetric $\underline{\alpha}$ of the form $\underline{\alpha} = \underline{A}\underline{J}\underline{R} + \underline{R}\underline{J}\underline{A}$, where we again chose A to be diagonal. Applying (89) to this α, yields the result that $\underline{A}$ must be of the form

$$\underline{A} = z_1\underline{I} - z_2\underline{R} \qquad (94)$$

and

$$\underline{\alpha} = z_1(\underline{J}\underline{R}+\underline{R}\underline{J}) - 2z_2\underline{R}\underline{J}\underline{R}$$

where z_1 and z_2 are constants.

Instead of solving the preceding (85), we again prefer to construct $\underline{\underline{\xi}}(r)$ immediately. We expect the following analytic form

$$\xi_{ij} = [\alpha_{ij}(R_{ij}-r) + \beta_{ij}(R-R)^2]\theta[r - \tfrac{1}{2}(R_j-R_i),R_i] \qquad (87)$$

where $\underline{\underline{\alpha}}$ and $\underline{\underline{\beta}}$ are matrices to be determined. Since $\int_r^\infty r'\underline{\underline{C}}(r')dr'$ is continuous and has a continuous first derivative, the possibility of having all polynomials the same is immediately eliminated. Moreover, on the left hand side of (86) the jump in the first derivative of the ij element of the convolution term at $\frac{1}{2}(R_j-R_i)$ must be cancelled by the jump in the first derivative of $\xi_{ij}(r) + \xi_{ji}(-r)$. Thus the structure of α and β is certainly not trivially simple. We proceed by inserting (87) in (86) and imposing the symmetry condition $\underline{\underline{C}}(r) = \underline{\underline{\tilde{C}}}(r) = \underline{\underline{\tilde{C}}}(-r)$.

After discarding some redundant ones, the conditions obtained are the following:

$$\underline{\underline{\alpha}} - \underline{\underline{\tilde{\alpha}}} = 0 \qquad (88)$$

$$(\underline{\underline{\alpha}}+\underline{\underline{R}}\underline{\underline{\beta}})\underline{\underline{\rho}}\underline{\underline{\beta}} - \underline{\underline{\beta}}\underline{\underline{\rho}}(\underline{\underline{\tilde{\alpha}}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}) = 0 \qquad (89)$$

$$(\underline{\underline{R}}\underline{\underline{\alpha}}+\underline{\underline{R}}^2\underline{\underline{\beta}})\underline{\underline{\rho}}\underline{\underline{\tilde{\alpha}}} - \underline{\underline{\alpha}}\underline{\underline{\rho}}(\underline{\underline{\tilde{\alpha}}}\underline{\underline{R}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}^2) - 2\underline{\underline{\beta}} + 2\underline{\underline{\tilde{\beta}}} = 0 \qquad (90)$$

$$(\underline{\underline{R}}\underline{\underline{\alpha}}+\underline{\underline{R}}^2\underline{\underline{\beta}}) - (\underline{\underline{\tilde{\alpha}}}\underline{\underline{R}}+\underline{\underline{\tilde{\beta}}}\underline{\underline{R}}^2) = 0 \qquad (91)$$

$$(\underline{\alpha}+2\underline{R}\underline{\beta}) + (\underline{\tilde{\alpha}}+2\underline{\tilde{\beta}}\underline{R}) = -(\underline{R}\underline{\alpha}+\underline{R}^2\underline{\beta})\underline{\rho}(\underline{\tilde{\alpha}}\underline{R}+\underline{\tilde{\beta}}\underline{R}^2) \qquad (92)$$

$$- 2(\underline{R}\underline{\alpha}+\underline{R}^2\underline{\beta})\underline{\rho}\underline{\tilde{\beta}} - 2\underline{\beta}\underline{\rho}(\underline{\tilde{\alpha}}\underline{R}+\underline{\tilde{\beta}}\underline{R}^2)$$

$$+ (\underline{\alpha}+2\underline{R}\underline{\beta})\underline{\rho}(\underline{\tilde{\alpha}}+2\underline{\tilde{\beta}}\underline{R}) - \underline{\alpha}\underline{\rho}\underline{\tilde{\alpha}} = 0 \qquad (93)$$

We note that the conditions of continuity (91) and of continuous derivative (92) are automatically included.

The simultaneous solution of these equations appears difficult. Fortunately the problem can again be simplified by using other information and by judicious guessing. The equality of all coeffic ents of r^3 in (85) tells us that $\underline{\beta}$ is of the form $\underline{\beta} = \underline{J}\underline{B}$, where B may be chosen as diagonal with no loss of generality. We expect that $\underline{\alpha}$ contains the ubiquitous $\underline{J}$, too, and guess a symmetric $\underline{\alpha}$ of the form $\underline{\alpha} = \underline{A}\underline{J}\underline{R} + \underline{R}\underline{J}\underline{A}$, where we again chose A to be diagonal. Applying (89) to this α, yields the result that $\underline{A}$ must be of the form

$$\underline{A} = z_1\underline{I} - z_2\underline{R} \qquad (94)$$

and

$$\underline{\alpha} = z_1(\underline{J}\underline{R}+\underline{R}\underline{J}) - 2z_2\underline{R}\underline{J}\underline{R}$$

where z_1 and z_2 are constants.

Imposing (91) next, we obtain the result that $\underline{\underline{B}}$ must be of the form

$$\underline{\underline{B}} = -z_1\underline{\underline{I}} + 2z_2\underline{\underline{R}} + z_3\underline{\underline{R}}^2 \tag{95}$$

Proceeding to impose (92), we conclude that z_3 must be zero, since the LHS of (92) contains no term $\underline{\underline{R}}^2\underline{\underline{J}}\underline{\underline{R}}^2$. The surviving terms yield $4z_2\underline{\underline{R}}\underline{\underline{J}}\underline{\underline{R}} = -z_1^2T_2\underline{\underline{R}}\underline{\underline{J}}\underline{\underline{R}}$, or

$$4z_2 = -z_1^2T_2$$

where

$$T_n = \mathrm{Tr}(\underline{\underline{\rho}}\underline{\underline{R}}^n).$$

We may now verify by substitution that the $\underline{\underline{\alpha}}$ and $\underline{\underline{\beta}}$ just constructed,

$$\underline{\underline{\alpha}} = z_1(\underline{\underline{R}}\underline{\underline{J}}+\underline{\underline{J}}\underline{\underline{R}}) - 2z_2\underline{\underline{R}}\underline{\underline{J}}\underline{\underline{R}}$$

$$\underline{\underline{\beta}} = -z_1\underline{\underline{J}} + 2z_2\underline{\underline{J}}\underline{\underline{R}}$$

also satisfy the remaining conditions (90) and (93).

As in one dimension, the symmetry condition has reduced the problem to that of evaluating a single constant. The appropriate condition for accomplishing the latter is again the relation between $\underline{\underline{C}}(0)$ and $\underline{\underline{I}} - \underline{\underline{\rho}} \int \underline{\underline{C}}$. We note that

$$\underline{\underline{C}}_{ij}(0) = \left[\underline{\underline{J}} - \underline{\underline{J}}\underline{\underline{\rho}} \int_{-\infty}^{\infty} \underline{\underline{C}}(r) r^2 dr \right]_{ij} \qquad (i \geq j) \qquad (96)$$

and proceed to express both sides of (96) in terms of $\underline{\underline{\xi}}$. Thus we find

$$\underline{\underline{C}}_{ij}(0) = \left\{ \underline{\underline{J}} \left[\int \underline{\underline{\xi}}(r) dr + \int \underline{\underline{\xi}}(-r) dr + \int \underline{\underline{\xi}}(r) \underline{\underline{\rho}} \tilde{\underline{\underline{\xi}}}(-r) dr \right] \right\}_{ij} \qquad (i \geq j) \qquad (97)$$

and

$$\underline{\underline{C}}_{ij}(0) = \tilde{\underline{\underline{\xi}}}_{ij}(0) + \left[\int \underline{\underline{\xi}}(r) \underline{\underline{\rho}} \tilde{\underline{\underline{\xi}}}''(-r) dr \right]_{ij} \qquad (i \leq j) \qquad (98)$$

From (97) we obtain

$$\underline{\underline{C}}_{ij}(0) = \left\{ \left(1 + \frac{1}{6} z_1 T_3\right) \left[\underline{\underline{J}} + \frac{1}{2} z_1 T_2 \underline{\underline{JR}} + \frac{1}{4} z_1 \left(2T_1 + z_1 T_2^2\right) \underline{\underline{JR}}^2 + \left(\frac{1}{6} z_1 T_o - \frac{1}{6} z_1 z_2 T_2^2 - \frac{2}{3} z_2 T_1\right) \underline{\underline{JR}}^3 \right] \right\}_{ij} \quad (i \leq j) \tag{99}$$

and from (98) it follows that

$$\underline{\underline{C}}_{ij}(0) = +2z_1 \left[\underline{\underline{J}} + \frac{1}{2} z_1 T_2 \underline{\underline{RJ}} + \left(\frac{1}{2} z_1 T_1 - z_2 T_2\right) \underline{\underline{R}}^2 \underline{\underline{J}} + \left(\frac{1}{6} z_1 T_o - \frac{2}{3} z_2 T_1 - \frac{1}{6} z_1 z_2 T_2^2\right) \underline{\underline{R}}^3 \underline{\underline{J}} \right]_{ij} \quad (i \geq j) \tag{100}$$

so that the obvious solution is

$$z_1 = \frac{1}{2} \left(1 - \frac{1}{12} T_3\right)^{-1} \tag{101}$$

Thus the final forms for $\underline{\underline{\alpha}}$ and $\underline{\underline{\beta}}$ are

$$\begin{aligned} \underline{\underline{\alpha}} &= \frac{1}{2} \left(1 - \frac{1}{12} T_3\right)^{-1} (\underline{\underline{RJ}} + \underline{\underline{JR}}) + \frac{1}{8} \left(1 + \frac{1}{12} T_3\right)^{-2} T_2 \underline{\underline{RJR}} \\ \underline{\underline{\beta}} &= -\frac{1}{2} \left(1 - \frac{1}{12} T_3\right)^{-1} \underline{\underline{J}} - \frac{1}{8} \left(1 - \frac{1}{12} T_3\right)^{-2} T_2 \underline{\underline{JR}} \end{aligned} \tag{102}$$

It is immediately clear that (102) reduces to the correct form in the limiting case of a single component. By some further tedious but straightforward algebra one may further verify that (85) is satisfied. This then completes the proof that we have found the correct solution.

Explicit expressions for the pressure as obtained from the compressibility as well as the virial equation of state are contained in the paper of Lebowtiz.[15] Further results are given by Lebowitz and Rowlinson.[16] In the latter paper it is shown that the excess volume of mixture is always negative and that there is no phase separation.

REFERENCES

7. G. Stell, Physica 29, 517 (1963).

8. J. K. Percus, unpublished.

9. Percus, Lebowitz, and Zucker, Bull. Am. Phys. Soc. 7, 415-1962

10. Reiss, Frisch, and Lebowitz, J. Chem. Phys. 31, 369 (1959).

11. B. J. Alder and T. E Wainwright, J. Chem. Phys. 27, 1208 (1957).

12. W. W. Wood and J. D. Jacobson, Proceedings of the Western Joint Computer Conference, San Francisco Institute of Radio Engineers, New York (1959), p. 261.

13. J. K. Percus and G. Stell, to be published.

14. First pointed out by J. K. Percus.

15. J. Lebowitz, Phys. Rev. 133, A895 (1964).

16. J. Lebowitz and J. S. Rowlinson, J. Chem. Phys. 41, 133 (1964).

Reiss, Frisch, Lebowitz Theory for Mixtures of Hard Spheres

We shall describe here in a simple way a generalization to mixtures[1] of the method developed by Reiss, Frisch and Lebowitz[2] for a single component hard sphere fluid. This theory yields an equation of state for these systems which coincides (in one and three dimensions[3]) with that obtained from the solution of the Percus-Yevick equation.[4] Agreement between theory and machine computations on hard spheres are very good.[5]

The basic idea of the R.F.L. theory is to consider the work $w(R,\rho_1,\ldots,\rho_m,R_1,\ldots,R_m)$ of putting in a single particle of diameter R into an m-component mixture of hard spheres of diameters $R_1,\ldots,R_m$ and densities $\rho_1,\ldots,\rho_m$. R has the meaning that the 'R-particle' can come only as close as $(R+R_i)/2$ to a particle of the i^{th} species. Hence R can be less than zero (it has only to be greater than $-R_1$, where $R_1 \leq R_2 \leq \ldots R_m$). Now w is equal to the change in the configurational part of the Helmholtz free energy upon adding an R-particle to the system. This can be readily computed for $R \leq 0$, (in which case the exclusion spheres for the R-particle, centered on each fluid particle, do not overlap),

$$w(R) = -\beta^{-1}\ln\left[1 - \sum_{i=1}^{m} \frac{4}{3}\pi\left(\frac{R+R_i}{2}\right)^3 \rho_i\right], \quad R \leq 0. \qquad (1)$$

It is also easy to show[2] that w(R) and its first two derivatives are continuous at R = 0. Furthermore for R very large, $R \gg R_m \geq \ldots \geq R_1$, the leading term in w(R) must be the pressure-volume term $\frac{4}{3}\pi p\left(\frac{R}{2}\right)^3$. The R.F.L. approximation now consists in assuming that w(R) can be approximated for all values $R \geq 0$ by a cubic polynomial.

$$w(R) = w_o + w_o' R + \tfrac{1}{2} w_o'' R^2 + \frac{\pi}{6} p R^3, \qquad R \geq 0 \tag{2}$$

where w_o, w_o' and w_o'' are the values of w and its first two derivatives at R = 0. Now, when $R = R_i$, $i = 1,\ldots,m$, then the R-particle is just another particle of the i^{th} species and $w(R_i)$ is therefore equal to the configurational part of the chemical potential μ_i. We therefore have,

$$\begin{aligned}\beta\mu_i(\rho_1,\ldots,\rho_m) &= \ln[\rho_i h^3/2\pi m_i kT]^{3/2} + w(R_i) \\ &= \ln[\rho_i h^3/2\pi m_i kT]^{3/2} - \ln(1-\xi) + \frac{3X}{(1-\xi)} R_i \\ &\quad + \frac{3}{2}\left[\frac{3X^2}{(1-\xi)^2} + \frac{2Y}{(1-\xi)}\right] R_i^2 + \frac{\pi}{6}\beta p R_i^3,\end{aligned} \tag{3}$$

where

$$\xi = \frac{\pi}{6}\sum \rho_i R_i^3, \qquad X = \frac{\pi}{6}\sum \rho_i R_i^2, \qquad Y = \frac{\pi}{6}\sum \rho_i R_i.$$

Combining (3) with the thermodynamic relation

$$\frac{\partial p}{\partial \rho_j} = \sum \rho_i \frac{\partial \mu_i}{\partial \rho_j} \tag{4}$$

yields,

$$\beta p = \left\{\rho(1+\xi+\xi^2) - \frac{\pi}{2} \sum_{i<j} \rho_i \rho_j (R_i - R_j)^2 [R_i + R_j + R_i R_j X]\right\} / (1-\xi)^3 \tag{5}$$

with $\rho = \sum_{i=1}^{m} \rho_i$. Equation (5) is identical with the pressure obtained from the exact solution of the generalized P.Y. equation via the Ornstein-Zernike compressibility relation, (Eq. (46) of Ref. 4).

REFERENCES

1. J. Lebowitz, E. Praestgaard and E. Helfand, Bull. Am. Phys. Soc. 9, 278 (1964); article to be published.
2. H. Reiss, H. L. Frisch, J. L. Lebowitz, J. Chem. Phys., 31, 369 (1959). cf. also Section VII of article by J. Percus.
3. E. Helfand, H. L. Frisch and J. L. Lebowitz, J. Chem. Phys. 34, 1037 (1961).
4. J. L. Lebowitz, Phys. Rev. 133, A895 (1964), cf. also article by M. Wertheim.
5. B. Alder, J. Chem. Phys. 40, 2724 (1964).

THE
PHYSICAL REVIEW

A journal of experimental and theoretical physics established by E. L. Nichols in 1893

SECOND SERIES, VOL. 85, NO. 5 MARCH 1, 1952

Radial Distribution Function of a Gas of Hard Spheres and the Superposition Approximation*

B. R. A. NIJBOER
Institute for Theoretical Physics, University of Utrecht, Utrecht, Holland

AND

L. VAN HOVE†
University of Brussels, Brussels, Belgium

(Received November 8, 1951)

The term $g_2(r)$ proportional to the square of the density in the expansion of the radial distribution function $g(r)$ of an imperfect gas in powers of the density is calculated exactly in the case of a gas consisting of hard spheres. The result is checked by means of Boltzmann's value of the 4th virial coefficient of such a gas. The integral equation for $g(r)$, obtained on applying the superposition approximation introduced by Kirkwood and by Born and Green, can also be solved by an expansion in powers of the density. For the case of hard spheres the approximate $g_2'(r)$ found in this way is compared with the exact $g_2(r)$. As a further application of our result on $g_2(r)$ a certain integral is discussed, which is of interest in the treatment of interference effects in neutron scattering problems.

I. INTRODUCTION

As has been shown by Yvon, Kirkwood, de Boer and others,[1] the radial distribution function of a compressed gas may be expanded in powers of the density

$$g(r)=\exp\{-V(r)/kT\}\cdot\{1+\rho g_1(r)+\rho^2 g_2(r)+\cdots\}. \quad \text{(I.1)}$$

The distance between molecules has been denoted by r, the intermolecular potential by $V(r)$ and by ρ the number of molecules per unit volume. The function $g(r)$ is normalized to the value 1 for large distances. As is known, $g_1(r)$, $g_2(r)$, $\cdots$ can be expressed by cluster integrals in which the position of two particles is kept fixed. In classical statistical mechanics, and on the assumption of additivity of intermolecular forces, one has[2]

$$g_1(r_{12})=\int f(r_{13})f(r_{23})d\mathbf{r}_3, \quad \text{(I.2)}$$

$$g_2(r_{12})=\tfrac{1}{2}\{g_1(r_{12})\}^2+\varphi(r_{12})+2\psi(r_{12})+\tfrac{1}{2}\chi(r_{12}), \quad \text{(I.3)}$$

where r_{ik} is the distance $|\mathbf{r}_i-\mathbf{r}_k|$ between particles i and k, where the function $f(r)$ is related to the intermolecular potential by

$$f(r)=\exp\{-V(r)/kT\}-1 \quad \text{(I.4)}$$

and

$$\varphi(r_{12})=\int f(r_{13})f(r_{24})f(r_{34})d\mathbf{r}_3 d\mathbf{r}_4, \quad \text{(I.5)}$$

$$\psi(r_{12})=\int f(r_{13})f(r_{23})f(r_{24})f(r_{34})d\mathbf{r}_3 d\mathbf{r}_4, \quad \text{(I.6)}$$

$$\chi(r_{12})=\int f(r_{13})f(r_{23})f(r_{14})f(r_{24})f(r_{34})d\mathbf{r}_3 d\mathbf{r}_4. \quad \text{(I.7)}$$

* Part of this work was done at the Institute for Advanced Study, Princeton, New Jersey. The authors are indebted to the Institute for grants which made their stay possible, and to ONR for making available computational help.

† Associé du Fonds National de la Recherche Scientifique (Belgium).

[1] For a survey we refer to J. de Boer, "Reports on Progress in Physics," Phys. Soc. (London) 12, p. 305 (1949).

[2] J. de Boer and A. Michels, Physica 6, 97 (1939).

Kirkwood was the first to evaluate $g_1(r)$ in the simple case of hard spheres without attractive forces, for which, if the diameter of the spheres is taken as unit of length,

$$V(r)=\begin{cases}\infty & \text{for } 0\leqslant r<1\\ 0 & \text{for } r\geqslant 1\end{cases},\quad f(r)=\begin{cases}-1 & \text{for } 0\leqslant r<1\\ 0 & \text{for } r\geqslant 1\end{cases}. \tag{I.8}$$

He found[3]

$$g_1(r)=\begin{cases}\frac{2}{3}\pi(2-\frac{3}{2}r+\frac{1}{8}r^3) & \text{for } r\leqslant 2\\ 0 & \text{for } r\geqslant 2\end{cases}. \tag{I.9}$$

The function $g_1(r)$ was also calculated by numerical integration by de Boer and Michels[2] for the case of a Lennard-Jones potential.

For the simple case of hard spheres again we give in this work the calculation of the ρ^2-term $g_2(r)$, in the expansion (I.1). Our result enables us to test, at least for small densities, the validity of the so-called "superposition approximation" proposed by Kirkwood in the theory of liquids.[1] This test includes as a special case the recent work of Hart, Wallis, and Pode[4] as well as that of Rushbrooke and Scoins[5] who compared the exact value of the 4th virial coefficient in the case of hard spheres with the value obtained from the superposition assumption.

The superposition assumption, when applied as proposed by Born and Green,[6] leads for the $g(r)$-function to the approximate integral equation

$$kT\partial \ln g(r_{12})/\partial \mathbf{r}_1=-\partial V(r_{12})/\partial \mathbf{r}_1 - \rho\int[\partial V(r_{13})/\partial \mathbf{r}_1]g(r_{13})g(r_{23})d\mathbf{r}_3, \tag{I.10}$$

the solution of which from now on we will call $g'(r)$, in distinction from the exact $g(r)$. The function $g'(r)$ can also be expanded in powers of the density

$$g'(r)=\exp\{-V(r)/kT\}\cdot\{1+\rho g_1'(r)+\rho^2 g_2'(r)+\cdots\}. \tag{I.11}$$

As will be shown in Sec. IV, $g_1'(r)$ is identical with the corresponding $g_1(r)$-function of the exact expansion (I.1), whereas $g_2'(r)$ is different from $g_2(r)$. The calculation of $g_2'(r)$ and its comparison with the exact $g_2(r)$ will provide the test of the superposition approximation mentioned above. A short account of the results was published elsewhere.[7]

[3] J. G. Kirkwood, J. Chem. Phys. **3**, 300 (1935).
[4] Hart, Wallis, and Pode, J. Chem. Phys. **19**, 139 (1951).
[5] G. S. Rushbrooke and H. I. Scoins, Nature **167**, 366 (1951); Phil. Mag. **42**, 582 (1951).
[6] M. Born and H. S. Green, Proc. Roy. Soc. (London) **A188**, 10 (1946).
[7] B. R. A. Nijboer et L. Van Hove, Proc. Koninkl. Nederlandse Akad. Wetenschappen **B54**, 256 (1951). The relation for the compressibility integral [Eq. (III.3) in the present paper] was quoted incorrectly here.

As a second application of the exact expression of $g_2(r)$, we use it to determine for small densities the value of the following integral

$$K=\int_0^{\infty}\{1-g(r)\}dr,$$

which is of interest in the discussion of interference effects in neutron scattering problems and has been discussed from quite a different standpoint in the region of large densities.[8]

II. CALCULATION OF $g_2(r)$

For the calculation of φ defined by (I.5) it turns out to be convenient first to calculate

$$\frac{d\varphi(r_{12})}{dr_{12}}=\int\frac{df(r_{13})}{dr_{13}}\frac{\mathbf{r}_{12}\cdot\mathbf{r}_{13}}{r_{12}r_{13}}f(r_{24})f(r_{34})d\mathbf{r}_3d\mathbf{r}_4 =\int\frac{df(r_{13})}{dr_{13}}\frac{\mathbf{r}_{12}\cdot\mathbf{r}_{13}}{r_{12}r_{13}}g_1(r_{23})d\mathbf{r}_3.$$

In the case of hard spheres, df/dr is the δ-function $\delta(r-1)$, and, when applying (I.9) and noting that φ vanishes for $r\geqslant 3$, we find

$$\varphi(r)=\begin{cases}\pi^2[-r^6/1260+r^4/20-r^3/6-r^2/4 \\ \quad+(9/5)r-9/4+(27/70)(1/r)] & \text{for } 1\leqslant r\leqslant 3,\\ 0 & \text{for } r\geqslant 3.\end{cases} \tag{II.1}$$

The function $\psi(r)$ defined in (I.6) is calculated in the same simple way, with the result

$$\psi(r)=\begin{cases}\pi^2[r^6/1260-r^4/20+r^3/6+r^2/4-(97/60)r \\ \quad+16/9-(9/35)(1/r)] & \text{for } 1\leqslant r\leqslant 2,\\ 0 & \text{for } r\geqslant 2.\end{cases} \tag{II.2}$$

Now as to $\chi(r)$ given by (I.7), we may remark first that $\chi(r)=0$ for $r\geqslant 2$. Further, in the region $\sqrt{3}\leqslant r\leqslant 2$, the factor $f(r_{34})$ in (I.7) is -1 and hence in this region $\chi(r)$ reduces to $-\{g_1(r)\}^2$. The evaluation of $\chi(r)$ in the remaining region $1\leqslant r\leqslant\sqrt{3}$, proves to be much more complicated. We could perform it by applying the following somewhat indirect procedure, making use of the theory of Fourier transforms.

Let us introduce

$$F(h)=\int f(r)\exp(2\pi i\mathbf{h}\cdot\mathbf{r})d\mathbf{r}=-h^{-1}J_{\frac{1}{2}}(2\pi h), \tag{II.3}$$

where $J_{\frac{1}{2}}$ is the Bessel function of order $\frac{1}{2}$, and

$$G(\mathbf{h})=\int f(r_{13})f(r_{23})\exp\{2\pi i\mathbf{h}\cdot[\mathbf{r}_3-\tfrac{1}{2}(\mathbf{r}_1+\mathbf{r}_2)]\}d\mathbf{r}_3.$$

[8] Placzek, Nijboer and Van Hove, Phys. Rev. **82**, 392 (1951).

We make use of cylindrical coordinates z, ρ, φ with the axis passing through the points 1 and 2, and the origin midway between these points. We denote accordingly by h_z and h_ρ the components of $\mathbf{h}$ respectively parallel and perpendicular to the axis. The integration with respect to φ gives

$$G(\mathbf{h})=4\pi\int_0^{1-\frac{1}{2}r_{12}}\cos(2\pi h_z z)dz \times\int_0^{\{1-(\frac{1}{2}r_{12}+z)^2\}^{\frac{1}{2}}} J_0(2\pi h_\rho\rho)\rho d\rho \tag{II.4}$$

$$=2h_\rho^{-1}\int_0^{1-\frac{1}{2}r_{12}}\cos(2\pi h_z z)\{1-(\tfrac{1}{2}r_{12}+z)^2\}^{\frac{1}{2}} \times J_1\{2\pi h_\rho[1-(\tfrac{1}{2}r_{12}+z)^2]^{\frac{1}{2}}\}dz.$$

Known properties of the Fourier transformation give

$$\chi(r_{12})=\int F(h)\{G(\mathbf{h})\}^2 d\mathbf{h}. \tag{II.5}$$

Introducing (II.3) and (II.4) into (II.5) and noticing that

$$2\cos(2\pi h_z z)\cos(2\pi h_z z')=\cos\{2\pi h_z(z+z')\}+\cos\{2\pi h_z(z-z')\}$$

and that[9] (for $|z\pm z'|\leqslant 1$)

$$\int_{-\infty}^{+\infty}(h_\rho^2+h_z^2)^{-\frac{1}{2}}J_1\{2\pi(h_\rho^2+h_z^2)^{\frac{1}{2}}\}\cos\{2\pi h_z(z\pm z')\}dh_z =h_\rho^{-1}\{1-(z\pm z')^2\}^{\frac{1}{2}}J_1\{2\pi h_\rho[1-(z\pm z')^2]^{\frac{1}{2}}\},$$

we find ($r=r_{12}$)

$$\chi(r)=-4\pi\int_0^{1-\frac{1}{2}r}dz\{1-(\tfrac{1}{2}r+z)^2\}^{\frac{1}{2}} \times\int_0^{1-\frac{1}{2}r}dz'\{1-(\tfrac{1}{2}r+z')^2\}^{\frac{1}{2}} \times\int_0^{\infty}J_1\{2\pi h_\rho[1-(\tfrac{1}{2}r+z)^2]^{\frac{1}{2}}\} \times J_1\{2\pi h_\rho[1-(\tfrac{1}{2}r+z')^2]^{\frac{1}{2}}\}\cdot[\{1-(z+z')^2\}^{\frac{1}{2}} \times J_1\{2\pi h_\rho[1-(z+z')^2]^{\frac{1}{2}}\}+\{1-(z-z')^2\}^{\frac{1}{2}} \times J_1\{2\pi h_\rho[1-(z-z')^2]^{\frac{1}{2}}\}]h_\rho^{-2}dh_\rho. \tag{II.6}$$

Thus we are led to the calculation of an integral having the form

$$I=\int_0^\infty J_1(\alpha x)J_1(\beta x)J_1(\gamma x)x^{-2}dx.$$

Integrating by parts and using the relation

$$(d/dx)J_1(\alpha x)=x^{-1}J_1(\alpha x)-\alpha J_2(\alpha x),$$

we obtain

$$I=\tfrac{1}{2}\alpha\int_0^\infty J_2(\alpha x)J_1(\beta x)J_1(\gamma x)x^{-1}dx +\tfrac{1}{2}\beta\int_0^\infty J_1(\alpha x)J_2(\beta x)J_1(\gamma x)x^{-1}dx +\tfrac{1}{2}\gamma\int_0^\infty J_1(\alpha x)J_1(\beta x)J_2(\gamma x)x^{-1}dx.$$

This expression reduces to elementary functions when we apply the formula due to Sonine and Dougall (see reference 9, p. 37),

$$\int_0^\infty J_2(\alpha x)J_1(\beta x)J_1(\gamma x)x^{-1}dx=(4\pi\alpha^2)^{-1}\beta\gamma(2A-\sin 2A),$$

where

$$A=\begin{cases}0 & \text{for } \alpha^2<(\beta-\gamma)^2\\ \arccos\{(2\beta\gamma)^{-1}(\beta^2+\gamma^2-\alpha^2)\} & \text{for } (\beta-\gamma)^2<\alpha^2<(\beta+\gamma)^2\\ \pi & \text{for } (\beta+\gamma)^2<\alpha^2.\end{cases}$$

Hence according to (II.6) the calculation of $\chi(r)$ is now reduced to the evaluation of a double integral. This turns out to be a straightforward though very lengthy process. The result is finally given by

$$\chi(r)=-\{g_1(r)\}^2+\pi\{-(3/280)r^4+(41/420)r^2\}(3-r^2)^{\frac{1}{2}} +\pi\{-(23/15)r+(36/35)(1/r)\} \times\arccos\{r[3(4-r^2)]^{-\frac{1}{2}}\} +\pi\{(3/560)r^6-(1/15)r^4+\tfrac{1}{2}r^2+(2/15)r -(9/35)(1/r)\}\arccos\{(r^2+r-3) \times[3(4-r^2)]^{-\frac{1}{2}}\}+\pi\{(3/560)r^6-(1/15)r^4 +\tfrac{1}{2}r^2-(2/15)r+(9/35)(1/r)\} \times\arccos\{(-r^2+r+3)[3(4-r^2)]^{-\frac{1}{2}}\} \qquad \text{for } 1\leqslant r\leqslant\sqrt{3},$$

$$\chi(r)=-\{g_1(r)\}^2 \qquad \text{for } \sqrt{3}\leqslant r\leqslant 2,$$

$$\chi(r)=0 \qquad \text{for } r\geqslant 2. \tag{II.7}$$

It is seen that $\chi(r)$ as well as $\varphi(r)$ and $\psi(r)$ can be expressed in terms of elementary functions.

III. NUMERICAL RESULTS AND CHECKS

The expressions occurring in (II.1), (II.2), and (II.7) have been computed numerically in the relevant interval $1\leqslant r\leqslant 3$ for the values of r indicated in the first column of Table I. Then from (I.3) the function $g_2(r)$ could be computed; the result is given in the second column of Table I. $g_2(r)$ is represented graphically in Fig. 1, a plot of $g_1(r)$ may be found in de Boer's article.[1]

[9] See W. Magnus und F. Oberhettinger, *Formeln und Sätze für die speziellen Funktionen der mathematischen Physik* (Julius Springer, Berlin, 1943), p. 119.

TABLE I. Numerical values of the term proportional to ρ^3 in the expansion of the radial distribution function of a gas of hard spheres in powers of the density ρ. $g_2(r)$ represents this term when calculated exactly, $g_2'(r)$ when the superposition approximation of Kirkwood and $g_2''(r)$ when another approximation mentioned in IV is used.

r	$g_2(r)$	$g_2'(r)$	$g_2''(r)$
1.00	1.258702	0.987940	1.096623
1.08	0.823710	0.575533	0.723392
1.16	0.453089	0.234207	0.394808
1.24	0.143406	−0.042523	0.112278
1.32	−0.108740	−0.260480	−0.123544
1.40	−0.307092	−0.425349	−0.313051
1.48	−0.455149	−0.542248	−0.456999
1.56	−0.556369	−0.615922	−0.556722
1.64	−0.613934	−0.650636	−0.613956
1.72	−0.630774	−0.650097	−0.630776
1.80	−0.608749	−0.617356	−0.608749
1.88	−0.552826	−0.554682	−0.552826
1.96	−0.463399	−0.463475	−0.463399
2.04	−0.350613	−0.350613	−0.350613
2.12	−0.254150	−0.254150	−0.254150
2.20	−0.178107	−0.178107	−0.178107
2.28	−0.119833	−0.119833	−0.119833
2.36	−0.076678	−0.076678	−0.076678
2.44	−0.046047	−0.046047	−0.046047
2.52	−0.025451	−0.025451	−0.025451
2.60	−0.012563	−0.012563	−0.012563
2.68	−0.005265	−0.005265	−0.005265
2.76	−0.001704	−0.001704	−0.001704
2.84	−0.000344	−0.000344	−0.000344
2.92	−0.000022	−0.000022	−0.000022
3.00	−0.000000	−0.000000	−0.000000

The result obtained for $g_2(r)$ may be checked in two ways by means of the known value of the fourth virial coefficient B_4 for a gas of hard spheres. First, as a consequence of the virial theorem we have generally (see e.g., reference 1, p. 327)

$$B_4=-2\pi(3kT)^{-1}\int_0^\infty \exp\{-V(r)/kT\} \times[dV(r)/dr]g_2(r)r^3dr.$$

In the case of hard spheres this becomes

$$B_4=\tfrac{2}{3}\pi\int_0^\infty \delta(r-1)g_2(r)r^3dr=\tfrac{2}{3}\pi g_2(1). \quad \text{(III.1)}$$

On substituting $r=1$ into our result for $g_2(r)$ we find an expression for B_4 which is analytically identical with the one obtained by Boltzmann.[10] Its numerical value is

$$B_4=0.2869b^3, \quad \text{(III.2)}$$

where according to the customary notation $\frac{1}{4}b$ is the volume of each particle.

An additional check which includes the values of $g_2(r)$ for all r can be based upon the well-known relation, due to Ornstein and Zernike, that connects the relative compressibility with the radial distribution function (see reference 1, p. 365)

$$kT(\partial\rho/\partial p)_T=1+4\pi\rho\int_0^\infty\{g(r)-1\}r^2dr. \quad \text{(III.3)}$$

Expanding both members of this relation into powers of ρ and equating the coefficients of equal powers, we are led to the following equality holding for hard spheres

$$B_4=-\tfrac{5}{8}b^3-\pi\int_1^\infty g_2(r)r^2dr. \quad \text{(III.4)}$$

Numerical integration of the function $g_2(r)$ gave again the result (III.2) for B_4.

IV. THE SUPERPOSITION APPROXIMATION

As has been mentioned in the introduction, the superposition assumption of Kirkwood when applied in the form as discussed by Born and Green leads to the integral Eq. (I.10) for the approximate radial distribution $g'(r)$. If we now substitute

$$g'(r)=v(r)\exp\{-V(r)/kT\},$$

then $v(r)$ obeys the integral equation

$$\frac{\partial v(r_{12})}{\partial \mathbf{r}_1}=-\frac{\rho v(r_{12})}{kT}\int v(r_{13})v(r_{23})\frac{\partial V(r_{13})}{\partial \mathbf{r}_1} \times\exp\left\{-\frac{V(r_{13})}{kT}-\frac{V(r_{23})}{kT}\right\}d\mathbf{r}_3. \quad \text{(IV.1)}$$

In accordance with the expansion (I.11) we try to solve this equation by the expansion

$$v(r)=1+\rho g_1'(r)+\rho^2 g_2'(r)+\ldots.$$

Introducing this expansion into (IV.1) and equating

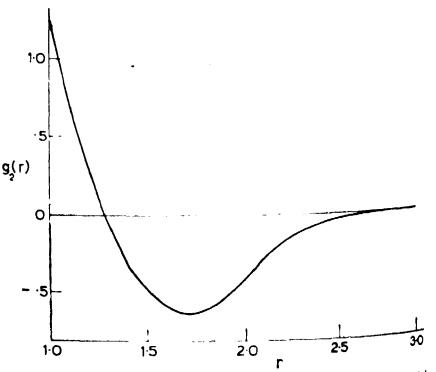


FIG. 1. The coefficient $g_2(r)$ of the term proportional to ρ^3 in the expansion of the radial distribution function for a gas of hard spheres into powers of the density ρ.

[10] L. Boltzmann, Verslag. Gewone Vergadering Afd. Natuurk. Nederlandse Akad. Wetensch. 7, 484 (1899), see also H. Happel, Ann. Physik 21, 342 (1906); R. Majumdar, Bull. Calcutta Math. Soc. 21, 107 (1929).

the coefficients of ρ on both sides we find, making use of the fact that $g_1'(r)\to 0$ as $r\to\infty$

$$g_1'(r)=\int f(r_{13})f(r_{23})d\mathbf{r}_3. \qquad \text{(IV.2)}$$

Comparison of (IV.2) with (I.2) shows that

$$g_1'(r)=g_1(r), \qquad \text{(IV.3)}$$

so that the approximate $g'(r)$ is exact to first order in ρ.

In a similar way it is found that

$$g_2'(r)=\tfrac{1}{2}\{g_1(r)\}^2+\varphi(r)+2\psi(r)+\tfrac{1}{2}\chi'(r), \qquad \text{(IV.4)}$$

so $g_2'(r)$ is identical with the exact $g_2(r)$ given by (I.3) except that $\chi(r)$ must be replaced by $\chi'(r)$ where

$$\partial\chi'(r_{12})/\partial\mathbf{r}_1=-2\int f(r_{13})f(r_{23})f(r_{34})[\partial f(r_{14})/\partial\mathbf{r}_1]d\mathbf{r}_3d\mathbf{r}_4,$$

or also

$$\frac{d\chi'(r_{12})}{dr_{12}}=-2\int f(r_{13})f(r_{23})\frac{\mathbf{r}_{12}\cdot\mathbf{r}_{13}}{r_{12}r_{13}}\frac{dg_1(r_{13})}{dr_{13}}d\mathbf{r}_3. \qquad \text{(IV.5)}$$

Thus it is seen that on the assumption of superposition $g_2(r)$ is given incorrectly. An analogous result was derived by Rushbrooke and Scoins for the 4th virial coefficient.[5] For the special case of hard spheres $\chi'(r)$ can be evaluated easily with the help of (I.9). One finds[11]

$$\chi'(r)=\begin{cases}\pi^2[-r^6/630+r^4/10-(19/72)r^3-r^2/2\\ \quad+(12/5)r-22/9+(18/35)(1/r)]\\ \qquad\qquad\qquad\qquad\text{for } 1\leqslant r\leqslant 2,\\ 0 \text{ for } r\geqslant 2.\end{cases} \qquad \text{(IV.6)}$$

Numerical values for the approximate function $g_2'(r)$ in the case of hard spheres are given in the third column of Table I. The difference $g_2(r)-g_2'(r)$ is plotted as a function of r in Fig. 2 (full curve). It is seen that for $r=1$ the Kirkwood approximation $g_2'(r)$ deviates considerably (by more than 20 percent) from the exact $g_2(r)$; for larger values of r the deviation decreases quickly and for $r\geqslant 2$ $g_2'(r)$ is equal to $g_2(r)$.

It is now very easy to calculate an approximate value for the 4th virial coefficient B_4 on the assumption of superposition. Formula (III.1), derived from the virial theorem, gives

$$B_4'=0.2252b^3, \qquad \text{(IV.7)}$$

a value that was found recently in a different way by Hart, Wallis and Poole[4] as well as by Rushbrooke and Scoins.[5] However, B_4' may also be calculated from (III.4), that is from the relation of Ornstein and Zernike. Then one obtains

$$B_4'=0.3424b^3. \qquad \text{(IV.8)}$$

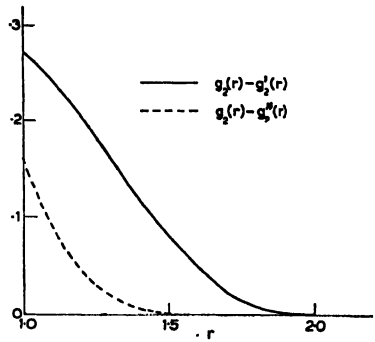


FIG. 2. Difference of the exact function $g_2(r)$ for hard spheres and the approximate functions $g_2'(r)$ (superposition approximation) and $g_2''(r)$.

We notice that the superposition approximation destroys the consistency between the equation of state as derived from $g(r)$ through the virial theorem and the equation of state resulting from the compressibility integral. This consistency requires some special analytical property of the exact $g(r)$, which is not compatible with the superposition assumption. In view of this fact one may think it questionable whether one is entitled to attach any physical meaning to analytical peculiarities exhibited by the approximate $g(r)$-function obtained under the superposition assumption. In particular conclusions drawn from Kirkwood's integral equation about a kind of condensation phenomenon for a fluid consisting of hard spheres should be accepted only with caution.[12,13]

It may be worth while to remark here that a better and simpler approximation for $g_2(r)$ is obtained in the case of hard spheres when in formula (I.7) for $\chi(r_{12})$ the factor $f(r_{34})$ in the integrand is replaced by -1, so that $\chi(r_{12})$ becomes $-\{g_1(r_{12})\}^2$. Then $g_2(r_{12})$ becomes simply [see Eq. (I.3)]

$$g_2''(r_{12})=\varphi(r_{12})+2\psi(r_{12}). \qquad \text{(IV.9)}$$

The numerical values of $g_2''(r)$ are given in the fourth column of Table I; the difference $g_2(r)-g_2''(r)$ is plotted as a dotted line in Fig. 2. The values for the 4th virial coefficient derived in this approximation from the virial theorem and from the relative compressibility integral are now respectively

$$B_4''=0.2500b^3 \quad\text{and}\quad B_4''=0.2969b^3.$$

This simple approximation is therefore much more satisfactory than that of Kirkwood, at least for the term $g_2(r)$ dealt with here. However, until now we have not

[11] This result can of course also be easily obtained directly from the integral equation (IV.1), after it has been reduced to the particularly simple form valid for hard spheres. Starting from that one could also easily calculate higher terms in the expansion of $g'(r)$ in powers of the density.

[12] Kirkwood, Maun, and Alder, J. Chem. Phys. 18, 1040 (1950).

[13] See also R. J. Riddell, thesis (Ann Arbor, 1951).

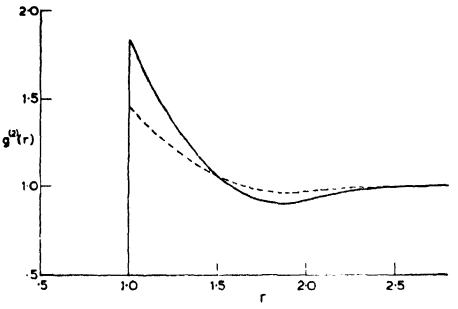


FIG. 3. Radial distribution function $g^{(2)}(r)$ (expansion into powers of the density ρ up to the term with ρ^2) for hard spheres as a function of r for values of $\rho=0.442$ (full line) and 0.275 (dotted line).

been able to extend it into a more general scheme superior to the superposition assumption.

Up to now we have discussed only the term $g_2(r)$ in the expansion (I.1) of the radial distribution function $g(r)$ into powers of the density. Evaluation of this term for hard spheres both exactly and on application of the superposition approximation made the deficiencies of the latter appear clearly. Let us consider now for a moment the radial distribution function $g(r)$ itself. Only for small densities does the part of the expansion up to the ρ^2-term, which we will call $g^{(2)}(r)$

$$g^{(2)}(r)=\exp\{-V(r)/kT\}\cdot\{1+\rho g_1(r)+\rho^2 g_2(r)\} \quad \text{(IV.10)}$$

give an adequate approximation for the function $g(r)$, because for larger densities the neglected terms would give an appreciable contribution. $g^{(2)}(r)$ has been plotted for the case of hard spheres in Fig. 3 for two values of the density: $\rho=0.442$ (full line) and $\rho=0.275$ (dotted line). It is for these densities and a few higher ones that Kirkwood, Maun, and Alder[12] recently calculated the total $g'(r)$-function for hard spheres by numerical integration of the integral Eq. (I.10). It is interesting to compare our graphs of $g^{(2)}(r)$ with the corresponding ones given by these authors for $g'(r)$; for these relatively low densities the deviations are only small. To obtain some insight in the behavior of the total $g'(r)$-function for the higher densities, one might reason as follows.[14]

From (IV.10) in combination with (I.9), (III.1), and (III.2) we have

$$g^{(2)}(1)=1+1.851\gamma+2.517\gamma^2, \quad \text{(IV.11)}$$

where $\gamma=2^{-\frac{1}{2}}\rho$ is the ratio of the density ρ to the density for close packing of the hard spheres. If in analogy with (IV.10) the function $g^{(2)\prime}(r)$ is defined by

$$g^{(2)\prime}(r)=\exp\{-V(r)/kT\}\cdot\{1+\rho g_1(r)+\rho^2 g_2'(r)\}, \quad \text{(IV.12)}$$

we find from (III.1) and (IV.7)

$$g^{(2)\prime}(1)=1+1.851\gamma+1.976\gamma^2. \quad \text{(IV.13)}$$

Table II compares, for various values of γ, the quantities $g^{(2)}(1)$ and $g^{(2)\prime}(1)$ with $g'(1)$ as obtained numerically by Kirkwood, Maun, and Alder.[12] We see that not only is

$$g'(1)<g^{(2)}(1), \quad \text{(IV.14)}$$

but also, for the higher densities

$$g'(1)<g^{(2)\prime}(1). \quad \text{(IV.15)}$$

From (IV.15) it follows that some of the virial coefficients obtained in the superposition approximation from the virial theorem are negative. It would seem extremely likely, however, that for hard spheres all virial coefficients are positive, though, as far as we are aware, this has not yet been proved in a rigorous way. Accepting this conjecture, one would conclude from (IV.14) that the pressure obtained in the superposition approximation from the virial theorem is too low, and for the higher densities, very much too low.

TABLE II. Values of the approximate radial distribution functions $g^{(2)}(r)$, $g^{(2)\prime}(r)$, and $g'(r)$ in the point $r=1$ for various densities in the case of hard spheres.

γ	$g^{(2)}(1)$	$g^{(2)\prime}(1)$	$g'(1)$
0.1942	1.455	1.434	1.45
0.3125	1.825	1.772	1.80
0.476	2.45	2.23	2.36
0.578	2.91	2.73	2.66
0.654	3.29	3.06	2.85
0.676	3.40	3.15	2.90

V. THE INTEGRAL $\int_0^\infty\{1-g(r)\}dr$

In a previous paper[8] it was shown that, when interference is taken into account, the total cross section for scattering of neutrons by a system of heavy nuclei is given asymptotically in the region of small wavelength by

$$(\sigma_{\text{tot}})_{\lambda=0}=1-\rho\lambda^2/2\pi\int_0^\infty\{1-g(r)\}dr. \quad \text{(V.1)}$$

Here the total cross section of an isolated particle is taken as the unit of cross section and the radial distribution function $g(r)$ which has the same meaning as above, is obtained from the possibly angle dependent pair distribution function $g(\mathbf{r})$ by averaging over all directions. The value of the integral was investigated extensively for dense systems of particles in reference 8.

For the special case of hard spheres without attractive forces (we will take the diameter of the spheres as unit of length as we did in the foregoing) it was shown that in the region of high density the coefficient

$$K=\int_0^\infty\{1-g(r)\}dr \quad \text{(V.2)}$$

[14] The following argument was kindly pointed out to us by Dr. G. Placzek.

can approximately be represented by

$$K \approx 0.643\gamma^{-\frac{1}{2}}. \qquad (V.3)$$

where γ again is the ratio of the density ρ to the density for close packing of the spheres. Actually (V.3) does represent an upper bound, but the deviations of the true K from (V.3) may for high densities be expected to be small. On the other hand our calculation of the term $g_2(r)$ in the expansion (I.1) now enables us to evaluate K in the region of small densities. Indeed, the introduction of (I.1) into (V.2) gives

$$K = 1 - \sqrt{2}\gamma \int_1^\infty g_1(r)dr - 2\gamma^2 \int_1^\infty g_2(r)dr - \cdots. \qquad (V.4)$$

Computation of the integrals, the first from (I.9) and the second numerically from Table I, leads to

$$K = 1 - 0.648\gamma + 0.537\gamma^2 - \cdots. \qquad (V.5)$$

It may be noted that, if we had replaced the exact $g_2(r)$ in (V.4) by $g_2'(r)$ (superposition approximation) or by $g_2''(r)$, the coefficient of γ^2 in (V.5) would have been 0.741 or 0.583, respectively.

It is rather remarkable that the integral K appears to be now pretty well known for all densities, because the two representations (V.3) and (V.5) valid for high and for small densities respectively join very nicely in the intermediate region. Both approximations namely yield for $\gamma = 0.5$ the same value $K = 0.810$ (which is of course accidental) and the slopes of the two $K(\gamma)$-curves are not too different at this junction. This fact is illustrated in Fig. 4, where the function $K(\gamma)$ is plotted using Eq. (V.5) for $0 \leqslant \gamma \leqslant 0.5$ and Eq. (V.3) for $0.5 \leqslant \gamma \leqslant 1$.

For comparison we have calculated values of K for the approximate functions $g'(r)$, mentioned already in IV, which were obtained by Kirkwood, Maun, and Alder by numerical integration of the integral equation (I.10).[12] Actually in their work to every solution $g'(r)$, characterized by the value of a certain parameter λ, belong two different values of the quantity γ, depending on whether the Born-Green or the slightly different Kirkwood version of the theory is used. By a rough numerical integration of their results for the $g'(r)$-functions we find the values of K, given in Table III.[16] In this table we have also given the respective values for $I = 4 \cdot 2^{1/6}\gamma^{1/3}K$, because it was this quantity that, at least for high densities, was discussed extensively in an earlier paper.[8] The results for K are included in Fig. 4. The circles correspond to the γ-values according to Born and Green, the crosses to those according to Kirkwood. It is seen that in particular the circles (densities according to Born and Green) appear to lie rather closely on a curve joining smoothly the low and high density parts of our $K(\gamma)$-curve. In the opinion of the authors, however, the values of K do not readily lend themselves to a discussion of the validity of the superposition approximation; without a more precise investigation of accuracy the fact, that the K-values resulting from the superposition assumption lie in the neighborhood of our approximate $K(\gamma)$-curve, would hardly seem significant. As to the question of the validity (of the Born-Green form) of the superposition assumption we must rather refer to IV. In the light of the investigation in reference 8 one could only state that the values of I (and hence also of K) resulting from the Kirkwood form of the superposition assumption are definitely too high for the higher densities.

In conclusion the authors wish to thank Dr. G. Placzek of the Institute for Advanced Study, Princeton, New Jersey, who suggested the present investigation, for his stimulating interest.

TABLE III. Values of the integrals $K = \int_0^\infty \{1 - g(r)\}dr$ and $I = 4.2^{1/6}\gamma^{1/3}K$ for various values of γ computed for the approximate radial distribution functions for hard spheres given by Kirkwood *et al.* (reference 12).

λ	K	$\gamma(K.)$	$I(K.)$	γ(B.G.)	I(B.G.)
5	0.89	0.211	2.38	0.194	2.31
10	0.84	0.353	2.67	0.313	2.56
20	0.78	0.562	2.91	0.476	2.75
33	0.72	0.775	2.98	0.654	2.81

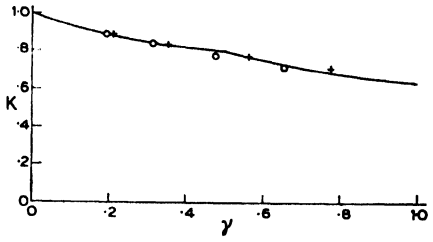


FIG. 4. The integral $K = \int_0^\infty \{1 - g(r)\}dr$ for hard spheres as a function of the ratio γ of density to density for close packing. Values obtained from integration of the numerical solutions by Kirkwood *et al.* (reference 12) of the integral equation (I.10) are indicated by circles when the Born-Green, by crosses when the Kirkwood version of the theory is adopted.

[16] The values of K for the higher densities had also been computed by Kirkwood some time ago (private communication).

Comment

The improved approximation suggested by Nijboer and Van Hove for the fourth virial coefficient, Eq. (IV.9) and subsequent discussion, is the one that comes out naturally from the Percus-Yevick equation, cf. G. Stell, J. Chem. Phys. 36, 1817 (1962).

Reprinted from THE JOURNAL OF CHEMICAL PHYSICS, Vol. 40, No. 4, 939-950, 15 February 1964
Printed in U. S. A.

Fifth and Sixth Virial Coefficients for Hard Spheres and Hard Disks*

FRANCIS H. REE AND WILLIAM G. HOOVER
Lawrence Radiation Laboratory, University of California, Livermore, California

(Received 12 August 1963)

New expressions for the fourth, fifth, and sixth virial coefficients are obtained as sums of modified star integrals. The modified stars contain both Mayer f functions and $\bar{f}$ functions ($\bar{f}=f+1$). It is shown that the number of topologically distinguishable graphs occurring in the new expressions is about half the number required in previous expressions. This reduction in the number of integrals makes numerical calculation of virial coefficients simpler and more nearly accurate. For particles interacting with a hard-core potential, values of the modified star integrals are shown to depend strongly on dimension; for example, several modified star integrals are identically zero for hard disks (two dimensions), but give nonzero values for hard spheres (three dimensions). Of all the modified star integrals contributing to the fourth, fifth, and sixth virial coefficients, the complete star integrals are shown to be the largest. Mayer's expressions for these coefficients made the complete star integrals the smallest contributing integrals.

The fifth (B_5) and sixth (B_6) virial coefficients of hard-sphere and hard-disk systems are obtained by Monte Carlo integration of the modified star integrals. The resulting values are

spheres: $B_5/b^4=0.1103\pm0.0003$; $B_6/b^5=0.0386\pm0.0004$

disks: $B_5/b^4=0.3338\pm0.0005$; $B_6/b^5=0.1992\pm0.0008$

where b is the second virial coefficient.

Estimated values of B_7 obtained from a Padé approximation to $PV^2/(N^2kT)-V/N$ are $B_7/b^6=0.0127$ for hard spheres and 0.115 for hard disks. For hard spheres virial series calculations including terms through the sixth virial coefficient give values of $PV/(NkT)$ which agree closely, for densities less than half of closest-packing, with the molecular dynamics data of Alder and Wainwright. Furthermore the approximate Padé expression agrees within 2% with the molecular dynamics data for all densities on the fluid side of the solid-fluid transition. This agreement indicates convergence of the virial series along the entire fluid branch of the hard-sphere equation of state.

I. INTRODUCTION

THE virial expansion of the pressure P of an imperfect gas is a power series expansion in the number density[1] $\rho(\equiv N/V)$,

$$PV/(NkT)=1+B_2\rho+B_3\rho^2+B_4\rho^3+\cdots, \quad (1)$$

where N is the number of particles in the volume V at a temperature T and k is Boltzmann's constant.

The nth virial coefficient B_n for a gas with the pairwise additive interaction potential ϕ_{ij} between Particles i and j can be expressed in terms of Mayer f functions[2]:

$$B_n=\frac{1-n}{n!}\lim_{V\to\infty}V^{-1}\int\cdots\int d\mathbf{r}_1\cdots d\mathbf{r}_n V_n, \quad (2)$$

$$V_n\equiv\sum_{\{S_n\}}\prod_{i>j}^{n}f_{ij}, \quad (3)$$

$$f_{ij}\equiv\exp(-\phi_{ij}/kT)-1; \quad (4)$$

where the sum in (3) includes all labeled stars with n points.

Because the number of terms in V_n, as well as the difficulty in evaluating them, grows rapidly with n, only the first few virial coefficients have been evaluated for "realistic" potentials. For the hard-sphere gas, the following exact results are known[3,4]:

spheres:

$$B_2\equiv b=(2\pi/3)\sigma^3, \qquad B_3/b^2=\tfrac{5}{8}, \qquad B_4/b^3=0.28695, \quad (5)$$

where σ is the sphere diameter. An approximate value of $B_5/b^4=0.115\pm0.005$ was obtained by the Rosenbluths,[5] who used Monte Carlo integration to evaluate the 10 types of star integrals occurring in B_5. For a two-dimensional gas composed of hard disks, the first three coefficients are known exactly[6]:

disks:

$$B_2\equiv b=(\pi/2)\sigma^2, \quad \text{and} \quad B_3/b^2=\tfrac{4}{3}-\sqrt{3}/\pi=0.78200, \quad (6)$$

where σ is the disk diameter. $B_4/b^3=0.5327\pm0.0005$

* This work was performed under the auspices of the U.S. Atomic Energy Commission.

[1] For an excellent discussion of this subject, refer to G. E. Uhlenbeck and G. W. Ford, in *Studies in Statistical Mechanics*, edited by J. de Boer and G. E. Uhlenbeck (North-Holland Publishing Company, Amsterdam, The Netherlands, 1962), Vol. 1, Part B. We shall follow the graph theoretical terminologies used by these authors.

[2] J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940).

[3] L. Boltzmann, Verslag Gewone Vergader. Afdel. Natuurk. Koninkl. Ned. Akad. Wetenschap. **7**, 484 (1899); H. Happel, Ann. Physik **21**, 342 (1906).

[4] B. R. A. Nijboer and L. Van Hove, Phys. Rev. **85**, 777 (1952).

[5] M. N. Rosenbluth and A. W. Rosenbluth, J. Chem. Phys. **33**, 1439 (1960).

[6] See M. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953). B_4 is calculated by L. Tonks, Phys. Rev. **50**, 955 (1936). We have checked B_4 for disks, using 10^7 Monte Carlo trial configurations, and find $B_4/b^3=0.5324\pm0.0003$.

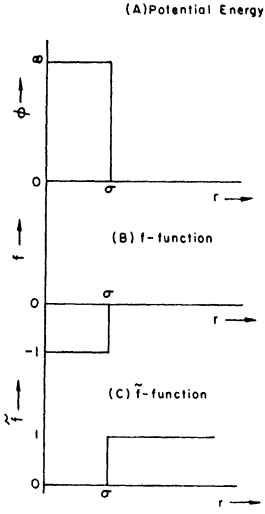


FIG. 1. (A) The hard-core potential ϕ; (B) The corresponding f function; (C) $\tilde{f}$ function.

and $B_5/b^4=0.312\pm0.016$ were obtained for disks by Metropolis *et al.*[6] using Monte Carlo integration.

As was observed by the above authors and also by Hoover and De Rocco[7] in calculating $B_{n\leq7}$ for the parallel hard-square and hard-cube models, the contribution of star integrals with positive values to $B_{n\geq4}$ is approximately equal to the contribution of star integrals with negative values. The final value for B_n is about the same order of magnitude as the contribution of a complete star integral alone; the complete star integral has the smallest absolute value of all the star integrals involved in calculating B_n. We present, in the following section, an alternative way of evaluating the virial coefficients. This new approach is particularly useful in the numerical calculation of virial coefficients beyond the third. In Secs. 3 and 4, this method is used to evaluate B_5 and B_6 for hard spheres and hard disks. It is found that the main contribution to B_5 and B_6 comes from the complete star integral, while other graphs, some positive and some negative, give smaller corrections. It is also shown that (depending on the dimensionality of the particles) some of the modified star integrals in this formalism are identically zero. In Sec. 5, we estimate B_7 for hard spheres and hard disks, and discuss the convergence of the virial series along the fluid branch of the equation of state.

2. MODIFIED STARS

We introduce the function $\tilde{f}_{ij}$ defined by the equation

$$\tilde{f}_{ij}\equiv\exp(-\phi_{ij}/kT) \tag{7}$$

and related to f_{ij} by the equation

$$1=\tilde{f}_{ij}-f_{ij}. \tag{8}$$

Whenever two points i and j in a star are not connected by f_{ij}, we use (8) to introduce $\tilde{f}_{ij}-f_{ij}$ into the star. When this procedure is carried out for all unconnected pairs of points in any particular star and the resulting expression is expanded, we find that the star can be written as a sum of modified stars composed of two kinds of lines, f_{ij} (denoted by a straight line between i and j), and $\tilde{f}_{ij}$ (denoted by a wiggly line between i and j). When this procedure is applied to all of the labeled stars occurring in a particular B_n, many of the graphs cancel out identically. Details of the expansion for the four-, five-, and six-point stars are given in Appendix I. The final expressions for V_4, V_5, and V_6 are as follows:

$V_4=3$  $-2\varnothing,$ (9)

$V_5=12$  $+10$  -60  $+45$  $-6\varnothing,$ (10)

$V_6=60$  $+180$  $+15$  -360 

-90  -720  -180  $+1080$ 

$+360$  $+240$  $+40$  $+540$

-288  $+180$  -360  -360 

-240  -900  -360  $+1440$ 

$+240$ -540  $+24\varnothing.$ (11)

For clarity, we have omitted drawing the f functions in (9), (10), and (11). This means, for example, that

the graph  denotes

 $=$  in V_4, in V_5, and in V_6. (12)

As a second example, the graph $\varnothing$ in (9), (10), and (11) denotes complete stars of four, five, and six points, respectively. For V_4, V_5, and V_6 there are, respectively, 3, 10, and 56 topologically different types of stars,

[7] W. G. Hoover and A. G. De Rocco, J. Chem. Phys. **36**, 3141 (1962).

while here we see that the corresponding modified expressions (9), (10), and (11) contain, respectively, 2, 5, and 23 topologically different types of modified stars. It is not possible at present to predict, in the general case, the type of modified stars which will occur in V_n, nor can we predict the multiplicative factors which will be associated with these modified stars. [Since this work was finished, we have obtained a new formalism, which gives all multiplicative factors corresponding to each modified star appearing in the expression of B_n. This new formalism will be reported in a later paper.] It is, however, possible to calculate the multiplicative factors for the complete star graph and the graph  , in general, by using some identities obtained by Riddell and Uhlenbeck.[8] The multiplicative factors are

$$M(\varnothing)=\sum_{k=n}^{\frac{1}{2}n(n-1)}(-)^{\frac{1}{2}n(n-1)-k}S(n,k)$$

$$=-(-)^{\frac{1}{2}n(n-1)}(n-2)!, \tag{13}$$

$$M\left(\begin{array}{c}\circ\\|\\\circ\end{array}\right)=\sum_{k=n}^{\frac{1}{2}n(n-1)-1}(-)^{\frac{1}{2}n(n-1)-k}\left[\tfrac{1}{2}n(n-1)-k\right]S(n,k)=0, \tag{14}$$

where $S(n, k)$ denotes the number of labeled stars of n points and k lines (f functions). We also notice that the sum of the positive coefficients of modified stars contributing to V_n is always one greater than the sum of the negative coefficients of such modified stars.

3. HARD SPHERES AND HARD DISKS

For particles interacting with the hard-core potential of Fig. 1(A), the corresponding f and $\bar{f}$ functions are shown in Figs. 1(B) and 1(C). In this section we consider the contribution of the modified stars to B_n in one-, two-, and three-dimensional systems. First, let us consider a system of N hard lines of length σ in a one-dimensional volume V. In 1934, Herzfeld and Mayer[9] obtained the equation of state for this system:

$$PV/(NkT)=(1-\sigma\rho)^{-1}. \tag{15}$$

We can easily evaluate the contribution of the complete star integrals in the new formalism to the B_n's for this system. From (2) and (13), these contributions are

$$B_n(\varnothing)=(-)^{\frac{1}{2}n(n-1)}(\varnothing)_n/n, \tag{16}$$

$$(\varnothing)_n=\lim_{V\to\infty}V^{-1}\int\cdots\int d\mathbf{r}_1\cdots d\mathbf{r}_n\prod_{i>j}^{n}f_{ij}. \tag{17}$$

The expression (17) can be easily evaluated[7]:

$$(\varnothing)_n=(-)^{\frac{1}{2}n(n-1)}n\sigma^{n-1}. \tag{18}$$

Therefore, we note from (16) and (18) that the B_n obtained by expanding (15) are identical to the $B_n(\varnothing)=\sigma^{n-1}$. This implies that the other modified star integrals contributing to the B_n sum to zero. In one dimension all of the four-, five-, and six-point modified star integrals contributing to B_4, B_5, and B_6 (except the complete star integrals) are zero for geometrical reasons. We conjecture that this is true for all of the higher virial coefficients as well.

We consider next a two-dimensional system of N hard disks of diameter σ. Several modified star integrals give zero contributions to the corresponding virial coefficients. We introduce a notation, $(\ )_n$, for a linear integral operator for the n particles of any graph given inside the parentheses; for example,

$$\left(\begin{array}{c}4\\|\\1\end{array}\quad\begin{array}{c}3\\|\\2\end{array}\right)_4\equiv\lim_{V\to\infty}V^{-1}\iiint d\mathbf{r}_1d\mathbf{r}_2d\mathbf{r}_3d\mathbf{r}_4\ \begin{array}{c}3\\\times\\1\end{array}\begin{array}{c}4\\\times\\2\end{array}. \tag{19}$$

The following modified star integrals represent geometrically inaccessible configurations for disks in two dimensions, and consequently vanish (see Appendix II):

$$\left(\begin{array}{c}\\\circ\!\!-\!\!\circ\end{array}\right)_{n\geq5}=0, \tag{20}$$

$$\left(\begin{array}{c}\\\times\end{array}\right)_n=\left(\begin{array}{c}\\\square\end{array}\right)_n=\left(\begin{array}{c}\\\square\end{array}\right)_n=\left(\begin{array}{c}\\\square\end{array}\right)_n$$

$$=\left(\begin{array}{c}\\\square\end{array}\right)_n=\left(\begin{array}{c}\\\square\end{array}\right)_n=0,\quad n\geq6, \tag{21}$$

$$\left(\begin{array}{c}\\\hexagon\end{array}\right)_{n\geq6}=0. \tag{22}$$

Notice that the diagrams in (21) contain at least one triangular set of $\bar{f}$ functions $\bar{f}_{ij}\bar{f}_{jk}\bar{f}_{ki}$. If, in addition to such a triangle, a modified star contains any wiggly line not linked to the triangle by fewer than two intermediate wiggly lines, the corresponding modified star integral vanishes for hard disks. Among the higher-point graphs many will contain the graphs in (20), (21), or (22) as disjoint subgraphs, and will therefore give zero integrals. The existence of integrals with zero values can be related directly to the values of Mayer's star integrals. Evidently there are linear relations among some of these integrals; for example,

[8] R. J. Riddell, Jr., and G. E. Uhlenbeck, J. Chem. Phys. **21**, 2056 (1953).
[9] K. F. Herzfeld and M. G. Mayer, J. Chem. Phys. **2**, 38 (1934).

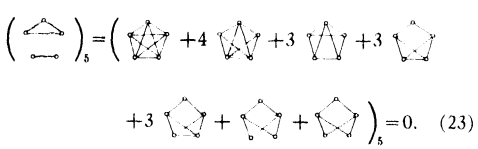
TABLE I. Values of B_5 and the five-point modified star integrals for hard spheres and hard disks.

Star	Coefficient in B_5	Unlabeling factor[a]	Values of the integrals[b]/b^4 Spheres	Disks	Contribution to B_5/b^4 Spheres	Disks
	6/30	1	$+(7.11\pm0.01)\times10^{-1}$	$+(1.809\pm0.002)$	+0.1422	+0.3618
	−45/30	6	$+(2.092\pm0.009)\times10^{-2}$	$+(1.77\pm0.01)\times10^{-2}$	−0.0314	−0.0266
	60/30	7	$-(8.25\pm0.05)\times10^{-3}$	$-(5.11\pm0.05)\times10^{-3}$	−0.0165	−0.0102
	−10/30	1	$+(7.1\pm0.4)\times10^{-4}$	0	−0.0002	0.0000
	−12/30	1	$-(4.05\pm0.03)\times10^{-2}$	$-(2.15\pm0.03)\times10^{-2}$	+0.0162	+0.0086
Values for B_5/b^4:			+0.1103±0.0003	+0.3338±0.0005	+0.1103	+0.3336

[a] The unlabeling factor is the number of ways a modified star graph can be labeled and still satisfy the Monte Carlo trial configuration conditions, $f_{12}=f_{23}=f_{34}=f_{45}=-1$.

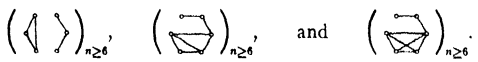
[b] B_5 and the modified star integrals for spheres (disks) are calculated from 81 (50) independent batches, each of which contains 100 000 Monte Carlo trial configurations. Each modified star integral is the average of a number (unlabeling factor) of topologically identical but differently labeled star integrals. This is equivalent to evaluating a particular labeled star integral using a number of trial configurations equal to (100 000×unlabeling factor) for each batch.

for hard disks, (20) implies the following linear relation:



(23)

The linear relations among Mayer's six-point star integrals can be obtained from Table VII, Appendix I. We can use Relation (23) to evaluate the complete star integral if the values of the other integrals in (23) are known. It must be emphasized that (20), (21) and (22) depend on the nature of the interparticle potential ϕ_{ij} (disks); other forms of ϕ_{ij} can lead to other relations.[10] In addition to the integrals in (20), (21), and (22), there are modified integrals which do not appear in the virial expansion that nonetheless vanish and lead to linear relations such as (23). For disks we cite three examples:



[10] For the hard-square model used by Hoover and De Rocco,[7] the modified star integral,



is zero, in addition to the integrals appearing in (20), (21), and (22).

Next, we consider a three-dimensional system of hard spheres with diameter σ. All contributing modified star integrals with four or five points are geometrically allowed. Of the six-point modified diagrams, one diagram contributing to B_6 is zero for geometrical reasons (see Appendix II),

(24)

For graphs with larger $n>6$, more modified stars have zero valued integrals.

4. MONTE CARLO CALCULATIONS

According to Relations (20), (21), and (22), some of the integrals required for evaluating B_5 and B_6 are zero. Therefore, we can limit our attention to the remaining integrals. For spheres, it is necessary to evaluate 5 and 22 modified star integrals for B_5 and B_6, respectively; for disks the corresponding numbers are 4 and 15. These integrals present formidable geometrical problems in 8-, 10-, 12-, and 15-dimensional spaces. We therefore evaluate them by a Monte Carlo method using an IBM 7090 computer. To make a "trial configuration" we place Particle 1 at the origin and Particles 2, $\cdots$, n randomly within a circle or sphere of diameter $(2n-2)\sigma$, with the conditions $f_{i,i+1}=-1$ for $i=1, \cdots, n-1$. Next, the remaining distances between pairs of particles are checked to see if any modified stars occurring in B_5 or B_6 correspond to this

TABLE II. Values of B_6 and the six-point modified star integrals for hard spheres and hard disks.

Star	Coefficient in B_6	Unlabeling factor[a]	Values of the integrals[b]/b^5		Contribution to B_6/b^5	
			Spheres	Disks	Spheres	Disks
⌀	−24/144	1	$-(3.53\pm0.01)\times10^{-1}$	$-(1.375\pm0.003)$	+0.0588	+0.2292
	540/144	21	$-(5.66\pm0.04)\times10^{-3}$	$-(7.28\pm0.06)\times10^{-3}$	−0.0212	−0.0273
	−240/144	5	$+(2.0\pm0.1)\times10^{-4}$	$+(2.0\pm0.7)\times10^{-5}$	−0.0003	−0.0000
	−1440/144	54	$+(1.87\pm0.01)\times10^{-3}$	$+(1.91\pm0.02)\times10^{-3}$	−0.0187	−0.0191
	360/144	19	$-(4.2\pm0.1)\times10^{-4}$	$-(3.6\pm0.1)\times10^{-4}$	−0.0011	−0.0009
	900/144	35	$-(3.47\pm0.08)\times10^{-4}$	$-(1.59\pm0.07)\times10^{-4}$	−0.0022	−0.0010
	240/144	10	$-(1.05\pm0.07)\times10^{-4}$	0	−0.0002	0.0000
	360/144	10	$-(1.07\pm0.02)\times10^{-3}$	$-(7.7\pm0.3)\times10^{-4}$	−0.0027	−0.0019
	360/144	45	$+(3.39\pm0.07)\times10^{-4}$	$+(1.59\pm0.05)\times10^{-4}$	+0.0008	+0.0004
	−180/144	5	$+(5.3\pm0.3)\times10^{-4}$	$+(4.0\pm0.3)\times10^{-4}$	−0.0007	−0.0005
	288/144	8	$+(4.93\pm0.05)\times10^{-3}$	$+(4.51\pm0.07)\times10^{-3}$	+0.0099	+0.0090
	−540/144	16	$+(6.6\pm0.5)\times10^{-5}$	0	−0.0002	0.0000
	−40/144	1	0	0	0.0000	0.0000
	−240/144	5	$-(1.03\pm0.08)\times10^{-4}$	-10^{-6}	+0.0002	+0.0000
	−360/144	24	$-(6.2\pm0.4)\times10^{-5}$	0	+0.0002	0.0000
	−1080/144	24	$-(1.76\pm0.02)\times10^{-3}$	$-(1.17\pm0.02)\times10^{-3}$	+0.0132	+0.0088
	180/144	3	$-(2.0\pm0.6)\times10^{-5}$	0	−0.0000	0.0000
	720/144	12	$+(2.46\pm0.10)\times10^{-4}$	$+(1.3\pm0.3)\times10^{-5}$	+0.0012	+0.0001
	90/144	4	$+(1.3\pm0.5)\times10^{-5}$	0	+0.0000	0.0000
	360/144	8	$+(4.83\pm0.06)\times10^{-3}$	$+(3.09\pm0.07)\times10^{-3}$	+0.0121	+0.0077
	−15/144	0	negligible	0	−0.0000	0.0000
	−180/144	2	$-(2.8\pm0.3)\times10^{-4}$	0	+0.0004	0.0000
	−60/144	1	$+(2.62\pm0.04)\times10^{-3}$	$+(1.23\pm0.04)\times10^{-3}$	−0.0109	−0.0051
Values for B_6/b^5:			+0.0386±0.0004	+0.1992±0.0008	+0.0386	+0.1994

[a] The unlabeling factor is the number of ways a modified star graph can be labeled and still satisfy the Monte Carlo trial configuration conditions, $f_{12}=f_{23}=f_{34}=f_{45}=f_{56}=-1$.

[b] B_6 and the modified star integrals for spheres (disks) are calculated from 60 (38) independent batches, each of which contains 100 000 Monte Carlo trial configurations. Each modified star integral is the average of a number (unlabeling factor) of topologically identical but differently labeled star integrals. This is equivalent to evaluating a particular labeled star integral using a number of trial configurations equal to (100 000×unlabeling factor) for each batch.

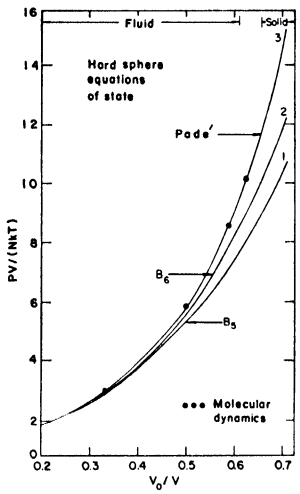
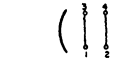


FIG. 2. Plot of $PV/(NkT)$ versus V_0/V for hard spheres. V_0 is the volume at closest-packing $N\sigma^3/\sqrt{2}$. The curves are: (1) virial series including B_5, (2) virial series including B_6, and (3) a Padé approximant. Molecular dynamics results of Alder and Wainwright (Ref. 12) are indicated by ●●●.

particlar trial configuration. For example,



can be evaluated according to the following recipe:



$$= (2b)^5(-)^{13}$$

$$\times\left(\text{number of occurrences of } \right.$$  $$\left. \Big/ \text{ number of trials}\right). \qquad (25)$$

In (25), b is B_2, $(2\pi/3)\sigma^3$ for spheres and $(\pi/2)\sigma^2$ for disks; the factor $(-)^{13}$ corresponds to the number of f functions in (25). In practice the (20) other possible labelings of  consistent with $f_{12}=f_{23}=f_{34}=f_{15}=f_{56}=-1$ are counted too, and the corresponding integrals are averaged. We call the number of ways a wiggly-line graph can be labeled subject to the restriction $f_{12}=f_{23}=f_{34}=f_{45}=f_{50}=-1$ the "unlabeling factor" for that graph. The corresponding unlabeling factors for the other modified star graphs are listed in Tables I and II. The unlabeling improves the efficiency of the Monte Carlo method. Of the modified stars contributing to B_6, only the configuration  cannot be constructed with the type of trial configuration we have just described. We attempted to evaluate this integral separately. Trial configurations were selected with the following restriction:

$$f_{13}=f_{14}=f_{15}=f_{16}=f_{23}=f_{24}=f_{25}=f_{26}=-1;$$

$$1\leq r_{12}<\sqrt{2}(\sigma\equiv 1). \qquad (26)$$

Out of 210 000 such configurations none satisfied the required configuration. From this we conclude that this modified star integral is orders of magnitude smaller than those which we evaluated. We omit this type of wiggly-line integral from the B_6 calculation with an error negligible relative to the uncertainty in our final result.

Tables I and II give the results and expected errors of the Monte Carlo calculations. The error for any Monte Carlo integral I is estimated by the following formula[11]:

$$\text{error}=\left[\sum_{i=1}^{q}(\langle I_i\rangle-\langle I\rangle)^2/(q(q-1))\right]^{\frac{1}{2}}, \qquad (27)$$

where $\langle I\rangle$ is the final Monte Carlo average of I as obtained from q independent Monte Carlo averages $\langle I_i\rangle$ $(i=1,\cdots,q)$ over batches of trial configurations. The number of independent batches, q, and the number of random Monte Carlo trial configurations in each batch are given in Tables I and II. The estimated errors are essentially independent of q for the same total number of Monte Carlo trial configurations. Each of these Monte Carlo trial configurations satisfies the restrictions of at most one of the modified stars contributing to the virial coefficient expressions (10) and (11).

5. DISCUSSION

Several aspects of the modified star integrals given in Tables I and II are notable. First, the complete star integrals have the largest absolute values of all the star integrals shown. The complete star integrals always make positive contributions to the virial coefficients. Second, the next seven modified star integrals contributing to B_6 (Table II) give nonpositive corrections. The net negative correction made by these terms

[11] The mean-square expected deviation of a quantity from the exact value I is $\langle(\langle I\rangle-I)^2\rangle$, where $\langle\ \rangle$ denotes the expectation operator, and $\langle I\rangle$ is the final average value over q independent values, $\langle I_i\rangle$, in the present problem. However, this deviation is equal to σ^2/q, where σ is the variance of I. For a finite number of batches, σ can be approximated by $\bar{\sigma}$ defined in a finite number of batches, i.e., $\bar{\sigma}^2=[(q-1)/q]\sigma^2$ [see P. G. Hoel, *Introduction to Mathematical Statistics* (John Wiley & Sons, Inc., New York, 1954), 2nd ed., p. 198].

contributes significantly to B_6, and may become large enough to give negative B_n for larger values of n, because the factors multiplying these graphs increase rapidly with n. Third, the next largest integrals among the modified star integrals are the ring graphs, which are formed by n f functions and $\frac{1}{2}n(n-3)f$ functions. These integrals give positive (negative) contributions to B_n if n is odd (even). Fourth, we see that the three-dimensional modified star integrals which are zero in two dimensions have much smaller values than the other three-dimensional modified star integrals. If only the complete star graphs are used to calculate B_5 and B_6 for disks ("one-dimensional approximation"), the values of B_5/b^4 and B_6/b^5 are, respectively, 8.4% and 15% larger than the values given in Tables I and II. If we include for spheres only those integrals which are not zero for disks ("two-dimensional approximation") B_5/b^4 and B_6/b^5 are, respectively, 0.21% larger and 0.17% smaller than our calculated values.

If the virial series converges to the true pressure in the density range of the first-order fluid-solid phase transition,[12,13] some of the higher virial coefficients must necessarily be negative in order to describe a flat or looped isotherm in the P–V diagram. It is interesting therefore to know the density range within which the five- or six-term virial series is a good approximation to the complete infinite series. In Figs. 2 and 3,

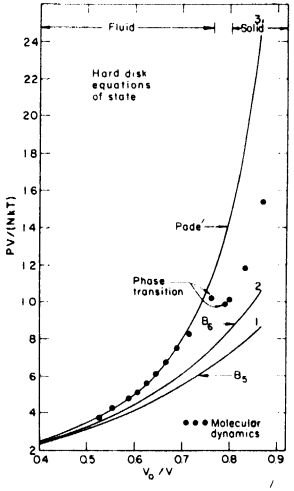


FIG. 3. Plot of $PV/(NkT)$ versus V_0/V for hard disks. V_0 is the volume at closest-packing, $N\sigma^2\sqrt{3}/2$. The curves are: (1) virial series including B_5, (2) virial series including B_6, and (3) a Padé approximant. Molecular dynamics results of Alder and Wainwright (Ref. 13) are indicated by ●●●.

[12] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **33**, 1439 (1960).

[13] B. J. Alder and T. E. Wainwright, Phys. Rev. **127**, 359 (1962). The additional data given in Fig. 3 and Table IV are kindly supplied to us by these authors.

TABLE III. Values of $PV(NkT)$ obtained by using the five- and the six-term virial series, a Padé approximant, and the molecular dynamics results (MD) on the fluid branch of the equation of state for hard spheres ($V_0=N\sigma^3/\sqrt{2}$).

V/V_0	Five-term series	Six-term series	Padé	MD[a]
1.60	8.11	8.95	10.11	10.17[b]
1.70	7.17	7.79	8.55	8.59
2.00	5.31	5.59	5.83	5.89
3.00	2.98	3.01	3.03	3.05
10.00	1.36	1.36	1.36	1.36

[a] These data for systems of 108 spheres were kindly furnished by B. J. Alder and T. E. Wainwright.

[b] At this density, both solid and fluid phases can occur for a system with a finite number of particles. The phase transition from the fluid to the solid for a system with an infinite number of particles is estimated to start at $V/V_0\approx 1.63$.

these series are plotted together with the molecular dynamics data of Alder and Wainwright.[12,13] Figures 2 and 3 also show the plots obtained by Padé approximants (see Appendix III), $P(3,3)$, to $PV^2/(N^2kT)-V/N$, using the known values of B_2 through B_6:

$$\text{spheres: } \frac{V}{N}\left[\frac{PV}{NkT}-1\right]\doteq P(3,3)$$

$$\equiv\frac{b(1+0.063507b\rho+0.017329b^2\rho^2)}{(1-0.561493b\rho+0.081313b^2\rho^2)}; \quad (28)$$

$$\text{disks: } \frac{V}{N}\left[\frac{PV}{NkT}-1\right]\doteq P(3,3)$$

$$\equiv\frac{b(1-0.196703b\rho+0.006519b^2\rho^2)}{(1-0.978703b\rho+0.239465b^2\rho^2)}. \quad (29)$$

In the case of disks, the six-term virial series pressure is considerably below (25%) the molecular dynamics pressure on the fluid side of the two-phase region ($V=1.312V_0$). We therefore expect the next several coefficients for disks to be positive. It is interesting, but probably not significant, to note that the sphere B_n from (28) are given by

$$\text{spheres: } B_{n+3\geq 3}/b^{n+2}=0.28515^n[0.62500\cos(0.17606n)+2.2603\sin(0.17606n)]. \quad (30)$$

From (30) we see that B_{20} has the first negative sign, and the sign of the sphere B_n changes roughly every 16 terms, while the disk B_n from (29) are all positive,

$$\text{disks: } B_{n+3\geq 3}/b^{n+2}=0.48935^n[0.78200+0.30597^n]. \quad (31)$$

Strangely enough, the denominator of (29) is (to six significant figures), a perfect square $(1-0.489351b\rho)^2$. We note that values of the six-term virial series agree within 1% with the molecular dynamics results for hard spheres and disks at volumes greater than twice

TABLE IV. Values of $PV/(NkT)$ obtained by using the five- and six-term virial series, a Padé approximant, and the molecular dynamics results (MD) on the fluid branch (Ref. 13) of hard disks. ($V_0 = N\sigma^2\sqrt{3}/2$).

V/V_0	Five-term series	Six-term series	Padé	MD
1.312[a]	6.50	7.51	11.11	10.13
1.40	5.71	6.44	8.45	8.25
1.45	5.33	5.95	7.45	7.47
1.50	5.01	5.52	6.66	6.67
1.55	4.72	5.16	6.03	6.08
1.60	4.47	4.84	5.52	5.56
1.65	4.24	4.56	5.10	5.13
1.70	4.04	4.31	4.74	4.76
1.80	3.69	3.90	4.18	4.24
1.90	3.41	3.57	3.76	3.78
2.00	3.17	3.30	3.43	3.39

[a] The phase transition from the fluid to the solid for a system with an infinite number of particles is estimated to start at $V/V_0 = 1.312$.

closest-packed. The values of the Padé approximants agree even better with the dynamics data. The agreement with the fluid branch is good even at phase-transition densities (Tables III and IV). However, the Padé approximants do not show any maxima or minima on the P–V diagram.

The Padé approximant method has proved to be very accurate in estimating critical parameters for Ising lattice problems.[14–17] If the Padé approximants (28) and (29) are used to estimate B_7 and B_8, the following values are obtained[18]:

$$\text{spheres: } B_7/b^6 = 0.0127, \quad \text{and} \quad B_8/b^7 = 0.0040, \tag{32}$$

$$\text{disks: } B_7/b^6 = 0.115, \quad \text{and} \quad B_8/b^7 = 0.065. \tag{33}$$

By subtracting the six-term virial series from the molecular dynamics $PV/(NkT)$ (Tables III and IV) and assuming the remainder can be represented by the single term $B_7\rho^6$, we find:

$$\text{Spheres: } B_7/b^6 \approx 0.03, \tag{34}$$

$$\text{Disks: } B_7/b^6 \approx 0.3. \tag{35}$$

The exact evaluation of B_7 by integrating the modified star integrals occurring in V_7 is now in progress for both spheres and disks.

[14] G. A. Baker, Jr., Phys. Rev. **124**, 768 (1961).
[15] J. W. Essam and M. E. Fisher, J. Chem. Phys. **38**, 802 (1963).
[16] P. Heller and G. B. Benedek, Phys. Rev. Letters **8**, 428 (1962).
[17] C. Domb and M. F. Sykes, J. Math. Phys. **2**, 63 (1961); **3**, 586 (1962).
[18] The Padé approximant, $P(1, 2)$, gives $B_6/b^6 = 0.0364$ for spheres and $B_6/b^6 = 0.1974$ for disks. These values agree well with the exact values obtained in this paper: 0.0386 and 0.1992, respectively.

6. ACKNOWLEDGMENTS

We wish to thank Dr. Donald H. Davis for illuminating discussions on the Monte Carlo calculations and Mr. Warren G. Cunningham for programming the Monte Carlo problem for the IBM 7090. We would like to acknowledge interesting discussions with Dr. S. Katsura, Dr. J. E. Kilpatrick, and Dr. J. S. Rowlinson on the general problem of hard-sphere virial coefficient calculations. Katsura's calculation of the hard-sphere B_5 has been published in *The Journal of Chemical Physics*.

APPENDIX I

This appendix gives the detailed transformation of the star integrals into modified star integrals. To the right of each labeled star type are listed the modified stars arising from it. The number of ways each type of star can be labeled is taken into account so that the entries in Tables V–VII give the total number of times each modified star integral appears when *all* of the Mayer labeled stars are expanded and the contributions added together.

TABLE V. Transformation of the four-point labeled stars to modified stars.

Labelings	Star			∅
3	□	3	−6	3
6	⊠		6	−6
1	⊠			1
Totals		3	0	−2

APPENDIX II

In this appendix we prove that three distinct modified star integrals are zero.

Proof 1:



For convenience we take the disk diameter to be unity. In Fig. 4 circles with unit radius, centered on Particles 1 and 2 are drawn. Particle 3 must lie outside these circles in order to satisfy $f_{13}f_{23} \neq 0$. Particles 4 and 5 ($f_{45} \neq 0 \rightarrow r_{45} > 1$) must satisfy the conditions

$$r_{14}, \quad r_{15}, \quad r_{24}, \quad \text{and} \quad r_{25} \leq 1, \tag{36}$$

$$r_{34} \quad \text{and} \quad r_{35} \leq 1. \tag{37}$$

TABLE VI. Transformation of the five-point labeled stars to modified stars.

Labelings	Star										Ø
12		12	−60			60	60	−60	−60	60	−12
60			60			−120	−120	180	180	−240	60
10				10	−10		−30	30	30	−40	10
10					10				−30	30	−10
60						60		−60	−120	180	−60
30							30	−60	−30	90	−30
15								15		−30	15
30									30	−60	30
10										10	−10
1											1
Totals		12	0	10	0	0	−60	45	0	0	−6

Equation (36) restricts 4 and 5 to be within the area common to the two unit circles, and the condition $r_{45}>1$ implies $r_{12}<\sqrt{3}$. The optimum position for Particle 3 which will still satisfy (37) is denoted by A (or A′) in Fig. 4. Particles 4 and 5 must then be located in the shaded area ABCA to satisfy (36) and (37). However, this optimum configuration cannot satisfy $r_{45}>1$. Consequently, the corresponding integral is zero. This proof is equally valid for any $n\geq 5$. Direct applications of the proof lead easily to other identities such as (21) involving the present integral.

Proof 2:

$$\left(\quad\right)_{n\geq 6}=0.$$

Referring to Fig. 4, we have the condition $1\leq r_{12}<\sqrt{3}$ because $r_{34}(>1)$ must lie within the area A′BACA′ common to the two unit circles. The restrictions r_{15}, r_{16}, r_{25}, and $r_{26}\geq 1$ place Particles 5 and 6 outside both unit circles in Fig. 4. Evidently 5 and 6 must both lie below or above the unit circles to satisfy $r_{56}<1$. We place them below, in the vicinity of A. But, since both r_{36} and r_{45} must be less than unity, Particles 3 and 4 must lie inside the shaded area ABCA. This violates the condition $r_{34}>1$. Therefore, the integral corresponding to the above configuration is zero. The proof is equally valid for any $n>6$.

Proof 3:

$$\left(\quad\right)_{n\geq 6}=0.$$

It can be shown that the area of any cross section within the volume common to three unit spheres centered at the points 1, 2, and 3 (and satisfying r_{12}, r_{13}, and $r_{23}>1$) is contained in an area common to two unit circles whose centers are separated by at least unit distance. However, we cannot place the triangle 456 (with r_{45}, r_{46}, and $r_{56}>1$) inside such an area. Therefore, the corresponding integral is zero. This proof is equally valid for any $n>6$.

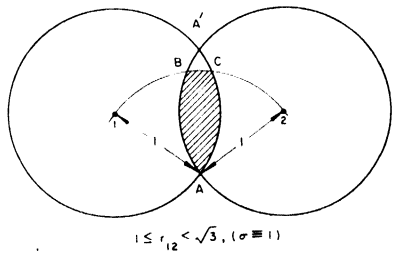


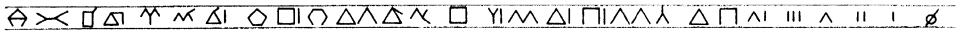
FIG. 4. Disk geometry. Unit circles are centered on Particles 1 and 2 with $1\leq r_{12}<\sqrt{3}$.

TABLE VII. Transformation of the six-point labeled stars to

Labeling																																										
60	1	60	-360	-180				720	360	180		360	180	360				-120	-360	-360	-360	-360	-720			-720	-1080	-720	-180	-60		
360	2		360					-720	-720	-360		-720		-360				360	720	720	720	360	1440			2160	2520	720	360			
180	3			180				-720					-360	-360					360	360	360	720	720				720	1440	180	180		
180	4				180	-180					-360	-180		-720		360					180		1440	180	360	1080	1080	360		-180	
180	5					180										-360					-180		-720							180	
15	6						15								-15									-90							90
720	7							720											-720	-720	-720	-720	-720			-720	-720					
360	8								360								-360			-360			-720		-720	-360					
180	9									180								-360							-720			-180			
360	10										360					-360							-720	-360	-360		-720			360	
180	11											180									-180				-360	-720					
90	12												90									-180					-360		-90		
360	13													360									-720			-720	-720	-360			
15	14														15															-90	
360	15															360														-360	
120	16																120														
360	17																	360													
360	18																		360												
180	19																				180										
180	20																					180									
720	21																						720								
90	22																							90							-90
360	23																								360						
360	24																									360					
360	25																										360				
60	26																											60			
10	27																												10		
90	28																													90	
360	29																														
90	30																														
360	31																														
360	32																														
360	33																														
360	34																														
180	35																														
72	36																														
45	37																														
360	38																														
60	39																														
180	40																														
360	41																														
45	42																														
60	43																														
360	44																														
60	45																														
180	46																														
90	47																														
60	48																														
20	49																														
180	50																														
180	51																														
15	52																														
60	53																														
45	54																														
15	55																														
1	56																														
Totals		60	0	0	180	0	15	0	0	0	0	-360	-90	-720	0	0	0	0	0	0	0	0	-180	0	1080	360	240	40	0		

[a] The above 56 types of six-point stars are listed in the same order as in Appendix I of Ref. 7.

modified stars. The lines in the graphs represent wiggly lines.[a]



360	180	720	720	1080	1440	360	360	180	1800	360	-360	-1440	-180	-360	-2520	-360	-1440	-900	360	120	1800	2520	240	-1080	-1080	540	-60
-1080	-360	-2160	-1440	-3600	-3600	-720	-1080	-720	-5040	-360	1080	5040	720	1080	8640	720	5400	2520	-1440	-360	-7560	-9720	-1080	5040	5040	-2880	360
-360	-360	-720	-1440	-720	-2160	-720	-360		-2160	-1080	720	2160	180	720	3960	1080	1800	1980	-720	-360	-3240	-5400	-360	2520	2520	-1440	180
-720		-360	-1080	-1080	-1800	-1080	-360	-180	-2880	-360	1080	1800	180	720	3960	720	2700	1440	-720	-360	-3420	-5040	-540	2520	2520	-1440	180
720		360	720	360	720	360			360		-1080	-1440	-180	-360	-1800	-360	-720	-360	720	360	2520	2520	180	-2160	-1620	1260	-180
						180		45			-180		-45	-60		-60	-180		60	60	180	180	45	-180	-135	105	-15
720	720	2160	2160	1440	3600	720	720		2160	720	-1440	-5760	-720	-1440	-8640	-1440	-2880	-2880	2160	720	9360	12240	720	-7920	-7200	5040	-720
1080		720	720	1800	720	360	360	360	1440		-1080	-2880	-360	-360	-3960	-360	-2880	-720	1080	360	5040	5400	720	-3960	-3600	2520	-360
	180	720		1080				360	1440			-1440	-360		-1800		-2160	-540	360		2520	2880	540	-1800	-1980	1260	-180
720		360	1080	720	720	1440		360	1440	360	-1800	-1800	-360	-720	-2880	-1080	-2880	-1080	1080	720	4320	5760	720	-3960	-3600	2520	-360
		360		720	1080		360	180	1080			-1080	-180	-360	-2520		-1440	-720	360		2340	3240	360	-1800	-1980	1260	-180
	90		360		360	180			360	540	-180	-360		-180	-720	-540	-360	-810	180	180	720	1980	90	-900	-990	630	-90
360			720	720	1440	720	360		2880	360	-720	-1440		-720	-4320	-720	-2880	-1800	720	360	3960	6840	720	-3600	-3960	2520	-360
											180		45						-60	-60	-180			180	45	-90	15
-720		-360	-720								1800	1440	360		1440	360			-1080	-720	-3600	-1800		3600	1800	-2160	360
-360				-360							360	720			360		360		-360	-120	-1080	-720	-120	1080	720	-720	120
	-360	-720		-720					-360			2160	360		1440		1080	360	-720		-3240	-2880	-360	2880	2520	-2160	360
-360		-720	-720				-360				720	1440	360		2520	360			-720	-360	-3600	-2520		3240	2160	-2160	360
		-360			-720							720	180	360	1080			360	-360		-1440	-1800		1440	1260	-1080	180
	-180		-360		-360					-180	180	720		180	720	360		540	-360	-180	-1080	-1980		1440	1260	-1080	180
-720			-720	-720	-720	-720			-720		1440	2160		720	2880	720	2160	720	-1440	-720	-5040	-6480	-720	5760	5040	-4320	720
						-360		-90			360		90	180		180	540		-180	-180	-540	-720	-180	720	630	-540	90
		-360		-720				-360	-720			1080	360		1440		2160	360	-360		-2880	-3240	-720	2520	2880	-2160	360
				-360	-720		-360		-720			720		360	2520		1080	720	-360		-2520	-3960	-360	2520	2880	-2160	360
			-360		-360	-360			-720	-360	360	360		360	1440	720	1080	1080	-360	-360	-1800	-4320	-360	2520	2880	-2160	360
									-360						360		360	180			-360	-720	-120	360	540	-360	180
										-60						60		90		-20		-180		60	90	-60	30
											-360		-90						180	180	540			-720	-180	450	-90
360											-720	-720			-360				720	360	1800	720		-2520	-1080	1800	-360
	90											-360						-90	180		360	360		-540	-360	450	-90
		360										-720	-360		-720				360		2160	1080		-2160	-1440	1800	-360
			360								-360	-360			-720	-360			360	360	1440	1440		-2160	-1440	1800	-360
				360								-720			-360		-720		360		1440	1440	360	-1800	-1800	1800	-360
					360							-360			-720			-360	360		1080	2160		-1800	-1800	1800	-360
						180					-180			-180		-180	-360		180	180	360	900	180	-900	-900	900	-180
							72								-360						360	360		-360	-360	360	-72
								45					-45				-180				180	180	90	-180	-270	225	-45
									360						-720		-720	-360			1080	2160	360	-1440	-2160	1800	-360
										60						-120		-180		60		540		-240	-360	300	-60
											180								-180	-180	-360			900	180	-720	180
												360							-360		-720	-360		1440	720	-1440	360
													45								-180			180	90	-180	45
														60					-60			-180		180	180	-240	60
															360						-720	-720		1080	1080	-1440	360
																60				-60		-180		180	180	-240	60
																	180				-180	-360	-180	360	720	-720	180
																		90				-360		180	360	-360	90
																			60					-180		180	-60
																				20				-60		60	-20
																					180			-360	-180	540	-180
																						180		-180	-360	540	-180
																							15		-45	45	-15
																								60		-120	60
																									45	-90	45
																										15	-15
																											1
0	0	0	0	0	0	540	-288	180	-360	0	0	0	0	-360	0	-240	-900	-360	0	0	0	1440	240	0	-540	0	24

APPENDIX III

This appendix indicates the general method for obtaining Padé approximants to the virial series.

We define the Padé approximant to $PV^2/(N^2kT)-V/N$ by the relation

$$P(n,m)=\sum_{i=1}^{n}a_i\rho^{i-1}/\sum_{i=1}^{m}\alpha_i\rho^{i-1}, \quad a_1\equiv B_2, \quad \alpha_1\equiv 1, \tag{38}$$

where n and m are positive integers and the expansion of (38) reproduces the virial coefficients up to B_{n+m}. The virial expansion of $PV^2/(N^2kT)-V/N$ is given by the following equation,

$$[PV/(NkT)-1](V/N)=\sum_{i=1}^{\infty}B_{i+1}\rho^{i-1}. \tag{39}$$

The coefficients a_i and α_i can be evaluated from (38) and (39), which can be written in the equivalent form:

$$\begin{array}{c} \\ \uparrow \\ n-1 \\ \downarrow \\ \uparrow \\ m-1 \\ \downarrow \end{array}
\overbrace{\qquad\qquad}^{n-1}\ \overbrace{\qquad\qquad\qquad}^{m-1}
\begin{pmatrix}
1 & 0 & \cdots & 0 & -B_2 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & -B_3 & -B_2 & \cdots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & 1 & -B_n & -B_{n-1} & \cdots & \cdots \\
0 & \cdots & \cdots & 0 & -B_{n+1} & -B_n & \cdots & \cdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
0 & \cdots & \cdots & 0 & -B_{n+m-1} & -B_{n+m-2} & \cdots & -B_n
\end{pmatrix}
\begin{pmatrix} a_2 \\ a_3 \\ \cdots \\ a_n \\ \alpha_2 \\ \cdots \\ \alpha_m \end{pmatrix}
=
\begin{pmatrix} B_3 \\ B_4 \\ \cdots \\ B_{n+1} \\ B_{n+2} \\ \cdots \\ B_{n+m} \end{pmatrix} \tag{40}$$

The recursion relation

$$B_k=-\sum_{i=2}^{m}B_{k+1-i}\alpha_i, \qquad k\geq n+2 \tag{41}$$

can be used to estimate higher virial coefficients.

Approximate Integral Equations for g(r) and the Equation of State

It is apparent from this section that one of the principal ways of obtaining information about the structure of fluids is via the method of integral equations for the radial distribution function g(r). After solving such an integral equation, exactly or approximately, one is still left with the problem of obtaining the thermodynamic properties of the fluid from g(r). The simplest of these is the equation of state which may be obtained directly from g(r) either via the virial theorem,

$$\beta p = \rho - \frac{2\pi}{3}\rho^2 \int_0^\omega \varphi'(r) g(r) r^3 dr \qquad (1)$$

or via the compressibility relation,

$$\frac{d(\beta p)}{d\rho} = \left\{1 + \rho \int [g(r) - 1] dr\right\}^{-1} = 1 - \rho \int C(\underline{r}_{12}) d\underline{r}_2 \qquad (2)$$

where C(r) is the direct correlation function. For the exact g(r) Eqs. (1) and (2) yield the same pressure,[1] but for an approximate g(r) the two pressures, which may be denoted by p^v and p^c will be different. Thus, for example, the solution of the Percus-Yevick equation for hard spheres of diameter a, yields[2]

$$\beta p^v = \beta p^c - \frac{18}{\pi} \eta^2/(1-\eta)^3 \tag{3}$$

$$\beta p^c = \rho[1+\eta+\eta^2]/(1-\eta)^3. \tag{4}$$

where

$$\eta = \frac{\pi}{6} \rho a^3.$$

The difference between p^v and p^c for an approximate $g(r)$ may be used as a check on the reliability of the approximation. The question then arises of whether it is possible to find an approximate integral equation for $g(r)$ which will make p^c equal to p^v. We will show here that if $g(r)$, and hence $C(r)$, are expressible as a sum of graphs with ρ-vertices and f-bonds, which is the case in the Percus-Yevick and Hypernetted Chain approximation, then this is possible only for the exact $g(r)$.

Starting with βp^c as a function of ρ and a functional of the Mayer f-function, $\beta p^c(\rho,\{f\})$ we determine a Helmholtz free energy per particle $\beta F^c(\rho,\{f\})$,

$$\beta p^c(\rho,\{f\}) = \rho^2 \frac{\partial \beta F^c(\rho,\{f\})}{\partial \rho}. \tag{5}$$

We then construct a new $\hat{g}(r,\rho:\{f\})$ by the relation

$$\rho\hat{g}(r_{12},\rho,\{f\}) = 2[1 + f(r_{12})]\,\frac{\delta[\beta F^c(\rho,\{f\})]}{\delta f(r_{12})}. \qquad (6)$$

The virial pressure obtained from $\hat{g}(r)$ via Eq. (1) $\hat{p}^V$ is then guaranteed, by a straightforward generalization of the proof given by Uhlenbeck and Ford,[1] to be the same as p^c. Now the graphs which enter into $\hat{g}(r_{12})$ are obtained, according to (6), by deleting any f-bond in βF^c and replacing it by $1 + f$. The graphs which occur in βF^c are obtained according to (2) and (5), by replacing the two labeled points (white circles) $\underline{r}_1$ and $\underline{r}_2$ in $C(\underline{r}_{12})$ by unlabeled points (black circles), (c.f. article by G. Stell), and dividing the result by ρV, where V is the volume of a cubical box in which we imagine our system enclosed and is eventually let to approach infinity. The graphs which enter $g(r_{12})$ are obtained on the other hand, by a convolution of all the graphs which appear in $C(r_{12})$,

$$g(r_{12}) - 1 = C(r_{12}) + \int C(\underline{r}_{13})\rho C(\underline{r}_{32})d\underline{r}_3 + \int C(\underline{r}_{13})\rho C(\underline{r}_{34})\rho C(\underline{r}_{42})d\underline{r}_3 d\underline{r}_4 + \cdots \qquad (7)$$

A little thought now shows that the only way in which the graphs appearing in $\hat{g}(r_{12})$ can be identical with the graphs appearing in $g(r_{12})$ is for $C(r_{12})$ to consist of all doubly connected graphs with two root points, i.e. for C and g to be exact.

A different proof of this result not utilizing graphs directly has been given by J. Percus (unpublished).

(1) G. E. Uhlenbeck and G. W. Ford in "Studies in Statistical Mechanics" vol. 1, edited by J. De Boer and G. E. Uhlenbeck, North Holland, Amsterdam, 1961.

(2) M. Wertheim, Phys. Rev. Lett. 8, 321 (1963): E. Thiele, J. Chem. Phys. 38, 1959 (1963).

SECTION III

Accidental Deviations of Density and Opalescence at the Critical Point of a Single Substance

L. S. Ornstein
and
F. Zernike

Reprinted from
PROC. AKAD. SCI. (Amsterdam)
Volume 17, pp. 793–806 (1914)

Physics. — "*Accidental deviations of density and opalescence at the critical point of a single substance.*" By Dr. L. S. ORNSTEIN and F. ZERNIKE. (Communicated by Prof. H. A. LORENTZ.)

(Communicated in the meeting of September 26, 1914).

1. The accidental deviations for a single substance as well as for mixtures have been treated by SMOLUCHOWSKI [1]) and EINSTEIN [2]) with the aid of BOLTZMANN'S principle; by ORNSTEIN [3]) with the aid of statistical mechanics. It appears as if the considerations used and the results obtained remain valid in the critical point. SMOLUCHOWSKI has applied the formula found for the probability of a deviation to the critical point itself, and has found for the average deviation of density

$$\bar{\sigma} = \frac{1.13}{\sqrt[4]{v}}.$$

He has used this formula to express in terms of the mean density

[1]) M. SMOLUCHOWSKI, Theorie Cinétique de l'opalescence. Bull. Crac 1907 p. 1057. Ann. der Phys. Bd. 25, 1908, p. 205. Phil. Mag. 1912. On opalescence of gases in the critical state. W. H. KEESOM, Ann der Phys. 1911 p. 591.

[2]) A. EINSTEIN. Ann. der Phys. Bd. 33 1910, p. 1276.

[3]) ORNSTEIN, These Proc., **15**, p. 54 (1912).

the accidental deviations in a cube, the side of which is equal to the wave-length of the light used in the experiments on opalescence.

Now there is a difficulty with this formula, to which, indeed, lead also the considerations of EINSTEIN as well as statistical mechanics when worked out in an analogous way for the critical point. In all these cases the mutual independence of the elements of volume is presupposed. Now, let there be given for the element of volume v the mean square of deviation *viz.* $\overline{(n-\bar{n})^2}$. Consider p equal contiguous elements of volume v_1, v_2, etc., in which n_1, n_2, etc. particles are situated, $\bar{n}_1, \bar{n}_2$ etc. indicating the mean values of these numbers.

Hence in the volume $V = v_1 + v_2 + \ldots$ there are $N = n_1 + n_2 + \ldots$ particles.

For the mean value of N we have

$$\bar{N} = \bar{n}_1 + \bar{n}_2 + \ldots$$

subsequently

$$\overline{(N-\bar{N})^2} = \overline{\{(n_1 - \bar{n}_1) + (n_2 - \bar{n}_2) + \ldots\}^2} = p\,\overline{(n-\bar{n})^2}$$

since, the elements of volume being supposed independent of each other, the means of the double products vanish. So we find for the deviation of density that the product of volume and mean square of deviation must be a constant.

Indeed the above-mentioned formula of probability for the deviations of density is so far inexact, as the terms of higher order appearing in it are at variance with the mutual independence of the elements of volume, which underlies the deduction of the frequency-law. In fact this deduction is only valid for such large elements of volume that these terms are no more of any influence. It is easily seen that this limit, above which the formula is valid, increases indefinitely in approaching the critical point. This explains also mathematically the wrong dependence on c found for the mean deviation in the critical point itself.

Now one could try to deduce the formula to a farther approximation. However, also the supposition of independence of the elements of volume is inexact in case these are small, and it would thus be impossible to ascertain how far the formula would yet differ from reality. [1])

[1]) A deduction of the inequalities in which the inexact terms of higher order do not at all appear, is given by ZERNIKE in his thesis, which will shortly appear. As this deduction too uses the independence alluded to, the objection mentioned holds here also.

The remark of EINSTEIN (l.c. p. 1285) that there would be no principal difficulty

2. Now, in order to avoid the difficulties mentioned, it is necessary to take into account the influence of deviations in the one element on the state in another. Let us divide the system into infinitely small elements of volume. A molecule is considered to lie in the element when its centre is situated in it. We consider an element dv_0 in the origin of coordinates. Around this element we imagine the sphere of attraction i.e. the region in which a molecule must lie when it is to have any influence on the state in dv_0. We determine the numbers of molecules for the elements of the sphere of attraction in giving the deviations v_1, v_2 etc. from the mean number of molecules per unit of volume.

We suppose the mean value of the density $\overline{v_0}$, when v_1 etc. are given, to be a linear function of the deviations v_1 etc., i.e. we put[1])

$$\overline{v_0} = C + f_1 v_1 dv_1 + f_2 v_2 dv_2 + \ldots \ldots \quad (1)$$

Taking the mean value of v_0 over all possible values of v_1, it appears immediately that $C = 0$, hence

$$\overline{v_0} = f_1 v_1 dv_1 + f_2 v_2 dv_2 + \ldots \ldots \quad (2)$$

The coefficients f denote the coupling of the elements, they only depend on the relative coordinates, i.e. here, on $x\,y\,z$. That the influence of an element, when the density is given, must be proportional to its size is immediately seen by considering the influence of uniting two elements in (2).

We shall now write the sum (2) as an integral. For the density in the element $dx\,dy\,dz$ we put v_{xyz}; further, we can dispose of f in such a way that $f(0,0,0) = 0$. Then for (2) we get

$$\overline{v_0} = \int_{-\infty}^{+\infty}\!\!\!\int\!\!\!\int v_{xyz} f(x,y,z)\, dx dy dz \quad \ldots \ldots \quad (3)$$

The integration may be extended here from $-\infty$ to $+\infty$, f being zero outside the sphere of action[2]).

---

in extending his deduction to a further approximation, is therefore mistaken. On the contrary, the consideration of higher terms so long as the independence is made use of, will not lead to anything.

[1]) Putting things more generally, we could write a series in v_1 etc. instead of (1). However, for the purpose we have in view, (1) is sufficient.

[2]) The quantity v can only take the values $1-adv$ and $-adv$, hence v is a discontinuous function of the coordinates. One might be inclined therefore, to continue writing a sum instead of the integral (3) and to solve the problem dealt with in the text with the aid of this sum. In doing so one gets sum-formulae which are wholly analogous to the integrals we used. However, we prefer introducing the integral, as the discontinuous function v has entirely disappeared from formula (6) only the function g appearing in it, which is continuous when the function f is

On the contrary, if v_0 is given, $\bar{v}$ has another value for the surrounding elements, than if $v_0 = 0$. Be in the element at xyz

$$\bar{v}_{xyz} = g(x, y, z, v_0 dv_0) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and let us try to determine the function g, the function f being given.

Now take the mean of formula (3), a fixed value v_1 being ascribed to v in a certain element $dx_1\, dy_1\, dz_1$.

In x, y, z, according to (4)

$$v_{xyz} = g(x-x_1, y-y_1, z-z_1, v_1\, dx_1 dy_1 dz_1) \quad . \quad . \quad . \quad . \quad (5)$$

For the first member we therefore get

$$g(x_1, y_1, z_1, v_1 dx_1 dy_1 dz_1)$$

as f and g do not depend on the direction of the line joining the elements. In the integral, (5) cannot be applied to the element $dx_1\, dy_1\, dz_1$; however, this element gives

$$v_1 f(x_1, y_1, z_1)\, dx_1 dy_1 dz_1$$

Further taking $g(0, 0, 0)$ zero, as it may arbitrarily be chosen, we get

$$g(x_1,y_1,z_1,v_1 dx_1 dy_1 dz_1) = \iiint_{-\infty}^{+\infty} g(x-x_1, y-y_1, z-z_1, v_1 dx_1 dy_1 dz_1) f(xyz)\, dx dy dz + \\ + v_1 f(x_1,y_1,z_1)\, dx_1 dy_1 dz_1.$$

This is true for all values of $v_1\, dx_1\, dy_1\, dz_1$, hence g must contain this quantity as a factor, and we obtain

$$g(x_1,y_1,z_1) - \iiint_{-\infty}^{+\infty} g(x-x_1, y-y_1, z-z_1) f(xyz)\, dx dy dz = f(x_1,y_1,z_1)$$

Now put $x-x_1 = \xi$, $y-y_1 = \eta$, $z-z_1 = \zeta$, and omit the index, then for g we get the integral equation

$$g(x,y,z) - \iiint_{-\infty}^{+\infty} f(x+\xi, y+\eta, z+\zeta)\, g(\xi\eta\zeta)\, d\xi d\eta d\zeta = f(xyz). \quad . \quad (6)$$

For g we have

$$\bar{v}_{xyz} = g(xyz)\, v_0\, dv_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

from which it appears immediately that

$$\overline{v_{xyz}\, v_0} = g(xyz)\, \overline{v_0^2}\, dv_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

---

continuous. The integral-formulae obtained in this way are easier to deal with mathematically, and besides the integral equation (6) has been solved, this being not so easily found from the analogous sum-formula.

Now let us consider more closely the coefficient of g in (8).

Let a molecules be present in the unit of volume, then the mean number of molecules in dv is equal to adv. If we take dv very small, there may be no or one molecule in it. The chance for one molecule is, therefore, adv; for none $1 - adv$. In the first case $v = \frac{1}{dv} - a$, in the latter it is $-a$, thus

$$\overline{v^2} = \frac{a}{dv} - a^2$$

or

$$\overline{v^2}dv = a \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Introducing this into (8), we find for the two elements $x_\sigma y_\sigma z_\sigma$ and $x_\tau y_\tau z_\tau$

$$\overline{v_\sigma v_\tau} = a\eta(x_\sigma - x_\tau, y_\sigma - y_\tau, z_\sigma - z_\tau) \quad . \quad . \quad . \quad . \quad . \quad (10)$$

This result can be used to indicate the values of $\overline{(N-\bar{N})^2} = \overline{\Delta N^2}$ for any volume.

We have

$$\Delta N = \int v dv$$

$$\overline{\Delta N^2} = \iint_{VV} \overline{v_\sigma^2}\, dv_\sigma\, dv_\tau + \iint_{VV} \overline{v_\sigma v_\tau}\, dx_\sigma\, dy_\sigma\, dz_\sigma\, dx_\tau\, dy_\tau\, dz_\tau$$

from which applying (9) and (10)

$$\overline{\Delta N^2} = aV + a\iint_{VV} g\,(x_\sigma - x_\tau, y_\sigma - y_\tau, z_\sigma - z_\tau)\, dx_\sigma\, dy_\sigma\, dz_\sigma\, dx_\tau\, dy_\tau\, dz_\tau .$$

This holds for every size and form of V. Elaborating it for a cube with side l the dependence on V is seen more clearly. Putting $x_\sigma - x_\tau = \xi$, $y_\sigma - y_\tau = \eta$, $z_\sigma - z_\tau = \zeta$, and integrating only for $\xi\eta\zeta$ positive, by which $^1/_8$ of the integral in question is found then, we get

$$\overline{\Delta N^2} = \bar{N} + 8a \int_0^l\int_0^l\int_0^l g\,(\xi\eta\zeta) \int_\xi^l\int_\eta^l\int_\zeta^l dx_\sigma\, dy_\sigma\, dz_\sigma$$

$$= \bar{N} + 8a \int_0^l\int_0^l\int_0^l (l^3 - l^2\,(\xi+\eta+\zeta) + l\,(\xi\eta+\eta\zeta+\zeta\xi) - \xi\eta\zeta)\, g\, d\xi\, d\eta\, d\zeta.$$

Hence

$$\frac{\overline{\Delta N^2}}{\bar{N}} = 1 + \iiint_{-l}^{+l} g\,(xyz)\,dxdydz - 3\iiint_{-l}^{+l} \frac{|x|}{l}\, g\,dxdydz + 3\iiint_{-l}^{+l} \frac{x^2}{l^2}\, g\,dx\,dy\,dz - \iiint_{-l}^{+l} \frac{|x^3|}{l^3}\, g\,dx\,dy\,dz.$$

Every integral in this formula is always smaller than the preceding one. If l is large with respect to the distance for which g has an appreciable value,.. there remains only the first integral. For any great volume we have

$$\frac{\overline{\Delta N^2}}{\bar{N}} = 1 + \iiint_{-\infty}^{+\infty} g\,dx\,dy\,dz \quad . \quad . \quad . \quad . \quad . \quad (11)$$

3. In trying to determine the function f by means of statistical mechanics, we meet with difficulties. Still something may be found about the quantities $\overline{\nu_r \nu_\tau}$ by applying the statistic-mechanical method to our problem. Indeed statistical mechanics permit to introduce a mutual action of the elements of volume.

We will avail ourselves of a canonical ensemble. We suppose the molecules to be spherical and rigid, and to attract each other for distances which are great with respect to their size. Elements small with respect to the sphere of attraction therefore may still contain a great number of molecules. But now we drop the supposition of the sphere of attraction being homogeneously filled for all systems (or at least for by far the greater part of them) [1]).

In calculating the number of the various distributions, we must, for the potential energy of attraction, take into account the mutual action of the elements; whereas, in calculating the exclusion of definite configurations of centres, we may neglect the fact that there is some correlation on the borders of the elements. For the dimensions of the elements have been supposed large with respect to the molecular diameter.

The mutual potential energy of the $r + \tau$ molecules contained in an element dv, will be represented by

$$-\tfrac{1}{2}\frac{(v+\tau)^2}{dv}\varphi_{00}$$

in this formula v represents the number of molecules contained in the volume dv for the most frequent system. In this system the distribution is homogeneous.

[1]) Cf ORNSTEIN, Toepassing der Statistische mechanica van GIBBS op molekulair-theoretische vraagstukken. Diss. Leiden 1908, p. 43 and p. 110.

Of course, the potential energies will not strictly be the same for different configurations within the elements, but we shall neglect this complication. Further we will represent the mutual potential energy for the two elements σ and ϱ by

$$\frac{-(v+\tau_\sigma)(v+\tau_\rho)\varphi_{\rho\sigma}}{dv}$$

all elements of volume being put equal.

For the total potential energy we find, in this way

$$-\frac{1}{2dv}\Sigma_\sigma(v+\tau_\sigma)\Sigma_\rho(v+\tau_\rho)\varphi_{\rho\sigma}.$$

For the frequency ξ of a system with the given distribution of molecules we find

$$\zeta = C\frac{n!}{(v+\tau_1)!(v+\tau_2)!\ldots}(\omega_1 dV)^{v+\tau_1}\ldots(\omega_1 dV)^{v+\tau_2}\ldots e^{\frac{1}{2\Theta dv}\Sigma_\sigma(v+\tau_\sigma)\Sigma_\rho(v+\tau_\rho)\varphi_{\rho\sigma}}.$$

Here ω is the function defined in the quoted dissertation on p. 48. Supposing $\tau \ll v$ and developing, we get,

$$\zeta = C\,\omega^n a^{-n} e^{\frac{1}{2\Theta}na\Sigma_\rho\varphi_{\rho\sigma}+\frac{1}{2}\Sigma\Sigma\left(-\frac{1}{v}+\frac{1}{v}\frac{d}{da}a^2\frac{d\log\omega}{da}+\frac{\varphi_{00}}{\Theta dv}\right)\tau_\sigma^2+}$$

$$+\frac{\varphi_{\sigma\rho}}{\Theta dv}\tau_\sigma\tau_\rho+\ldots\ldots\ldots\ldots\ldots\quad(12)$$

The number of molecules per unit of volume represented there by n, has been put a in this paper. The function ω and the faculties are developed in the same way as in the quoted dissertation. The double sum in the exponent gives the forms $\Sigma_\sigma v\Sigma_\rho\tau_\rho\varphi_{\rho\sigma}$ and $\Sigma\tau_\sigma\Sigma v\varphi_{\rho\sigma}$. These forms are identical, as they consist of the same terms differently arranged, further $\Sigma\varphi_{\rho\sigma}$ is the same for all molecules and $\Sigma\tau_\sigma = 0$, consequently both sums vanish.

The constant C contains the factor $Ne^{\Psi/\Theta}$ along with quantities which do not depend on the volume by summing up (12) over all possible values of τ (and taking into account that $\Sigma\tau_\sigma = 0$) we get N, the total number of systems in the ensemble. So we find

$$e^{-\frac{\Psi}{\Theta}} = \frac{C'}{\sqrt{\Delta}}\omega^n V^n e^{\frac{n^2}{2\Theta v}\Sigma\varphi_{\rho\sigma}},$$

the quantity Δ being the discriminant of the quadratic form in the exponent.

When we write $\Sigma_\rho\varphi_{\rho\sigma} = a$, we find for the pressure $p = -\frac{\partial\psi}{\partial v}$

$$\frac{p}{\Theta} = \frac{n}{V} + n\frac{d\log\omega}{dV} - \frac{a}{2\Theta}\frac{n^2}{V^2}$$

$\Delta^{-1/3}$ being very small with respect to the other factors, we may neglect its influence in ψ [1]). The equation of state has the same form as VAN DER WAALS' equation. However, the correlation is sensible in the accidental deviation; for it changes the value of $\overline{\tau^2}$; and $\overline{\tau_\sigma\tau_\rho}$, which vanish if the correlation is neglected, obtain values deviating from zero.

Denoting by $\Delta_{\sigma\sigma}$ and by $\Delta_{\sigma\tau}$ the minors of the discriminant, we have

$$\overline{\tau_\sigma^2} = \frac{\Delta_{\sigma\sigma}}{\Delta}(k-1)$$

$$\overline{\tau_\sigma\tau_\rho} = \frac{\Delta_{\sigma\rho}}{\Delta}(k-1)\ ^2)$$

where k is the number of elements into which the volume is divided.

The condition $\Delta = 0$ is equivalent to the condition $\frac{dp}{dv} = 0$. For if we write down the determinant in some arrangement, and if we add all rows to the first row, we get a determinant of which all terms of the first row have the form

$$-\frac{1}{v} + \frac{1}{v}\frac{d}{da}a^2\frac{d\lg\omega}{da} + \frac{1}{\Theta dv}\Sigma\varphi_{\sigma\rho}.$$

Strictly speaking, this is not true for some terms at the end of the row, but as we have neglected the action on the borders, we may neglect this fact too. In reality our considerations are only true for an infinitely great volume, where this difficulty disappears, as Δ is then an infinite determinant.

Now if

$$-\frac{1}{v} + \frac{1}{v}a^2\frac{df\omega}{da} + \frac{1}{\Theta dv}\Sigma\varphi_{\sigma\rho} = 0$$

then $\Delta = 0$.

Or if $\Delta = 0$

$$1 - \frac{d}{da}a^2\frac{df\omega}{da} + \frac{a}{\Theta}a = 0$$

which therefore agrees with $\frac{dp}{dv} = 0$.

---

[1]) Cf. l. c. p. 129.

[2]) Cf. ORNSTEIN, Accidental deviations in mixtures. These Proceedings **15**, p. 54, (1912).

The quantities $\overline{\tau_\sigma \tau_\rho}$ etc. here found are related to those mentioned above. And though a statistical deduction of the function f entering into details may lead to difficulties, yet it is clear that statistical mechanics yield a correlation analogous to that expressed in g.

If we should wish to continue the deduction of the conditions of the critical point, we should have to use higher powers of τ_σ, which can be done without difficulty; we then find for the second condition $\frac{d^2p}{dv^2} = 0$.

If we drop the supposition that the sphere of attraction is large, we can use the function ϑ, defined in the quoted dissertation. In order to take into account the correlation, we must suppose the integrals

$$\int e^{-\varepsilon_q/\Theta} \, dw_1 \ldots dz_n \, \tau_\varkappa = \vartheta(n_\varkappa)$$

defining ϑ, to depend on n_k for the element in question and also on the numbers of molecules in the surrounding elements. Therefore, in general, the numbers of molecules of all elements will appear in $\vartheta n_\varkappa$, but the influence of distant elements is so small that $\frac{\partial \vartheta n_\varkappa}{\partial n_\sigma}$ can be put zero.

By considerations analogous to those used in the quoted dissertation, we can show that $\vartheta(n_\varkappa)$ has the form

$$V_\varkappa^{n_\varkappa} (\omega \, n_\varkappa, n_{\varkappa'}, n_{\varkappa''})^{n_\varkappa}$$

in which $n_{\varkappa'}$ $n_{\varkappa''}$ denote the densities (molecular), in the elements with which $V_\varkappa$ is in mutual action. The values of all $n_\varkappa$ are equal for the most frequent system.

Now we find for ζ

$$\zeta = C \, V^n (\omega \, n, n, n, n \ldots)^n \, e^{-P}$$

where P is a quadratic form in the deviations for the various elements, containing squares as well as double products. The form might be easily indicated, but we will omit it, as it is only our purpose to show how in general the statistic-mechanical considerations, changed in the sense of a correlation of elements of volume, lead to formulae analogous to those given in § 2. Here too the mean square of deviation and the means of double products are represented by quotients of minors of the discriminant of P and this quantity itself. Here too for $\frac{dp}{dv} = 0$ the discriminant vanishes.

4. The above considerations can be applied in calculating the critical opalescence. For that purpose we use the simple method indicated by LORENTZ[1]), which consists in superposing the light-vectors caused by the influence of every individual molecule in a point at great distance.

Consider in the substance through which a beam of light passes, a volume V great with respect to the wave-length, and take a distant point P, the direction VP forming an angle φ with the incident ray.

All molecules lying in one plane perpendicular to the line which bisects the angle φ, will cause equal phase in P. Take therefore a system of axes with the Z-axis parallel to this line, then the contribution of one molecule will be

$$\beta \sin \frac{2\pi}{\mu\lambda}(ct + 2z \cos \tfrac{1}{2}\varphi)$$

where β depends only on the kind of molecules, on λ and on the distance VP, μ being the index of refraction.

The number of molecules in $dx\,dy\,dz$ amounts to

$$(a + v)\,dx\,dy\,dz.$$

The total light-vector in P thus becomes

$$\beta \int_V (a + v) \sin \frac{2\pi}{\mu\lambda}(ct + 2z \cos \tfrac{1}{2}\varphi)\,dx\,dy\,dz.$$

and the intensity

$$\beta^2 \frac{c}{\mu\lambda} \int_0^{\mu\lambda/c} dt \iint_{VV} (a + v_\sigma)(a + v_\tau) \sin \frac{2\pi}{\mu\lambda}(ct + 2z_\sigma \cos \tfrac{1}{2}\varphi) \sin \frac{2\pi}{\mu\lambda}(ct + 2z_\tau \cos \tfrac{1}{2}\varphi)\,dx_\sigma\,dy_\sigma\,dz_\sigma\,dx_\tau\,dy_\tau\,dz_\tau.$$

Integrating with resp. to t, we get

$$\frac{1}{2}\beta^2 \iint_{VV} \{a^2 + a(v_\sigma + v_\tau) + v_\sigma v_\tau\} \cos \left\{\frac{4\pi}{\mu\lambda}(z_\tau - z_\sigma) \cos \tfrac{1}{2}\varphi\right\} dx_\sigma\,dy_\sigma\,dz_\sigma\,dx_\tau\,dy_\tau\,dz_\tau.$$

The mean value of this must be calculated. The term with $v_\sigma + v_\tau$ vanishes, and that with a^2 yields no contribution proportional to V. We introduce the value of $\overline{v_\sigma v_\tau}$ from formula (10), and for $\sigma = \tau$ from form. (9). This gives

---

[1]) H. A. LORENTZ, On the scattering of light by molecules. These Proceedings **13** p. 92 (1910).

$$\frac{1}{2}\beta^2 aV + \frac{1}{2}\beta^2 a \iint_V\int_V g(x_\tau - x_\tau,\ y_\tau - y_\tau,\ z_\tau - z_\tau)\cos\frac{4\pi}{\mu\lambda}(z_\tau - z_\tau)\,dx_\tau \ldots dy_\tau$$

For a great volume one integration over V can be performed (compare the deduction of formula (11)); further we put $aV = \overline{N}$ and for the sake of brevity $\frac{4\pi}{\mu\lambda}\cos\frac{1}{2}\varphi = C$, then we get

$$\frac{1}{2}\beta^2\,\overline{N}\left[1 + \int_{-\infty}^{+\infty}\iiint \cos Cz\, g(x, y, z)\,dxdydz\right] \quad . \quad . \quad . \quad (13)$$

The integral appearing here will be represented by G_c, that of formula (11) by G. It will be seen that the deductions criticised in § 1 yield an opalescence proportional to $\overline{v^2}$, a quantity which according to the above is proportional to $1 + G$, whereas the opalescence is proportional to $1 + G_c$.

With the aid of the integral-equation (6) we can express G and G_c in the corresponding integrals of the function f, which we will indicate by F and F_c.

Integrating (6) with resp. to xyz from $-\infty$ to $+\infty$, we find

$$\int_{-\infty}^{+\infty}\iiint g(xyz)\,dx\,dy\,dz - \int_{-\infty}^{+\infty}\iiint g(\xi\eta\zeta)\,d\xi\,d\eta\,d\zeta \int_{-\infty}^{+\infty}\iiint f(x+\xi, y+\eta, z+\zeta)\,dxdydz =$$

$$\int_{-\infty}^{+\infty}\iiint f(xyz)\,dxdydz$$

or

$$G = \frac{F}{1-F} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

Multiplying (6) by $\cos Cz$ and again integrating, we get

$$G_c - \int_{-\infty}^{+\infty}\iiint g(\xi\eta\zeta)\,d\xi d\eta d\zeta \int_{-\infty}^{+\infty}\iiint \{\cos C(z+\zeta)\cos C\zeta + \sin C(z+\zeta)\sin C\zeta\}$$

$$f(x+\xi, y+\eta, z+\zeta)\,dxdydz = F_c.$$

The integral with the sines disappears because f and g are even functions; we find

$$G_c = \frac{F_c}{1-F_c} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

In order to apply the results obtained and to test them experimentally, one might try to deduce f from molecular theory. This would at best be possible under very simplifying suppositions and

even then only an approximation can be obtained. Therefore we will take another way. As remarked in § 1, the exact value of $\overline{v^2}$ for *very great volumes* was already known. In our notation we have

$$\overline{v^2}\, V = -\frac{RT}{N}\frac{a^2}{v\dfrac{dp}{dv}}$$

where N is the number of AVOGRADO, v the molecular volume.

According to formula (4) we have

$$\overline{v^2}\, V = a\,(1+G) = \frac{a}{1-F}.$$

Putting these results equal, we get

$$1 - F = -\frac{v^2}{RT}\frac{dp}{dv}.$$

In the critical point $F = 1$. [1])

The formula of opalescence first arrived at by KEESOM and EINSTEIN

---

[1]) There appears to exist a closer correspondence between the given statistic-mechanical method and the method using general considerations of probability, than perhaps might be expected. The elements of the discriminant (which is an infinite determinant in the former) agree with the function f in the latter. The former finds from this the value of $\overline{\nu_\tau\, \nu_\tau}$ as the quotient of a minor with that discriminant, the latter deduces this value from an integral-equation. In the critical point the discriminant vanishes, corresponding to this the FREDHOLM determinant of the integral-equation is likewise zero. That this is the case when $F = 1$, appears by more closely studying the equation

$$g(\xi\eta\zeta) - \lambda \int g(\xi\eta\zeta)\, f(x+\xi, y+\eta, z+\zeta)\, d\xi d\eta d\zeta = 0$$

which only permits appropriate solutions if $\lambda = \frac{1}{F}$, (i e. this is the only proper value). For $F = 1$ this is therefore the case with the equation (6) without second member.

From the formula (15) it will be seen that form. (6) can be solved by a FOURIER integral. Putting

$$\iiint_{-\infty}^{+\infty} \cos mx \cos ny \cos lz\, f(xyz)\, dxdydz = \varphi(m, n, l)$$

we have

$$g(xyz) = \frac{1}{8\pi^3} \iiint_{-\infty}^{+\infty} \frac{\varphi(m, n, l)}{1-\varphi(m, n, l)} \cos mx \cos ny \cos lz\, dmdndl.$$

$$\frac{I_{Op.}}{I} = \frac{2\pi^2 V}{D^2}\frac{RT}{N}\frac{\mu v\left(\frac{d\mu}{dv}\right)^2}{-\frac{dp}{dv}}\lambda^{-4}\sin^2\psi. \quad . \quad . \quad . \quad . \quad (16)$$

in which represent

D distance of observation

μ index of refraction,

ψ angle of electric force in incident light with direction of observation,

will likewise be found by using in (13) the value found for F instead of F_c. The exact formula then will result by multiplying by the factor $\frac{1+G_c}{1+G} = \frac{1-F}{1-F_c}$.

Developing the cosine in F_c we find

$$F - F_c = \frac{C^2}{2}\iiint_{-\infty}^{+\infty} z^2 f(x,y,z)\, dxdydz.$$

Representing this integral by ε^2 and introducing the value of C, we get

$$F - F_c = 4\pi^2(1 + \cos\varphi)\left(\frac{\varepsilon}{\mu\lambda}\right)^2.$$

The formula of opalescence then will be:

$$\frac{I_{Op.}}{I} = \frac{2\pi^2\frac{V}{D^2}\frac{RT}{N}v\mu\left(\frac{d\mu}{dv}\right)^2\lambda^{-4}\sin^2\psi}{-\frac{dp}{dv} + 4\pi^2\frac{RT}{v^2}(1+\cos\varphi)\left(\frac{\varepsilon}{\mu\lambda}\right)^2}. \quad . \quad . \quad . \quad (17)$$

In the critical point itself it therefore is

$$\frac{I_{Op.}}{I} = \frac{V}{D^2}\frac{v^3\mu^3\left(\frac{d\mu}{dv}\right)^2\sin^2\psi}{N\varepsilon^2(1+\cos\varphi)}\lambda^{-2} \quad . \quad . \quad . \quad . \quad (18)^1)$$

The greater exactness of form. (17) as compared with (16) is confirmed by the measurements of one of us (Z.). According to these measurements, which however bear upon a mixture of liquids the

[1]) According to this formula the proportionality of the opalescence to λ^{-4}, which holds for higher temperatures, changes continuously in the immediate neighbourhood of the critical point, into proportionality with λ^{-2}. This *real* "getting whiter" of the opalescence should not be confused with the *apparent* changing of colour which is always observed much farther from the critical point. The latter indeed is only a result of the method of observation, as is clearly proved by the measurements of one of us (cf. ZERNIKE thesis).

reciprocal value of a quantity proportional to the opalescence changes linearly with the difference of temperature $T-T_k$, but by extrapolation does not vanish for $T=T_k$, but for $T-T_k=0,^\circ 012$. When therefore for this value of $T-T_k$ the denominator of (17) is equal to zero, we can find from this, using VAN DER WAALS' equation, an estimation for ε/λ. The calculation yields:

$$\frac{\varepsilon}{\lambda}=0{,}0022 \text{ or } \varepsilon=1{,}2.10^{-7}\ cm.$$

The quantity ε is a measure for the size of the sphere of attraction. For

$$\varepsilon^2=\frac{1}{3}\int_{-\infty}^{+\infty}\!\!\iiint \varrho^2 f(xyz)\,dxdydz$$

(ϱ distance to origin) whereas in the critical point

$$\int_{-\infty}^{+\infty}\!\!\iiint f(xyz)\,dx\,dy\,dz=1.$$

If f were constant within a sphere with radius R, then ε^2 would be $^1/_6\,R^2$, and the above estimation would give

$$R=2{,}7\,.\,10^{-7}\ cm.$$

SUMMARY.

1. The known formulae of critical opalescence give an infinite value at the critical point. Efforts to escape from this difficulty have furnished formulae for the deviations of density with a dependence upon the volume, at variance with the assumed mutual independence of the elements of volume.

2. In order to obtain formulae applicable in the critical point, it is found necessary to take into account the mutual influence of the elements of volume, it being shown that near the critical point this influence is sensible for distances large in comparison with the radius of the sphere of attraction.

3. Two functions are introduced, one relating to the direct interaction of molecules, the other to the mutual influence of two elements of volume. An integral equation gives the relation between the two functions.

4. Corrected values are found for the mean deviations, and in the formula of opalescence a correction is introduced. The latter depends upon the sphere of attraction which can thus be calculated from observations.

5. Further it is shown that the same results may be arrived at by taking into account the mutual influence of the elements of volume in the deductions of statistical mechanics.

Groningen, Sept. 1914.

The Clustering-Tendency of the Molecules in the Critical State and the Extinction of Light Caused Thereby

F. Zernike

Reprinted from
PROC. AKAD. SCI. (Amsterdam)
Volume 18, pp. 1520–1527 (1916)

Physics. — "*The clustering-tendency of the molecules in the critical state and the extinction of light caused thereby*". By Dr. F. ZERNIKE. (Communicated by Prof. LORENTZ).

(Communicated in the meeting of February 26, 1916).

1. In a former paper by ORNSTEIN and the author [4]) an improved formula was derived for the scattering of light by a single substance in the neighbourhood of the critical point. In that paper the inaccuracy of former approximate formulae, which gave an infinitely strong

[1]) G. A. F. MOLENGRAAFF. Overschuivingen in overschuivingsbladen op de eilanden Timor en Leti. Voordracht gehouden in de Bestuursvergadering van 18 Januari 1913. Tijdschr. Kon. Ned. Aardr. Gen. 2. XXX. p. 273. 1913.

[2]) J. WANNER. Geologie von West-Timor, Geologische Rundschau Bd. IV. Heft 2.

[3]) A. TOBLER. l.c. p. 29.

[4]) These Proceedings XVII, p. 793. This paper will here be indicated by l. c.

scattering for the critical state, had been corrected by exactly taking into account the influence of the arrangement of the molecules in space, after a new probability-method. In fact, the intensity of the scattered light could be given for the critical point itself.

The former difficulty however has not quite disappeared in that way. For when we compute from the given formula the total scattered energy, in order to find in this way the extinction of the incident light, then a logarithmic expression arises which still becomes infinite in the critical point.

The performance of a better approximation also appeared of practical importance, because it has already been shown l.c. that the magnitude of the sphere of attraction might experimentally be found from measurements of the opalescence. Therefore a theoretically accurate formula for the extinction is of great importance.

It could be seen beforehand that the inaccuracy of the given derivation had to be ascribed to the optical treatment, which was far too rough. Presently I have succeeded in finding a better value of the extinction, by deriving it directly from the theory of electrons. To this end it appeared necessary to calculate explicitly the formerly introduced function g, which indicates the course of the mean density of the molecular clusters. I will start with this calculation. The result thereof is by itself interesting, as I think the problem of the "Neigung zur Schwarmbildung", which SMOLUCHOWSKI [1]) posed in 1904, is solved definitively by it.

2. Two different functions were introduced l.c.: the first, $f(x,y,z)$, represents the influence which a known deviation of density in the point x, y, z has on the mean density in the origin, if the mean density obtains at the same time at all other points of the neighbourhood. The second $g(x, y, z)$ represents the mean density in (x, y, z), when it is only known that there exists a certain deviation of density in the origin. Of these functions it has been shown l. c. that they are connected by an integral-equation, which can be put in the following form

$$g(x_1,y_1,z_1) - \int_{-\infty}^{+\infty}\iiint g(x_1-x,y_1-y,z_1-z)\,f(x,y,z)\,dx\,dy\,dz = f(x_1,y_1,z_1) \quad (1)$$

From this integral-equation we also derived a simple relation between the volume-integrals of g and f, for which we respectively write G and F:

[1]) M. SMOLUCHOWSKI. Boltzmann-Festschrift 1904.

$$1 + G = \frac{1}{1 - F} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

It will be seen from the meaning of f and g that these functions will not depend upon the direction of the radius-vector $r = \sqrt{x^2 + y^2 + z^2}$. It gives, however, no simplification to introduce this in (1). Now develop g under the integral sign, according to powers of x, y and z. The differential quotients of g can then be brought before the integral-sign, and there remain integrals of the general form

$$\iiint_{-\infty}^{+\infty} x^r y^s z^t f(x,y,z)\, dxdydz$$

As f only depends upon r, these integrals will vanish when one or more of the numbers r, s, t are odd; and their value will remain unchanged by transposition of r, s and t. If we only go to terms of the second order, besides the value F for $r = s = t = 0$, we also meet the integral with $r = 2$ $s = t = 0$, which has already been represented l. c. by ε^2. (The function f will only differ from zero for the small value of r, for which the molecular attraction is yet sensible. Therefore we could call the magnitude $\sqrt{3}.\varepsilon$, in analogy with the "mean error", the mean radius of the sphere of attraction). The just mentioned manipulation makes the integral-equation pass into a differential equation for g:

$$g(x_1,y_1,z_1) - Fg - \frac{\varepsilon^2}{2}\left(\frac{\partial^2 g}{\partial x_1^2} + \frac{\partial^2 g}{\partial y_1^2} + \frac{\partial^2 g}{\partial z_1^2}\right) = f(x_1, y_1, z_1)$$

or, when we now introduce r

$$\frac{d^2 g}{dr^2} + \frac{2}{r}\frac{dg}{dr} - \frac{2(1-F)}{\varepsilon^2} g = -\frac{2}{\varepsilon^2} f \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The general solution of the equation (3) without second member is

$$Ar^{-1} e^{-\varkappa r} + Br^{-1} e^{\varkappa r}$$

in which $\varkappa^2 = \frac{2(1-F)}{\varepsilon^2}$ and from this we easily find by variation of constants the solution of (3). The two constants in this solution may be determined from the two following conditions. In the first place g must remain finite for $r = \infty$, a condition which was already required with the integral-equation (1). If we take for the second condition, that g remains finite also for $r = 0$, then we shall find

$$g = \frac{2e^{-\varkappa r}}{\varepsilon^2 \varkappa r} \int_0^r s f(s) \sinh \varkappa s\, ds + \frac{2 \sinh \varkappa r}{\varepsilon^2 \varkappa r} \int_r^\infty s f(s) e^{-\varkappa s} ds \quad . \quad . \quad (4)$$

It may be verified that the second condition has been well chosen by integrating this expression (4) from 0 to ∞. Reducing the integral by partial integration, it will be found that the condition (2) is accurately satisfied. In the critical point F becomes 1, as found l.c., whence $\varkappa = 0$. We have chosen the form of (4) in such a way that in it we can easily take $\varkappa = 0$. Thence

$$g = \frac{2\,r^{-1}}{\varepsilon^2}\int_0^r s^2 f(s)\,ds + \frac{2}{\varepsilon^2}\int_r^\infty s f(s)\,ds$$

a solution which we might have written down directly from (3) after well-known theorems of the potential-theory. Now nearly all the simplifications which arise for $\varkappa = 0$, can also be used approximately for small values of $\varkappa$. Besides, we must keep in mind that the equation (3) is also but an approximate one, which holds the more accurately, as will be seen by a closer consideration, as $\varkappa$ is smaller. Therefore I used the following formula for the further calculations

$$g = \frac{2}{\varepsilon^2}\,r^{-1}\,e^{-\varkappa r}\int_0^r s^2 f\,ds + \frac{2}{\varepsilon^2}\int_r^\infty s f\,ds. \quad . \quad . \quad . \quad . \quad (5)$$

If we call σ the radius of the sphere outside which $f = 0$, then (5) gives for $r > \sigma$

$$g = \frac{F}{2\pi\varepsilon^2}\,r^{-1}\,e^{-\varkappa r}.$$

It appears clearly from this form that the function g is sensible over a far greater distance than f, and that outside the sphere of attraction the way in which g vanishes for increasing r no longer depends upon the course of f.

3. The considerations from the theory of electrons to be given now, may be understood as an extension of the treatment which LORENTZ [1]) gave for the scattering of light by an ideal gas. In the first place it appears that this treatment may be applied unchanged to our case, if it is allowed to consider separately a space with dimensions small with respect to the wave-length, surrounding a given molecule (and especially if we may assume the mean density in the space outside). According to the above calculation, this will be the case if the quantity $\varkappa$, and therefore also the distance from the critical point, is not too small. In the final formula of LORENTZ

[1]) H. A. LORENTZ, These Proc. XIII, p. 92.

there appears in the damping-term a coefficient which represents the excess of the mean number of molecules in the small space considered, if it is known that one molecule lies within that space. This number is 1 for a random distribution, but in our case it is, as will be seen easily, $1 + G$. Thus this way also leads to the known formula for the extinction, according to which this quantity is inversely proportional to dp/dv.

Thence it is clear that we must take into account the influence of the molecule considered on the density of its surroundings, even at distances comparable with the wave-length. To this end it will no longer be possible to use the developments in power-series, which LORENTZ applies repeatedly. In order to make the calculation possible without that, I introduce the following simplification.

The electrons which are brought into vibration in the molecules by the incident light, (for simplicity we imagine one electron in every molecule) will in reality perform a somewhat irregular vibration, with an amplitude not always equal for everyone. I neglect these differences in calculating the resistance which the neighbouring oscillators exert on the electron particularly considered. LORENTZ also makes use of this approximation at a later point. Of course this is much more acceptable, however, for a diluted gas than for the rather dense nearly critical state.

First consider one molecule in the origin of coordinates, which bears a variable moment in the direction of z:

$$\mathfrak{p}_z = p_0 \sin kt \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

With the aid of well-known formulae for the potentials φ and $\mathfrak{a}$ we then find for a point (x, y, z) at a distance r

$$\frac{\partial \varphi}{\partial z} = \frac{p_0}{4\pi}\left[\left(\frac{k^2 z^2}{c^2 r^3} + \frac{r^2 - 3z^2}{r^5}\right)\sin k\left(t - \frac{r}{c}\right) + \frac{k}{c}\frac{r^2 - 3z^2}{r^4}\cos k\left(t - \frac{r}{c}\right)\right]$$

$$\frac{\mathfrak{a}}{c} = -\frac{p_0}{4\pi}\frac{k^2}{c^2 r}\sin k\left(t - \frac{r}{c}\right)$$

The sum of these two expressions with the negative sign gives the z-component of the electric force. Now if the incident light comes from the negative x-axis, the phase of the variable moment of a molecule in (x, y, z) will be so much in advance of one for which $x = 0$, as corresponds to a difference of way μx, in which μ is the refractive index of the material. Thus

$$\mathfrak{p}_{x,y,z} = p_0 \sin k\left(t - \frac{\mu x}{c}\right).$$

The electric force which this moment exerts in the origin is evi-

dently equal to the already calculated one in (x, y, z) from a moment in O. Terms with $\sin kt$ in that force will not yield a resistance to the motion (6). Therefore we further only need consider the terms with $\cos kt$. For the coefficient of $\frac{p_0}{4\pi}\cos kt$ we find

$$\left[-\frac{k^2}{c^2}\frac{r^2-z^2}{r^3}+\frac{r^2-3z^2}{r^5}\right]\sin\frac{k}{c}(\mu x+r)-\frac{k}{c}\frac{r^2-3z^2}{r^4}\cos\frac{k}{c}(\mu x+r) \quad (7)$$

4. We might now add the expression (7) for every molecule, except the one in the origin. We can also find at once the mean value of that sum, by multiplying by the mean density in (x, y, z) and afterwards integrating through the whole space. But here the difficulty arises that the integrals do not converge. Therefore we make use of the known fact that no extinction arises, and no resistance against the motion of the electrons, if the oscillators are quite regularly distributed through the whole space. The phenomenon which we want to calculate can thus only be caused by the deviation from the homogeneous distribution, as it is caused by the action of the molecule in the origin. The mean deviation of density at a distance r is exactly $g(r)$. So we multiply (7) by g and by the element of volume $dx\,dy\,dz$ and integrate through the whole space. To this end we introduce spherical coordinates, such that

$$x = r\cos\vartheta \qquad y = r\sin\vartheta\sin\varphi \qquad z = r\sin\vartheta\cos\varphi$$

then the integration over φ may be performed without difficulty and there remains

$$-\pi\int_0^{\pi}\sin\vartheta\,d\vartheta\int_0^{\infty}g\,dr\left[\left\{\frac{k^2}{c^2}r(1+\cos^2\vartheta)-\frac{3\cos^2\vartheta-1}{r}\right\}\sin\left\{\frac{k}{c}(1+\mu\cos\vartheta)r\right\}+\right.$$
$$\left.+\frac{k}{c}(3\cos\vartheta-1)\cos\left\{\frac{k}{c}(1+\mu\cos\vartheta)r\right\}\right] \quad (8)$$

In this I put for g the value (5) and then performed the integrations over r. About the rather complicated calculations I will only mention that the integrations from 0 to σ — the region for which the integrals appearing in (5) are variable — were separately performed, and that approximate values for a small σ could be used for the integrals from σ to ∞, which are of the type of the cosine-integral.

The expression in $\cos\vartheta$ thus obtained, contains many terms which vanish afterwards by the integration with respect to ϑ. The other terms can be reduced to an integral of a rational fraction in $\cos\vartheta$, and these give finally the following result for (8) ·

$$-\frac{F}{16\varepsilon^2}\frac{k}{c}\Big[-4-4\mu^{-2}-4a\mu^{-2}+$$

$$+\Big\{\mu+6\mu^{-1}+\mu^{-3}+2a(\mu^{-1}+\mu^{-3})+a^2\mu^{-3}\Big\}l\,\frac{(\mu+1)^2+a}{(\mu-1)^2+a}\Big] \quad . \quad (9)$$

in which we have put $a=\frac{c^2\varkappa^2}{k^2}$.

The form (9) yields the force of resistance by multiplying by $\frac{p_0}{4\tau}\cos kt$ and from this we further find the energy used up. The quantity p_0, however, still remains unknown. We can find this by considering that the refractive index μ also depends upon the vibrations performed by the electrons. Thus we find e.g. for the critical point where $a=0$ and $F=1$:

$$h=\frac{\pi v}{N\varepsilon^2\lambda^2}\frac{(\mu^2-1)^2}{16\mu}\Big[-4-4\mu^{-2}+2(\mu+6\mu^{-1}+\mu^{-3})\,l\,\frac{\mu+1}{\mu-1}\Big]$$

being an extinction proportional to λ^2, and depending directly upon the quantity ε, just as was found l. c. for the opalescence. Now I especially wanted to find the way in which the extinction-coefficient increases on approaching to the critical point. For this it will suffice to consider the variability of the expression between square brackets in (9). The quantity

$$a=\frac{c^2}{k^2}\,\frac{2(1-F)}{\varepsilon^2}$$

may be used as a measure of the "distance to the critical point."

For large values of a^2 we may expand the logarithm in that expression into descending powers of a^2. On reduction it is seen that the term with a^{-2} is the first that has a coefficient differing from zero, and that for large values of a^2 we can take for the whole expression

$$\frac{64}{3}a^{-1}-\frac{64}{3}a^{-2}(\mu+\mu^{-1}) \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Only preserving the first term, this will duly give an extinction inversely proportional to 1—F.

In my thesis I have graphically represented extinction-measurements of a liquid mixture, by plotting the reciprocal value of h as a function of $T-T_k$. The points lay rather accurately on a straight line, which cut the temperature-axis below T_k. Using this same method with the quantity (10), we find for the development of the reciprocal value of (10)

$$\frac{3}{64}a+\frac{3}{64}(\mu+\mu^{-1})+ - \ldots\ldots\ldots \quad (11)$$

This is a straight line which cuts the axis for $a = -(\mu+\mu^{-1})$.

If the difference of temperature which corresponds to this value is found by extrapolating from observations, the quantity ε can thus be deduced. The expansion used in (10) and (11) would nearly always be sufficient in practice. This appears from the following values P of the expression in (9), computed rigorously with $\mu = 1{,}1$, which is close to the true value for many substances. In the next column the reciprocal value has been inserted and in the fourth

Extinction P (arbitrary unit).

a	P	$\frac{10}{P}$	$10\times$ form. (11)
0	18.58	0.54	1.88
0.05	12.34	0.81	1.93
0.1	10.38	0.96	1.98
1	4.244	2.36	2.82
2	2.871	3.48	3.76
3	2.213	4.52	4.70
7	1.186	8.43	8.45
10	0.885	11.30	11.26
20	0.483	20.72	20.63

column the same quantity, as it would be deduced from the straight line (asymptote of the curve). If necessary one could, of course, also use the rigorous value for comparison with experimental data.

Groningen, February 24, 1916.

Reprinted from *Printed in U.S.A.*
JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 2 FEBRUARY 1963

On the van der Waals Theory of the Vapor–Liquid Equilibrium. I. Discussion of a One-Dimensional Model

M. KAC, G. E. UHLENBECK, AND P. C. HEMMER
The Rockefeller Institute, New York, New York
(Received 17 September 1962)

For a one-dimensional fluid model where the pair interaction potential between the molecules consists of a hard core and an exponential attraction, Kac has shown that the partition function can be determined exactly in the thermodynamic limit. In Sec. II this calculation is reviewed and further discussed. In Sec. III, we show that in the so-called van der Waals limit when the range of the attractive force goes to infinity while its strength becomes proportionally weaker, a phase transition appears which is described exactly by the van der Waals equation plus the Maxwell equal-area rule. In Sec. IV the approach to the van der Waals limit is discussed by an appropriate perturbation method applied to the basic integral equation. The perturbation parameter is the ratio of the size of the hard core to the range of the attractive force. It is seen that the phase transition persists in any order of the perturbation. The two-phase equilibrium is characterized by the fact that in this range of density, the maximum eigenvalue of the integral equation is doubly degenerate and that the corresponding two eigenfunctions do not overlap. In Sec. V we comment on the relevance of our results for the three-dimensional problem.

I. INTRODUCTION

IN this series of papers we intend to present a new discussion of the old theory of van der Waals[1] of the continuity of the gaseous and liquid states of matter. It is well known that the great merit of this theory lies in the fact that it gave the first qualitative kinetic interpretation of condensation phenomena and of the existence of a critical point. On the other hand, it has proved very difficult to make the theory more rigorous and as a result the modern theory[2] of the equation of state of a nonideal gas, has followed more the idea of Kamerlingh Onnes to represent all properties of the gas as power series in the density—the so-called virial expansion. In this way one can take successively into account the interaction of the molecules in pairs, triples, quadruples, etc., and one can derive precise expressions for the successive deviations from the ideal gas laws in terms of the intermolecular potential. Many attempts,[3] thus far unsuccessful, have been made to construct a rigorous theory of condensation phenomena from such expansions. In fact we believe that such a construction is very difficult, if not impossible,[4] and it therefore seems worthwhile to try to reformulate in a more rigorous way, the basic ideas of van der Waals.

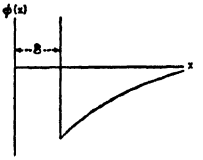


FIG. 1. Interaction potential $\varphi(x)$.

We have attempted to do this starting always from a one-dimensional gas model, first proposed by M. Kac,[5] for which all calculations can be carried out exactly. The model consists of N particles moving on a line of length L and interacting in pairs through a potential $\varphi(x)$ which consists of a hard core of length δ and an exponential attraction (see Fig. 1.) For this model it is possible to give an exact discussion of the partition function in the *thermodynamic limit* $L \to \infty$, $N \to \infty$, $l = L/N$ finite. As shown already by Kac, the problem in this limit can be reduced to the discussion of a linear integral equation with a positive definite, Hilbert–Schmidt kernel of which the maximum eigenvalue determines the

[1] J. D. van der Waals, Dissertation Leiden, 1873. This was expanded in the book: *Die Kontinuität des gasförmigen und flüssigen Zustandes* (Johann A. Barth, Leipzig, Germany, 1899), 2 volumes. Compare also the monograph by J. P. Kuenen, Die Zustandsgleichung (Vieweg, Braunschweig, 1907).

[2] Due mainly to J. E. Mayer. A summary is given in the book by J. E. Mayer and M. G. Mayer, *Statistical Mechanics* (John Wiley & Sons, Inc., New York, 1940), Chaps. 13 and 14. For a recent account see the monograph by G. E. Uhlenbeck and G. W. Ford in *Studies in Statistical Mechanics* (North-Holland Publishing Company, Amsterdam, 1962), Vol. 1, Part B.

[3] These also go back to J. E. Mayer; (see reference 2, Chap. 14). For a more recent discussion, see K. Ikeda, Progr. of Theoret. Phys. (Kyoto) 19, 653 (1958), and 26, 173 (1961). Much of the motivation came from the analogy with the Bose-Einstein condensation, which was pointed out by B. Kahn and G. E. Uhlenbeck [Physica 5, 399 (1938)]. However, we now believe that this analogy is superficial and that there is no real connection between the Bose-Einstein condensation and ordinary condensation phenomena.

[4] The difficulties become especially clear in the formulation of the condensation problem according to C. N. Yang and T. D. Lee, Phys. Rev. 87, 404 (1952). Compare also the discussion by G. E. Uhlenbeck and G. W. Ford in Chapter 3 of the book *Lectures in Statistical Mechanics* (Proceedings of the Summer Seminar, Boulder, Colorado, 1960, published by the American Mathematical Society, Providence, Rhode Island, 1963.)

[5] M. Kac, Phys. Fluids 2, 8 (1959).

thermodynamic potential (Gibbs free energy) of the system. These results will be recapitulated in Sec. II.

For finite γ, that is for a finite range of the attractive force, the system does *not* show a phase transition in agreement with all previously known results[6] for one-dimensional systems. However, if one sets $\alpha = \alpha_0\gamma$, and then lets $\gamma \to 0$, (i.e., for a weak but very long-range force) so that the integral

$$\int_0^\infty dx\, \varphi_{\text{attr.}}(x)$$

is finite (we will call this the *van der Waals limit*), a phase transition appears which is described exactly by the van der Waals equation

$$p = kT/(l - \delta) - \alpha_0/l^2, \tag{1}$$

together with the well known Maxwell rule.[7] This will be shown in detail in Sec. III.

In Sec. IV, all the eigenfunctions and eigenvalues of the Kac integral equation are found in this limit by a perturbation procedure with $\gamma\delta$ as the expansion parameter. The phase transition appears as a double degeneracy of the maximum eigenvalue, and the phase transition *persists* when the perturbation calculation is carried further to any finite power of $\gamma\delta$. One should also note that this entire development breaks down near the critical point. For the critical region, a separate discussion is required which will be given in Part III of this series of papers.

In the last section, we will comment on the relevance of our results for the three-dimensional problem, and on the relation to the usual derivation of the van der Waals equation.

II. THE KAC INTEGRAL EQUATION

The partition function for our one-dimensional gas is given by

$$Z(L, T, N) = \frac{1}{\Lambda^N}\cdot\frac{1}{N!}\int_0^L \cdots \int_0^L dt_1 \cdots dt_N \times \exp\left[-\frac{1}{kT}\sum_{i<j}\varphi(|t_i - t_j|)\right]$$

$$= \frac{1}{\Lambda^N}\cdot\frac{1}{N!}\int_0^L \cdots \int_0^L dt_1 \cdots dt_N \times \exp\left[\nu \sum_{i<j} e^{-\gamma|t_i - t_j|}\right] \prod_{i<j} S(|t_i - t_j|), \tag{2}$$

with $\Lambda^2 = h^2/2\pi mkT$, $\nu = \alpha/kT$, and where the stepfunction $S(x)$ is defined by

$$S(x) = \begin{cases} 0 & \text{for } |x| < \delta \\ 1 & \text{for } |x| > \delta. \end{cases}$$

Since the integrand is symmetric in $t_1, t_2 \cdots t_N$, and because the hard cores impose a linear order of the molecules in L, one can write (2) in the form

$$Z(L, T, N) = \frac{e^{-\frac{1}{2}N\nu}}{\Lambda^N} \int\cdots\int_{0<t_1<t_2\cdots<t_N<L} dt_1 \cdots dt_N \times \exp\left[\frac{\nu}{2}\sum_{i=1}^N\sum_{j=1}^N e^{-\gamma|t_i - t_j|}\right]\cdot\prod_{i=1}^{N-1} S(|t_{i+1} - t_i|). \tag{3}$$

Also the attractive part of the integrand can be ordered by making use of the identity

$$\exp\left[\frac{\nu}{2}\sum_{i,j=1}^N \exp(-\gamma|t_i - t_j|)\right] = \int_{-\infty}^{+\infty}\cdots\int dx_1 \cdots dx_N \times \exp[\nu^{\frac{1}{2}}(x_1 + x_2 + \cdots + x_N)] \cdot W(x_1)\prod_{i=1}^N P(x_i \mid x_{i+1}, t_{i+1} - t_i), \tag{4}$$

where

$$W(x) = \frac{1}{(2\pi)^{\frac{1}{2}}}\exp[-\tfrac{1}{2}x^2],$$

$$P(x \mid y, t) = [\{2\pi(1 - e^{-2\gamma t})\}^{\frac{1}{2}}]^{-1} \times \exp[-(y - xe^{-\gamma t})^2/2(1 - e^{-2\gamma t})]. \tag{5}$$

The motivation and the proof of the identity (4) follows from the observation that $\exp(-\gamma|t_i - t_j|)$ is the covariance of a one-dimensional, Gaussian Markoff process (the so-called O–U process), but of course one can also prove (4) directly. By substituting (4) into (3) and by making a Laplace transformation in L, the integrals over the t_i can be separated in pairs since clearly

$$\int_0^\infty e^{-sL} Z(L, T, N)\, dL$$

is of the form

$$\int_0^\infty dL e^{-sL}\int_0^L dt_1 \int_{t_1}^L dt_2 \times \int_{t_2}^L dt_3 \cdots \int_{t_{N-1}}^L dt_N \prod_{i=1}^{N-1} F_i(t_{i+1} - t_i)$$

[6] That there is no phase transition for a one-dimensional system with only nearest-neighbor interaction, was shown first by F. Gursey, Proc. Cambridge Phil. Soc. 46, 182 (1950). This was generalized (with the same result) to the case where each molecule interacts with a finite number of neighbors by L. Van Hove, Physica 16, 137 (1950). Compare also the discussion given by A. Munster in his book *Statistische Thermodynamik* (Springer-Verlag, Berlin, Germany, 1956), Secs. 7.7 and 8.8, where one also finds further references.

[7] Also called the equal-area rule. See Maxwell, *Collected Works*, Dover reprint, Vol. II, p. 425.

$$= \int_0^\infty dt_1 \int_{t_1}^\infty dt_2 \cdots \int_{t_{N-1}}^\infty dt_N \times \int_{t_N}^\infty dL e^{-sL} \prod_{i=1}^{N-1} F_i(t_{i+1} - t_i). \quad (6)$$

Setting

$$t_1 = \tau_1,$$

$$t_2 = \tau_1 + \tau_2, \cdots, t_N = \tau_1 + \tau_2 + \cdots + \tau_N,$$

$$L = \tau_1 + \tau_2 + \cdots + \tau_N + \tau_{N+1},$$

one easily shows that (6) becomes

$$\frac{1}{s^2} \prod_{i=1}^{N-1} \int_0^\infty d\tau\, e^{-s\tau} F_i(\tau),$$

so that by putting in the appropriate form for the $F_i(z)$, one obtains (suppressing the temperature T from now on)

$$\int_0^\infty dL e^{-sL} Z(L, N) = \frac{1}{\Lambda^N} \frac{e^{-N\nu/2}}{s^2} \int_{-\infty}^{+\infty} \cdots \int dx_1 \cdots dx_N \times \exp[\nu^{\frac{1}{2}}(x_1 + \cdots + x_N)] W(x_1) \prod_{i=1}^{N-1} p_s(x_i \mid x_{i+1}) \quad (7)$$

with

$$p_s(x \mid y) = \int_0^\infty d\tau\, e^{-s\tau} P(x \mid y, \tau). \quad (8)$$

The ordering of the x_i in successive pairs suggests the introduction of the kernel

$$K_s(x, y) = \frac{W(x) p_s(x \mid y)}{[W(x)W(y)]^{\frac{1}{2}}} \exp[(\nu^{\frac{1}{2}}/2)(x + y)] \quad (9)$$

and the corresponding Kac integral equation

$$\int_{-\infty}^{+\infty} dy\, K_s(x, y)\psi(y) = \lambda\psi(x). \quad (10)$$

It is easy to see that $K_s(x, y)$ is symmetric and in addition one can show (for the proof, see Kac[5]):

(a) $K_s(x, y)$ is positive definite, which means that

$$\iint K_s(x, y)\psi(x)\psi(y)\, dx\, dy$$

is always positive, whenever $\psi(x)$ is not identically 0;

(b) $K_s(x, y)$ is a Hilbert–Schmidt kernel, which means that

$$\iint K_s^2(x, y)\, dx\, dy < \infty.$$

From these facts, one can conclude that Eq. (10) has a *discrete* set of *positive* eigenvalues $\lambda_i(s)$ starting from a maximum eigenvalue $\lambda_0(s)$ and converging to zero as $i \to \infty$, that the corresponding eigenfunctions $\psi_i(x, s)$ form a complete orthonormal set, and that the kernel $K_s(x, y)$ can be expanded in the convergent series

$$K_s(x, y) = \sum_i \lambda_i(s)\psi_i(x, s)\psi_i(y, s). \quad (11)$$

Writing Eq. (7) in the form

$$\int_0^\infty dL e^{-sL} Z(L, N) = \frac{e^{-\frac{1}{2}N\nu}}{\Lambda^N s^2} \int_{-\infty}^{+\infty} \cdots \int dx_1 \cdots dx_N \exp\left[\frac{\nu^{\frac{1}{2}}}{2}(x_1 + x_N)\right] \times [W(x_1)W(x_N)]^{\frac{1}{2}} \prod_{i=1}^{N} K_s(x_i, x_{i+1}),$$

and using Eq. (11), one can integrate over x_2, $x_3 \cdots x_{N-1}$ and obtain

$$\int_0^\infty dL e^{-sL} Z(L, N) = \frac{e^{-\frac{1}{2}N\nu}}{\Lambda^N s^2} \sum_{i=0}^{\infty} \lambda_i^{N-1}(s) A_i^2 \quad (12)$$

with

$$A_i = \int_{-\infty}^{+\infty} dx\, \psi_i(x, s) W^{\frac{1}{2}}(x) e^{\frac{1}{2}\nu x}. \quad (13)$$

If one now forms the grand partition function

$$G(L, z) = \sum_{N=1}^{\infty} Z(L, N)(\Lambda z)^N, \quad (14)$$

it follows from Eq. (12), that

$$\int_0^\infty dL e^{-sL} G(L, z) = \frac{z e^{-\frac{1}{2}\nu}}{s^2} \sum_i \frac{A_i^2}{1 - \lambda_i z e^{-\frac{1}{2}\nu}}, \quad (15)$$

provided $z < e^{\frac{1}{2}\nu}/\lambda_0(s)$. It is clear therefore that the abscissa of convergence of the Laplace transform of $G(L, z)$ is that value of s, for which

$$z = e^{\frac{1}{2}\nu}/\lambda_0(s). \quad (16)$$

Since on the other hand this abscissa is also

$$\lim_{L\to\infty} \frac{1}{L} \ln G(L, z), \quad (17)$$

which has the thermodynamic meaning of p/kT, it follows that the relation between z and the pressure of the gas is given by

$$\lambda_0(p/kT) = e^{\frac{1}{2}\nu}/z. \quad (18)$$

Furthermore, since the thermodynamic meaning of z is the fugacity, which is related to the chemical potential (Gibbs free energy per particle) $\mu(p, T)$ by

$$\mu = kT \ln (\Lambda z), \tag{19}$$

the equation of state follows from

$$l = (\partial\mu/\partial p)_T = -\lambda_0'(s)/\lambda_0(s), \quad s = p/kT, \tag{20}$$

where the prime denotes differentiation after s.

The following statements can be proved concerning the maximum eigenvalue $\lambda_0(s)$ and the eigenvalue spectrum $\lambda_i(s)$ (see Appendix 1):

(a) *All* eigenvalues are monotonically decreasing functions of s and approach zero for $s \to \infty$.

(b) The maximum eigenvalue $\lambda_0(s)$ goes to infinity for $s \to 0$. For no value of s can it be degenerate, and it is an analytic function of s for all real $s > 0$. Finally there is the inequality

$$\lambda_0'(s)^2 - \lambda_0(s)\lambda_0''(s) < 0, \tag{21}$$

which implies that $\lambda_0'(s)$ is always positive.

(c) All other eigenvalues are finite for $s = 0$.

It seems very likely that the curves $\lambda_i(s)$, $i > 0$ will also not cross each other, in which case the spectrum as a function of s may look as shown in Fig. 2.

One can conclude that for any value of z there is an *unique* value s fulfilling Eq. (16), and from Eqs. (20) and (21) it then follows that l is a monotonically decreasing and analytic function of $s = p/kT$ for all positive values of s. There is therefore *no* phase transition. Note that the theory gives directly [see Eqs. (18) and (19)], the Gibbs free energy as a function of the pressure, and that therefore one obtains the specific volume l as a function of the pressure instead of the pressure as a function of l as in the Mayer theory. The connection with the virial expansion of the equation of state is therefore quite complicated. One can show that the maximum eigenvalue $\lambda_0(s)$ can be expanded in the form

$$\lambda_0(s) = e^{\frac{1}{2}\nu}(s^{-1} + a_0 + a_1 s + a_2 s^2 + \cdots). \tag{22}$$

The coefficients a_i can be calculated successively and they are related to the Mayer cluster integrals b_l. One finds

$$a_0 = b_2; \qquad a_1 = b_3 - b_2^2; \tag{23}$$

and so on. For the proof we refer again to Appendix I. From Eqs. (22) and (23) one can then verify that the equation of state (20) is, for small pressures, simply the inversion of the usual virial expansion

$$pl/kT = 1 + B/l + C/l^2 + \cdots,$$

with $B = -b_2$, $C = -2b_3 + 4b_2^2$, etc. We have been unable to find a simple "graphological" characterization of the coefficients a_i.

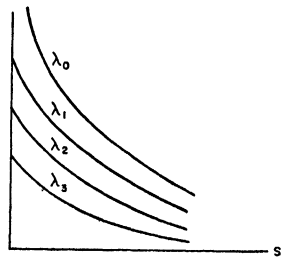


FIG. 2. Eigenvalue spectrum as a function of s.

III. THE VAN DER WAALS LIMIT

We will now study what happens to the maximum eigenvalue $\lambda_0(s)$, if one puts $\alpha = \alpha_0\gamma$ and then goes to the limit $\gamma \to 0$. To do this, we investigate the moments $\sum_i \lambda_i^n(s)$ of the distribution of the eigenvalues $\lambda_i(s)$. One has

$$\begin{aligned}\sum \lambda_i^n &= \int_{-\infty}^{+\infty}\cdots\int dx_1 \cdots dx_n \\ &\quad \times K_s(x_1, x_2)K_s(x_2, x_3) \cdots K_s(x_n, x_1) \\ &= \int_{-\infty}^{+\infty}\cdots\int dx_1 \cdots dx_n \\ &\quad \times \exp\left[-(\nu_0\gamma)^{\frac{1}{2}}(x_1 + x_2 \cdots + x_n)\right] \\ &\quad \times \int_0^{\infty}\cdots\int d\tau_1 \cdots d\tau_n \\ &\quad \times \exp\left[-s(\tau_1 + \cdots + \tau_n)\right] \prod_{i=1}^{n} P(x_i \mid x_{i+1}, \tau_i),\end{aligned}$$

where $\nu_0 = \alpha_0/kT$ and $x_{n+1} \equiv x_1$. The integrals over the x_i can be carried out and one can then investigate the limit $\gamma \to 0$. Some of the details are given in Appendix II. The result is the following theorem. For $\gamma \to 0$,

$$\lim_{\gamma\to 0} \gamma \sum_i \lambda_i^n = \frac{1}{2\pi}\iint_{-\infty}^{+\infty} d\xi\, d\eta \tag{24}$$
$$\times \left(\exp\left[\eta(2\nu_0)^{\frac{1}{2}}\right] \int_0^{\infty} d\tau \exp\left[-s\tau - \tfrac{1}{2}(\xi^2 + \eta^2)\tau\right]\right)^n.$$

One now can reason heuristically as follows. Call

$$f(\xi, \eta) = \exp\left[\eta(2\nu_0)^{\frac{1}{2}}\right] \int_0^{\infty} d\tau \times \exp\left[-s\tau - \tfrac{1}{2}(\xi^2 + \eta^2)\tau\right]; \tag{25}$$

then since (24) holds for all n, one can expect that for any reasonable function $g(x)$,

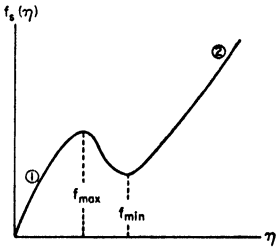


FIG. 3. The function $f_s(\eta)$ in the two-phase region.

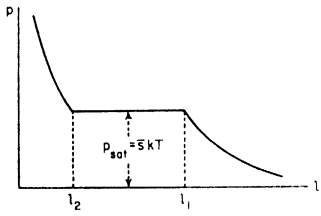


FIG. 5. Isotherm corresponding to Fig. 4.

$$\lim_{\gamma\to 0} \gamma \sum_i g(\lambda_i) = \frac{1}{2\pi} \iint_{-\infty}^{+\infty} d\xi\, d\eta\, g[f(\xi, \eta)]. \qquad (26)$$

Take now for $g(x)$ the step function

$$g(x) = \begin{cases} 1 & \text{for } \alpha < x < \beta \\ 0 & \text{otherwise,} \end{cases}$$

and let $N_\gamma(\alpha, \beta)$ be the number of eigenvalues between α and β; then one concludes from Eq. (26) that

$$\lim_{\gamma\to 0} \gamma N_\gamma(\alpha, \beta) = \frac{1}{2\pi} \times [\text{Area in } (\xi, \eta) \text{ plane for which } \alpha < f(\xi, \eta) < \beta]. \qquad (27)$$

Suppose that $f(\xi, \eta)$ has an *absolute* maximum $\omega(s)$, then it is clear that for $\gamma \to 0$, ω must be a limit point for the series of eigenvalues, because in any interval between $\omega - \epsilon$ and ω there must be in the limit $\gamma \to 0$, an infinite number of eigenvalues because the area $(\omega - \epsilon) < f(\xi, \eta) < \omega$ is finite. It seems therefore reasonable to expect that for $\gamma \to 0$, $\omega(s)$ is the maximum eigenvalue, or in other words that

$$\lim_{\gamma\to 0} \lambda_0(s, \gamma) = \omega(s) = \max_{(\xi,\eta)} f(\xi, \eta) = \max_{(\eta)} F(\eta), \qquad (28)$$

with

$$F(\eta) \equiv f(0, \eta) = \exp[-\delta(s + \tfrac{1}{2}\eta^2) - \ln(s + \tfrac{1}{2}\eta^2) + \eta(2\nu_0)^{\frac{1}{2}}],$$

since the maximum of $f(\xi, \eta)$ will always occur for $\xi = 0$.

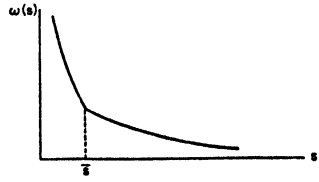


FIG. 4. The maximum eigenvalue, in the van der Waals limit, for a temperature below the critical.

Note that our argument from Eq. (27) does not really prove Eq. (28), since (27) does not exclude that a few discrete eigenvalues remain above $\omega(s)$ for $\gamma \to 0$. However, Eq. (28) is correct and the argument can be made completely rigorous. This is not done, because in the next section, the detailed investigation of the approach to the limit $\gamma \to 0$ implies a rigorous proof of Eq. (28).

To discuss Eq. (28) further, first note that the maximum will occur for that value $\eta(s)$ for which

$$f_s(\eta) \equiv \eta[\delta + 1/(s + \tfrac{1}{2}\eta^2)] = (2\nu_0)^{\frac{1}{2}}. \qquad (29)$$

Since $df_s/d\eta = 0$ implies

$$s + \tfrac{1}{2}\eta^2 = (1/2\delta)[1 \pm (1 - 8s\delta)^{\frac{1}{2}}], \qquad (30)$$

one must distinguish two cases.

(a) $8s\delta > 1$; $f_s(\eta)$ is then a monotonically increasing function of η, so that (29) has always one solution $\eta(s)$ corresponding to an unique maximum for $F(\eta)$. The equation of state follows from

$$l = -\omega'(s)/\omega(s) = +\delta + [1/(s + \tfrac{1}{2}\eta^2(s))] = +(2\nu_0)^{\frac{1}{2}}/\eta(s), \qquad (31)$$

using (29). Putting $s = p/kT$ and $\eta(s) = (2\nu_0)^{\frac{1}{2}}/l$ in Eq. (29), one obtains the van der Waals equation

$$p = kT/(l - \delta) - \alpha_0/l^2.$$

(b) $8s\delta < 1$; $f_s(\eta)$ has then a maximum and a minimum (see Fig. 3), and one easily verifies that for $s = 1/8\delta$, $f_{min} = f_{max} = (\frac{3}{2})(3\delta)^{\frac{1}{2}}$, and that for decreasing s, f_{min} and f_{max} increase monotonically; f_{min} remains finite, while f_{max} goes to infinity for $s \to 0$.

If now $(2\nu_0)^{\frac{1}{2}} < f_{min}$, then again there is only one solution of Eq. (29) corresponding to an unique maximum of $F(\eta)$. This is also the case if $(2\nu_0)^{\frac{1}{2}} > f_{max}$. However if $f_{min} < (2\nu_0)^{\frac{1}{2}} < f_{max}$, then Eq. (29) has *three* roots, the outer two corresponding to local maxima of $F(\eta)$ and the inner one to a local minimum. One must now decide which of the two maxima is the absolute maximum. Clearly if one varies ν_0 at fixed s, or if one varies s (keeping $8s\delta < 1$) at

fixed ν_0, there must be a value $\bar{s}(\nu_0)$ for which the two maxima are equal. It is at this value that $\omega(s)$ changes its analytic behavior abruptly, since $\eta(s)$ will suddenly jump from the branch ① to the branch ② of the $f_s(\eta)$ curve (see Fig. 3). In fact it is easy to show that for $s = \bar{s}$, $\omega(s)$ has a discontinuity in slope, and since $l = -\omega'(s)/\omega(s)$ this corresponds to an isotherm with a horizontal piece (see Figs. 4 and 5). Furthermore, since (31) remains valid, *both* the liquid and the vapor part are *still* described by the van der Waals equation. Finally, the equality of the maxima of $F(\eta)$ for $s = \bar{s}$ or of the two values of $\omega(s)$ means the equality of the Gibbs free energy of the two phases for $s = \bar{s}$, which in turn is equivalent to the Maxwell rule.

The critical point is determined by $8s_c\delta = 1$ and $(2\nu_c)^{\frac{1}{2}} = f_{\min}(s_c) = f_{\max}(s_c) = (\frac{3}{2})(3\delta)^{\frac{1}{2}}$, and this leads to the well-known formula

$$v_c = 3\delta; \quad p_c = \alpha_0/27\delta^2; \quad kT_c = 8\alpha_0/27\delta. \tag{32}$$

IV. THE EIGENFUNCTIONS AND EIGENVALUES FOR SMALL γ

Clearly the reason that, for small γ, the eigenvalues $\lambda_n(s, \gamma)$ of the Kac equation crowd together near the value $\omega(s)$, must be related to the fact that for $\gamma \to 0$, $P(x \mid y, t)$ approaches the Dirac δ function $\delta(x - y)$. However if one assumes that for $\gamma \to 0$, the eigenfunctions $\psi_n(x, s, \gamma)$ would remain centered around $x = 0$, as in the usual eigenvalue problems, then it follows from Eq. (10) that for $\gamma \to 0$, the eigenvalues would cluster around $e^{-s\delta}/s$ which is the maximum eigenvalue for the gas of hard rods. This contradiction can only be removed if, for $\gamma \to 0$, the eigenfunctions are centered farther and farther away from the origin. This in fact is the clue for constructing a consistent successive-approximation method for the eigenfunctions and eigenvalues if γ is small.

In the basic integral equation (10), substitute

$$x = x' + \eta(2/\gamma)^{\frac{1}{2}}, \qquad y = y' + \eta(2/\gamma)^{\frac{1}{2}}, \tag{33}$$

where $\eta \equiv \eta(s)$ is the value of η for which the function $F(\eta, s)$ defined by (28) has an absolute maximum. Think *first* of the *one-phase region*, so that $\eta(s)$ is unique. Let

$$h(x') \equiv \psi[x' + \eta(2/\gamma)^{\frac{1}{2}}]; \tag{34}$$

then Eq. (10) can be written in the form

$$\exp[\eta(2\nu_0)^{\frac{1}{2}}] \int_{-\infty}^{+\infty} dy' \int_\delta^\infty d\tau \left[\frac{W(x')}{W(y')}\right]^{\frac{1}{2}} P(x' \mid y', \tau) \times \exp\left[-s\tau - \frac{\eta^2}{\gamma}\tanh\left(\frac{\gamma\tau}{2}\right) + (x' + y')\frac{(\nu_0\gamma)^{\frac{1}{2}}}{2}\right.$$
$$\left. - \frac{\eta}{2}\left(\frac{2}{\gamma}\right)^{\frac{1}{2}}\tanh\left(\frac{\gamma\tau}{2}\right)(x' + y')\right] h(y') = \lambda h(x'). \tag{35}$$

Clearly now for $\gamma \to 0$, $P(x' \mid y', \tau) \to \delta(x' - y')$ if $h(x')$ is centered around $x' = 0$, the eigenvalues will approach

$$\exp[\eta(2\nu_0)^{\frac{1}{2}}] \int_\delta^\infty d\tau \exp\left[-s\tau - \frac{\eta^2\tau}{2}\right] = \omega(s),$$

as expected. One may say that by the substitution (33) one has "tamed" the integral equation for a perturbation expansion of the form

$$\lambda = \omega(s)\{1 + \mu^{(1)}\gamma + \mu^{(2)}\gamma^2 + \cdots\},$$
$$h(x') = h^{(0)}(x') + \gamma^{\frac{1}{2}}h^{(1)}(x') + \gamma h^{(2)}(x') + \cdots. \tag{36}$$

Since the algebra is involved, we simply indicate how this is done. In the left-hand side of Eq. (35), set

$$y' = x'e^{-\gamma\tau} + \zeta(1 - e^{-2\gamma\tau})^{\frac{1}{2}},$$

and develop the whole integrand in powers of $\gamma^{\frac{1}{2}}$. The integrals over ζ are Gaussian integrals and the integrals over τ can all be expressed in terms of $\omega(s)$ and its derivatives. Up to order $\gamma^{\frac{1}{2}}$, and except for a common factor $\omega(s)$, the left-hand side of (35) becomes

$$h(x') + \gamma l[h''(x') + h(x)(\tfrac{1}{2} - \tfrac{1}{4}x'^2 + \tfrac{1}{2}x'^2\nu_0(l - \delta)^2/l^3)]$$
$$+ [\gamma^{\frac{1}{2}}\nu_0^{\frac{1}{2}}(l - \delta)^2/l][\tfrac{1}{4}x'^3h(x')(1 - 4\nu_0(l - \delta)/l^2)$$
$$- \tfrac{1}{2}x'h(x') - h'(x') - x'h''(x')],$$

where (31) has been used to express $\eta(s)$ in terms of l. If one now introduces the expansions (36) on both sides of Eq. (35) and equates equal powers of $\gamma^{\frac{1}{2}}$, one obtains, in zeroth approximation,

$$\frac{d^2H^{(0)}}{dz^2} + \left[\frac{1}{B}\left(\frac{l}{2} - \mu^{(1)}\right) - \frac{z^2}{4}\right]H^{(0)}(z) = 0, \tag{37}$$

where we have set

$$B^2 = l^2 - 2\nu_0(l - \delta)^2/l, \qquad x' = z(l/B)^{\frac{1}{2}},$$
$$h^{(0)}(x') = H^{(0)}(z). \tag{38}$$

This is the equation for the Weber functions and hence one concludes that

$$\mu_n^{(1)} = \tfrac{1}{2}[l - (2n + 1)B],$$
$$H_n^{(0)}(z) = N_nD_n(z), \tag{39}$$

where N_n is the normalization factor $(B/2\pi l)^{\frac{1}{2}}(n!)^{-\frac{1}{2}}$. Note that the quantity B is related to the compressibility of the gas, since, from the van der Waals equation,

$$B^2 = [-l^2(l - \delta)^2/kT](\partial p/\partial l)_T. \tag{40}$$

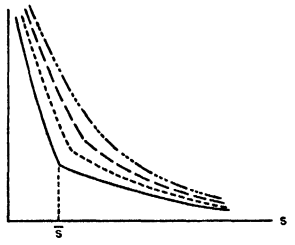


FIG. 6. ——— $\omega(s)$; —·—·— $\lambda(s,\gamma)$ (exact); - - - - - - - $\lambda(s,\gamma)$ to order $\gamma\delta$; — — — $\lambda(s,\gamma)$ to order $(\gamma\delta)^2$.

In the one-phase region $\partial p/\partial l$ is always negative, and therefore B can be taken positive. The maximum eigenvalue corresponds therefore to $n = 0$, and is given up to first order by

$$\lambda_0(s,\gamma) = \omega(s)[1 + \tfrac{1}{2}\gamma(l - B)]. \tag{41}$$

This proves the statement in Sec. III, that the maximum eigenvalue is $\omega(s)$ in the van der Waals limit.

In first approximation one obtains

$$\frac{d^2H_n^{(1)}}{dz^2} + (n + \tfrac{1}{2} - \tfrac{1}{4}z^2)H_n^{(1)} = \frac{\nu_0^{\frac{1}{2}}(l-\delta)^2}{l^{\frac{1}{2}}B^{\frac{1}{2}}}\frac{dH_n^{(0)}}{dz} + \frac{\nu_0^{\frac{1}{2}}(l-\delta)^2}{l^{\frac{1}{2}}B^{\frac{1}{2}}}\left(\frac{l}{2B} - n - \tfrac{1}{2}\right)zH_n^{(0)} - \frac{\nu_0^{\frac{1}{2}}(l-\delta)^3(l-3\delta)}{6l^{5/2}B^{5/2}}z^3H_n^{(0)}. \tag{42}$$

It can be easily verified that the right-hand side is orthogonal to $H_n^{(0)}$, so that (42) has an unique solution. For the explicit form of the solution and also for the second approximation, which will be needed in Part II of this series of papers, we refer to Appendix III.

Let us now consider the *two-phase region*.

It is clear that in this case the argument must be revised, since the $\eta(s)$ in the shift (33) is no longer unique. In fact, since $F(\eta, s)$ has now two equal maxima, there are *two* values of $\eta(s)$ corresponding to the volumes l_1 and l_2 of the saturated vapor and liquid phase (see Fig. 5). For each of these two values one can make the shift (33), and repeat the the argument. Clearly in both cases one will get the *same* zeroth-order eigenvalue $\omega(s)$, but one will get *another* eigenfunction, since the quantity l/B will be different at l_1 and l_2. One must say therefore that for $l_2 < l < l_1$, i.e. in the two-phase region, the maximum eigenvalue in the van der Waals limit is *doubly degenerate*, and in lowest order, the eigenfunction will have the form [since $D_0(z) = \exp(-\frac{1}{4}z^2)$],

$$\psi_0(x,s) = \alpha_1\left(\frac{B_1}{2\pi l_1}\right)^{\frac{1}{4}}\exp\left[-\frac{B_1}{4l_1}\left(x - \eta_1\left[\frac{2}{\gamma}\right]^{\frac{1}{2}}\right)^2\right] + \alpha_2\left(\frac{B_2}{2\pi l_2}\right)^{\frac{1}{4}}\exp\left[-\frac{B_2}{4l_2}\left(x - \eta_2\left[\frac{2}{\gamma}\right]^{\frac{1}{2}}\right)^2\right], \tag{43}$$

where $\eta_1(\bar{s}) = (2\nu_0)^{\frac{1}{2}}/l_1$ and $\eta_2(\bar{s}) = (2\nu_0)^{\frac{1}{2}}/l_2$.

The question now arises as to the appropriate linear combination corresponding to a specific length $l = \xi_1 l_1 + \xi_2 l_2$, (where ξ_1, ξ_2 are the mole fractions of the vapor and liquid phase), so that $\xi_1 + \xi_2 = 1$. We show that the answer is given by

$$\xi_1 = \alpha_1^2 \qquad \xi_2 = \alpha_2^2. \tag{44}$$

To prove this, note first that for small γ, *the two Gaussian functions in* (43) *do not overlap*. Hence from the normalization of $\psi(x,\bar{s})$, one obtains $\alpha_1^2 + \alpha_2^2 = 1$. Furthermore, for any finite γ, it follows from the Kac equation that

$$\lambda_0(s,\gamma) = \iint dx\,dy\,K_s(x,y) \times \psi_0(x,s,\gamma)\psi_0(y,s,\gamma), \qquad \lambda_0'(s,\gamma) = -\lambda_0(s,\gamma)l = \iint dx\,dy\,\frac{\partial K_s}{\partial s} \times \psi_0(x,s,\gamma)\psi_0(y,s,\gamma) \tag{45}$$

[since the differentiation of the eigenfunctions after s gives zero because the normalization of $\psi_0(x,s,\gamma)$ implies

$$\int \psi_0(x,s,\gamma)\frac{\partial}{\partial s}\psi_0(x,s,\gamma)\,dx = 0].$$

Now take $s = \bar{s}$, and let $\gamma \to 0$. Since for each of the two functions in Eq. (43) one has an equation like (45) with the same s and the same $\lambda_0(s) = \omega(s)$, but with l successively l_1 and l_2, and since one can again neglect the overlap of the two functions, it follows from Eq. (45) that for $\gamma \to 0$,

$$l = \alpha_1^2 l_1 + \alpha_2^2 l_2,$$

which proves Eq. (44).

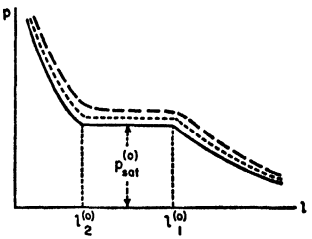


FIG. 7. ——— The van der Waals isotherm ($\gamma = 0$); — — — The exact isotherm; ····· The isotherm to order $\gamma\delta$.

We conclude this section with the following two remarks.

(a) It is clear that the entire development breaks down near the critical point because at that point, $B \to 0$ and $l_1 \cong l_2$. The overlap of the two eigenfunctions in the two-phase region can therefore no longer be neglected. In the critical region, one has to "tame" the integral equation in a different way, which will be explained in the third part of this series.

(b) It follows from Eq. (41) that up to first order in γ, $\lambda_0(s, \gamma)$ *still* shows a discontinuity in slope, and one can be convinced that the discontinuity *remains* if the perturbation calculation is continued up to *any* finite order in γ. For finite γ, $\lambda_0(s, \gamma)$ is an analytic function of s for $s > 0$. This means that for small γ, the successive-approximation method which we have developed, approximates the analytic function $\lambda_0(s, \gamma)$ by a series of functions which have a discontinuity in slope in the region of s, where the slope of $\lambda_0(s, \gamma)$ varies very quickly (see Fig. 6). This implies that *the phase transition persists* up to any finite order in γ, although the equation of state is then of course different from the van der Waals equation. In fact it is not difficult to show from Eq. (41), that, up to first order in γ, the equation of state in the one-phase region is given by

$$p = \frac{kT}{l - \delta} - \frac{\alpha_0}{l^2} + \frac{\gamma}{2}\left\{kT - \frac{l}{B}\left[kT - \frac{\alpha_0(l^2 - \delta^2)}{l^3}\right]\right\}.$$

The saturated vapor pressure and the densities of the saturated vapor and liquid also change accordingly. One finds

$$p_{\text{sat}} = p_{\text{sat}}^{(0)} + \gamma \frac{kT}{2}\left(\frac{B_1 - B_2}{l_1 - l_2} - 1\right)$$

$$l_i = l_i^{(0)} + \frac{\gamma}{2}(l_i - \delta)^2\left(\frac{l_i}{B}\right)^2 \times \left[\frac{B_1 - B_2}{l_1 - l_2} - \frac{l_i}{B_i}\left(1 - \frac{\alpha_0(l_i^2 - \delta^2)}{kTl_i^3}\right)\right]$$

with $i = 1, 2$.

V. CONCLUDING REMARKS

To judge the relevance of our results for the three-dimensional case, let us first recall the usual derivation[8] of the van der Waals equation from the partition function. One can argue as follows: If the range of the attractive force is very large so that there are many molecules in the action sphere of each molecule, then the potential energy of the attractive forces should for almost all configurations of the molecules, be equal to the average value

$$\Phi_{\text{attr}} = -CN^2/2V, \tag{46}$$

where

$$C = -\int d\mathbf{r}\, \varphi_{\text{attr}}(r).$$

Equation (46) follows because on the average, the potential energy of *one* molecule will be proportional to the number density N/V and is in fact equal to $-CN/V$. The factor two in (46) is of course needed because each molecule interacts with all other molecules. For long-range attractive forces it seems reasonable therefore to approximate the partition function by

$$Z(V, T, N) = \frac{1}{\Lambda^{3N}}\frac{1}{N!}\exp\left[CN^2/2kTV\right] \times \int_V \cdots \int_V d\mathbf{r}_1 \cdots d\mathbf{r}_N \prod_{i<j} S(|\mathbf{r}_i - \mathbf{r}_j|), \tag{47}$$

where $S(x)$ is the same step function as used in Sec. II. Now if V is very large compared to the total proper volume Nv_0, $v_0 = \frac{4}{3}\pi r_0^3$ of the molecules, then clearly the integral approaches V^N, while if V is of the same order as Nv_0 the integral will become strongly zero about as $(V - Nv_0)^N$. Approximation of the integral by the "interpolation" formula $(V - b)^N$ with $b \cong Nv_0$, then leads immediately to the van der Waals equation

$$(p + a/V^2)(V - b) = RT, \tag{48}$$

with $a = \frac{1}{2}CN^2$.

This derivation has been criticized mainly for two reasons:

1. A correct evaluation of the partition function in the thermodynamic limit should *always* give a *stable* isotherm for which $\partial p/\partial v \leq 0$. This has been proved rigorously by Van Hove[9] for an intermolecular force of finite range and with a hard core. Since there is an unstable part in the van der Waals equation below the critical temperature, the equation is in conflict with this theorem.

2. The approximation $(V - b)^N$ for the repulsive

[8] This derivation goes back to L. S. Ornstein, Dissertation Leiden 1908. Van der Waals himself, and then later also Lorentz [*Collected Papers* (Martinus Nyhoff, The Hague, Netherlands, 1935), Vol. 6, p. 40], started from the virial theorem of Clausius. We will come back to this in Part II in this series of papers.

[9] L. Van Hove, Physica 15, 951 (1949). Recently Dr. N. G. van Kampen has pointed out a gap in Van Hove's argument. Apparently the theorem can be proved even under less restrictive conditions by a different method. See a recent paper by D. Ruelle, which will appear in the Helv. Phys. Acta. (to be published).

part of the partition function must be very rough. Even if b is fixed so that the correct second virial coefficient (which makes $b = 4Nv_0$) is obtained, the higher virial coefficients which can be computed exactly, do *not* agree with the van der Waals' values.

Let us now return to our one-dimensional system. Note that by the same reasoning as used above, one would obtain

$$Z(L, T, N) = \frac{1}{\Lambda^N} \frac{1}{N!} \exp\left[\frac{\alpha_0 N^2}{kTL}\right] \times \int_L \cdots \int_L dt_1 \cdots dt_N \prod_{i<j} S(|t_i - t_j|), \quad (49)$$

since C becomes $2\alpha_0$, with the factor two because of the two sides of the attractive force $-\alpha_0\gamma \exp[-\gamma|t|]$. The *great difference* between Eqs. (47) and (49) is that the integral in Eq. (49) can be carried out *exactly* because of the linear order which the hard cores impose. The result is $(L - N\delta)^N$. For a gas of hard rods of length δ, the equation of state as noted by Tonks[10] has precisely the van der Waals form

$$p = kT/(l - \delta). \quad (50)$$

The second objection against the van der Waals equation disappears in one dimension, and it is therefore *no* wonder that for our model in the van der Waals limit, Eq. (48) with $b = N\delta$, $a = \alpha_0 N^2$ is obtained exactly. The only question which remains is the conflict of the form (49) for Z with the Van Hove theorem. From the point of view of the development in Secs. II and III, it is clear that this conflict is due to the fact that in the derivation given above, the thermodynamic and the van der Waals limits have been coupled together, while they should be taken separately and in the proper order. Since we go *first* to the thermodynamic limit for a finite range of the attractive force, and *then* allow the range to go to infinity, we do not conflict with the Van Hove theorem, and we always obtain the stable isotherm[11]). It seems to us therefore, that the first objection against the van der Waals equation is *not* as serious as we have been led to believe.

To summarize this discussion, we conclude with the following statements.

(a) It seems a good idea, following the example of van der Waals, to try to separate the effects of the repulsive and the attractive forces on the equation of state, and to develop all thermodynamic functions in the ratio

$$\rho = \frac{\text{range of repulsive force}}{\text{range of attractive force}}. \quad (51)$$

One expects that for $\rho \ll 1$ one will get a van der Waals-*like* equation of state which would lead to the condensation phenomena and the existence of a critical point. The thermodynamic equilibrium conditions will follow automatically by going first to the thermodynamic limit.

(b) *Only* in *one* dimension does one obtain in the limit $\rho \to 0$, *exactly* the van der Waals equation of state for hard-core repulsion. This should be independent of the precise form of the attractive force. We have been able to confirm this by generalizing the Kac model to the case where the attractive force is of the form

$$\varphi_{attr}(x) = -\sum_{i=1}^{m} \alpha_i\gamma \exp[-\gamma\sigma_i x]. \quad (52)$$

The calculation of the partition function can be reduced as in Sec. II to an integral equation in m variables and the behavior of the maximum eigenvalue in the limit $\gamma \to 0$ may be considered in the same way as in Sec. III: the *same* equation (28) is obtained, except that $2\nu_0$ is replaced by

$$\frac{1}{kT} \sum_{i=1}^{m} \frac{2\alpha_i}{\sigma_i} = -\frac{2}{kT} \int_0^\infty \varphi_{attr}(x)\, dx. \quad (53)$$

Thus, again the van der Waals equation is obtained and the van der Waals a is again proportional to the area of the attractive potential.

(c) In *three* dimensions, the equation of state of a gas of hard spheres is certainly *not* $p = kT(v - b)$, although qualitatively the behavior may be similar. Even for very long-range attractive forces one can only hope to get a van der Waals-*like* equation of state. In a sense, the problem of the gas of hard spheres becomes the central problem. Whether such a gas would show a phase transition for densities near close packing (the so-called Kirkwood transition), is a famous open question. We believe that there are strong indications for the existence of such a transition and that this transition might be an idealization (or caricature!) of the fluid-solid transition. Clearly, in one dimension [see Eq. (50)], such a transition does *not* exist, and we believe therefore that even with long-range forces, to enforce cooperation, *one phase transition at most* can occur in one dimension.

Finally, we would like to mention two questions

[10] L. Tonks, Phys. Rev. **50**, 955 (1936).

[11] For our model, the Van Hove theorem is expressed by the inequality (21) for the maximum eigenvalue $\lambda_0(s, \gamma)$, which holds for any value of γ. The isotherm is monotonically decreasing for any γ, and therefore also in the limit $\gamma = 0$, $\partial p/\partial l \leq 0$.

which we have so far been unable to answer in a satisfactory way.

1. What is the relation of our theory with the Yang–Lee theory of condensation? From the existence of a hard core, it follows that the grand partition function $G(L, z)$ must be a polynomial of degree $M = L/\delta$ in z with positive coefficients. Setting

$$G(L, z) = \exp [L\chi(L, z)],$$

one can write

$$\chi(L, z) = \frac{1}{L} \sum_{i=1}^{M} \ln \left(1 - \frac{z}{z_i}\right),$$

where the z_i are the zero's of $G(L, z)$, none of which can be on the positive real axis. One can interpret $\chi(L, z)$ as the complex logarithmic potential of M point charges of strength $1/L$ situated at the points z_i. In the limit $L \to \infty$, the strength of the charges becomes smaller and smaller while the number M increases. Suppose now that in the limit $L \to \infty$, a number of the charges concentrate in a single layer *which crosses the positive real axis*, say at $z = z_0$. The limit function

$$\bar{\chi}(z) = \lim_{L\to\infty} \chi(L, z),$$

which surely exists, will then for positive and real z, consist of two analytic pieces, one for $z < z_0$ and the other for $z > z_0$. At z_0 these two pieces will be continuous but the first derivative will have a discontinuity. Since $\bar{\chi}(z) = p/kT$ one would therefore obtain a curve as in Fig. 4 (with abscissa and ordinates interchanged), and condensation would occur. It is clear that something like this must happen for our model in the van der Waals limit. It must be that for finite γ, the single layer will *not* cross the positive real axis but leave a gap, say, of order γ. For finite γ, $\bar{\chi}(z)$ will then be analytic for all real and positive z, but in the van der Waals limit when the gap closes, condensation occurs. However, we have been unable to substantiate this picture, because it is difficult to discuss the behavior of the eigenvalues of the Kac equation for complex values of s.

2. What is the relation of our theory with the Ursell–Mayer theory? We have already mentioned the problem of the "graphological" interpretation of the coefficients a_l in the expansion of $\lambda_0(s)$ [see Eq. (22)], but there are also other questions. For instance it should be possible to introduce Mayer f functions both for the repulsive and for the attractive parts of the intermolecular potential, which would lead to graphs with two types of lines. The question is then, which simplifications occur if the attractive force is long range, and whether one can learn from the one-dimensional model how to characterize better the van der Waals-like equation in three dimensions.

APPENDIX 1

For the proof of statement (a) see Kac (reference 5, p. 11). There one also finds the proof that

$$\lambda_0(s) \geq e^{\frac{1}{2}\nu} \int_\delta^\infty d\tau \exp [-s\tau + \nu e^{-\gamma\tau}], \qquad (A.1)$$

and since for $s = 0$ the integral diverges, $\lambda_0(s) \to \infty$ for $s \to 0$. Furthermore, since the right hand side of (A.1) can be shown to correspond to the "nearest-neighbor" approximation, one should expect that for small s (i.e. small pressure), $\lambda_0(s)$ becomes equal to it, which implies that for $s \to 0$, $\lambda_0(s) \sim e^{\frac{1}{2}\nu}/s$, and this will be confirmed later when we derive the expansion (22).

The statement that $\lambda_0(s)$ cannot be degenerate for any value of s, follows from the fact that the kernel $K_s(x, y)$ is positive for all x, y and for $s > 0$. Since

$$\lambda_0(s) \geq \frac{\iint dx\, dy\, \varphi(x) K_s(x, y) \varphi(y)}{\int dx\, \varphi^2(x)} \qquad (A.2)$$

for any function $\varphi(x)$, and since the maximum is reached when $\varphi(x) = \psi_0(x)$, $K_s(x, y) > 0$ implies that $\psi_0(x)$ must have the same sign for all x, and we may assume therefore $\psi_0(x) \geq 0$. Furthermore, $\psi_0(x)$ cannot be zero, unless for some particular value of x, say x_0, $K_s(x_0, y) \equiv 0$ in y, which is excluded by $K_s > 0$. Now suppose that for some value of s, $\lambda_0(s)$ was degenerate and corresponded to the two eigenfunctions $\psi_0^{(1)}(x)$ and $\psi_0^{(2)}(x)$. Taking $\psi_0^{(1)}(x) > 0$, one then can construct a linear combination $\varphi(x)$ of $\psi_0^{(1)}$ and $\psi_0^{(2)}$ which is *orthogonal* to $\psi_0^{(1)}$, and which fulfills the equation

$$\lambda_0(s) = \frac{\iint dx\, dy\, \varphi(x) K_s(x, y) \varphi(y)}{\int dx\, \varphi^2(x)}. \qquad (A.3)$$

Since $\psi_0^{(1)}(x)$ cannot be zero, $\varphi(x)$, being orthogonal to $\psi_0^{(1)}$, must be both positive and negative over some intervals, and this yields a contradiction. The absence of a degeneracy of $\lambda_0(s)$ implies that $\lambda_0(s)$ is an analytic function of s for all real $s > 0$. This is because it is known that the eigenvalues $\lambda_l(s)$ are the zero's in λ of the Fredholm determinant

$D(x, \lambda)$, which is an entire function of λ, and is analytic for all real $s > 0$. Therefore the $\lambda_\epsilon(s)$ must be analytic functions of s except at those values of s where two or more eigenvalues cross.

Finally the inequality (21) is proved as follows: With normalized eigenfunctions, one finds from

$$\lambda_0(s) = \iint dx\, dy\ \psi_0(x, s) K_s(x, y) \psi_0(y, s),$$

that the quadratic form

$$\begin{aligned} Q(z) &\equiv \lambda_0(s)z^2 + 2\lambda_0'(s)z + \lambda_0''(s) \\ &= \iint_{-\infty}^{+\infty} dx\, dy\ \psi_0(x, s)\psi_0(y, s) \\ &\quad \times \left[z^2 K_s(x, y) + 2z \frac{\partial K_s}{\partial s} + \frac{\partial^2 K_s}{\partial s^2}\right] \\ &\quad + \iint_{-\infty}^{+\infty} dx\, dy \frac{\partial K_s}{\partial s}\left[\psi_0(x, s) \frac{\partial \psi_0(y, s)}{\partial s}\right. \\ &\quad \left. + \psi_0(y, s) \frac{\partial \psi_0(x, s)}{\partial s}\right]. \end{aligned} \tag{A.4}$$

From Eq. (9) for $K_s(x, y)$, it can be seen that the variable s enters only through $p_s(x, y)$, and from Eq. (8) it follows that the first term in Eq. (4) is positive for all z. Also the second term in Eq. (4) is positive, since by using

$$K_s(x, y) = \sum_0^\infty \lambda_n(s)\psi_n(x, s)\psi_n(y, s),$$

and the fact that for any s the $\psi_n(x, s)$ form an orthonormal set, it can be transformed into

$$\begin{aligned} 2\lambda_0(s) &\int_{-\infty}^{+\infty} dx \left[\frac{\partial \psi_0(x, s)}{\partial s}\right]^2 \\ &- 2\sum_{n=1}^\infty \lambda_n(s)\left[\int_{-\infty}^{+\infty} dx \frac{\partial \psi_0(x, s)}{\partial s} \psi_n(x, s)\right]^2, \end{aligned}$$

and since $\lambda_n(s) < \lambda_0(s)$ for $n \geq 1$, this is bigger than

$$2\lambda_0(s)\left[\int_{-\infty}^{+\infty} dx \left(\frac{\partial \psi_0}{\partial s}\right)^2 - \sum_0^\infty \left(\int_{-\infty}^{+\infty} dx\, \psi_n \frac{\partial \psi_0}{\partial s}\right)^2\right],$$

which is zero because of Parseval's relation applied to the function $\partial\psi_0/\partial s$. Hence $Q(z) > 0$ for all z, which implies Eq. (21).

Statement (c), Sec. II follows from the relation

$$\begin{aligned} \sum_0^\infty \lambda_i &= \int_{-\infty}^{+\infty} dx\, K_s(x, x) \\ &= \int_s^\infty \frac{d\tau}{1 - e^{-\gamma\tau}} \exp\left[-s\tau + \frac{\nu}{2}\frac{1 + e^{-\gamma\tau}}{1 - e^{-\gamma\tau}}\right], \end{aligned}$$

which shows that for $s \to 0$, the sum of all the eigenvalues goes to infinity as $e^{\frac{1}{2}\nu}/s$, and since this is also the behavior of $\lambda_0(s)$, all other eigenvalues must be finite for $s = 0$.

To find the expansion (22) for $\lambda_0(s)$, one must make a perturbation calculation around $s = 0$ for the eigenvalue problem

$$\int_{-\infty}^{+\infty} dy\ K_s(x, y)\psi_0(y, s) = \lambda_0(s)\psi_0(x, s). \tag{A.5}$$

First we expand $K_s(x, y)$ in powers of s. To do this we write $K_s(x, y)$ in the form

$$\begin{aligned} K_s(x, y) &= \exp\left[\tfrac{1}{2}\nu^{\frac{1}{2}}(x + y)\right]\left\{\frac{e^{-s\delta}}{s}\left[W(x)W(y)\right]^{\frac{1}{4}}\right. \\ &\quad + \sum_{n=0}^\infty \frac{(-s)^n}{n!}\int_s^\infty d\tau\, \tau^n\left(\left[\frac{W(x)}{W(y)}\right]^{\frac{1}{4}} P(x \mid y, \tau)\right. \\ &\quad \left.\left. - \left[W(x)W(y)\right]^{\frac{1}{4}}\right)\right\}. \end{aligned} \tag{A.6}$$

Now put

$$\begin{aligned} \lambda_0(s) &= (e^{\frac{1}{2}\nu}/s)(a_{-1} + a_0 s + a_1 s^2 + \cdots), \\ \psi_0(x, s) &= \psi^{(0)}(x) + s\psi^{(1)}(x) + \cdots, \end{aligned} \tag{A.7}$$

and let $\psi^{(0)}(x)$ be normalized to one. Introducing (6) and (7) into (5) and equating the terms proportional to $1/s$, one obtains

$$\begin{aligned} &\int_{-\infty}^{+\infty} dy\ [W(x)W(y)]^{\frac{1}{4}} \\ &\quad \times \exp\left[\tfrac{1}{2}\nu^{\frac{1}{2}}(x + y)\right]\psi^{(0)}(y) = a_{-1}\psi^{(0)}(x)e^{\frac{1}{2}\nu}. \end{aligned} \tag{A.8}$$

One easily verifies that this has the solution

$$a_{-1} = 1; \quad \psi^{(0)}(x) = (W(x))^{\frac{1}{4}} \exp\left[-\tfrac{1}{4}\nu + \tfrac{1}{2}\nu^{\frac{1}{2}}x\right]. \tag{A.9}$$

Notice that if we try to determine in the same way the values of the other eigenvalues $\lambda_n(s)$ at $s = 0$, we would find in this approximation that $a_{-1}^{(n)} = 0$, so that all $\lambda_n(0)$ for $n > 0$ are degenerate, and to find the actual values, we would have to solve an infinite secular determinant. *Only* $\lambda_0(s)$ becomes infinite for $s \to 0$, and only for $\lambda_0(s)$ can we carry the perturbation calculation further.

Equating in Eq. (5) the terms independent of s, one finds

$$\begin{aligned} &\int_{-\infty}^{+\infty} dy \exp\left[\tfrac{1}{2}\nu^{\frac{1}{2}}(x + y)\right]\left\{\left(-\delta[W(x)W(y)]^{\frac{1}{4}}\right.\right. \\ &\quad + \int_s^\infty d\tau \left[\frac{W(x)}{W(y)}\right]^{\frac{1}{4}} \\ &\quad \times [P(x \mid y, \tau) - W(y)]\Big)\psi^{(0)}(y) \\ &\quad \left. + [W(x)W(y)]^{\frac{1}{4}}\psi^{(1)}(y)\right\} \\ &\quad = e^{\frac{1}{2}\nu}[\psi^{(1)}(x) + a_0\psi^{(0)}(x)]. \end{aligned} \tag{A.10}$$

Multiply with $\psi^{(0)}(x)$ and integrate over x. Using Eq. (9), the fact that $\psi^{(1)}(x)$ is orthogonal to $\psi^{(0)}(x)$, and the formula

$$\iint_{-\infty}^{+\infty} dx\,dy\,\exp[\nu^{\frac{1}{2}}(x+y)]W(x)P(x\mid y,\tau) = \exp[\nu(1+e^{-\gamma\tau})],$$

which is a special case of Eq. (4), one obtains

$$a_0 = -\delta + \int_\delta^\infty d\tau\,\{\exp(\nu e^{-\gamma\tau} - 1)\}. \qquad (A.11)$$

The right-hand side can be written as

$$\lim_{L\to\infty}\frac{1}{2L}\iint_L dt_1\,dt_2\left\{\exp\left[\frac{1}{kT}\varphi(|t_1-t_2|)\right]-1\right\},$$

which is the definition of Mayer's b_2. One also easily verifies from Eq. (10), that

$$\psi^{(1)}(x) = -(a_0+\delta)\psi^{(0)}(x) + e^{-\frac{1}{2}\nu}\int_{-\infty}^{+\infty} dy\,\psi^{(0)}(y)\exp\left[\frac{\nu^{\frac{1}{2}}}{2}(x+y)\right] \times \int_\delta^\infty d\tau\left[\frac{W(x)}{W(y)}\right]^{\frac{1}{2}}[P(x\mid y,\tau)-W(y)].$$

We can go on in this way, but the calculation soon becomes very lengthy. We have verified that in next approximation, one obtains $a_1 = b_3 - b_2^2$, where Mayer's b_3 is defined by

$$b_3 = \lim_{L\to\infty}\frac{1}{6L}\iiint_L dt_1\,dt_2\,dt_3 \times [f_{12}f_{13}+f_{12}f_{23}+f_{13}f_{23}+f_{12}f_{13}f_{23}]$$

with

$$f_{ij} = \exp\left[\frac{1}{kT}\varphi(|t_i-t_j|)\right]-1.$$

The general expression for the a_i in terms of the Mayer b_l is lacking. To obtain the correct equation of state up to the fourth virial coefficient, a_2 should be equal to $b_4 - 3b_2b_3 + 2b_2^3$, but we have not verified this.

APPENDIX II

To indicate the proof of Eq. (24), it will be sufficient to consider the case $n = 3$. One has

$$\sum_i \lambda_i^3 = \iiint_\delta^\infty d\tau_1\,d\tau_2\,d\tau_3\,e^{-\delta(\tau_1+\tau_2+\tau_3)} \times \iiint_{-\infty}^{+\infty} dx_1\,dx_2\,dx_3\exp[-(\nu_0\gamma)^{\frac{1}{2}}(x_1+x_2+x_3)] \times P(x_1\mid x_2,\tau_1)P(x_2\mid x_3,\tau_2)P(x_3\mid x_1,\tau_3),$$

where $P(x\mid y,\tau)$ is given by Eq. (5). Introduce the new variables

$$\xi_1 = (x_2 - x_1e^{-\gamma\tau_1})(1-e^{-2\gamma\tau_1})^{-\frac{1}{2}}$$
$$\xi_2 = (x_3 - x_2e^{-\gamma\tau_2})(1-e^{-2\gamma\tau_2})^{-\frac{1}{2}}$$
$$\xi_3 = (x_1 - x_3e^{-\gamma\tau_3})(1-e^{-2\gamma\tau_3})^{-\frac{1}{2}},$$

with the Jacobian

$$\frac{\partial(\xi_1,\xi_2,\xi_3)}{\partial(x_1,x_2,x_3)} = \{1-e^{-\gamma(\tau_1+\tau_2+\tau_3)}\} \times \prod_{i=1}^3(1-\exp[-2\gamma\tau_i])^{-\frac{1}{2}}$$

The integral over the x_i then becomes

$$\{(2\pi)^{\frac{1}{2}}[1-e^{-\gamma(\tau_1+\tau_2+\tau_3)}]\}^{-1} \times \iiint_{-\infty}^{+\infty} d\xi_1\,d\xi_2\,d\xi_3\exp[-\tfrac{1}{2}(\xi_1^2+\xi_2^2+\xi_3^2) \times \exp\left[\frac{(\nu_0\gamma)^{\frac{1}{2}}}{e^{-\gamma(\tau_1+\tau_2+\tau_3)}-1}\sum_{i=1}^3 A_i\xi_i\right],$$

with

$$A_i = (1-e^{-2\gamma\tau_i})^{\frac{1}{2}}[e^{-\gamma(\tau_{i+1}+\tau_{i+2})}+e^{-\gamma\tau_{i+1}}+1],$$

and $\tau_4 \equiv \tau_1$, $\tau_5 \equiv \tau_2$.

In this form it is convenient to go to the limit $\gamma \to 0$. Since the A_i approach $3(2\gamma\tau_i)^{\frac{1}{2}}$, the integrand of the ξ_i integral becomes independent of γ, and one obtains

$$\lim_{\gamma\to0}\gamma\sum_i\lambda_i^3 = \iiint_\delta^\infty d\tau_1\,d\tau_2\,d\tau_3\,e^{-\delta(\tau_1+\tau_2+\tau_3)} \times [(2\pi)^{\frac{1}{2}}(\tau_1+\tau_2+\tau_3)]^{-1}\iiint_{-\infty}^{+\infty} d\xi_1\,d\xi_2\,d\xi_3 \times \prod_{i=1}^3\exp\left[-\tfrac{1}{2}\xi_i^2 - \frac{3(2\nu_0\tau_i)^{\frac{1}{2}}\xi_i}{\tau_1+\tau_2+\tau_3}\right]. \qquad (a)$$

To uncouple the τ_i integrals, introduce the auxiliary variable

$$\eta = \frac{1}{\tau_1+\tau_2+\tau_3}\sum_{i=1}^3\xi_i\tau_i^{\frac{1}{2}},$$

and multiply (a) with

$$1 = \int_{-\infty}^{+\infty} d\eta\,\delta\left(\eta - \frac{1}{\tau_1+\tau_2+\tau_3}\sum_{i=1}^3\xi_i\tau_i^{\frac{1}{2}}\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\eta\int_{-\infty}^{+\infty} dv \times \exp\left[iv\left(\eta - \frac{1}{\tau_1+\tau_2+\tau_3}\sum_{i=1}^3\xi_i\tau_i^{\frac{1}{2}}\right)\right],$$

using the Fourier representation of the δ function. Set $v = w(\tau_1 + \tau_2 + \tau_3)$, and integrate first over the ξ_i. We then obtain

$$\lim_{\gamma\to 0} \gamma \sum_i \lambda_i^3 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\eta \int_{-\infty}^{+\infty} dw \times \left\{\int_0^\infty d\tau \exp\left[-\eta(2\nu_0)^{\frac{1}{2}} + i\tau\eta w - \frac{\tau}{2}w^2\right]\right\}^3.$$

Finally, setting $w = \xi + i\eta$, one obtains Eq. (24) for $n = 3$. The proof for arbitrary n follows the same procedure.

APPENDIX III

The solution of Eq. (42) can be written in the form

$$H_n^{(1)}(z) = N_n[P_n D_{n+3}(z) + Q_n D_{n+1}(z) + R_n D_{n-1}(z) + S_n D_{n-3}(z)],$$

where

$$P_n \equiv P = \frac{\nu_0^{\frac{1}{2}}(l-\delta)^3(l-3\delta)}{18 l^{5/2} B^{5/2}},$$

$$Q_n = 9(n+1)P + (n+1)V - W,$$

$$R_n = -9n^2 P - n^2 V + nW,$$

$$S_n = -n(n-1)(n-2)P,$$

with

$$V = \nu_0^{\frac{1}{2}}(l-\delta)^3/l^{\frac{1}{2}}B^{\frac{1}{2}}, \qquad W = \nu_0^{\frac{1}{2}}(l-\delta)^2/2l^{\frac{1}{2}}B^{\frac{1}{2}}.$$

To obtain the equation for the next approximation, the left-hand side of Eq. (35) must be expanded to terms of order γ^2. Using the expansions (36), and equating terms of order γ^2, one obtains

$$\frac{d^2H_n^{(2)}}{dz^2} + \left(n + \tfrac{1}{2} - \frac{z^2}{4}\right)H_n^{(2)} = \frac{\mu_n^{(2)}}{B} H_n^{(0)} + \frac{\nu_0^{\frac{1}{2}}(l-\delta)^2 l^{\frac{1}{2}}}{4B^{5/2}}\left(\frac{4\nu_0}{3l^2} - 1\right)z^3 H_n^{(1)} + \frac{\nu_0^{\frac{1}{2}}(l-\delta)^2}{l^{\frac{1}{2}}B^{\frac{1}{2}}} z \frac{d^2H_n^{(1)}}{dz^2} + \frac{\nu_0^{\frac{1}{2}}(l-\delta)^2}{l^{\frac{1}{2}}B^{\frac{1}{2}}} \times \left\{\tfrac{1}{2}zH_n^{(1)} + \frac{B}{l}\frac{dH_n^{(1)}}{dz}\right\} + H_n^{(0)}(z)\left[\frac{\nu_0^2(l-\delta)^4(l^2-2l\delta-\delta^2)}{8l^4B^3} z^4 + \frac{\nu_0(l+\delta)(l-\delta)^3}{2l^3B}\left(n + \tfrac{1}{2} - \frac{l}{2B}\right)z^2 - \frac{\nu_0}{12l^2B}(15l^3 - 36l^2\delta + 30l\delta^2 - 8\delta^3) + \frac{1}{2B}(2l^2 - 2l\delta + \delta^2) \times \left\{\tfrac{1}{2} + (n+\tfrac{1}{2})\frac{B}{l} - (n^2+n+\tfrac{3}{4})\frac{B^2}{l^2}\right\}\right] - \frac{\nu_0(l-\delta)^3(l+\delta)}{l^3B} z \frac{dH_n^{(0)}}{dz},$$

from which $\mu_n^{(2)}$ and then $H_n^{(2)}$ can be determined in the usual way.

Reprinted from *Printed in U.S.A.*
JOURNAL OF MATHEMATICAL PHYSICS VOLUME 4, NUMBER 2 FEBRUARY 1963

On the van der Waals Theory of the Vapor–Liquid Equilibrium. II. Discussion of the Distribution Functions

G. E. UHLENBECK, P. C. HEMMER, AND M. KAC
The Rockefeller Institute, New York, New York
(Received 17 September 1962)

For the same one-dimensional fluid model discussed in Part I, we have derived general expressions for the two- and three-particle distribution functions. It is seen that these distribution functions depend on all the eigenvalues and eigenfunctions of the basic Kac integral equation, and the dependence is so transparent that the generalization to s particles is obvious. The fluctuation and virial theorems are discussed and shown to be consequences of our general formula. In the van der Waals limit, the behavior of the two-point distribution function is discussed, both for distances of the order of the hard core and for distances of the order of the range of the attractive force. The long-range behavior is, in first approximation, equivalent to the one-dimensional version of the Ornstein–Zernike theory, but only in the one-phase region and not too near the critical point. In the two-phase region, all distribution functions are linear combinations of the two corresponding distribution functions of the saturated vapor and liquid, with coefficients proportional to the mole fractions of vapor and liquid. This is shown for our model; we also give arguments for our belief that these relations are general, and express the geometrical separation of the two phases. The relation to the Ornstein–Zernike theory is discussed in more detail, especially in connection with a recent formulation of this theory by Lebowitz and Percus. We conclude with some comments on the relevance of our results for the three-dimensional problem.

Reprinted from *Printed in U.S.A.*
JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 1 JANUARY 1964

On the van der Waals Theory of the Vapor–Liquid Equilibrium. III. Discussion of the Critical Region

P. C. HEMMER,* M. KAC, AND G. E. UHLENBECK
The Rockefeller Institute, New York, New York
(Received 20 September 1963)

The discussion of the properties of the Kac one-dimensional fluid model presented in Parts I and II of this series of papers breaks down near the critical point. In Sec. II of the present paper we develop a new successive-approximation method for the eigenvalues and eigenfunctions of the Kac integral equation which is valid in the critical region and which connects smoothly with the developments in the one- and two-phase regions given in Part I. The perturbation parameter is $(\gamma\delta)^{\frac{1}{2}}$ where $\gamma\delta$ is the ratio of the ranges of the repulsive and attractive forces. The main physical consequence is that in the critical region the long-range behavior of the two-point distribution function is represented by an infinite series of decreasing exponentials with ranges all of order $1/\gamma(\gamma\delta)^{\frac{1}{2}}$, and with amplitudes of order $(\gamma\delta)^{\frac{1}{2}}$. This leads to deviations from the Ornstein–Zernike theory and to a specific heat anomaly which are discussed in Sec. V. We conclude with some comments on the possible relevance of our results for the three-dimensional problem.

COMMENT ON PAPERS (3.3)-(3.5)

The work of Kac, Uhlenbeck and Hemmer has been extended recently by Lebowitz and Penrose (to be published) to higher dimensions and to more general interparticle potentials of the form $v(r)=q(r)+\gamma^{\nu}\varphi(\gamma r)$, where ν is the dimensionality of the space. In this work, which is also related to recent work by van Kampan (Phys. Rev. 135, 362 (1964)), it is proven rigorously for a wide class of q's and φ's that in the 'van der Waals limit' $\gamma\to 0$, the equation of state assumes the form

$$\lim_{\gamma\to 0} p(\rho,\gamma)=p^{o}(\rho)+\frac{\alpha}{2}\rho^{2};\ \underline{\text{plus Maxwell's rule}}$$

where $p^{o}(\rho)$ is the pressure in the reference system, $\varphi(r)=0$, and $\alpha=\int\varphi(y)dy$.

STATISTICAL MECHANICS OF SYSTEMS WITH LONG-RANGE INTERACTIONS

by

Eugene Helfand
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

1. INTRODUCTION

Interest in systems with long-range forces is stimulated by the success of even simple approximate treatments in exhibiting a phase transition. In recent years much effort has been devoted to the development of systematic and precise methods for understanding such systems and hopefully from this phase transitions in general. These efforts have proceeded in two directions. The first, our major interest here, is the specialization to a type of model which is particularly amenable to analysis. This is a model with exponential interactions, introduced by Kac,[1] and studied in depth by Kac, Uhlenbeck, and Hemmer (KUH),[2,3,4] Baker,[5] and Kac and Helfand.[6] Other efforts have concentrated on treating more general interactions by developing series in powers of the inverse range of interaction.[7-15] This work will be surveyed more briefly.

As we shall see the long-range force model has some rather unusual properties. The paramount question of whether these findings are general or artifacts of the model is still largely unsettled. Therefore it is hoped that this report will serve as an introduction and stimulant for further work in the field.

There is some gain in clarity and ease of exposition if we discuss not the continuum model of a one-dimensional system of hard rods with exponential interactions[1-4] but rather the analogous lattice model.[1,5,6]

First, it is of interest to develop in general the basic functional integral technique, which is done in Section 2. Then the specific exponential model, which leads to an essential simplification, is introduced. Calculations of the thermodynamic properties, paralleling those in KUH, are performed. In particular the manner in which the second order phase transition is manifest as one decreases the temperature at the critical density is discussed. The first order phase transition is also analyzed with special emphasis on the role of the long-range nature of the potential in bringing it about. In one-dimension it is shown that the phase transition is approached in an essentially singular fashion as the range of interaction becomes infinite.[16]

The lines along which the critical properties of this model must be analyzed are somewhat novel and worthy of study. It is suggested that there is a critical region inside which the qualitative nature of the dependence of the properties of the system on temperature and density differ from the behavior outside.

The pair correlation function is a sensitive indicator of changes which occur in the system. It is studied in detail for the Kac exponential model in Section 5.

For the one-dimensional model the long-range force plays an essential a role in bringing about the transition. A preliminary survey of the attempts to remove this difficulty by considering certain two-dimensional systems is presented in Section 6.

Finally, two more general approaches to the problem of systems with long-range forces are examined. One is a diagrammatic analysis and the other stems from the functional integration formalism.

2. BASIC FORMALISM

Let us begin with a discussion of a model ferromagnet. The system is composed of a lattice of N sites each of which is occupied by an up or down spin denoted by $\mu_i = \pm 1$, respectively (i may be a vector). Let the pair interaction be of the form

$$- J\mu_i\mu_j\gamma^d\varphi(\gamma|i-j|), \qquad (2.1)$$

where the minus sign leads to a tendency to align spins, the constant J provides units of energy, d is the number of dimensions of the system, and the parameter $1/\gamma$ determines the effective range of the interaction. The interaction of each spin with an external field, H, is $-H\mu_i$. Thermodynamic properties are derivable from the partition function

$$Q_N = \sum_{\mu_1 \cdots \mu_N = \pm 1} \exp\left[\tfrac{1}{2}\nu\gamma^d \sum_{\substack{i,j \\ i \neq j}} \mu_i \mu_j \varphi(\gamma|i-j|) + \zeta \sum_i \mu_i \right], \quad (2.2)$$

where $\nu = J/kT$ and $\zeta = H/kT$.

There is a corresponding lattice gas, as discussed in the paper by Lee and Yang.[17] Consider N sites, each of which may be singly occupied or vacant. An attraction, $-\alpha\gamma^d\varphi(\gamma|i-j|)$ exists between particles on sites i and j. If we assign $\mu = +1$ to an occupied site and $\mu = -1$ to a vacant site, the interaction energy, now associated with the pair of sites i and j, is

$$-\alpha\gamma^d \cdot \tfrac{1}{2}(\mu_i+1) \cdot \tfrac{1}{2}(\mu_j+1)\varphi(\gamma|i-j|). \quad (2.3)$$

Multiplying this out we get $\mu_i\mu_j$ terms, like the interactions in the ferromagnet problem, μ_i terms giving effects like those caused by external fields, and constants. In addition, to calculate the grand partition function we must include factors of the fugacity, z, raised to the power $\frac{1}{2}\sum_i (\mu_i+1)$. As a result the grand partition function is

$$G_N = \exp\left\{N\zeta - \tfrac{1}{2}\nu\gamma^d \sum_k \varphi(\gamma|k|)\right\} \sum_{\mu_1 \cdots \mu_N = \pm 1} \exp\left[\tfrac{1}{2}\nu\gamma^d \sum_{ij} \mu_i\mu_j\varphi(\gamma|i-j|) + \zeta \sum_i \mu_i \right], \quad (2.4)$$

where N plays the role of a volume, $\nu = \frac{1}{4}\alpha/kT$, and ζ is related to the fugacity by

$$\zeta = \tfrac{1}{2} \ln z + \nu\gamma^d \sum_{k\neq 0} \varphi(\gamma|k|), \tag{2.5}$$

so that $\zeta = 0$ corresponds to the critical density $\rho = \frac{1}{2}$. The computational problem of evaluating the magnet Q_N and lattice gas G_N are identical.

The innovation in the evaluation of the partition function which we wish to consider involves utilization of the mathematical identity

$$\exp\left[\tfrac{1}{2}\nu\gamma^d \sum_{ij} \mu_i\mu_j \varphi(\gamma|i-j|)\right]$$

$$= \int_{-\infty}^{\infty} dx_1 \ldots dx_N \exp\left[(\nu\gamma^d)^{\frac{1}{2}} \sum_i \mu_i x_i\right] \frac{|B|^{\frac{1}{2}}}{(2\pi)^{N/2}} \exp\left[-\sum_{ij} B_{ij}x_i x_j\right], \tag{2.6}$$

$$\equiv E\left\{\exp\left[(\nu\gamma^d)^{\frac{1}{2}} \sum_i \mu_i X(i)\right]\right\}. \tag{2.7}$$

The matrix $\underset{\approx}{B}$ is the inverse of the matrix with ij element $\varphi(\gamma|i-j|)$. The existence of Eq. (2.6) depends on $\underset{\approx}{B}$ being positive definite. The second way of expressing the identity

is a convenient shorthand borrowed from probability theory. The integral may be regarded as an expectation value, $E\{\ \}$, (or functional integral), defined in terms of a Gaussian random process, $X(i)$. The probability distribution function for such a Gaussian process is completely specified by the mean $E\{X(i)\} = 0$, and the covariance

$$E\{X(i)X(j)\} = \varphi(\gamma|i-j|). \tag{2.8}$$

If Eq. (2.7) is inserted in the partition function then the sum over μ may be performed first to yield

$$Q_N = E\Big\{ \prod_i \cosh[(v\gamma^d)^{\frac{1}{2}}X(i) + \zeta]\Big\}[2 \exp \tfrac{1}{2}v\gamma^d\varphi(0)]^N. \tag{2.9}$$

The question arises as to what the physical significance of the stochastic process $X(i)$ might be. The answer as yet has not been provided (cf., however, Lenard[18]). Note that the partition function of a system of independent spins in an inhomogeneous field H_i is $2^N \prod_i \cosh \beta H_i$, so that $\beta^{-1}(v\gamma^d)^{\frac{1}{2}}X(i)$ is a stochastic field whose effect is to essentially duplicate the action of the internal interaction in determining the free energy. This point of view is merely a tautology unless it can be shown that such a stochastic field has a physical origin in the model.

3. ONE-DIMENSIONAL SYSTEM WITH LONG-RANGE EXPONENTIAL INTERACTION

In a number of cases the approach outlined above has proved to be particularly enlightening and it is to these systems that we now turn attention. The analogue of the continuum model studied by KUH[2] in the previous paper is a one-dimensional lattice system with pair potential[6] $-J\gamma\mu_i\mu_j\exp(-\gamma|i-j|)$. The virtue of the system is that the stochastic process involved in Eq. (2.9) is Markoffian, the Ornstein-Uhlenbeck[19] process familiar from Brownian motion theory. Alternatively stated, the matrix $\underset{\approx}{B}$ of Eq. (2.6) has only diagonal and first off-diagonal elements. Equation (2.9) thus has the form

$$Q_N = (2e^{-\frac{1}{2}\nu\gamma})^N \int dx_1 \ldots dx_N \prod_k \cosh[(\nu\gamma)^{\frac{1}{2}} x_k + \zeta]$$

$$W(x_1)P(x_1|x_2,\gamma)P(x_2|x_3,\gamma)\ldots P(x_{N-1}|x_N,\gamma), \quad (3.1)$$

where

$$W(x) = (2\pi)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}x^2), \quad (3.2)$$

$$P(x|x';\gamma) = [2\pi(1-e^{-2\gamma})]^{-\frac{1}{2}}\exp[-\tfrac{1}{2}(x'-xe^{-\gamma})^2/(1-e^{-2\gamma})]. \quad (3.3)$$

The integrand is a chain of two-variable functions and as such the integral may be regarded a repeated operation of an integral operator, thus expressible in terms of the spectrum.

In particular, if we define the operator, eigenvalues, and eigenfunctions by

$$\mathcal{K}\varphi(x) \equiv \int K(x|x')\varphi(x')dx = \lambda\varphi(x). \tag{3.4}$$

$$K(x|x') = \cosh^{\frac{1}{2}}[(\nu\gamma)^{\frac{1}{2}}x + \zeta]W^{\frac{1}{2}}(x)P(x|x';\gamma)W^{-\frac{1}{2}}(x')\cosh^{\frac{1}{2}}[(\nu\gamma)^{\frac{1}{2}}x + \zeta], \tag{3.5}$$

then

$$Q_N = (2e^{-\frac{1}{2}\nu\gamma})^N \int_{-\infty}^{\infty} dx\, \cosh^{\frac{1}{2}}[(\nu\gamma)^{\frac{1}{2}}x + \zeta]W^{\frac{1}{2}}(x)\mathcal{K}^{N-1}W^{\frac{1}{2}}(x)\cosh^{\frac{1}{2}}[(\nu\gamma)^{\frac{1}{2}}x + \zeta]. \tag{3.6}$$

Making a spectral decomposition of the operator we obtain

$$Q_N = (2e^{-\frac{1}{2}\nu\gamma})^N \sum_j \lambda_j^{N-1} A_j^2, \tag{3.7}$$

with

$$A_j = \int_{-\infty}^{\infty} W^{\frac{1}{2}}(x)\cosh^{\frac{1}{2}}[(\nu\gamma)^{\frac{1}{2}}x + \zeta]\varphi_j(x)dx. \tag{3.8}$$

For N large the contribution of the largest eigenvalue, λ_o, dominates and one obtains

$$\lim_{N\to\infty} N^{-1}\log Q_N = \ln 2 - \tfrac{1}{2}\nu\gamma + \ln \lambda_o. \tag{3.9}$$

It can be shown[6] that the integral equation, (3.4) can be converted into a more convenient quasi-differential form

$$\exp\{\tfrac{1}{2}\log\cosh[(\nu\gamma)^{\frac{1}{2}}x + \zeta]\}\exp\{\gamma[d^2/dx^2 - \tfrac{1}{4}x^2 + \tfrac{1}{2}]\}$$

$$\cdot\exp\{\tfrac{1}{2}\log\cosh[(\nu\gamma)^{\frac{1}{2}}x + \zeta]\}\varphi(x) = \lambda\varphi(x). \qquad (3.10)$$

So far the results are independent of the interaction range. Our further discussion will concentrate on the case of long-range interactions, i.e., $\gamma \to 0$. In this limit the operator exponents in Eq. (3.10) combine merely by addition. Corrections, involving commutators of the operators, are higher order in γ. The eigenvalue problem which must be considered for the long-range problem is

$$[-d^2/dx^2 + U(x,\zeta)]\varphi(x) = \kappa\varphi(x), \qquad (3.11)$$

$$U(x,\zeta) = \tfrac{1}{4}x^2 - \gamma^{-1}\log\cosh[(\nu\gamma)^{\frac{1}{2}}x + \zeta]. \qquad (3.12)$$

λ_o is related to the smallest eigenvalue, κ_o, by

$$\lambda_o = \exp[\gamma(\tfrac{1}{2}-\kappa_o)]. \qquad (3.13)$$

For $\zeta = 0$ (external field of zero or critical density) and high enough temperature one may seek solutions to this eigenvalue problem by expanding the log cosh. The equation becomes

$$\left[-d^2/dx^2 + \frac{1}{4}(1-2\nu)x^2 + \frac{1}{12}\nu^2\gamma x^4 + \ldots\right]\varphi = \kappa\varphi, \qquad (3.14)$$

which to lowest order in γ has smallest eigenvalue $\kappa_o = \frac{1}{2}(1-2\nu)^{\frac{1}{2}}$. Eq. (3.14) is Schrödinger's equation for the harmonic oscillator. The eigenfunctions are the Weber functions. The combination $\gamma\kappa$ appears in Eq. (3.13) so that corrections to the partition function of a set of noninteracting spins are $O(\gamma)$.

Difficulty develops for $\nu > \frac{1}{2}$ since the coefficient of x^2 becomes negative. One must return to the full $U(x,0)$ and note that now U has a pair of minima at $\pm|\eta_o|(2/\gamma)^{\frac{1}{2}}$, where η_o satisfies

$$\eta_o(2\nu)^{-\frac{1}{2}} = \tanh[(2\nu)^{\frac{1}{2}}\eta_o], \qquad (3.15)$$

(cf. Fig. 1). For temperatures a finite amount below the critical ($\nu_c = J/kT_c = \frac{1}{2}$) the wells, which are of equal depth, are separated by $O(\gamma^{-\frac{1}{2}})$ in x, and are of depth $O(\gamma^{-1})$. This becomes an infinite barrier as $\gamma \to 0$. The lowest eigenvalue corresponds to two eigenfunctions, one centered about the positive well, the other the negative. This degeneracy leads to the manifestations of two phases. Note now that since κ_o is of the order of the depth of the well, $\gamma\kappa_o$ is of $O(1)$, so that there are finite changes in the noninteracting spin partition function. The values of x which are significant in the eigenfunction $\varphi(x)$ have $O(1)$ spread about $\pm|\eta_o|(2/\gamma)^{\frac{1}{2}}$; using the interpretation of x as proportional to a stochastic field, we find that the stochastic field fluctuates about the

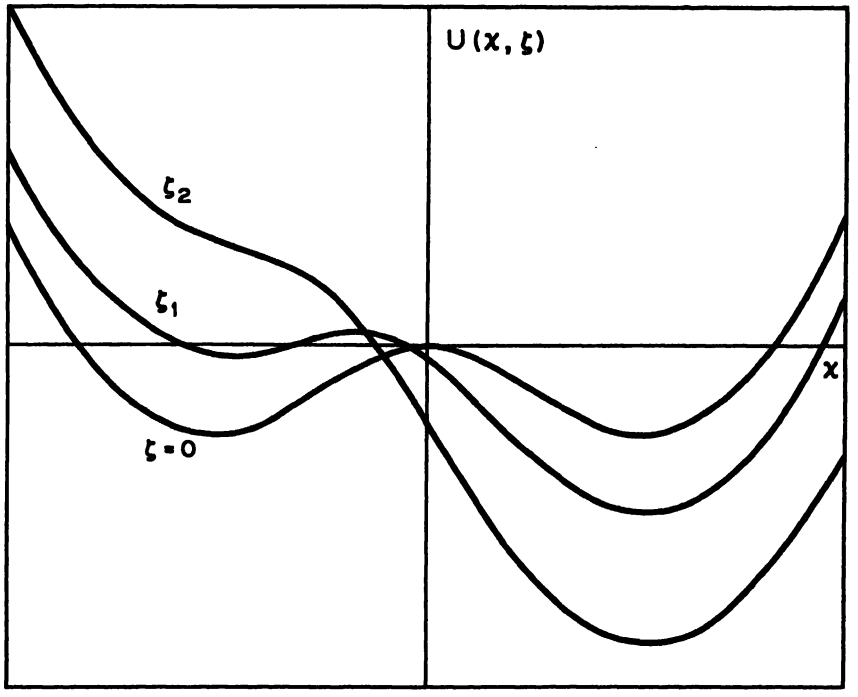


FIG. 1

Fig. 1. $U(x,\zeta)$ vs x for three chemical potentials, such that $0 < \zeta_1 < \zeta_2$, and at a fixed subcritical temperature. The scale of x is proportional to $\gamma^{-\frac{1}{2}}$, while that of U is proportionsl to γ^{-1}

mean field of the Weiss theory. Thus, corrections to mean field theory of higher order in the inverse range of the interaction can be made.

In order to attain an understanding of the lattice gas, or any of a variety of systems exhibiting first order phase transitions, we must return to the case of nonvanishing ζ (chemical potential such as to cause deviations from the critical density, or an external field for the magnetic problem). As has been seen, the uniform bias, ζ, must be added to the stochastic term, $(\nu\gamma)^{\frac{1}{2}}x$, in the log cosh of Eq. (3.12). $U(x,\zeta)$ is no longer symmetric in x (cf. Fig. 1). Except near the critical point, as $\gamma \to 0$, the eigenfunction corresponding to the lowest eigenvalue is localized about the absolute minimum, at the point $\eta_\zeta(2/\gamma)^{\frac{1}{2}}$, where η_ζ is one of the solutions of

$$\eta_\zeta(2\nu)^{-\frac{1}{2}} = \tanh[(2\nu)^{\frac{1}{2}}\eta_\zeta + \zeta]. \qquad (3.16)$$

Straightforward manipulation of the grand partition function then yields as the equation of state (in the one phase region) a van der Waals-like equation

$$p = p_{SR} - \alpha\rho^2 + O(\gamma) \qquad (3.17)$$

where p_{SR} is the equation of state of a lattice gas interacting only with the short-range forces; in this case the short-range force is that which leads to the exclusion of multiple occupancy, and the equation of state is

$$\beta p_{SR} = - \ln(1-\rho). \qquad (3.18)$$

To achieve a better insight into how the first order phase transition is manifest consider varying the chemical potential at some fixed subcritical temperature. For a range of ζ, Eq. (3.16) will lead to two minima, but one is lower than the other except for $\zeta = 0$ when the above mentioned degeneracy occurs (cf. Fig. 1). For $\gamma \to 0$, as ζ passes through zero the eigenfunction will shift from one minimum to the other. This results in a change in the rate of variation of κ with ζ; or, in thermodynamic terms, a finite change in the density at the transition.

One can make higher order γ corrections to the results so far reported here. This involves retention of higher order commutators in the combination of the exponential operators in Eq. (3.10) and of more terms in the expansion of $U(x,\zeta)$ about its minima. The results are long, and conceptually rather uninteresting. What is important is that the phase transition is retained in each order of expansion in inverse interaction range. A theorem of van Hove[20] tells us that a one-dimensional system with finite range interactions cannot exhibit a phase transition. Thus we must consider the nature of the singularity at $\gamma = 0$ which is causing this paradox.[16]

The source of the difficulty is the assumption made earlier that the two wells of the function $U(x,\zeta)$ are completely separated. In reality the barrier is $O(\gamma^{-\frac{1}{2}})$ in width and $O(\gamma^{-1})$ in height. The lowest eigenfunction is not truly degenerate. A symmetric and an antisymmetric state have eigenvalues "split" by an amount Δ, proportional to the exponential of minus a quantity of the order of the barrier width times the square root of the barrier height: i.e., $\Delta \propto \exp[-c/\gamma]$. In order to get away from density $\rho = \frac{1}{2}$ one varies the chemical potential via ζ. The most important effect that the change of ζ from 0 has on $U(x,\zeta)$ is to raise one of the wells and lower the other by an amount we will call b, which is proportional to ζ. When ζ varies about zero over a range such that $b/\Delta = O(1)$ the eigenfunction still has significant weight in both wells. Quantitatively, the ratio of the integral over one well of the eigenfunction to the integral over the other well is

$$R = [1 + (b/\Delta)^2]^{\frac{1}{2}} - (b/\Delta). \qquad (3.19)$$

Qualitatively, instead of ρ (which is linearly related to $d\kappa/d\zeta$) changing suddenly as ζ goes through zero, ρ changes over a span of ζ (or of p) of the order of Δ. This implies, for instance, that as $\gamma \to 0$ the isotherm approaches flatness in the "two phase region" as $\exp(-c/\gamma)$. This essentially singular law cannot be expressed in powers of γ.

This calculation also presents an alternative to the method proposed by KUH[2] for proving that the square root of the volume fractions are the weights to be assigned to the two portions of the lowest "degenerate" eigenfunction in the two phase region.[16] Namely, one first calculates the true nondegenerate eigenfunction for finite but small γ so that each over-all density, even those which will ultimately represent two phases, corresponds to a distinct ζ. One uses this ζ to determine the relative distribution of the eigenfunction via Eq. (3.19), and then passes to the $\gamma \to 0$ limit.

4. CRITICAL PHENOMENA

Of particular interest are the phenomena which occur in the neighborhood of the critical point for the above model.[4,6] It is found that there is a region surrounding the critical point, in which region special changes occur in the nature of the thermodynamic properties. For example, above the critical region, according to our earlier calculations, it can be shown that the specific heat of the field-free ferromagnet is zero as $\gamma \to 0$, but the γ correction goes as $\gamma(T-T_c)^{-3/2}$. Before this has an opportunity to diverge, however, we shall see that the critical region is entered, and one goes smoothly over to the value $(3/2)k$ per particle appropriate to temperatures below the critical.

Our investigation might begin with the question: how small must $(1-2\nu)$ be before the quartic term in the expansion of $U(x,0)$ is as important as the quadratic term

(cf., Eq. (3.9))? Dimensional analysis indicates that the answer is $1 - 2\nu = O(\gamma^{\frac{2}{3}})$. A related question is: how small must ζ be for this quartic term to still be important? The equation to be considered is:

$$\left[-\frac{d^2}{dx^2} + \tfrac{1}{4}(1-2\nu)x^2 + \frac{\nu^2 x^4}{12} - \sqrt{\frac{\nu}{\gamma}}\,\zeta x \right] \varphi = \kappa\varphi. \quad (4.1)$$

The answer is $\zeta = O(\gamma)$ which corresponds to a range of densities of $O(\gamma^{\frac{1}{3}})$ about the critical.

Thus there is, in a sense, a critical region in temperature and density surrounding the critical point. Within this region the dependence of the properties of the system on temperature and density are significantly modified. Physically this is due to the broadening of the distribution of certain fluctuations in the system, which appears in the present model as a broadening of the range of values of x (or the stochastic field) for which $\varphi(x)$ is significant. The actual critical point is somewhat arbitrary since there are no actual singularities in thermodynamic properties. HKU[4] propose to define the temperature at which $\kappa(\nu, \zeta = 0) = 0$ as the critical point, which corresponds to a temperature lower than the mean field theory for finite γ, in agreement with known results. This temperature marks a change in the eigenfunction of the eigenvalue equation from singly to doubly peaked.

One can go on to an analysis of the thermodynamic properties in this region but many of the quantitative results and some of the qualitative results are peculiar to the one-dimensional system. The main conclusion to be drawn is that the emergence of these quartic terms leads to important qualitative changes in the nature of the thermodynamic functions and the correlation functions.

5. CORRELATION FUNCTIONS

The method employed above to evaluate the partition function also serves for a determination of the correlation functions.[3,6] Consider, for instance, the pair correlation function,

$$f(t) \equiv (1/Q_N) \sum_{\substack{\mu_1 \cdots \mu_N \\ =\pm 1}} \mu_\ell \mu_{\ell+t} e^{-\beta H}. \qquad (5.1)$$

For the evaluation of the sum one first converts to an integral like Eq. (3.1). For zero field or at the critical density the factor $\cosh(\nu\gamma)^{\frac{1}{2}}x_\ell$ is replaced by

$$\sinh(\nu\gamma)^{\frac{1}{2}}x_\ell \equiv \cosh(\nu\gamma)^{\frac{1}{2}}x_\ell \tanh(\nu\gamma)^{\frac{1}{2}}x_\ell.$$

due to the minus sign introduced by μ_ℓ; and likewise with $\cosh(\nu\gamma)^{\frac{1}{2}}x_{\ell+t}$. Thus the structure of the integrand as a chain of identical integral kernels is interrupted by the

$\tanh(v\gamma)^{\frac{1}{2}}x$ factors. From site 0 to ℓ and $\ell + t$ to N the operator makes a large number of repeated operations and thus may be represented, as before, by (Dirac notation)

$$\lambda_0 |0\rangle\langle 0| . \tag{5.2}$$

Between the two tanh's, however, the operator only acts a finite number of times, t, so that the full spectral representation

$$\sum_j \lambda_j |j\rangle\langle j| \tag{5.3}$$

is needed. It is then not difficult to show that the correlation function is given by

$$f(t) = \sum_j \langle 0|\tanh(v\gamma)^{\frac{1}{2}}x|j\rangle(\lambda_j/\lambda_0)^t \langle j|\tanh(v\gamma)^{\frac{1}{2}}x|0\rangle, \tag{5.4}$$

except for f(0) which is equal to unity. Nonzero field results are given by a simple extension of this formula, as are three particle correlations, etc.

For temperatures above the transition, $\tanh(v\gamma)^{\frac{1}{2}}x$, may be expanded to $(v\gamma)^{\frac{1}{2}}x$, and since to lowest order the φ_j are Weber functions, the only nonvanishing matrix element is $j = 1$. Thus f(t) is found to be

$$f(t) = [v\gamma/(1-2v)^{\frac{1}{2}}]\exp[-\gamma t(1-2v)^{\frac{1}{2}}], \quad t \neq 0. \tag{5.5}$$

The compressibility equation of state may now be verified. For the lattice gas it is

$$\frac{1}{4}\sum_{t} f(t) - (\rho-\tfrac{1}{2})^2 = \rho^2(\partial \ln \rho/\partial\beta p), \qquad (5.6)$$

from which the compressibility is calculated to be ($\rho = \frac{1}{2}$)

$$\kappa\, kT = \partial \ln \rho/\partial\beta p = (1-2v)^{-1} \qquad (5.7)$$

in agreement with Eq. (3.12).

Qualitatively the lattice gas results are analogous to those of UHK[3] for the one-dimensional continuum model. The short-range behavior, to lowest order in γ, is that of a gas with only the short-range forces. Both models have identical $O(\gamma)$ corrections when expressed appropriately in terms of compressibility. The long-range behavior is as given by the Ornstein-Zernike theory as adapted by Lebowitz and Percus[9] for long-range forces, but this is true only outside the critical region. The theory inside the critical region must be based on considerations similar to those of the previous section.

In application of Eq. (5.4) to the critical region the eigenfunctions appropriate to the full quartic eigenvalue equation (4.1) must be used. This introduces two major effects. The first is that the pair correlation function becomes quite long ranged, viz., of range $O(\gamma^{-4/3})$. The fact

that the range does not become even longer, in fact strictly infinite, is associated with the lack of a singularity in the compressibility, and is thus probably an artifact of the model. The second effect is that the pair correlation function has the form of a series of exponentials rather than just one. To illustrate this we exhibit the pair correlation for $\zeta = 0$ and $\nu = \frac{1}{2}$:

$$f(t) = \gamma^{2/3}[0.66 \exp(-0.75\ \gamma^{4/3}t) + 0.0019 \exp(-2.9\ \gamma^{4/3}t) + 0.00006 \exp(-5.6\ \gamma^{4/3}t) + \ldots]. \quad (5.8)$$

While the other exponentials are theoretically significant in that their contributions are of the same γ order, the first is numerically dominant.

Finally, let us consider the correlation functions in the two phase region.[3] In an earlier paper by Mayer,[21] reprinted here, it is asserted that the correlation functions in the two phase region should be linear combinations of those appropriate to the pure phases, with the volume fraction as weightings. In the present model this result comes out explicitly as a consequence of the linear combination occurring in the "degenerate" lowest eigenfunction in the two phase region (cf., end of Section 3).

6. TWO-DIMENSIONAL SYSTEMS

The extraction of even qualitative physical information about phase transitions from the one-dimensional system which we have been discussing is difficult. Only by passing

to the $\gamma \to 0$ limit can the transition be made to occur. A manifestation of this is that the critical region vanishes as $\gamma \to 0$ whereas the transition vanishes if γ does not go to zero. One may attempt to make something of the difference between the $\gamma^{\frac{2}{3}}$ effects and the $e^{-c/\gamma}$ effects but it is not obvious how or why this should be done.

A number of two-dimensional systems have been treated by Kac and Helfand,[6] but we will concentrate here on what is in some ways the most interesting of these. Consider a square array of N columns and M rows with interaction between site (k,ℓ) and (k',ℓ') given by

$$- J\gamma^2 \mu_{k,\ell}\mu_{k',\ell'} \exp[-\gamma(|k-k'| + |\ell-\ell'|)], \qquad (6.1)$$

i.e., exponential in the distance but measured along the bonds. The way that the intrinsically one-dimensional concept of a Markoff process can be extended for the treatment of this two-dimensional problem is to employ a set of Markoff processes, one for each row of the system. Thus, the general formalism of Section 2 requires that we find a Gaussian process $X(k,\ell)$ with mean zero and covariance

$$E\{X(k,\ell)X(k',\ell')\} = \exp[-\gamma(|k-k'| + |\ell-\ell'|)]. \qquad (6.2)$$

Define $X(k,\ell)$ as a linear combination of real, independent Ornstein-Uhlenbeck processes $U_m(k)$:

$$X(k,\ell) = \sum_m c_{\ell-m} U_m(k), \qquad (6.3)$$

$$E\{U_m(k)U_{m'}(k')\} = \delta_{mm'}\exp(-\gamma|k-k'|). \tag{6.4}$$

To satisfy Eq. (6.2) the matrix $\underset{\approx}{C}$ must be defined by

$$\sum_{\ell} c_{p-\ell}c_{p'-\ell} = e^{-\gamma|p-p'|}. \tag{6.5}$$

There are several solutions for $\underset{\approx}{C}$ in the limit as $M \to \infty$ but for the most part we will be interested in the combination (6.5) so we need not exhibit $\underset{\approx}{C}$ explicitly here. The functional integral for the partition function can be evaluated in terms of the highest eigenvalue of an eigenvalue equation, but now one involving M variables. In the limit of long-range forces the equation to be considered is

$$\left\{\sum_{\ell=1}^{M}\left(-\,d^2/dx_\ell^2\right) + U(x_1 \ldots x_M, \zeta)\right\}\varphi(\underset{\sim}{x}) = \kappa\varphi(\underset{\sim}{x}), \tag{6.5}$$

$$U(\underset{\sim}{x},\zeta) = \sum_{\ell}\left\{\tfrac{1}{4}x_\ell^2 - \gamma^{-1}\log\cosh\left[\sum_{p}{}' c_{\ell-p}x_p + \zeta\right]\right\}. \tag{6.6}$$

Methods applicable to the one-dimensional problem outside the critical region may be extended. Namely, U may be expanded about the absolute minimum to quadratic terms. One must then perform a normal mode analysis to solve the eigenvalue problem, and this is possible since all the matrices involved are Toeplitz as $M \to \infty$.

For $\zeta = 0$ a transition is manifest for temperatures near that for which the lowest normal mode "frequency" vanishes. To lowest order in γ this temperature again corresponds to that predicted by mean field theory, viz., $J/kT_c = \frac{1}{4}$. For lower temperatures the U surface is extremely complicated possessing many minima. A solution may be sought in the neighborhood of the one or two absolute minima. The investigation of the effect of the other minima is still incomplete.

One should again question how the phase transition is approached. Van Hove's theorem[20] is easily extended to tell us that for a finite-range interaction and a finite number of rows, M, (number of columns $\rightarrow \infty$), the system can exhibit no phase transition. Efforts to resolve the problem center once more on the general theory of multiple-well eigenvalue problems. The full theory has not been worked out, but the results that are indicated are so important that perhaps it is of value to state them here as a conjecture. If we assume that the eigenfunction is localized for the most part about the two lowest minima, or several minima in this neighborhood, then we may attempt to calculate the splitting by an application of the WKB approximation along contours marking the passes between these two valleys. The existence of two types of passes has been demonstrated. Some of the

passes are of length $O(M^{\frac{1}{2}}\gamma^{-\frac{1}{2}})$ and height $O(M\gamma^{-1})$. Others are of length $O(M\gamma^{-\frac{1}{2}})$ and height $O(\gamma^{-1})$. In both cases the contribution to the splitting is $\exp(-cM/\gamma)$. Thus if we first pass to the thermodynamic limit $M \to \infty$ this essential singularity in γ at zero disappears, and the transition is already induced for a finite γ, i.e., for a finite-range interaction, as it must be.

This result, if correct, is encouraging. The presence of the essential singularity in the one-dimensional calculation was a bad omen with respect to attempts to develop the properties of the system in powers of γ. If the above conjecture is correct then such efforts may still be worthwhile in higher dimensions, away from the critical region.

Study of the critical phenomena in these two-dimensional lattice systems is currently underway.

7. DIAGRAM THEORY

Some of the properties of systems with long-range forces can be treated by diagrammatic methods. Basically these ideas have their origin in Mayer's[22] treatment of systems with Coulomb interactions, where he was forced to develop this formalism to avoid divergences of the individual terms. Brout[7] adapted these methods to enable him to develop an expansion of the properties of lattice systems in the inverse of the effective number of neighbors with which each particle interacts, or the inverse range of interaction.

Horowitz and Callen[8] have carried these ideas further. Extension to continuous systems has been made by Lebowitz and Percus,[9] and Brout,[10] and Hemmer.[11] We will discuss the essential ideas for the continuous systems rather than lattice systems, since the diagram theory of the former is generally more familiar.

Consider a d dimensional system of particles interacting via a pair potential which is a sum of a short-range and long-range function,

$$v(r) = v_{SR}(r) + v_{LR}(r). \tag{7.1}$$

$v_{LR}(r)$ is scaled by a parameter γ so that it has the form

$$v_{LR}(r) = \gamma^d u(\gamma r), \tag{7.2}$$

which keeps the energy due to the long-range forces finite as $\gamma \to 0$. The Mayer f function can be written

$$f = e^{-\beta v} - 1 = f_{SR} + f_{LR} + f_{SR} f_{LR}. \tag{7.3}$$

The usual diagrams generalize in such a fashion that each diagram goes over into all the diagrams which result from the replacement of each bond by either the short-range f bond, the long-range f bond, or both.

The f_{LR} may actually be expanded into

$$f_{LR} \equiv \exp(-\beta v_{LR}) - 1 = -\beta\gamma^d u(\gamma r) + \tfrac{1}{2}\beta^2\gamma^{2d}u^2(\gamma r) + \ldots . \tag{7.4}$$

Thus, such a bond introduces a factor γ^d. When integrated freely over $\underset{\sim}{r}$, however, the integrand is significant over γ^{-d} so that the integral of f_{LR} is $O(1)$. Therefore, a diagram of k bonds and p points composed entirely of long-range bonds makes a contribution $O[(\gamma^d)^{k-p+1}]$. To be zeroth order such a diagram would have to have one more point than bond, but this would imply at least one free end. A free end signifies that the graph is not multiply connected as in general the graphs of a density series must be. The only diagram of a density series with a free end is the single bond diagram, which represents the second virial coefficient. The contribution to βp from this diagram is

$$- \tfrac{1}{2}\rho^2 \int f_{LR} d\underset{\sim}{r} = \tfrac{1}{2}\rho^2 \beta\gamma^d \int v_{LR}(\gamma r) d\underset{\sim}{r} + \ldots,$$

$$= \tfrac{1}{2}\beta\rho^2 \int u(x) d\underset{\sim}{x} + \ldots . \qquad (7.5)$$

Of course, all diagrams composed entirely of short-range bonds are $O(\gamma^0)$. Thus, to lowest order in γ we obtain the van der Waals-like equation

$$p = p_{SR} + \tfrac{1}{2}\rho^2 \int u(x) d\underset{\sim}{x}. \qquad (7.6)$$

The next order contribution (besides higher order terms of Eq. (7.5)) comes from a class of graphs with p points and p long-range bonds, namely the polygons or rings of long-range bonds. However, now it is possible to introduce more

than just a point between each of the long-range bonds. We may insert, rather, various combinations of short-range links. The sum over all the allowed combinations yields the short-range pair correlation function. Therefore the new long-range bond must be taken as representing a convolution of f_{LR} and the short-range correlation function. One must then sum over all rings composed of these bonds to obtain a γ^d term.

The general principles which allow extension to higher order terms and also to formulas for the pair correlation function are clear.

The diagrammatic expansion procedure is not appropriate in the neighborhood of the critical point. One finds as a criterion for the validity of the expansion that $\gamma^{2d/3}$ times the compressibility must be much less than unity.[11] The violation of this condition is what marks the critical region found for the one-dimensional Kac model.

8. MORE GENERAL LONG-RANGE INTERACTIONS

If one is to proceed via a tractable operator formalism the indication seems to be that some special type of potential such as exponential or nearest-neighbor on a lattice is necessary. On the other hand, the diagram technique discussed above in no way is restricted to special functional forms. Siegert[12,13] and Muhlschlegel and Zittartz[14] have gone even further than the diagram techniques proceeding

via the general functional integration expression for the partition function, Eq. (2.9). The method is too long to outline in any detail here. Suffice it to say that a procedure has been developed for removing from the partition function factors contributing higher and higher order γ^d contributions. Also below the transition a saddle point or variational calculation leads to a shift in the important region of the functional integration. This shift, like that in the eigenfunction of our earlier Sections, is related to the presence of Weiss mean fields. Difficulties develop in these analyses in the neighborhood of the critical point. There we know from our earlier studies that the expansion is not in γ^d.

The manner in which a long-range interaction effects the phase transition of systems with general short-range interactions has been studied along these lines by Baker.[23]

REFERENCES

1. M. Kac, Phys. Fluids 2, 8 (1959).
2. M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. 4, 216 (1963), referred to as KUH herein.*
3. G. E. Uhlenbeck, P. C. Hemmer, and M. Kac, J. Math. Phys. 4, 229 (1963), referred to as UHK.*
4. P. C. Hemmer, M. Kac, and G. E. Uhlenbeck, J. Math. Phys. 5, 60 (1964), referred to as HKU.*

* For a survey of these papers and many other topics in the theory of the phase transition see G. E. Uhlenbeck, in Statistical Physics 3, Brandeis Summer Institute 1962 (W. A. Benjamin, Inc., New York, 1963). Also see E. Helfand, Ann. Rev. Phys. Chem. 14, 117 (1963).

5. G. A. Baker, Jr., Phys. Rev. 122, 1477 (1961).
6. M. Kac and E. Helfand, J. Math. Phys. 4, 1078 (1963).
7. R. Brout, Phys. Rev. 118, 1009 (1960); 122, 469 (1961).
8. G. Horowitz and H. B. Callen, Phys. Rev. 124, 1757 (1961).
9. J. L. Lebowitz and J. K. Percus, J. Math. Phys. 4, 248 (1963).
10. M. Coopersmith and R. Brout, Phys. Rev. 130, 2539 (1964).
11. P. C. Hemmer, J. Math. Phys. 5, 75 (1964).
12. A. J. F. Siegert, Statistical Physics 3, Brandeis Summer Institute 1962 (W. A. Benjamin, Inc., New York, 1963).
13. A. J. F. Siegert, On the Ising Model with Long-Range Interaction, Northwestern University preprint, 1962.

References (Cont'd.) - 2

14. B. Muhlschlegel and H. Zittartz, Z. Physik. 175, 553 (1963).

15. G. A. Baker, Jr., Phys. Rev. 126, 2072 (1962).

16. E. Helfand, J. Math. Phys. 5, 127 (1964).

17. T. D. Lee and C N. Yang, Phys. Rev. 87, 410 (1952).

18. A. Lenard, J. Math. Phys. 4, 533 (1963).

19. G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. 36, 823 (1930).

20. L. van Hove, Physica 16, 137 (1950).

21. J. E. Mayer, J. Chem. Phys. 15, 187 (1947).

22. J. E. Mayer, J. Chem. Phys. 18, 1426 (1950).

23. G. A. Baker, Jr., Phys. Rev. 130, 1406 (1963).

PHYSICAL REVIEW VOLUME 127, NUMBER 2 JULY 15, 1962

Phase Transition in Elastic Disks*

B. J. ALDER AND T. E. WAINWRIGHT
University of California, Lawrence Radiation Laboratory, Livermore, California
(Received October 30, 1961)

The study of a two-dimensional system consisting of 870 hard-disk particles in the phase-transition region has shown that the isotherm has a van der Waals-like loop. The density change across the transition is about 4% and the corresponding entropy change is small.

A STUDY has been made of a two-dimensional system consisting of 870 hard-disk particles. Simultaneous motions of the particles have been calculated by means of an electronic computer as described previously.[1] The disks were again placed in a periodically repeated rectangular array. The computer program has been improved such that about 200 000 collisions per hour can be calculated by the LARC computer regardless of the number of particles in the system. This speed made it possible to follow large systems for several million collisions.

It became necessary to study larger systems in the phase transition region when for smaller ones in three dimensions, it did not seem to be possible for the two phases to exist together in equilibrium.[2,3] Even in the largest three-dimensional system investigated with the improved program (500 hard spheres), the particles were either all in the fluid phase or all in the crystalline phase. The system would typically remain in one phase for many collisions. The occasional shift from one phase to the other would be accompanied by a change of pressure. The equation of state was represented by two disconnected branches overlapping in the density range of the transition, since with the limited number of phase interchanges it was not possible to average the two branches.

Two-dimensional systems were then studied, since the number of particles required to form clusters of particles of one phase of any given diameter is less than in three dimensions. Thus, an 870 hard-disk system is effectively much larger than a 500 hard-sphere system. First, however, it was necessary to establish that small two-dimensional systems behave analogously to the three-dimensional systems. This is illustrated in Fig. 1 by the two disconnected branches drawn lightly through the triangular points for a 72-particle system. In that figure, the reduced pressure pA_0/NkT is plotted against the reduced area A/A_0, where A_0 is the area of the system at close packing. In the region of A/A_0 from 1.33 to 1.35 the system fluctuated infrequently between a high-pressure fluid branch and a low-pressure crystalline branch, while at A/A_0 of 1.31 and higher densities the solid phase was always stable.

For the larger 870-particle system, however, the two phases exist side by side. One piece of evidence for this coexistence is the cathode-ray tube pictures described earlier (see Fig. 2).[1] The trajectories of the particles plotted on the oscilloscope show regions where the particles are localized (crystallites) in between regions of mobile particles (fluid). Further evidence is the characteristically large pressure fluctuations in the phase transition region where two states can exist with almost equal probability. The extent of the fluctuations in a typical run of about 10 million collisions is obtained

* This work was performed under the auspices of the U. S. Atomic Energy Commission.
[1] B. J. Alder and T. E. Wainwright, J. Chem. Phys. 31, 459 (1959).
[2] B. J. Alder and T. E. Wainwright, J. Chem. Phys. 33, 1439 (1960).
[3] W. W. Wood, R. R. Parker, and J. P. Jacobson, Suppl. Nuovo cimento 9, 133 (1958).

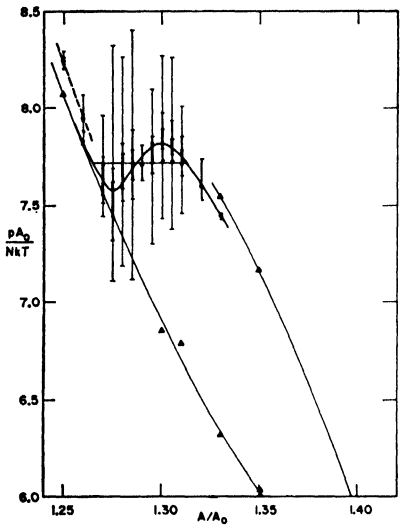


FIG. 1. The equation of state of hard disks in the phase transition region.

by breaking the run into intervals of 50 thousand collisions (about 100 collisions for each particle) and determining the average pressure in each interval. The light vertical lines at various densities (Fig. 1) extend from the maximum to the minimum pressure found among these intervals. As can be seen, the fluctuations are much larger in the phase transition region than in the pure solid ($A/A_0<1.26$) and pure fluid ($A/A_0>1.33$) phases and they are also larger in the middle of the phase transition region than near the ends. The medium vertical lines in Fig. 1 indicate the middle range of the fluctuations, that is, $\frac{1}{4}$ of the intervals show pressures above the top of the vertical line and $\frac{1}{4}$ show pressures below the bottom of the line. For the shorter runs at A/A_0 of 1.32 and 1.29 only this medium vertical line could be drawn.

The heavy vertical lines in Fig. 1 indicate the estimated accuracy of the average pressure determination at each density. These estimates are made by comparing several runs of 10 million collisions each with various starting conditions. The comparison was typically within 1% at a few selected densities. The fact that the pressures calculated with varying starting conditions agree rather closely is the only indication that phase space has been adequately sampled. Four different starting conditions have been used: (1) all the particles are located in lattice positions with only one particle in motion; (2) all particles are in lattice positions and in motion with randomly selected velocities; (3) the starting configuration for a run at one density is taken from an instantaneous configuration at a lower density, having been effectively changed by increasing the diameter of the particles; and (4) the same as (3) except a higher density configuration is used as an initial low-density one.

The smooth curve drawn through the heavy vertical sections in Fig. 1 clearly shows a van der Waals loop-like behavior for the equilibrium state of a finite system. To confirm this it was found that decreasing the density from A/A_0 of 1.29 to 1.30 by procedure (4) above, increased the average pressure. Similarly, increasing the density from A/A_0 of 1.285 to 1.280 and subsequently to 1.275 by procedure (3) decreased the average pressure in both cases, although, immediately after increasing the density, the pressure at 1.275 was higher for some millions of collisions. This shows that the density must be increased very slowly near the solid region or otherwise the particles will be locked into a disordered configuration. Thus, on increasing the density by procedure (3) from A/A_0 of 1.275 to 1.26 the crystalline region of phase space (lower value) became disconnected from the disordered region (upper value) as seen in Fig. 1, since the extremes of the pressure fluctuations no longer overlap. A further increase in density of the disordered or glass-like configuration at A/A_0 of 1.26 to 1.25 defines the dashed line in Fig. 1.

It has been shown that in an infinite system the isotherms for this system would always have to be of negative or at most zero slope.[4] Thus, the loop could not exist. The existence of a loop for finite systems probably derives from the fact that the constraint of constant density is imposed over a region occupied by however many particles are dealt with. This constraint restricts

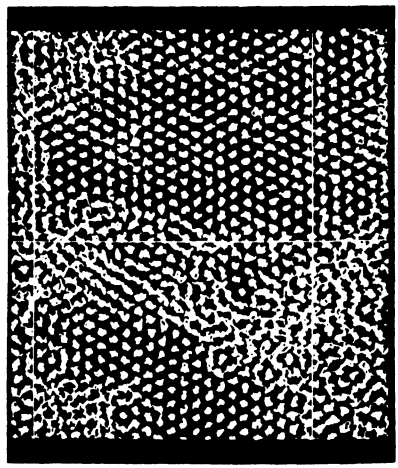


FIG. 2. The traces of the centers of particles in the phase-transition region showing fluid and crystalline regions. The horizontal and vertical lines represent an arbitrary grid.

[4] L. van Hove, Physica 15, 951 (1949).

the configurations which can be reached in a finite system, that is, for example, a fluid in equilibrium with crystallites of average size greater than the number of particles dealt with is impossible to achieve. This constraint for the small systems previously investigated resulted in stabilization of the predominant phase. Thus, the system was either all solid and when a rare fluctuation disordered enough of the system, it became completely fluid. For the 870-particle system the constraint again stabilizes the more abundant phase causing the pressure to be high on the fluid side and low on the crystal side. It thus seems that the phase separation which might occur in infinite systems is not complete in finite systems, since a sizable portion of the system lies in the fluid-crystal boundary region and this region is of intermediate density and evidently takes on more of the character of the predominant phase.

The horizontal line in Fig. 1 drawn at $pA_0/NkT=7.72$ and extending from A/A_0 of 1.266 to 1.312 corresponds to the usual "equal area" rule. If the phase transition for an infinite system is of first order at the pressure indicated by this straight line, then the resulting entropy change across the transition $\Delta S/Nk$ is $p\Delta A/NkT=0.36$. The change of entropy across the same density interval corresponding to the expansion of the one particle cell as calculated by the free volume theory is 0.30. This indicates that the change of communal entropy (0.06) across the transition is very much smaller than unity. This is hardly in accord with the view[5] that the difference between a dense fluid and a solid is one of the accessibility of the entire space in the fluid and localization of a molecule in a solid.

The complete equation of state and comparisons of it with the predictions of various theories will be the subject of further publications.

We are deeply indebted to Mary Ann Mansigh and Norman Hardy for their invaluable help in programming, and to Dr. Sidney Fernbach of the Livermore Computing Division for his cooperation.

[5] J. O. Hirschfelder, D. P. Stevenson, and H. Eyring, J. Chem. Phys. 5, 896 (1937); however, see also O. K. Rice, J. Chem. Phys. 6, 476 (1938).

Comment

For the general principles underlying the molecular dynamic calculations see B. J. Alder and T. E. Wainwright, Nuovo Cimento Suppl. No. 1 to Vol. 9, 116 (1958). Results for the three-dimensional fluid of hard spheres can be found in B. J. Alder and T. E. Wainwright, J. Chem. Phys., 31, 459 (1959) and ibid. 33, 1439 (1960). The reader should also consult the comment on the "N-dependence" at the end of the Yang and Lee paper of Section 1.

JOURNAL OF MATHEMATICAL PHYSICS VOLUME 5, NUMBER 7 JULY 1964

Correlation Functions and the Critical Region of Simple Fluids

MICHAEL E. FISHER*
The Rockefeller Institute, New York, New York
(Received 4 February 1964)

The "classical" (e.g. van der Waals) theories of the gas–liquid critical point are reviewed briefly and the predictions concerning the nature of the singularities of the coexistence curve, the specific heat and the compressibilities are compared critically with experiment and with the analytical and numerical results for lattice gas models.

The critical singularities are related to the behavior of the pair correlation function $G(r) = g(r) - 1$ and the Ornstein–Zernike theory of critical scattering is reviewed. Alternative derivations of the theory are discussed and its validity is assessed in relation to experiment and to more detailed theoretical calculations. The nature and magnitude of the expected deviations from the "classical" theory are described. The analogies with critical magnetic phenomena are mentioned briefly.

THIS is a review article about theories of the critical point and their experimental and theoretical validity. Recent work has revealed the shortcomings of the well-known approximate theories and it is hoped that this review, although in the main nonmathematical in character, will provide some stimulus to further theoretical (and experimental) work on this old, but still imperfectly understood problem.

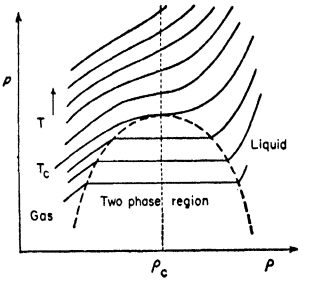


FIG. 1. Schematic isotherms for a simple fluid in the critical region.

1. THE CLASSICAL THEORY OF THE CRITICAL POINT

At temperatures below its critical temperature T_c a gas can be condensed by isothermal compression. At temperatures above T_c the transition from dense gas to liquid takes place without discontinuity of density or, as far as has been determined experimentally, without any higher-order singularities in the density or other variables. As the temperature increases to T_c the difference between the densities ρ_L and ρ_G of coexisting liquid and gas tends continuously to zero.[1] The limiting density ρ_c, and corresponding pressure p_c, define the *critical point* (see Fig. 1). In this article we shall be concerned with the properties of a simple fluid in the region of its critical point and in particular on the critical isochore $\rho = \rho_c$.

In many respects the behavior of binary fluid mixtures which undergo phase separation is closely analogous to the condensation of simple fluids and most of our remarks can be translated directly into such terms.[2] For simplicity, however, we refer in the main only to single-component systems.

As is well known, a qualitative account of condensation phenomena and the critical point is given by the classical equation of van der Waals

$$p/kT = \rho/(1 - b\rho) - a\rho^2/kT, \qquad (1.1)$$

provided this is supplemented by the "equal area rule" of Maxwell which ensures that the density ρ is a single-valued function of the pressure p.[3] The general appearance of the isotherms of a van der Waals gas in the critical region is similar to that of a real gas as shown in Fig. 1. In particular the p, ρ isotherms become flatter and flatter as T approaches T_c from above at the critical density and correspondingly the isothermal compressibility

$$K_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = \frac{1}{\rho}\left(\frac{\partial \rho}{\partial p}\right)_T \qquad (1.2)$$

* On leave of absence from The Wheatstone Physics Laboratory, King's College, London, W.C. 2, England.

[1] It now seems quite well established experimentally that for simple fluids, such as the noble gases, the coexistence curve does *not* have a significant "flat top." For a discussion of this suggestion see O. K. Rice, J. Phys. & Colloid Chem. **54**, 1293 (1950); and Ref. 60 below. See also D. R. Thompson and O. K. Rice quoted in Ref. 11 below.

[2] J. S. Rowlinson, *Liquids and Liquid Mixtures* (Butterworths Scientific Publications, Ltd., London, 1959), Chap. 5.

[3] Maxwell's thermodynamic derivation of the rule is unsatisfactory in that it is necessary to give thermodynamic significance to "unstable" (and "metastable") states on the original van der Waals isotherm. A formally equivalent but theoretically somewhat more convincing argument is to use the minimal properties of the Gibbs free energy (or chemical potential) to eliminate the unwanted parts of the van der Waals isotherm.

which is essentially the reciprocal of this slope, diverges to infinity at the critical point.

The derivation of the van der Waals equation given by Ornstein[4] suggests that it should be a reasonably good approximation when the pair interaction potential $\phi(r)$ of the molecules of the fluid has a short range strongly repulsive core and a very long range weakly attractive tail. Indeed Kac, Uhlenbeck, and Hemmer[5] have recently shown that the van der Waals isotherm (*with* flat part) follows rigorously for a one-dimensional gas of hard rods interacting with an attractive exponential potential *in the limit* that the exponential becomes infinitely long-ranged *and* infinitely weak [holding $\int_0^\infty \phi(r)\,dr$ constant]. It seems probable, however, that the behavior of a real gas *in its critical region* is crucially dependent on the finite or relatively *short* range of the attractive parts of more realistic potentials. This is supported by the comparison with real systems and exactly soluble models we present in the next section.

Three principal predictions concerning the critical region which follow from the van der Waals equation are:

(a) that the *coexistence curve* follows a square-root law, i.e., the difference between liquid and gaseous densities vanishes as

$$\rho_L - \rho_G \approx A(T_c - T)^{\frac{1}{2}} \qquad (T \to T_c-); \tag{1.3}$$

(b) that the *compressibility* along the critical isochore diverges as a simple pole,

$$K_T \approx B/|T - T_c| \qquad (\rho = \rho_c,\ T \to T_c+); \tag{1.4}$$

and

(c) that the *specific heat* (at constant volume) along the critical isochore rises to a maximum and then falls discontinuously as T increases through T_c, i.e.,

$$C_V(T) \approx C_c^{\pm} - D^{\pm}\,|T - T_c|, \qquad T \gtrless T_c, \tag{1.5}$$

with $C_c^- - C_c^+ = \Delta C > 0$.

The compressibility of the gas and of the liquid along the coexistence curve (i.e. at condensation) also diverges as a simple pole as $T \to T_c-$ according to the van der Waals equation but the amplitude corresponding to B in (1.4) is smaller. [The constants A, B, $C_c^{\pm}$ and $D^{\pm}$ can of course be written explicitly in terms of the van der Waals parameters a and b.]

It is important to note that these predictions are not peculiar to the van der Waals equation but follow from almost *all* approximate equations of state. Indeed they are essentially a direct consequence of the implicit or explicit assumption that the free energy and the pressure can be expanded in a Taylor series in density and temperature *at* the critical point: in other words that the critical point is not a singular point of the free energy expressed as a function of ρ and T (except in as far as Maxwell's rule is utilized *below* $T = T_c$).[6]

To demonstrate this put

$$\Delta p = p - p_c, \quad \Delta\rho = \rho - \rho_c, \quad \Delta T = T - T_c \tag{1.6}$$

and assume that

$$\Delta p = a(T) + b(T)\Delta\rho + c(T)\Delta\rho^2 + d(T)\Delta\rho^3 + \cdots. \tag{1.7}$$

By definition $a(T_c) = 0$ so we assume similarly that

$$a(T) = a_1\Delta T + a_2\Delta T^2 + \cdots. \tag{1.8}$$

Since the compressibility is infinite at the critical point, $b(T_c)$ must vanish so again we assume

$$b(T) = b_1\Delta T + b_2\Delta T^2 + \cdots. \tag{1.9}$$

Finally since $b(T_c) = 0$ and the pressure above T_c must be a monotonic increasing function of ρ we have $C(T_c) = 0$ and, presumably, $C(T)$ is small in the critical region. We thus obtain for small $\Delta\rho$ and ΔT the isotherms

$$\Delta p = a_1\Delta T + b_1\Delta T\Delta\rho + d_0\Delta\rho^3 + \cdots. \tag{1.10}$$

Near the critical point the compressibility is hence given by

$$K_T \simeq \frac{(b_1\rho c)^{-1}}{\Delta T + (3d_0/b_1)\Delta\rho^2}, \tag{1.11}$$

from which the prediction (b) follows. Application of the equal area rule to the isotherms (1.10) with negative ΔT yields the coexistence curve

$$\Delta\rho^2 \simeq (b_1/d_0)\,|\Delta T|, \qquad (T \leq T_c), \tag{1.12}$$

which implies the square-root law (a). The divergence

[4] L. S. Ornstein, Dissertation, Leiden 1908; see also Ref. 5.

[5] M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. 4, 216 (1963). It should be noted that the correction to ideal gas behavior, arising from the hard core and represented by the parameter b, is exact in one dimension but only approximate in two or three dimensions. On the other hand the accuracy of the correction represented by the parameter a depends mainly on the long-range nature of the attractive tail of the potential and not so directly on dimensionality. Consequently although the behavior of a three-dimensional model in the corresponding long-range limit should be van der Waals-*like* [in that Eqs. (1.3) to (1.5) should hold] one should still not expect van der Waals' equation (1.1) to hold precisely.

[6] See, for example, Landau's theory of the critical point and second-order phase transitions [L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, Ltd., London, 1958), pp. 259–268 and pp. 434–439].

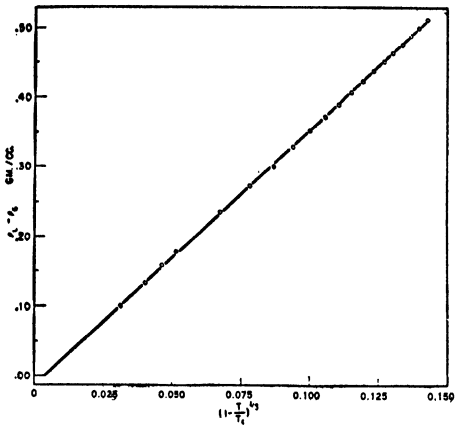


FIG. 2. Coexistence curve for xenon: plot of $(\rho_L - \rho_G)$ vs $[1 - (T/T_c)]^{\frac{1}{3}}$ (after Weinberger and Schneider[9]).

of the compressibility along the coexistence curve as $(T_c - T)^{-1}$ then follows from (1.10).

By integrating (1.9) with respect to Δp one finds the free energy $F(T, \rho)$. If one assumes that the additive constant of integration is also a nonsingular function of T and imposes the continuity of F on the coexistence curve (1.12) one finally derives the prediction (c) of the discontinuity in $C_V(T)$.

2. THE NATURE OF THE CRITICAL SINGULARITIES

Perhaps the most striking test of the predictions of the classical theories is provided by the data on the coexistence curves of simple gases. Some time ago Guggenheim[7] showed that the gases Ne, Ar, Kr, Xe, N_2, and O_2 obey closely a law of corresponding states of the form

$$(\rho_L - \rho_G)/2\rho_c = A(1 - T/T_c)^{\beta}, \quad (T \to T_c), \qquad (2.1)$$

with $\beta = \frac{1}{3}$. This relation was observed to hold with an accuracy of 0.5% (or better) in $\Delta\rho/\rho_c$ and in $\Delta T/T_c$ and over a range of temperatures from $T/T_c \simeq 0.6$ up to within $\frac{1}{2}$% of the critical temperature by when $(\rho_L - \rho_G)/2\rho_c$ had fallen to about 0.30. In earlier work on CO_2 Michels, Blaisse, and Michels[8] found that the coexistence curve could be fitted over a similar range but with somewhat greater accuracy by (2.1) with the index $\beta = 0.357$.

These results seem to be in clear disagreement with the classical prediction $\beta = \frac{1}{2}$ which should be applicable in range of $(T - T_c)$ and $(\rho - \rho_c)$ observed. They certainly show that the experimental coexistence curve is much flatter than the van der Waals curve. However, experiments near the critical region are very difficult to perform; a long time is needed to establish equilibrium and hysteresis phenomena are difficult to avoid; the system is very susceptible to minute amounts of impurities and, due to the large compressibility, highly sensitive to gravitational fields. Indeed as shown by Weinberger and Schneider[9] it is important to take special precautions to reduce the effects of gravity if the true shape of the coexistence curve is to be measured close to T_c. In a very careful study of xenon (see Fig. 2) they extended the density measurements down to temperatures differing from T_c by only 1 part in 30 000 (i.e., $\Delta T/T_c \simeq 0.003$%) and down to corresponding density differences of $(\rho - \rho_c)/2\rho_c \simeq 0.04$. [The temperature was controlled to within ± 0.001 °C.] Their data accurately obey the relation (2.1) with a nonclassical value of the index β over about three decades in $(T - T_c)$. Analysis of their measurements indicates[10] that

$$\beta = 0.345 \pm 0.015 \qquad (2.2)$$

which is not inconsistent with a value of exactly $\frac{1}{3}$ as may be seen in Fig. 2.

While it is always possible that measurements taken much closer still to the critical point might yet yield the value $\beta = \frac{1}{2}$ it seems reasonable to conclude that the classical theory does *not* provide the correct description of reality. Furthermore measurements of the phase boundaries of binary fluid mixtures near both their upper and lower critical points are also fitted well by the same cube-root law.[11] This suggests that the behavior close to a critical point is insensitive to the detailed nature of the intermolecular forces. To check how far β is really independent of the interaction potentials it would be desirable to have measurements on other systems of an accuracy matching the experiments on xenon. One should note, however,

[7] E. A. Guggenheim, J. Chem. Phys. **13**, 253 (1945).

[8] A. Michels, B. Blaisse, and C. Michels, Proc. Roy. Soc. (London) **A160**, 358 (1937). More accurate measurements on CO_2 have since been made by H. L. Lorentzen, Acta Chem. Scand. **7**, 1335 (1953) and later work.

[9] M. A. Weinberger and W. G. Schneider, Can. J. Chem. **30**, 422 (1952).

[10] Note that the plot of $\rho_L - \rho_G$ versus $[1 - (T/T_c)]^{1/3}$ in Fig. 2 is a good straight line down to $(\rho_L - \rho_G)/2\rho_c = 0.04$ but does not extrapolate exactly to the origin [$\rho_L = \rho_G$ at $T = T_c$] as it should. This suggests that the index β is not precisely $\frac{1}{3}$ and a log-log plot leads to the value quoted. The uncertainty reflects the spread of the experimental points about the best straight line.

[11] See Ref. 2, pp. 165–166 and especially the work of O. K. Rice referred to therein. See, also, D. R. Thompson and O. K. Rice, "Shape of the coexistence curve in the perfluoromethylcyclohexane-carbon tetrachloride system, II." J. Am. Chem. Soc. (1964) in press.

TABLE I. Critical indices.

Index	below T_c			above T_c ($\rho = \rho_c$)			
	α'	β	γ'	α	γ	ν	η
defined in equations	(2.7, 2.12)	(2.1)	(2.8)	(2.7, 2.12)	(2.8)	(3.10, 5.6)	(4.7, 5.3)
Classical Theory	$0_{\text{discon.}}$	$\frac{1}{2}$	1	$0_{\text{discon.}}$	1	$\frac{1}{2}$	0
Lattice Gases $d = 2$	$0_{\log}$	$\frac{1}{8}$	$1\frac{3}{4}$	$0_{\log}$	$1\frac{3}{4}$	1	$\frac{1}{4}$
Lattice Gases $d = 3$	≥ 0	$\simeq \frac{5}{16}$	$\geq 1\frac{1}{4}$	≥ 0, ≤ 0.2	$1\frac{1}{4}$	$\simeq 0.64$	$\simeq \frac{1}{16}$
Experiment	$\gtrsim 0_{\log}$	0.33–0.36	≥ 1.27(?)	≥ 0.1 ?	> 1.1 ?	> 0.55 ?	> 0(?)

Note: α', β and γ' are related by (2.20) and γ, ν and η by (5.7). The queries, ? and (?), indicate greater and lesser degrees of experimental doubt.

that any value of β between $\frac{1}{2}$ and $\frac{1}{4}$ is inconsistent with the assumption that the critical point is a nonsingular point of the free energy in the sense discussed in the previous section.

The shape of the coexistence curve may also be studied theoretically for more-or-less idealized models of a fluid. The only model so far sufficiently tractable to yield significant predictions in the critical region is the very simplest lattice gas in which each molecule occupies a site of a lattice to the exclusion of other molecules and interacts, attractively, only with nearest-neighboring molecules. This model is equivalent to the well known Ising model of ferromagnetism which has been studied intensively.[12,13]

The exact calculation of the free energy of the plane square lattice gas along its critical isochore (corresponding to zero magnetic field) was first achieved by Onsager.[14] In addition, however, Onsager[15] and Yang[16] were able to find the coexistence curve (corresponding to the spontaneous magnetization). They found the relation (2.1) but with the index

$$\beta = \tfrac{1}{8}. \tag{2.3}$$

This result, which implies a very flat coexistence curve, is, of course, quite inconsistent with classical theory.[17] It is important, furthermore, to note that the index β is *independent of lattice structure* for all soluble plane Ising lattices (including the triangular, honeycomb, kagomé, and checkerboard lattices[12,13]).

For three-dimensional lattice gases no rigorous theoretical results are available. Nonetheless on the basis of sufficiently long power series expansions[18] it has proved possible to draw quite accurate conclusions concerning the shape of the corresponding coexistence curves. [The coefficients are analyzed numerically with the aid of the recently introduced technique of Padé approximants.[19]]

The behavior again appears to be independent of lattice structure. [The simple, body-centered and face centered cubic lattices[20,21] and the tetrahedral (diamond) lattice[22] have been studied; the latter on the basis of the more sensitive ratio method.[23]] Dimensionality, however, is important since, in contrast to (2.3), the index β is found to lie in the range[21,22]

$$0.303 \leq \beta \leq 0.318 \tag{2.4}$$

which is consistent with the conjecture $\beta = \frac{5}{16} = 0.31250$.

It is remarkable, and perhaps unexpected, that a model as simple as a lattice gas with only nearest-neighbor interactions should yield a result for the shape of the coexistence curve so close to the experimental results (2.1) and (2.2). The agreement suggests that in the critical region the lattice gas represents rather adequately the pertinent features of a real gas. It appears that only the grosser features of the model—in particular the dimensionality and the short range of the forces—are really essential for obtaining a good description of critical behavior.

It seems probable, nonetheless, that the difference of about 0.025 between the experimental and theoretical values of β is a real discrepancy due, presumably, to the more artificial aspects of the Ising Hamiltonian which, in particular, restricts the molecules to the lattice positions. There remains

[12] A. comprehensive review of the Ising model is C. Domb, Advan. Phys. 9, Nos. 34, 35 (1960), while Ref. 13 is a brief review of more recent results.

[13] M. E. Fisher, J. Math. Phys. 4, 278 (1963).

[14] L. Onsager, Phys. Rev. 65, 117 (1944).

[15] L. Onsager, Nuovo Cimento Suppl. 6, 261 (1949); see also E. W. Montroll, R. B. Potts, and J. C. Ward, J. Math. Phys. 4, 308 (1963).

[16] C. N. Yang, Phys. Rev. 85, 808 (1952); T. D. Lee and C. N. Yang, Phys. Rev. 87, 410 (1952).

[17] Since β is the inverse of an integer it could still be possible for the free energy to be an analytic function of ρ at the critical point. In view of the logarithmic singularity in the specific heat, however, (Ref. 14 and the discussion below) this possibility seems rather remote.

[18] The expansion variable is $\exp[-V_0/kT]$ where V_0 is the depth of the well in the pair interaction potential. Details of the expansions are given in Ref. 11.

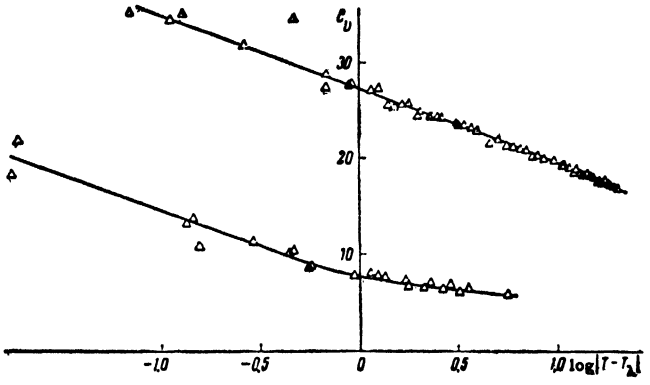
[19] G. A. Baker, Jr., J. L. Gammel, and J. G. Wills, J. Math. Anal. Appl. 2, 405 (1961).

[20] G. A. Baker, Jr., Phys. Rev. 124, 768 (1961).

[21] J. W. Essam and M. E. Fisher, J. Chem. Phys. 38, 802 (1963).

[22] J. W. Essam and M. F. Sykes, Physica 29, 378 (1963).

[23] C. Domb and M. F. Sykes, J. Math. Phys. 2, 63 (1961).



Variation of C_V of argon with $\log|T - T_c|$. ($T_c = 150.5°$ K).

FIG. 3. Variation of the constant-volume specific heat of argon along the critical isochore (after Bagastskii *et al.*[25]). [Note that the logarithm to the base ten of $|T - T_c|$ in °K is plotted.]

the theoretical problem of calculating β for more realistic continuum models. [For convenience the various results for β are collected in Table I.]

Turning now to the question of the specific heats, it has long been known that real gases exhibit a large "anomalous" specific heat maximum above T_c which lies near the critical isochore and which is not expected on classical theory.[24] Similarly in the two-phase region below the critical point specific heats rise much more rapidly than expected as T approaches T_c. From the earlier measurements one could not conclude with certainty that the specific heat actually became infinite at T_c but recent measurements by Bagatskii, Voronel', and Gusak[25] of $C_V(T)$ for argon along the critical isochore (see Fig. 3) suggest strongly that

$$C_V(T) \to \infty \quad \text{as} \quad T \to T_c\pm. \tag{2.5}$$

Such a result is again inconsistent with classical theory.

The measurements covered a range from 15% below to 5% above T_c at temperature intervals of 0.04 to 0.05°C (corresponding to $\Delta T/T_c \simeq 0.03\%$). Over a range of one or two decades in $|T - T_c|$ the specific heat could be fitted quite well[25] by a logarithmic singularity of the form

$$C_V(T) \simeq -A^* \log|1 - (T/T_c)| + B^*, \qquad (T \gtrless T_c), \tag{2.6}$$

where the marked asymmetry of the curve (see Fig. 3) indicates that $B^+ \ll B^-$ and possibly that $A^+ < A^-$. The specific heat curve in fact resembles quite closely the famous lambda anomaly displayed by liquid helium at its transition to the superfluid state.[26] [In this case the formula (2.6) is followed very accurately over four or more decades.[26] It should be noted, however, that the lambda point of helium has a quantum-mechanical origin and is not a critical point in the usual sense.]

The data for argon are not at present, accurate enough to confirm (2.6) as closely as might be wished.[26a] To avoid prejudicing the conclusions one should preferably consider a singularity of a form such as

$$C_V(T) \simeq (A^*/\alpha)\{|1 - (T/T_c)|^{-\alpha} - 1\} + B^*, \qquad (T \gtrless T_c) \tag{2.7}$$

and ask for the experimental value (and uncertainty) of the index α. [When $\alpha \to 0$ this expression reduces to the logarithmic singularity (2.6).] In particular the definite curvature of the plot of $C_V(T)$ versus $\log|T - T_c|$ for $T > T_c$ (see Fig. 3) suggests that the true value of α might be greater than say 0.1. If this is so it seems probable that below T_c the index has a somewhat different value, α' which is probably less than 0.1. It would be valuable to have similar and more extensive measurements on the other noble gases in order to test the relation

[24] See Ref. 2, pp. 100–101; Ref. 8; and A. Michels, J. M. H. Levelt, and W. de Graaff, Physica **24**, 769 (1958).

[25] M. I. Bagatskii, A. V. Voronel', and B. G. Gusak, Zh. Eksperim. i Teor. Fiz. **43**, 728 (1962) [English transl.: Soviet Phys.—JETP **16**, 517 (1963)].

[26] M. J. Buckingham and W. M. Fairbank, in *Progress in Low Temperature Physics III*, edited by C. J. Gorter (North-Holland Publishing Company, Amsterdam, 1961), Chap. 3.

[26a] More recent experiments on oxygen have given very similar results: A. V. Voronel', Yu. R. Chasshkin, V.A. Popov and V. G. Simkin, Zh. Eksperim. i Teor. Fiz. **45**, 828 (1963) [English transl.: Soviet Phys.—JETP **18**, 568 (1964)].

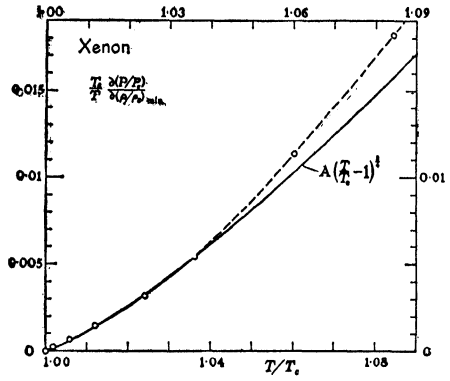


FIG. 4. Variation of the inverse of the maximum isothermal compressibility of xenon (based on the experiments of Habgood and Schneider[30]).

(2.7) more critically and to decide how far a logarithmic singularity might be truly "universal."[26a]

Significant theoretical predictions are again available only for the simple lattice gas models. Onsager's rigorous solution[14] for the plane square lattice (and subsequent results for *all* other soluble plane lattices[12,13]) yielded a *symmetric* logarithmic specific heat singularity, i.e., $\alpha = \alpha' = 0$ and $A^+ = A^-$, $B^+ = B^-$. Although this famous result demonstrates conclusively the weakness of the classical theory and is very suggestive in view of the experimental results it is, unfortunately, restricted to two-dimensional systems.

For three-dimensional Ising lattices the specific heat series expansions have been calculated both above and below T_c. Numerical analysis of these series indicates that $C_V(T)$ is almost certainly infinite at T_c but the precise nature of the divergence is more difficult to ascertain. Below the critical point the series can be fitted well by a logarithmic singularity (i.e., $\alpha' = 0$).[13,22,27,28] As with the experimental results, however, it is not easy to exclude the possibility of a slightly sharper singularity corresponding in (2.7) to, say, $\alpha' = 0.06$. On the high-temperature side of the transition the series may be fitted moderately well by a logarithmic singularity if $A^+/A^- \simeq \frac{1}{8}$.[27,13] However, careful analysis of the ratios of coefficients definitely suggests a sharper singularity of the form (2.7) with $\alpha \simeq 0.2$.[29,13] This would lead to a shallow parabolic curve for $C_V(T)$ versus $\log|T - T_c|$ not inconsistent with the corresponding (lower) experimental curve in Fig. 3.

[27] M. E. Fisher and M. F. Sykes, Physica **28**, 939 (1962).
[28] G. A. Baker, Jr., Phys. Rev. **129**, 99 (1963).
[29] C. Domb and M. F. Sykes, Phys. Rev. **108**, 1415 (1957).

One may hope that with further work based on longer series or on more rigorous arguments the theoretical conclusions will be drawn more firmly. The present results are summarized in Table I.

The accurate experimental measurement of the isothermal compressibility of a gas near its critical point is not easy and the classical prediction (1.4) that K_T should diverge as $(T - T_c)^{-1}$ along the isochore does not seem to have been properly tested. In practice plots of $1/K_T$ versus temperature are distinctly concave upwards in the critical region which suggests that the compressibility might diverge more sharply than a simple pole, i.e., as

$$K_T(T) \approx \frac{B}{|(T/T_c) - 1|^\gamma} \quad (\rho = \rho_c, T \to T_c) \qquad (2.8)$$

with $\gamma > 1$. By way of illustration a plot of $1/K_T$ (suitably normalized) versus T/T_c for xenon is shown in Fig. 4.[30] The experimental results indicated by circles and the broken line, were obtained by differentiation of the experimental isotherms, a procedure which is necessarily subject to appreciable uncertainty when K_T is large. The solid curve in Fig. 4 corresponds to the nonclassical prediction (2.8) with $\gamma = 1.25$ (see below) and is evidently quite consistent with the experimental points near T_c. This fit cannot be considered very significant, however, since estimates of γ based on the data alone are rather indefinite, although they do seem to indicate that γ is greater than 1.1.

More extensive and accurate experimental data would be extremely valuable since the compressibility is an important theoretical parameter and one which is usually somewhat easier to calculate than the specific heat or the coexistence curve. Indeed the theoretical situation for the simple lattice gases is quite unequivocal (see Table I). On the basis of Onsager and Kaufman's exact calculations[31] of the correlation functions it can be shown[32] that the compressibility of the plane square lattice gas (which is isomorphic to the magnetic susceptibility) should diverge as (2.8) with

$$\gamma = 1\tfrac{3}{4}. \qquad (2.9)$$

This represents a very large deviation from the classical prediction $\gamma = 1$. Numerical examination of the corresponding series expansions confirms (2.9) in two dimensions for all other lattices and leads,

[30] Figure 4 is based on the measurements of H. W. Habgood and W. G. Schneider, Can. J. Chem. **32**, 98 (1954) as presented in their Fig. 4.
[31] B. Kaufman and L. Onsager, Phys. Rev. **76**, 1244 (1949).
[32] See Sec. 5 and M. E. Fisher, Physica **25**, 521 (1959).

in three dimensions, to the estimate,

$$\gamma = 1.250 \tag{2.10}$$

which is again independent of lattice structure and accurate to ± 0.001 or better.[12,20,23]

The compressibility of the simple lattice gases can also be studied on the coexistence curve below T_c.[21] In two dimensions it is found that (2.8) holds with the index $\gamma' = \frac{7}{4}$ so that $\gamma' = \gamma$ although the amplitude B^- is much smaller than B^+. In three dimensions the analysis yields $\gamma' \simeq 1.25 = \gamma$ but cannot at present exclude the possibility that γ' exceeds γ (by perhaps 0.05).

On the basis of heuristic arguments related to the Frenkel–Bijl–Band picture of condensation[33] Essam and Fisher[21,34] conjectured that the indices for the specific heat, coexistence curve and compressibility below T_c were in general related by

$$\alpha' + 2\beta + \gamma' = 2, \tag{2.11}$$

where the index α' is defined, more precisely, by

$$\alpha' = \lim_{T \to T_c-} [\log C_V(T)/|\log (T_c - T)|], \tag{2.12}$$

and similarly for β and γ'. A logarithmic specific heat still corresponds to $\alpha' = 0$ so that the relation (2.11) is certainly verified for the two-dimensional lattice gases (see Table I). The formula remains true even for a van der Waals gas since a discontinuity in specific heat is now also equivalent to $\alpha' = 0$ [assuming only $C_V(T_c-) > 0$].

Rushbrooke[35] has shown that the relation (2.11) can be proved as an inequality (with $\geq$ replacing $=$) by purely thermodynamic reasoning. His argument as presented applies only to a ferromagnetic system but it may be adapted for a fluid as follows.

Firstly recall that in the two-phase region the specific heat at constant *total* volume is related to the properties of the system in its liquid and gaseous phases separately by[36]

$$C_V = x_L C_\sigma^L + x_G C_\sigma^G - T\left(\frac{\partial p}{\partial T}\right)_\sigma \left[x_L \left(\frac{\partial V_L}{\partial T}\right)_\sigma + x_G \left(\frac{\partial V_G}{\partial T}\right)_\sigma \right], \tag{2.13}$$

where the subscript σ denotes properties along the coexistence curve and where the mole fractions are given by

$$x_L = \frac{V_G - V}{V_G - V_L}, \qquad x_G = \frac{V - V_L}{V_G - V_L}. \tag{2.14}$$

Now by the standard argument used to relate C_p and C_V one may show that in a single phase

$$C_\sigma = C_V - T\left(\frac{\partial p}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_p \left(\frac{\partial V}{\partial T}\right)_\sigma. \tag{2.15}$$

On eliminating the factor $(\partial V/\partial T)_p$ through

$$\left(\frac{\partial V}{\partial T}\right)_\sigma = \left(\frac{\partial V}{\partial T}\right)_p + \left(\frac{\partial V}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_\sigma, \tag{2.16}$$

and substituting for C_σ^L and C_σ^G separately in (2.12) one expresses C_V in terms of C_V^L and C_V^G. Finally on introducing the coexisting densities ρ_L and ρ_G and the corresponding isothermal compressibilities K_T^L and K_T^G, one gets

$$C_V(T) = x_L C_V^L + x_G C_V^G + \frac{x_L T}{\rho_L^3 K_T^L}\left(\frac{\partial \rho_L}{\partial T}\right)^2 + \frac{x_G T}{\rho_G^3 K_T^G}\left(\frac{\partial \rho_G}{\partial T}\right)^2 \tag{2.17}$$

Now C_V^L and C_V^G are necessarily positive since they are essentially mean square energy fluctuations. Consequently all terms on the right of (2.17) are positive and by dropping the first three we obtain the inequality

$$C_V(T) \geq \frac{x_G T}{\rho_G^3 K_T^G}\left(\frac{\partial \rho_G}{\partial T}\right)^2. \tag{2.18}$$

As the critical point is approached at constant density x_G (and x_L) approaches the value $\frac{1}{2}$, ρ_G (and ρ_L) tends to ρ_c and $(\partial \rho_G/\partial T)$ diverges as $(T_c - T)^{-1+\beta}$.[37] If K_T^G, the compressibility at condensation, diverges as $(T_c - T)^{-\gamma'}$ we obtain

$$\log C_V(T) \geq (2 - 2\beta - \gamma')\,|\log (T_c - T)| + \cdots. \tag{2.19}$$

The higher order terms vanish on dividing by $|\log (T_c - T)|$ and taking the limit $T \to T_c-$ which, by (2.12), yields the index α'. We have thus proved quite generally

$$\alpha' + 2\beta + \gamma' \geq 2. \tag{2.20}$$

Various consequences follow from this inequality. For the two-dimensional lattice gases the rigorous results[14–16] $\alpha' = 0$ and $\beta = \frac{1}{8}$ show that $\gamma' \geq \frac{7}{4}$ thereby confirming the numerical estimates. [As before, the van der Waals gas corresponds to the case of equality.] If, for a three-dimensional lattice gas, the values $\beta = 0.3125$ and $\gamma' = 1.25$ are

[33] J. Frenkel, *Kinetic Theory of Liquids* (Oxford University Press, London, 1946), Chap. VII; J. Chem. Phys. **7**, 200, 538 (1939); A. Bijl, Doctoral Dissertation, Leiden, 1938; W. Band, J. Chem. Phys. **7**, 324, 927 (1939).

[34] Other conjectures relating the indices γ and β have been made by B. Widom [J. Chem. Phys. **37**, 2703 (1962)].

[35] G. S. Rushbrooke, J. Chem. Phys. **39**, 842 (1963).

[36] See, for example, Ref. 2, p. 41.

[37] We assume (as is true in reality and for the models considered) that as $T \to T_c$ $[\frac{1}{2}(\rho_L + \rho_G) - \rho_c]$ does not vanish as rapidly as does $(\rho_L - \rho_G)$.

accepted one must conclude that $\alpha' \simeq 0.125$. Conversely for a logarithmic specific heat singularity the compressibility index γ' would have to exceed 1.25 by about 0.1. At present it is difficult to judge between these alternatives.[31,35]

For a real gas the evidence of Fig. 3 suggests $0.1 > \alpha' \geq 0$ and the coexistence data indicate $\beta < 0.36$. Consequently we should certainly have

$$\gamma' \geq 1.27, \tag{2.21}$$

and probably $\gamma' \geq 2 - 2(0.345) = 1.31$. It would be most interesting to have an experimental test of this prediction since it differs appreciably from the classical result.

It should be mentioned that Widom and Rice[36] have observed that the critical isotherms of real gases also deviate significantly from the classical prediction being much flatter than the cubic (p, ρ) curve which follows from the van der Waals equation. As yet however, this feature has not been investigated theoretically for lattice gases.

3. PAIR CORRELATION FUNCTION AND CRITICAL SCATTERING

To obtain insight into the microscopic nature of a fluid in the critical region it is natural to consider the many-particle distribution functions $n_s(\mathbf{r}_1, \cdots \mathbf{r}_s)$ which describe the correlations between the constituent molecules. In particular the pair correlation function defined, for a uniform system, by

$$g(\mathbf{r}_{12}) = n_2(\mathbf{r}_1, \mathbf{r}_2)/n_1(\mathbf{r}_1)n_1(\mathbf{r}_2) = n_2(\mathbf{r}_{12})/\rho^2 \tag{3.1}$$

is of central importance. When the system is in one phase $g(r) \to 1$ as $r \to \infty$ and one may introduce the net correlation function

$$G(\mathbf{r}) = g(\mathbf{r}) - 1 \tag{3.2}$$

which decays to zero as $r \to \infty$. The deviation of $G(\mathbf{r})$ from zero is a direct measure of the influence of one molecule on another.

As is well known $G(\mathbf{r})$ is rather directly related to the thermodynamic variables of the system. For our purposes the most important result is the so called fluctuation theorem for the isothermal compressibility,

$$k_B T\left(\frac{\partial \rho}{\partial p}\right)_T = k_B T \rho K_T = 1 + \rho \int G(\mathbf{r})\, d\mathbf{r}, \tag{3.3}$$

which is a quite general consequence of the laws of statistical mechanics. (For a two-dimensional system the integral in (3.3) is restricted appropriately while for a lattice system it is replaced by a sum.)

Now $G(\mathbf{r})$ is essentially a bounded function (more precisely its integral over a finite region is bounded in virtue of the existence of a maximum density arising from the incompressibility of real molecules.) Hence the fact that K_T becomes infinite at the critical point can only be understood if the integral over $G(\mathbf{r})$ diverges at its upper limits. This means that at the critical point the net correlation function becomes *long-range* in the sense that it decays to zero more slowly than $1/r^3$ (or in d-dimensions than $1/r^d$). It is clearly of interest to know the precise nature of this critical decay and to understand the rate of approach of $G(\mathbf{r})$ to its long-range behavior as $T \to T_c$.

Fortunately the pair correlation function can also be studied directly by scattering waves off the system. In practice experiments are usually performed with light or with x rays but thermal neutrons may also be used.[39] The observed angular disymmetry is then a direct measure of the degree of correlation. To the extent that multiple scattering may be neglected (first Born approximation) we have for the scattering intensity

$$I(\mathbf{k})/I_0(\mathbf{k}) = \chi(\mathbf{k}) = 1 + \rho \hat{G}(\mathbf{k}), \tag{3.4}$$

where $I_0(\mathbf{k})$ is the scattering intensity in the absence of correlation (the molecular form factor), $\mathbf{k}$ is the wave vector, $k = (4\pi/\lambda) \sin \frac{1}{2}\theta$, and where

$$\hat{G}(\mathbf{k}) = \int e^{i\mathbf{k}\cdot\mathbf{r}} G(\mathbf{r})\, d\mathbf{r} \tag{3.5}$$

is the Fourier transform of $G(\mathbf{r})$; for an isotropic three-dimensional system

$$\hat{G}(k) = 4\pi \int_0^\infty \frac{\sin kr}{kr} G(r)\, r^2\, dr. \tag{3.6}$$

The ratio $\chi(k)$ may be regarded as a generalized "susceptibility" since it measures the response of the fluid to an impressed periodic potential of wave number k.[40]

Comparison of (3.5) and (3.4) with the fluctuation relation (3.3) shows that

$$\chi(0) = \lim_{k \to 0} I(k)/I_0(k) = 1 + \rho \hat{G}(0) = k_B T \rho K_T, \tag{3.7}$$

so that the scattering intensity extrapolated to zero angle is proportional to the isothermal compress-

[36] B. Widom and O. K. Rice, J. Chem. Phys. 23, 1250 (1955).

[39] L. Van Hove, Phys. Rev. 95, 249 (1954). To discuss neutron scattering fully one must also consider the time dependence of the pair correlation function. Near the critical point, however, the decay of fluctuations probably becomes slower and it is reasonable to neglect this aspect of the problem in first approximation.

[40] P. G. de Gennes, Nuovo Cimento 9, Suppl. 1, 240 (1958).

ibility. By virtue of the divergence of K_T at the critical point the low-angle scattering must thus become very large as the critical point is approached in the one phase region. This is the "anomalous" critical scattering long known with visible light as *critical opalescence*. In physical terms one may say that the large compressibility near the critical point allows long-wavelength density fluctuations to grow to large amplitude and these produce visible diffraction.[39]

The classical theory of critical scattering is that developed by Ornstein and Zernike.[41,42] Their results have since been rederived many times and in various ways. In order to discuss the validity of their conclusions we will outline two approaches which characterize most of the derivations: on the one hand the original method of Ornstein and Zernike[41,43] which is perhaps the more mathematical, and on the other hand a semithermodynamic method which concentrates attention on the fluctuations of the free energy and their relation to the gradients of the density deviations. This latter approach was initiated by Rocard[44-46] in the spirit of Einstein's semiphenomenological ideas, but we will follow the presentation of Landau.[46]

Ornstein and Zernike[41] argue on a heuristic basis that the correlation $G(\mathbf{r}_1 - \mathbf{r}_2)$ between molecules 1 and 2 can be regarded as caused by (i) a *direct* influence of 1 on 2 described by the so-called "direct correlation function" $C(\mathbf{r}_1 - \mathbf{r}_2)$ which should be short-ranged [essentially having the range of the pair potential $\phi(r)$], and (ii) an indirect influence propagated directly from 1 to a third molecule at $\mathbf{r}_3$ which in turn exerts its total influence on molecule 2. Integrating over $\mathbf{r}_3$ they thus write the relation

$$G(\mathbf{r}_1-\mathbf{r}_2)=C(\mathbf{r}_1-\mathbf{r}_2)+\rho\int C(\mathbf{r}_1-\mathbf{r}_3)G(\mathbf{r}_3-\mathbf{r}_2)\,d\mathbf{r}_3. \quad (3.8)$$

In the absence of an independent theory enabling one to calculate $C(\mathbf{r})$ in terms of the molecular parameters this relation is really only a *definition* of the direct correlation function: we will, in the main, adopt this attitude. However, Ornstein and Zernike regarded $C(\mathbf{r})$ as the more basic function (in its, presumably, closer relation to the intermolecular forces) and they contemplated the possibility of calculating $C(\mathbf{r})$ directly.

On introducing the Fourier transform of $C(\mathbf{r})$ the relation (3.8) (which states that $G(\mathbf{r}_1 - \mathbf{r}_2)$ and $C(\mathbf{r}_1 - \mathbf{r}_2)$ are reciprocal kernels in the sense of the theory of integral equations) can be solved to yield

$$1 + \rho\hat{G}(\mathbf{k}) = 1/[1 - \rho\hat{C}(\mathbf{k})]. \quad (3.9)$$

On substituting in (3.4) one finds for the inverse scattering intensity

$$1/\chi(\mathbf{k}) = 1 - \rho\hat{C}(\mathbf{k}). \quad (3.10)$$

Consequently the divergence of the compressibility $K_T = \chi(0)/k_BT\rho$ at the critical point is associated with the equation

$$1 - \rho\hat{C}(0) = 1 - \rho\int C(\mathbf{r})\,d\mathbf{r} = 0. \quad (3.11)$$

This shows that the integral of $C(\mathbf{r})$ (i.e., its zeroth moment) remains finite at the cirtical point. Thus $C(\mathbf{r})$ certainly decays to zero more rapidly than $G(\mathbf{r})$ thereby confirming the expectation that it should be relatively short-ranged. To develop the theory further, however, one makes the central assumption that $C(\mathbf{r})$ is strictly short-ranged *at* (and near) the critical point in the sense that its transform $\hat{C}(\mathbf{k})$ has a Taylor series expansion in powers of k^2. In particular, one assumes that the second moment

$$R^2 = \tfrac{1}{2}\rho\langle\cos^2\theta\rangle\int r^2C(\mathbf{r})\,d\mathbf{r}, \quad (3.12)$$

exists at the critical point and does not vary rapidly in the vicinity. [In three dimensions the average over angles yields $\langle\cos^2\theta\rangle = \frac{1}{3}$.]

We note again that unless $C(r)$ can be calculated in an independent way R^2 will have the status only of a semiphenomenological parameter. The short-range character of $C(r)$ and the existence of R^2 at the critical point constitute a major problem of the theory and we will return to it. We remark here, however, that the virial expansion of $C(r)$ may be obtained quite easily from that for $G(r)$[42,43] and indicates for low densities at least, that $C(r)$ is, in a definite sense, shorter ranged than $G(r)$.

This may be seen in terms of the graphical representations of the respective expansions. All the graphs required[47,48] are connected and have two

[41] L. S. Ornstein and F. Zernike, Proc. Acad. Sci. Amsterdam 17, 793 (1914); Physik. Z. 19, 134 (1918); *ibid.* 27, 761 (1926).

[42] F. Zernike, Proc. Acad. Sci. Amsterdam 18, 1520 (1916).

[43] See also J. Yvon, Nuovo Cimento 9, Suppl. 1, 144 (1958) and Ref. 40.

[44] Y. Rocard, J. Phys. Radium 4, 165 (1933).

[45] See also: M. Fierz in *Theoretical Physics in the Twentieth Century*, edited by M. Fiers and V. F. Weisskopf (Interscience Publishers, Inc., New York, 1960), pp. 175 *et seq.*; M. J. Klein and L. Tisza, Phys. Rev. 76, 1861 (1949).

[46] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Pergamon Press, Ltd., London, 1958), Sec. 116.

[47] See, for example, E. Meeron, J. Math. Phys. 1, 192 (1960), J. M. J. Van Leeuwen, J. Groeneveld, and J. De Boer, Physica 25, 792 (1959); T. Morita and K. Hiroike, Prog. Theoret. Phys. (Kyoto) 23, 1003 (1960).

[48] G. E. Uhlenbeck and G. W. Ford, in *Studies in Statistical Mechanics, I*, edited by J. De Boer and G. E. Uhlenbeck, (North-Holland Publishing Company, Amsterdam, 1962), Chap. B. III 4.

fixed points (corresponding to the two molecules fixed at distance $\mathbf{r}$ apart) and $n = 0, 1, 2, \cdots$ field points. Each point is associated with a power of the density ρ while the bonds are, as usual, associated with the Mayer f factors $f_{ij} = \exp[-\phi(r_{ij})/kT] - 1$. If the potential is of strictly finite range b [in the sense that $\phi(r) = 0$ for $r > b$] the factor f_{ij} vanishes unless $r_{ij} \leq b$. This fact restricts the distance up to which a graph of given type can "stretch." Thus in the expansion of $G(r)$ the longest-ranged term in a given order, say ρ^{2m+1}, comes from the open chain of $2m$ bonds which contributes up to distances $r = 2mb$ but not beyond. The graphs entering the expansion of $C(r)$, however, are restricted to be "nonnodal", i.e., none of the field points may be cutting points whose removal would separate the graph into two parts.[47] The open chain of bonds is thus excluded and the longest-ranged contribution to $C(r)$ in order ρ^{2m+1} comes from the graph consisting of two parallel chains of m bonds each. This graph, however, will make no contribution for $r > mb$.

We see therefore that in a given order the range of $C(r)$ is only half that of $G(r)$. It is clear, nevertheless, that one should not expect the virial series to converge well in the critical region (if it converges at all) so that the assumed short-range nature of $C(r)$ cannot be established in this way. For the present, however, we follow Ornstein and Zernike and accept the existence of the second moment R^2 and the possibility of a Taylor series expansion of $\hat{C}(k)$. From (3.10) and (3.12) we then obtain on neglecting terms of order k^4, the scattering formula

$$\chi(k) = 1 + \rho\hat{G}(k) \simeq R^{-2}/(\kappa^2 + k^2), \quad (k^2 \to 0), \tag{3.13}$$

where κ, which has the dimensions of an inverse length, is defined by

$$\kappa^2 = [1 - \rho\hat{C}(0)]/R^2. \tag{3.14}$$

Fourier inversion of the simple Lorentzian scattering curve (3.13) shows that the behavior of $G(r)$ for large r is given in *three* dimensions by the famous result

$$G(r) \simeq \frac{1}{4\pi\rho R^2}\frac{e^{-\kappa r}}{r} \qquad (r \to \infty). \tag{3.15}$$

From this one concludes that the correlations decay exponentially with an inverse range κ which, via the fluctuation theorem (3.3), should be related to the compressibility by

$$K_T = A/\kappa^2, \quad (T \to T_c). \tag{3.16}$$

The constant of proportionality is $A = 1/k_B T\rho R^2$ and is expected to be only slowly varying in the critical region. The divergence of K_T at the critical point thus implies

$$\kappa(T) \to 0 \quad \text{as} \quad T \to T_c, \tag{3.17}$$

so that the critical point correlation function is no longer exponentially damped but is predicted to follow the law

$$G_c(r) \simeq D/r \qquad (r \to \infty, T = T_c). \tag{3.18}$$

In as far as the relation $K_T \sim 1/\kappa^2$ is valid and the classical variation of $K_T(T)$ is accepted, the inverse range will go to zero along the critical isochore as

$$\kappa(T) \simeq \kappa^0 |1 - (T/T_c)|^\nu, \quad (T \to T_c), \tag{3.19}$$

with $\nu = \frac{1}{2}$. This conclusion was mentioned by Zernike[42] and is accepted by other authors.[40-52] More generally, however, if one recognizes deviations of the isothermal compressibility from van der Waals behavior one would get the "nonclassical" result $\nu = \frac{1}{2}\gamma > \frac{1}{2}$ [see Eq. (2.8)].

The principal alternative approach to the Ornstein–Zernike theory is based on considering the thermodynamic work, or change in free energy, required to establish a density fluctuation in the system,[44-46,49-52] i.e., a local inhomogeneity. One supposes that a local free-energy density $F(\mathbf{r})$ can be defined for an inhomogeneous system and considers the expansion of F (or its integral over a small but macroscopic volume) about its homogeneous mean value $\bar{F}$ in terms of the local deviation $\delta\rho(\mathbf{r})$ of the density from its mean value ρ. In the spirit of the calssical theory of the equation of state one assumes that a Taylor series exists even at the critical point. The first power of $\delta\rho$ may be dropped by virtue of the conservation of particles. The coefficient of $\delta\rho^2$ is, thermodynamically, proportional to $1/K_T$ and so this term must be retained.

Since the state of the system will be inhomogeneous, however, one must also expect terms dependent on $\nabla\rho$ the gradient of the density deviation, and on higher derivatives. The necessity for such terms can indeed be seen rather generally from the existence of surface tension which represents, of course, an additive contribution to the free energy directly associated with the density inhomogeneities at an interface. On the grounds of symmetry the leading term will be proportional to $(\nabla\rho)^2$. [The terms $\nabla^2\rho$ and $\rho\nabla^2\rho$ add nothing further after

[49] P. Debye, J. Chem. Phys. **31**, 680 (1959).
[50] E. W. Hart, J. Chem. Phys. **34**, 1471 (1961).
[51] M. Fixman, J. Chem. Phys. **33**, 1357 (1960).
[52] M. Fixman, J. Chem. Phys. **36**, 1965 (1962).

integrating over a small volume.[46]] Consequently, one writes the expansion

$$\Delta F(\mathbf{r}) = \tfrac{1}{2}c\delta\rho^2 + \tfrac{1}{2}d(\nabla\rho)^2 + \cdots \qquad (3.20)$$

and assumes that at least for small slowly varying deviations (i.e., small k), the higher-order terms may be neglected to a good approximation even at the critical point. By stability considerations c and d must be positive. Although c will vanish at the critical point d remains nonzero.

On introducing the Fourier components of the density deviation

$$\delta\hat{\rho}_{\mathbf{k}} = V^{-1}\int e^{i\mathbf{k}\cdot\mathbf{r}}\,\delta\rho(\mathbf{r})\,d\mathbf{r}, \qquad (3.21)$$

substituting in (3.20) and integrating over the volume of the system one obtains for the total free energy fluctuation

$$\Delta F_{\text{total}} = \tfrac{1}{2}V\sum_{\mathbf{k}}(c + dk^2)\,|\delta\hat{\rho}_{\mathbf{k}}|^2. \qquad (3.22)$$

We notice that each density mode contributes additively to the free energy so that the modes are statistically independent or, in other words, effectively noninteracting. This conclusion is, of course, a direct consequence of the truncation of (3.20) although it also has an immediate physical appeal for long wavelength modes.

Now the Boltzmann factor for a fluctuation $\delta\hat{\rho}_{\mathbf{k}}$ is $\exp[-\Delta F_{\mathbf{k}}/k_BT]$ and consequently the mean square fluctuation is predicted, at least for small k, to be

$$\langle|\delta\hat{\rho}_{\mathbf{k}}|^2\rangle = k_BT/V(c + dk^2). \qquad (3.23)$$

Now by the definition (3.21) of the Fourier coefficients we have

$$\langle|\delta\hat{\rho}_{\mathbf{k}}|^2\rangle = \langle\delta\hat{\rho}_{\mathbf{k}}\,\delta\hat{\rho}_{-\mathbf{k}}\rangle = V^{-1}\int e^{i\mathbf{k}\cdot\mathbf{r}}\langle\delta\rho(0)\,\delta\rho(\mathbf{r})\rangle\,d\mathbf{r},$$

where one integration over $\mathbf{r}$ space has been performed using the (approximate) translational invariance of the system. Since the mean-density fluctuation is zero

$$\langle\delta\rho(0)\,\delta\rho(\mathbf{r})\rangle/\rho = \langle[\rho + \delta\rho(0)][\rho + \delta\rho(\mathbf{r})]\rangle/\rho - \rho,$$
$$= \rho g(\mathbf{r}) + \delta(\mathbf{r}) - \rho,$$

where the second line follows from the definition (3.1) of $g(\mathbf{r})$, the delta function being added to allow for the identity of the correlated particles [which was implicitly excluded in (3.1)]. Substitution shows that

$$(V/\rho)\langle|\delta\hat{\rho}_{\mathbf{k}}|^2\rangle = 1 + \rho\hat{G}(\mathbf{k})$$
$$\simeq \frac{(k_BT/\rho)}{c + dk^2}, \qquad (3.24)$$

which is clearly equivalent to the Ornstein–Zernike result (3.13).[53]

The relationship between the two approaches to critical scattering theory is revealed by the more recent development of a complete formal theory of the statistical mechanics of nonuniform systems which shows the significance of the direct correlation function in constructing expansions of the thermodynamic variables of inhomogeneous systems.[54–58] In particular the existence of local thermodynamic variables and the convergence of the corresponding expansions turns out to be dependent on the short range nature of $C(r)$.

A fruitful method developed by Lebowitz and Percus[56,57] is to consider the deviations in density produced by an externally imposed potential $U(\mathbf{r})$ when this is chosen to be the potential $\phi(\mathbf{r})$ which would just correspond to a molecule of the fluid fixed at the origin. The induced singlet-density deviation is then related to the pair correlation function in the homogeneous fluid as follows:

$$n_1(\mathbf{r}|\phi) = \rho + \delta n_1(\mathbf{r}) = n_2^0(\mathbf{r})/n_1^0,$$
$$= \rho + \rho G(\mathbf{r}), \qquad (3.25)$$

where the superscripts zero denote the uniform system.[58]

To obtain an equation for $n_1(\mathbf{r}|\phi)$ and hence for $G(r)$ one assumes the inhomogeneous system can be represented by a grand canonical ensemble and one asks for the relation between the external potential $U(r) = \phi(r)$ and the induced-density deviation. To this end it might be natural to try to expand $n_1(\mathbf{r}|\phi)$ as a functional Taylor series in $\phi(r)$ but Lebowitz and Percus show, on the contrary, that it is possible, and more useful, to expand $\phi(r)$ (in combination with the chemical potential) in terms of the density deviation it produces.[55–57]

[53] J. L. Lebowitz (private communication) has pointed out that by the methods of Refs. 55–57 one may show that in an appropriate grand canonical ensemble the fluctuation of the free energy is given to *second order* in $\delta\hat{\rho}_{\mathbf{k}}/\rho$ exactly by

$$\Delta F_{\text{total}}/Nk_BT = \tfrac{1}{2}\sum_{\mathbf{k}}[1 - \rho\hat{C}(\mathbf{k})]\,|\delta\hat{\rho}_{\mathbf{k}}/\rho|^2.$$

Comparison with (3.22) shows even more directly the equivalence to the Ornstein–Zernike theory.

[54] F. H. Stillinger, Jr. and F. P. Buff, J. Chem. Phys. 37, 1 (1962).

[55] J. L. Lebowitz and J. K. Percus, Phys. Rev. 122, 1675 (1961); J. Math. Phys. 4, 116 (1963).

[56] J. K. Percus, Phys. Rev. Letters 8, 462 (1962).

[57] J. L. Lebowitz and J. K. Percus, J. Math. Phys. 4, 248 (1963).

[58] We use the notation $\delta n_1(\mathbf{r})$ rather than $\delta\rho(\mathbf{r})$ as previously, since in the thermodynamic arguments one really is considering a macroscopic or coarse-grained density fluctuation whereas here one refers directly to the microscopic distribution functions. By the same token one should replace $\delta\hat{\rho}_{\mathbf{k}}$ in Footnote 53 by $\delta\hat{n}_1(\mathbf{k})$ whereas in Eq. (3.24) and the preceding steps $g(\mathbf{r})$ and $\hat{G}(\mathbf{k})$ represent coarse-grained or macroscopic correlation functions.

One anticipates that such an expansion would converge most rapidly when taken at each point, about a uniform system with the same *local* density $n_1(\mathbf{r}|\phi)$. The successive terms of the expansion, which is conveniently derived by the technique of functional differentiation, are then found to be multiple integrals over the density deviation with kernels which involve $C(\mathbf{r}_1 - \mathbf{r}_2)$ and certain higher-order correlation functions.

Now when $C(\mathbf{r})$ is a short-ranged function one may expand these integrals as a "local" series in the spatial derivatives of the density deviation. The justification for this step must rest, of course, on showing that $C(r)$ decays rapidly to zero for large r [and that $n_1(\mathbf{r}|\phi)$ is not too rapidly varying]. If this is the case one finds,

$$\phi(\mathbf{r}) = \mu - \mu^0[n_1(\mathbf{r})] + (R^2 k_B T \rho^2/n_1^2)\nabla^2 G(\mathbf{r}) + \tfrac{1}{8}(l^2\rho^2/k_B T n_1^4 K_T^2)[\nabla G(\mathbf{r})]^2 + \cdots, \quad (3.26)$$

where μ is the chemical potential and R^2, defined already in equation (3.12), is the second moment of the direct correlation function. The length l is defined similarly in terms of $C(r)$ and the three-particle distribution function. Further terms of (3.26) can be written down explicitly and involve higher derivatives of $G(r)$ with coefficients depending on higher moments of $C(r)$ and the further many-particle distribution functions.[55-57]

If we now drop the terms in $(\nabla G)^2$ and higher-order derivatives and expand $n_1(r) = \rho[1 + G(r)]$ to first order in $G(r)$ we obtain

$$R^2 k_B T \nabla^2 G(\mathbf{r}) - \rho(\partial\mu^0/\partial\rho)G(\mathbf{r}) = \phi(\mathbf{r}), \quad (3.27)$$

which should be valid in the asymptotic region where $G(r)$ is small. The derivative of μ^0 arises from expanding $\mu^0[n_1(\mathbf{r})]$ and may be eliminated through the thermodynamic relation $\rho(\partial\mu^0/\partial\rho)_T = (\partial p/\partial\rho)_T = 1/\rho K_T$. On introducing the length Λ by

$$\Lambda^2 = R^2 k_B T \rho K_T = \tfrac{1}{2}\langle\cos^2\theta\rangle\rho \int r^2 G(\mathbf{r})\, d\mathbf{r}/[1 + \hat{G}(0)] \quad (3.28)$$

[where the last formula follows from (3.12), (3.10) and (3.9)] we obtain the equation

$$\nabla^2 G(r) - \Lambda^{-2} G(r) = \phi(r)/k_B T R^2 \quad (3.29)$$

first derived by Zernike.[42]

If the potential $\phi(r)$ is negligible for large r the asymptotic solution of (3.29) is, in three dimensions,

$$G(r) \approx D e^{-(r/\Lambda)}/r. \quad (3.30)$$

Comparison with (3.15) shows the equivalence to the previous theory and identifies $\Lambda = 1/\kappa$ as the range of correlation. Fourier transformation of Eq. (3.29) leads directly to the previous expression (3.13) for the scattering intensity for small κ^2 and k^2.

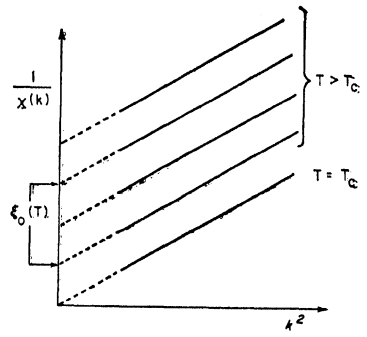


FIG. 5. Dependence of the inverse critical scattering on $k = (4\pi/\lambda)\sin\frac{1}{2}\theta$ according to the Ornstein-Zernike theory. The linear intercept $\xi_0(T)$ is proportional to $1/K_T$.

The advantage of this last derivation is that the neglected terms are explicitly displayed and that the central role of the direct correlation function is apparent. It does not, however, enable us to decide if the assumption that $C(r)$ is short-ranged at the critical point is justified. On the other hand *away* from the critical point, but in a region where K_T is moderately large the analysis indicates that the Ornstein-Zernike theory should be correct. Indeed the exponential part of the asymptotic decay law for $G(r)$ appears to have a rather wide range of validity in a one-phase fluid system since it depends essentially only on the short-range nature of the interactions.[59]

4. VALIDITY OF THE ORNSTEIN-ZERNIKE THEORY

The principal experimental predictions following from the classical theory of critical scattering [i.e., from Eq. (3.13)] are (a) that $1/\chi(k, T)$, the reciprocal of the relative scattering intensity, should vary *linearly* with k^2 with a temperature-independent coefficient of proportionality, and (b) that the extra-

[59] This conclusion follows from a formulation of statistical mechanics in which the system is taken to be in a cylinder of length L and cross section A. For forces with a hard core and strictly finite range b, a nonsingular integral kernel describes the addition of a layer of thickness b to the cylinder. The thermodynamics and correlation functions are related to the resolvent of this kernel: M. E. Fisher, abstract in Proc. Second Eastern Theoretical Physics Conf., University of North Carolina (October, 1963). D. Ruelle and, independently, J. Groeneveld (to be published) have also shown that in the region where the activity expansion can be proved to converge, $\hat{G}(r)$ [suitably smoothed] decays to zero at least exponentially fast for strictly short range potentials.

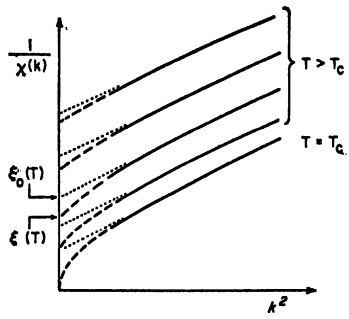


FIG. 6. Schematic variation of the inverse critical scattering expected in view of the limitations of the classical theory. (Compare with Fig. 5.) The apparent linear intercept $\xi_0(T)$ differs from the true intercept $\xi(T)$ which is proportional to $1/K_T$.

polated intercept with the $k^2=0$ axis, $\xi_0(T)$, should be proportional to $1/K_T(T)$ and hence go to zero as $T \to T_c$. The predicted behavior is shown schematically in Fig. 5.

Most modern tests of the theory have been made by light or x ray scattering measurements on binary fluid mixtures of relatively complex organic molecules.[60-65] For example, Zimm[60] studied a mixture of perfluoromethylcyclohexane in carbon tetrachloride, Brady and Frisch,[62] perfluoroheptane in iso-octane, while Debye and co-workers[63] and McIntyre, Wims and Green[64] have investigated polystryrene-cyclohexane solutions.

Earlier measurements on binary fluids and other systems have been reviewed by Rice.[65] Recently Thomas and Schmidt[66] have made an extensive series of x-ray measurements on argon at various constant pressures in the critical region. They also give references to other more recent work on single-component systems such as carbon dioxide, ethylene and neon.

Accurate experiments are not easy to perform near the critical point and the interpretation can be confused by multiple scattering. Qualitatively, however, the theory seems quite well confirmed by the binary fluid measurements. Plots of the reciprocal scattering intensity versus k^2 in the experimentally accessible range are well represented by sets of parallel straight lines whose intercepts $\xi_0(T)$ fall roughly linearly with $(T - T_c)$ as indicated in Fig. 5. Closer inspection, nevertheless, reveals certain "anomalies", although at present these are not much larger than the experimental uncertainties.[64,67,68] In particular one observes (*a*) a tendency for the scattering curves taken near T_c to be slightly curved and to dip downwards somewhat at the lowest values of k^2 and (*b*) the intercepts $\xi_0(T)$ obtained by extrapolation of the best straight line fits to the data (all necessarily lying above some k^2_{min}) do not seem to approach zero as T goes to T_c: rather plots of $\xi_0(T)$ versus T along the isochore are slightly concave upwards and tend to level off or to extrapolate to a nonzero value at T_c. These deviations are indicated in Fig. 6.

Green[68] reports that significant deviations from the Ornstein–Zernike theory were also found for the, presumably physically simpler, system of nitrogen at its critical point.[69] No deviations were observed by Thomas and Schmidt[66] in argon but they did not measure along the critical isochore and their lowest values of k^2 were relatively large.

As we show below these deviations are in the direction to be expected theoretically on the basis of an analysis of the limitations of the classical theory. Needless to say, however, it would still be desirable to have more accurate and extensive experimental data, especially for low values of k^2 and for simple systems like the noble gases, in order to elucidate fully the true nature of the critical scattering.

As we have seen the main theoretical problem in justifying the derivation of the Ornstein–Zernike result is to establish the short-range nature of the direct correlation function $C(r)$, or, what is equivalent, to show that its Fourier transform $\hat{C}(k)$ has a Taylor series expansion in powers of k^2 *at* the critical point. Our previous discussion of the thermodynamic variables at the critical point, has shown that the hypothesis of a Taylor expansion in temperature or density is probably not tenable. By analogy we should be prepared for a similar failure for the correlation functions.

An obvious defect of the theory can be seen by

[60] B. H. Zimm, J. Phys. & Colloid Chem. 54, 1306 (1950).
[61] Chow Quantie, Proc. Roy. Soc. (London) A224, 90 (1954); R. Fürth and C. L. Williams, Proc. Roy. Soc. (London) A224, 104 (1954).
[62] G. W. Brady and H. L. Frisch, J. Chem. Phys. 35, 2234 (1961).
[63] P. Debye, H. Coll, and D. Woermann, J. Chem. Phys. 32, 939 (1960); *ibid* 33, 1746 (1960); P. Debye, D. Woermann and B. Chu, J. Chem. Phys. 36, 851 (1962).
[64] D. McIntyre, A. Wims and M. S. Green, J. Chem. Phys. 37, 3019 (1962).
[65] O. K. Rice, "Critical Phenomena" Sec. E, in *Thermodynamics and Physics of Matter*, edited by F. D. Rossini (Princeton University Press, Princeton, New Jersey, 1955).
[66] J. E. Thomas and P. W. Schmidt, J. Chem. Phys. 39, 2506 (1963).
[67] H. L. Frisch and G. W. Brady, J. Chem. Phys. 37, 1514 (1962).
[68] M. S. Green, J. Chem. Phys. 33, 1403 (1960).
[69] R. L. Wild, J. Chem. Phys. 18, 1627 (1950).

considering its application to model systems of dimensionality d different from three. In any number of dimensions the classical expression for the (appropriate) Fourier transform is formally the same, namely

$$\chi(\mathbf{k}) = 1 + \rho\hat{G}(\mathbf{k}) \simeq A/(\kappa^2 + k^2). \qquad (4.1)$$

For fixed $\kappa(T) > 0$, that is *away* from the critical point, we find[70] by inverting (4.1) that as r becomes very large

$$G(r) \simeq B_d(e^{-\kappa r}/r^{\frac{1}{2}(d-1)})[1 + O(1/\kappa r)],$$
$$(r \to \infty,\ \kappa \text{ fixed} > 0). \qquad (4.2)$$

As we argued at the end of the previous section it seems probable that this result is generally valid for *fixed* $T > T_c$ and large enough r. However, as $\kappa \to 0$ for *fixed* large r we find a different result,[70] namely, for $d \geq 3$

$$G(r) \simeq D_d(e^{-\kappa r}/r^{d-2})[1 + O(\kappa r)],$$
$$(\kappa \to 0,\ r \text{ fixed}) \qquad (4.3)$$

while for $d = 2$

$$G(r) \simeq D_2(\log r)e^{-\kappa r}[1 + O(1/\log \kappa r)],$$
$$(\kappa \to 0,\ r \text{ fixed}). \qquad (4.4)$$

One notices that $d = 3$ is a rather special case in which both limits (4.2) and (4.3) agree and [by comparison with equation (3.15)] the higher-order terms vanish identically!

Now (4.4) implies that at the critical point of a two-dimensional system the correlation function will vary as $D_2 \log r$. This is clearly nonphysical for large r and shows that the assumptions of the theory are certainly not to be trusted for two-dimensional systems.

This defect of the theory can be repaired in an *ad hoc* fashion by retaining further (nonlinear) terms in the density expansions of the free energy [in Eq. (3.20)] or in the Taylor series expansion of $n_1(\mathbf{r})$ [in Eq. (3.26)].[71] It would seem difficult, however, to justify keeping nonlinear terms in $G(r)$ rather than, say nonlinear terms in ∇G or higher-order derivatives of $G(r)$: the more so as the whole question of the convergence of such an expansion is in doubt at the critical point.

The first author to question the theory in respect of the prediction $G(r) \sim 1/r$ at $T = T_c$, for three dimensions was Green[68] who based his arguments on an integral relation for the pair correlation function derived by cluster diagram summation techniques.[47] This so called, hypernetted-chain integral equation may be written

$$1 + G(r) = \exp[-\beta\phi(r) + G(r) - C(r) + E(r)], \quad (4.5)$$

where

$$E(r) = \mathcal{E}\{\rho; G(r)\} \qquad (4.6)$$

is a nonlinear integral functional of $G(r)$ known only as an expansion in powers of ρ representable in terms of certain, so called, "basic graphs."[47] [The leading term is of order ρ^2]. The same relations hold for a lattice system (with appropriate definition of the functional $\mathcal{E}$).

Green[68] considered the consequences of assuming that the term $E(r)$ involving the basic graphs, might be neglected at the critical point. Stillinger and Frisch extended his analysis to two-dimensional systems.[72]

To follow the argument let us suppose in greater generality that at the critical point of a d-dimensional system

$$G(r) \simeq D/r^{d-2+\eta} \qquad (r \to \infty), \qquad (4.7)$$

where the index η ($0 \leq \eta \leq 2$) measures the departure from the Ornstein–Zernike prediction. [As in equation (2.7) $1/r^0$ corresponds to $\log r$.] It follows that

$$\rho\hat{G}(k) \simeq \hat{D}/k^{2-\eta}, \qquad (k \to 0) \qquad (4.8)$$

so that at fixed density one has, through the relation (3.9),

$$\hat{C}(k) \simeq C(0)[1 - c_0 k^{2-\eta} + \cdots], \quad (k \to 0). \qquad (4.9)$$

If $\eta > 0$ asymptotic inversion yields

$$C(r) \simeq F/r^{d+2-\eta} \qquad (r \to \infty). \qquad (4.10)$$

Thus when $\eta > 0$, the direct correlation function is *also* "long ranged" in the sense that its second moment does not exist, although it certainly decays to zero more rapidly than $G(r)$, in fact by a factor $1/r^{4-2\eta}$. Notice that if $\eta = 0$ we find instead

$$C(r) \simeq F' e^{-\kappa' r}/r, \qquad (\kappa'^2 = 1/c_0) \qquad (4.11)$$

so that only in this special case is $C(r)$ short ranged in the sense that $\hat{C}(k)$ has a Taylor series expansion at $k^2 = 0$.

To analyze the hypernetted integral relation it is convenient to define

$$S(r) = G(r) - C(r). \qquad (4.12)$$

Clearly $S(r)$ must have the same asymptotic be-

[70] M. E. Fisher, Physica 28, 172 (1962).
[71] M. Fixman, J. Chem. Phys. 36, 1965 (1962).
[72] F. H. Stillinger, Jr. and H. L. Frisch, Physica 27, 751 (1961).

havior as $G(r)$. If the potential $\phi(r)$ is short ranged we may expand the exponential in (4.5) for large r to get

$$C(r) = E(r) + \tfrac{1}{2}[S(r) + E(r)]^2 + \cdots . \quad (4.13)$$

If we now suppose $E(r)$ can be neglected or, more weakly, that $E(r)$ decays faster than $[S(r)]^2$ it follows that

$$C(r) = \tfrac{1}{2}[S(r)]^2 + \cdots . \quad (4.14)$$

On substituting (4.7) and (4.10) we obtain the consistency relation

$$\eta = 2 - \tfrac{1}{2}d, \quad (4.15)$$

which fixes the asymptotic form of the correlation functions.

On this basis we would predict that at the critical point in a three-dimensional system $G(r) \sim 1/r^2$ rather than $1/r$.[68] Correspondingly $\hat{G}(k) \sim 1/k = 1/(k^2)^{\frac{1}{2}}$ so that a plot of inverse scattering versus k^2 at, and near $T = T_c$, should be significantly curved downwards for small k^2 (see Fig. 6).

In two dimensions (4.15) leads to $G(r) \sim 1/r^{4/3}$ ($T = T_c$) which is more reasonable than the Ornstein–Zernike result $\log r$. However as pointed out by Stillinger and Frisch,[72] this prediction can be tested against the rigorous result obtained by Onsager and Kaufman[31] for the correlation function of the nearest-neighbor plane square lattice gas at its critical point. [Note that $G(\mathbf{r})$ at $\rho = \rho_c$ is proportional to the spin pair correlation function $\langle S_0^z S_r^z \rangle$ of the Ising ferromagnet in zero field.] This exact result is[31,32,72,73]

$$G(r) \approx D/r^1 \qquad (T = T_c, d = 2) \quad (4.16)$$

so that the true value of the index η is $\frac{1}{4}$ rather than $\frac{4}{3}$ or zero.

We conclude that both the Ornstein–Zernike theory and Green's argument are incorrect for a two-dimensional lattice. Consequently both must be suspect for three-dimensional lattice systems.

It is interesting to note that if, as seems to be the case, the true value of η is less than $2 - \frac{1}{2}d$ one really has, as $r \to \infty$

$$E(r) \approx \tfrac{1}{2}[S(r)]^2 \approx \tfrac{1}{2}[G(r)]^2 \approx D^2/r^{2d-4+2\eta} \quad (4.17)$$

so that the contribution of the basic graphs is also long ranged although it decays faster than $G(r)$. Since after all, $E(r)$ is a functional of $G(r)$ this is not really surprizing.

Of course, the rigorous result (4.16) is known only for nearest-neighbor interactions and one should ask to what extent the behavior of lattices with interactions reaching to further neighbors would be similar. It seems plausible that the correct answer is that the behavior sufficiently near the critical point is qualitatively unchanged provided the range of the potential is finite [for example, if $\phi(r) = 0$ for $r > b$]. The reason for this surmise is that the range of correlation near the critical point becomes, as we have seen, very large compared to the lattice spacing and, indeed, very large compared to the potential range b. The asymptotic correlations are then determined by long chains of interactions and should thus be insensitive to the detailed variation of $\phi(r)$. [The same conclusion is really implicit in Ornstein and Zernike's and in Green's approach.]

The independence of the indices α, β, γ and η of the lattice structure is evidence for this conclusion. Further evidence comes from numerical studies of the compressibility of lattice gases with first and *second* neighbor interactions which indicate unchanged values for the index γ in two and three dimensions.[74]

It is more difficult to assess the relationship of the lattices gases to more realistic *continuum* models. At low temperatures and high densities the properties of a lattice gas will always deviate from continuum behavior but in the critical region the long range of the correlations again suggests an insensitivity to the details of the potential and hence, for a lattice subdivision sufficiently fine relative to the range of the potential, one would expect qualitatively very similar behavior.[75] This conclusion is supported by the apparently quite close resemblance between the critical singularities of real gases and of even the simplest nearest neighbor three-dimensional lattice gases discussed already in Sec. 2.

5. MORE GENERAL ANALYSIS OF CRITICAL SCATTERING

It is clear from the foregoing that the Ornstein–Zernike theory is probably not valid at the critical point of a three-dimensional system. On the contrary one should evidently expect that $G(r)$ behaves asymptotically as $1/r^{1+\eta}$ with $0 < \eta < 1$.[70,72]

A natural way to extend the Ornstein–Zernike theory within the same framework is to consider further powers of k^2 in the expansion of the Fourier

[72] See also Ref. 12, pp. 200–201 but notice that an exponent $\frac{1}{2}$ is missing on the left of Eq. (108).

[74] M. F. Sykes and N. Dalton (to be published).

[75] In one dimension one may verify explicitly that the behavior of the lattice gas approaches that of the continuum gas as the lattice spacing is made smaller relative to the scale of the potential.

transform of the direct correlation function $\hat{C}(k)$. Inclusion of such terms leads to a representation of $G(r)$ as a sum of increasingly more rapidly damped exponentials of the form (for $d = 3$)

$$G(r) \approx \frac{De^{-\kappa r}}{r} + \frac{D_1 e^{-\kappa_1 r}}{r} + \frac{D_2 e^{-\kappa_2 r}}{r} + \cdots \qquad (5.1)$$

with $\kappa < \kappa_1 < \kappa_2 \cdots$. At a fixed temperature, however, such an expansion leads to the same asymptotic behavior for $G(r)$ and to similar scattering at small k^2 as does the original theory. [For smaller enough r and large enough k^2 one must, of course, expect derivations from any general theory since the detailed nature of the potential must eventually make itself felt.]

As one considers T approaching T_c along $\rho = \rho_c$, however, it is possible that more and more exponentials in (5.1) become "excited," so that the first exponential is no longer a good approximation except for extremely large r.[76] This certainly represents essentially what happens for the two-dimensional lattice gas. Here the higher-order range parameters $\kappa_1, \kappa_2, \cdots$ obey the relation[12,14,77]

$$\kappa_n(T) - \kappa(T) \sim n\kappa_0 \, |1 - (T/T_c)|, \qquad (5.2)$$

so that the "spectrum" of exponentials closes up as $T \to T_c$ and in fact becomes dense at the critical point.

A similar behavior for a continuum model is suggested by the recent calculations of Hemmer, Kac, and Uhlenbeck[76] for a potential with a strongly repulsive core and a weakly attractive long-range exponential tail. A series similar to (5.1) can be derived but the amplitudes of successive terms are proportional to higher powers of the compressibility K_T and hence the expansion breaks down near the critical point.

These considerations [which really amount to a restatement of our previous conclusion that $\hat{C}(k)$ probably does not have a Taylor series expansion at the critical point although it does for $T > T_c$] indicate that (5.1) is not the best basis for analyzing the deviations from the Ornstein–Zernike theory. Indeed the asymptotic form $1/r^{1+\eta}$ at $T = T_c$ with $\eta \neq 0$ could only arise from (5.1) if the expansion broke down in some way. To investigate the possibilities more generally let us, therefore, extend (4.7) by writing[70] for $T \geq T_c$ and $\rho = \rho_c$

$$G(r) \simeq (De^{-\kappa r}/r^{d-2+\eta})[1 + Q(\kappa r)], \quad (r \to \infty) \qquad (5.3a)$$

where $D = D(T)$ is a relatively slowly varying function of temperature and where $Q(x) \to 0$ as $x \to 1$ and $Q(x)$ does not grow exponentially fast as $x \to \infty$. This expression is still in the spirit of the Ornstein–Zernike theory in as far as the main assumption implicit in (5.3a) is that in the critical region the correlation functions for large r can be described in terms of only *two* lengths: (i) the *range of correlation* $1/\kappa(T)$ which becomes infinite at the critical point, and (ii) an *effective range of direct interaction* $r_0(T)$ which remains finite at the critical point. In (5.3a) r_0 has been absorbed into the coefficient D. [Compare with R of the Ornstein–Zernike theory: Eqs. (3.12) and (3.15).] We could write in analogy with (3.15)

$$D = \bar{d}/\rho r_0^{2+\eta}, \qquad (5.3b)$$

where $\bar{d}$ is a dimensionless constant.

If the classical theory is valid away from the critical point, as concluded in Sec. 3, the formula (5.3a) should reduce to (4.2) when $\kappa > 0$ and r is very large, i.e. only the first term in (5.1) should remain. [For the plane square nearest-neighbor lattice gas at $\rho = \rho_c$ one may verify that (4.2) is indeed valid above T_c.[77]] This would imply that for large x

$$1 + Q(x) \approx qx^{\frac{1}{2}(d-3)+\eta}, \qquad (x \to \infty). \qquad (5.3c)$$

Sufficiently close to the critical point the nature of this behavior will not matter. It is clear, however, that if $Q(x)$ becomes of order unity for small x the region of significant deviation from classical theory may be rather small. [Stated alternatively, the further exponentials in (5.1) would be significantly excited only very close to the critical point.]

Accepting (5.3) we may calculate the fluctuation integral (3.3) (making the substitution $x = \kappa r$). For the divergence of the compressibility along the critical isochore this yields

$$\chi(0) = k_B T \rho K_T \simeq \hat{D}_0/\kappa^{2-\eta}, \qquad (T \to T_c), \qquad (5.4)$$

where $\hat{D}_0$ is a slowly varying function of T [of magnitude dependent on $Q(x)$]. Similarly a calculation of the Fourier transform of (5.3) for small k^2 yields, near T_c, the non-Lorentzian critical scattering formula

$$\chi(k) \simeq \hat{D}/(\kappa^2 + k^2)^{1-\frac{1}{2}\eta} \qquad (k^2 \to 0), \qquad (5.5)$$

where $\hat{D}$ is a slowly varying function of T and k^2.[78]

[76] P. C. Hemmer, M. Kac and G. E. Uhlenbeck, J. Math. Phys. 5, 60 (1964), P. C. Hemmer, J. Math. Phys. 5, 75 (1964).

[77] The correlation functions may be expressed as a sum of integrals over the complete set of eigenvalues of the basic transition matrix for an Ising lattice (see Ref. 31). The exact limiting density of these eigenvalues is known from Onsager's work (Ref. 14).

[78] If $Q(x)$ is neglected one has for small η

$$\hat{D}(k^2) = D_0[1 - \eta k^2/(\kappa^2 + k^2) + \cdots].$$

As in our discussion of the Ornstein–Zernike theory the way in which the inverse range of correlation $\kappa(T)$ vanishes as $T \to T_c$ for $\rho = \rho_c$ is related to the nature of the corresponding divergence of the compressibility as $1/(T - T_c)^\gamma$. Assuming, as in (3.19) that

$$\kappa(T) \simeq \kappa^0 |1 - (T/T_c)|^\nu \qquad (T \to T_c) \tag{5.6}$$

and substituting in (5.4) shows that the critical indices are related by

$$\gamma = (2 - \eta)\nu. \tag{5.7}$$

This relation may be checked for the plane square lattice gas since, as we have seen,[31,32,72,73] $\eta = \frac{1}{4}$ and $\gamma = 1\frac{3}{4}$. (see Table I). Consequently we should have

$$\nu = 1, \qquad (d = 2) \tag{5.8}$$

in contrast to the classical result $\nu = \frac{1}{2}$. Now the result $\nu = 1$ was in fact derived rigorously by Onsager in his original paper on the Ising model.[14] His derivation is based on the relation

$$e^{-\kappa a} = \lambda_1/\lambda_0, \tag{5.9}$$

where a is the lattice spacing, and where $\lambda_0(T)$ and $\lambda_1(T)$ are the largest and next largest eigenvalues of the basic matrix which adds a row to the lattice at temperature T.[77]

If we *assume*[79] that ν is still unity for the three-dimensional lattice gas and utilize the result[80] $\gamma = 1\frac{1}{4}$ we would predict $\eta = \frac{3}{4}$. Thence the critical point decay law would be $G_c(r) \sim 1/r^{7/4}$ in closer agreement with Green's result. However the assumption $\nu = 1$ in three dimensions is not much better justified a priori than the classical assumption $\nu = \frac{1}{2}$. Furthermore, examination of the other critical indices (see Table I) indicates that the classically predicted behavior is more closely approached the larger the dimensionality of the system.[81] Consequently one might anticipate that for the three-dimensional nearest-neighbor lattice gas ν lies between $\frac{1}{2}$ and 1 and η is *less* than $\frac{1}{4}$. [As mentioned previously the *assumption* $\eta = 0$ leads to $\nu = \frac{1}{2}\gamma = \frac{5}{8}$ when $d = 3$.]

To decide between these various speculations it is necessary to calculate $\kappa(T)$ or some other feature of the correlation functions. Fortunately the Ising model is again sufficiently tractable to allow some progress. It is possible at the critical density, to derive a diagrammatic expansion for the decay factor $e^{-\kappa a}$ in powers of $1/T$ via Eq. (5.9).[82] The required graphs consist of an infinite chain of connected bonds stretching right across the lattice together with nonoverlapping closed polygons, as occur in the expansion of the partition function.[12]

Evaluation of the series[83] for $\kappa(T)$ (with the aid of Padé approximants) reveals a behavior near T_c clearly intermediate between the two-dimensional and classical results in a region $T = T_c$ to $2T_c$. Direct estimation of the index ν, however, proves to be not very accurate but indicates a range $\nu = 0.6$ to 0.7. [The coefficients of the series are difficult to calculate and not very smooth.]

An alternative approach is to study the temperature dependence of the higher moments of the correlation function, namely

$$\mu_s(T) = \rho \int r^s G(\mathbf{r})\, d\mathbf{r}. \tag{5.10}$$

The zeroth moment is essentially the compressibility but the second moment is also a direct measure of the range of correlation. It is evident furthermore, that μ_2 is proportional to the curvature of $\rho\hat{G}(k)$ for small k and hence to the true limiting slope of the curve of $1/\chi(k)$ versus k^2 as $k^2 \to 0$.[84] From (5.3) we see that when $T \to T_c$, $\mu_2(T)$ diverges as

$$\begin{aligned} \mu_2(T) &\simeq M_2/\kappa^{4-\eta}, \qquad (T \to T_c) \\ &\simeq M_2'/|1 - (T/T_c)|^\delta, \end{aligned} \tag{5.11}$$

with

$$\delta = (4 - \eta)\nu. \tag{5.12}$$

The coefficients M_2 and M_2' are slowly varying functions of T. By comparing (5.12) with the index relation (5.7) we see that η and ν can both be determined if γ and δ are known. [In particular only if $\delta = 2\gamma$ would we have $\eta = 0$.]

For the two- and three-dimensional lattice gases expansions for $\mu_2(T)$ is series of powers of $1/T$ at $\rho = \rho_c$ are not too difficult to calculate.[83] [The labor and graphical analysis is similar to that for the compressibility.] The coefficients prove to be rather smoothly varying and for the plane lattices numerical analysis confirms quite accurately the relation (5.12)

[79] This assumption was made tentatively in Ref. 70 and tested experimentally by Frisch and Brady (Ref. 67). A good fit was obtained but although the data revealed definite nonclassical behavior, they were not sufficiently accurate to distinguish between $\nu = 1$ and some appreciably lower value.

[80] See the discussion in Sec. 2 and Refs. 23 and 28.

[81] From Table I we see $\gamma = 1\frac{3}{4}$, $1\frac{1}{4}$ for $d = 2$, $d = 3$ to which we may add $\gamma = 1.094$ ($d = 4$); see M. E. Fisher and D. S. Gaunt, Phys. Rev. 133, A224 (1964). The classical result $\gamma = 1$ corresponds to $d \to \infty$.

[82] M. E. Fisher (to be published).

[83] M. E. Fisher and R. J. Burford (to be published).

[84] The second moment μ_2 is closely related to the length Λ defined in (3.28) and to the "persistence length" L defined by Debye (Ref. 44) explicitly: $\Lambda^2 = \frac{1}{6}\langle\cos^2\theta\rangle\,\mu_2/(1 + \mu_0)$ and $L^2 = \mu_2/\mu_0$. From (3.28) one sees that $1/\chi(k) = 1/\chi(0)\{1 + \Lambda^2k^2 + O(k^4)\}$.

[which predicts $\delta = 3.75$]. This is an important result since it provides support for the original hypothesis (5.3).

Initial estimates for the simple cubic lattice gas[83] yield $\delta = 2.538 \pm 0.003$ and hence

$$\nu = 0.644 \pm 0.003, \quad \eta = 0.060 \pm 0.007. \quad (5.13)$$

Results for other three-dimensional lattices confirm these estimates with lower accuracy. [One might mention that (5.14) is not inconsistent with the conjecture $\eta = \frac{1}{16} = 0.0625$ which is rather natural in view of the estimate $\beta \simeq \frac{5}{16}$ discussed in Sec. 2.]

The values (5.13) are in accordance with our expectation that the three-dimensional results should be closer to the classical predictions.[81] The magnitude of η for the lattice gas is indeed rather close to zero but it is not unlikely that a more realistic continuum model would lead to somewhat larger value, say $\eta \simeq 0.1$.[85] Until more rigorous theories are developed and more realistic models become soluble we must content ourselves with these rough estimates. Accordingly let us review the nature of the critical scattering to be expected on the basis of our analysis.

The peak in the critical scattering is described by (5.5) and (for $\eta > 0$) should be narrower at a fixed temperature near T_c than the Lorentzian curve of same peak height. Correspondingly a plot of the inverse scattering intensity versus k^2 should be somewhat convex although, as suggested in Fig. 6, for larger values of $k^2 > k^2_{\min}$ the curves might *appear* to be reasonably linear. For small values of η, of the magnitude (5.13), it might indeed be rather difficult experimentally to detect the increasing curvature of the scattering plots as $k^2 \to 0$ even though the curve for $T = T_c$ will theoretically have an infinite slope at $k^2 = 0$. In practice a nonzero value of η can probably best be detected by the observation that the apparent linear intercepts $\xi_0(T)$ [see Fig. 6] would not approach zero when $T \to T_c$ as must the true intercepts $\xi(T) = 1/\chi(0)$. In particular the scattering plot taken *at* $T = T_c$ would tend to extrapolate to a small positive value at $k^2 = 0$. [Of course to detect this behavior it is important to have an independent measurement of T_c and *not* to judge T_c by extrapolating $\xi_0(T)$ to zero with T as would otherwise be tempting!] This behavior is reminiscent of the experimental "anomalies" described in the previous Section although, as yet, these can probably not be regarded as fully established.

The likely fact that γ exceeds unity for a real gas and the consequent curvature of $1/K_T$ versus T would indicate that a plot of the true intercepts, $\xi(T)$, versus T should flatten out as T approaches T_c (and theoretically have zero slope at $T = T_c$). This effect might well be less obvious in a plot of the apparent intercepts $\xi_0(T)$ although it does seem to have been observed.[86]

[85] For example the effectively more "continuum-like" Heisenberg model of ferromagnetism yields $\gamma \simeq 1\frac{1}{4}$ compared with the Ising value $\gamma = 1\frac{1}{4}$ ($d = 3$). (See Ref. 90, 91).

6. SUMMARY AND CONCLUSIONS

We have shown that the classical theories of the gas–liquid critical point are unsatisfactory both on experimental and theoretical grounds. Thus the coexistence curve must be described experimentally by

$$\rho_L - \rho_G \sim (T_c - T)^\beta, \quad (6.1)$$

with $\beta \simeq 0.33$ to 0.36 (rather than with $\beta = \frac{1}{2}$) while for three-dimensional lattice gas models one finds $\beta \simeq 0.31 \simeq \frac{5}{16}$. The specific heat $C_V(T)$ measured along the critical isochore of a fluid becomes infinite at T_c, diverging approximately as $\log|T - T_c|$. A similar result holds for the lattice gas. Theoretically one also expects that the compressibility above and below T_c should diverge as

$$K_T \sim 1/|T - T_c|^\gamma \quad (6.2)$$

with $\gamma > 1$. This prediction while qualitatively correct awaits quantitative experimental verification. (Values of these indices are given in Table I.)

On theoretical grounds the classical (Ornstein–Zernike) theory of critical scattering has been shown to be unsatisfactory close to the critical point (although it probably is valid away from the critical region when the compressibility is still moderately large). More generally one should expect the scattering intensity to vary as

$$I(k) \sim 1/(\kappa^2 + k^2)^{1-\frac{1}{2}\eta}, \quad (6.3)$$

where $\eta > 0$ and where the range parameter vanishes along the critical isotherm as

$$\kappa(T) \sim (T - T_c)^\nu \quad (6.4)$$

with $\nu(2 - \eta) = \gamma$. Theoretical analysis suggests that the index η might be no larger than 0.1 (see Table I). Consequently it is probably not easy experimentally to detect deviations from the classical theory (for which $\eta = 0$). Nevertheless there are experimental indications of the failure of the classical predictions for small k^2 near T_c and these are consistent with $\eta > 0$. Final confirmation of the theory must, however, rest on further, more extensive and accurate measurements on sufficiently simple fluid systems.

[86] D. McIntyre, Ref. 64 and a private communication.

In conclusion one should mention the rather close analogy between the gas–liquid critical point and the Curie point of a ferromagnetic crystal.[87] Indeed almost all our analysis and conclusions apply directly to ferromagnetic systems if appropriately translated. The density deviation $\rho - \rho_c$ should be identified with the magnetization M while the magnetic field H is isomorphic to the chemical potential of the fluid. The critical isochore, $\rho = \rho_c$, corresponds to zero magnetic field (since the mean magnetization then vanishes) and the coexistence curve corresponds to the curve of spontaneous magnetization $M_0(T)$. The specific heat $C_V(T)$ along the critical isochore of a fluid is isomorphic to the specific heat $C_H(T)$ of a ferromagnet in zero field and the compressibilities at condensation and above T_c for $\rho = \rho_c$ correspond essentially to the initial susceptibilities $\chi_0(T) = (\partial M/\partial H)_T$, $(H \to 0)$.

The van der Waals and equivalent classical theories find their precise parallel in the Weiss molecular field theory and its extensions.[88] Similarly the Ornstein–Zernike theory and its developments have been adapted to describe the critical scattering of neutrons by ferromagnets.[89,90] The net pair correlation function $G(\mathbf{r})$ is replaced by the spin–spin correlation functions $\Gamma_{\alpha\beta}(\mathbf{r}) = \langle S_0^\alpha S_\mathbf{r}^\beta \rangle$. Theoretically one is able to estimate the susceptibility index γ for the nearest neighbor Heisenberg model above T_c with the approximate result[91,92] $\gamma = \frac{4}{3}$ (independent of spin and lattice structure in three-dimensions). It is interesting that this nonclassical prediction has been quite accurately confirmed recently.[93–95] Furthermore modern neutron scattering experiments[95] have also given a definite suggestion of deviations from the Ornstein–Zernike theory consistent with a small positive value of the index η.

The spontaneous ferromagnetic moment $M_0(T)$ is not easy to measure near T_c but nuclear magnetic resonance experiments by Benedek and Heller[96] have shown that the somewhat analogous sublattice magnetization (or long range order) of an *anti*ferromagnet (actually MnF_2) varies as $(T - T_c)^\beta$ with $\beta = 0.335 \pm 0.010$. The relation holds with remarkable accuracy up to within millidegrees of the critical or Néel point $[\Delta T/T_c = 0.007\%]$.[97] Experiments on antiferromagnets also reveal specific heat infinities at T_c which are approached in approximately logarithmic fashion.[98] The close similarity of these results to the corresponding behavior of fluid systems presents a striking challenge to our theoretical understanding.

ACKNOWLEDGMENTS

I am most grateful to Professor G. E. Uhlenbeck for his criticisms of a first draft of this article and would like to thank Professor Mark Kac and Professor Joel Lebowitz for their comments.

[87] As mentioned in Sec. 2, this relationship is formally exact for a lattice gas and an Ising model ferromagnet. See, for example, T. D. Lee and C. N. Yang, Phys. Rev. **87**, 410 (1952).

[88] See, for example, J. H. Van Vleck, Rev. Mod. Phys. **17**, 27 (1945); P. W. Kasteleijn and J. Van Kranendonk, Physica **22**, 317 (1956).

[89] L. Van Hove, Phys. Rev. **95**, 1374 (1954).

[90] R. J. Elliott and W. Marshall, Rev. Mod. Phys. **30**, 75 (1958).

[91] C. Domb and M. F. Sykes, Phys. Rev. **128**, 168 (1962).

[92] J. L. Gammel, W. Marshall, and L. Morgan, Proc. Roy. Soc. (London) **A275**, 257 (1963).

[93] J. E. Noakes and A. Arrott, J. Appl. Phys. (Suppl.) **35**, 931 (1964).

[94] J. S. Kouvel [private communication based on an analysis of the measurements by P. Weiss and R. Forrer, Ann. Phys. **5**, 153 (1926).]

[95] L. Passell, K. Blinowski, T. Brun and P. Nielsen, J. Appl. Phys. (Suppl.) **35**, 933 (1964). Note that the linear intercepts with the $k^2 = 0$ axis of the inverse scattering appear to approach a positeve value at $T = T_c$.

[96] P. Heller and G. B. Benedek, Phys. Rev. Letters **8**, 428 (1962).

[97] More recent NMR experiments by Heller and Benedek (private communication) reveal a similar behavior for the spontaneous magnetization of the ferromagnet EuS.

[98] See, for example, W. K. Robinson, and S. A. Friedberg, Phys. Rev. **117**, 402 (1960).

THE JOURNAL
OF
CHEMICAL PHYSICS

VOLUME 13, NUMBER 7 JULY, 1945

The Principle of Corresponding States

E. A. GUGGENHEIM
Imperial College, London

(Received May 3, 1945)

The principle of corresponding states in modern form has been applied to the following properties: the critical state, the virial coefficient, the Boyle point, the densities of coexistent phases, the vapor pressure of the liquid, the entropies of evaporation and of fusion, the coefficient of thermal expansion of the liquid, the triple-point temperature and pressure, the heat capacity of the liquid, and the surface tension of the liquid. It has been shown that argon, krypton, xenon, and with less accuracy neon, follow the principle with respect to all these properties. It has further been shown that nitrogen, oxygen, carbon monoxide, and methane follow the principle with fair accuracy as vapors and as liquids, but not as solids. The relations between surface tension, temperature, and densities have been analyzed empirically. For the "ideal" substances under consideration Katayama's modification of Eötvos' relation holds good, but McLeod's relation does not; in the relation $\gamma \propto (\rho_l - \rho_c)^s$, the exponent s is not 4 but much more nearly $3\frac{1}{2}$.

1. INTRODUCTION

THE principle of corresponding states may safely be regarded as the most useful by-product of van der Waals' equation of state. Whereas this equation of state is nowadays recognized to be of little or no value, the principle of corresponding states correctly applied is extremely useful and remarkably accurate.

Pitzer[1] has stated a set of assumptions sufficient to lead to the principle of corresponding states and has shown that argon, krypton, and xenon have several properties in accordance with the principle. In the present paper, I shall draw attention to a number of further relations, not mentioned by Pitzer, all in accordance with the principle of corresponding states. I shall show that the principle is also followed accurately by nitrogen and oxygen, and somewhat less accurately by carbon monoxide and methane. For the sake of continuity and completeness, I shall at small cost of space repeat some of Pitzer's contributions to the subject.

2. ASSUMPTIONS

I shall with the utmost brevity recall Pitzer's list of assumptions and make comments on some of them. They are:

I. "Classical statistical mechanics will be used."
II. "The molecules are spherically symmetrical, either actually or by virtue of rapid and free rotation."
III. "Intramolecular vibrations will be assumed the same in the liquid and gas states."
IV. "The potential energy will be taken as a function only of the various intermolecular distances."
V. "The potential energy for a pair of molecules can be written $A\phi(R/R_0)$ where R is the intermolecular distance, A and R_0 are characteristic constants and ϕ is a universal function."

The precise meaning of assumption I is ambigu-

[1] K. Pitzer, J. Chem. Phys. 7, 583 (1939). See also Boer and Michels, Physica 5, 946 (1938).

ous. My interpretation of the correct condition is that

(a) any distinction between Fermi-Dirac statistics and Bose-Einstein statistics has a negligible effect;

(b) the effect of quantization of the translational degrees of freedom is negligible.

Both (a) and (b) are satisfied provided $(mkT)^{\frac{1}{2}}v^{\frac{1}{3}} \gg h$, where m denotes molecular mass, T absolute temperature, and v volume per molecule. As pointed out by Pitzer this condition excludes hydrogen and helium owing to their small molecular masses and to some extent limits the applicability of the principle to neon. I can however see no need to assume that degrees of freedom other than the translational must be classical. On the contrary, the vibrational degrees of freedom of diatomic molecules are practically unexcited, and yet it will be shown that such molecules can obey the principle of corresponding states.

The effect of assumption II is that some diatomic and polyatomic molecules may obey the principle in the gaseous and liquid state, but cannot be expected to obey it in the solid state. Assumption III might be more usefully worded: "The intramolecular degrees of freedom are assumed to be completely independent of the volume per molecule." Highly polar molecules are ruled out by conditions II, III, and IV. Metals and molecules capable of forming hydrogen bonds are ruled out by condition IV.

Even when conditions I to IV are fulfilled, there is no *a priori* reason why assumption V should be fulfilled. In fact there is good reason to suppose that it does not hold accurately. The assumption is however accurate for large values of R where the intermolecular potential energy ϵ is proportional to $-R^{-6}$. Moreover many macroscopic properties are insensitive to the precise form of the relation between ϵ and R for small R. Consequently assumption V, though not rigorously true, turns out to be a useful approximation for many non-polar molecules.

3. FORMULATION OF PRINCIPLE

From the above assumptions it follows that the free energy F of a system of N molecules at the temperature T obeys the relation

$$-\frac{F}{NkT}=\frac{3}{2}\log\frac{2\pi mkT}{h^2}+1+\log v+\log j(T, I, \nu_i)+\log\psi\left(\frac{kT}{\epsilon_0}, \frac{v}{R_0^3}\right), \quad (3.1)$$

where m is the molecular mass, v is the molecular volume, j is the partition function for the internal degrees of freedom of a molecule and depends only on the temperature T, the principal moments of inertia I, and the frequencies ν_i of the normal modes. Finally ψ is a complicated, but universal function of kT/ϵ_0 and v/R_0^3, where R_0 is the intermolecular distance at which the intermolecular energy has its minimum value $-\epsilon_0$. At infinite dilution ψ tends to unity and formula (3.1) becomes that of a perfect gas.

From the thermodynamic relation

$$P=-\left(\frac{\partial F}{\partial V}\right)_{T,N}=-\frac{1}{N}\left(\frac{\partial F}{\partial v}\right)_{T,N}, \quad (3.2)$$

where P is the pressure, V the volume of the whole system, and v the volume per molecule, it follows that Pv/kT is a universal function of kT/ϵ_0 and v/R_0^3. From the well known conditions for the critical state it follows at once that kT_c/ϵ_0 and v_c/R_0^3, where the subscript c denotes critical values, are both universal constants. The equation of state is therefore of the form

$$\frac{P}{P_c}=u\left(\frac{T}{T_c}, \frac{v}{v_c}\right), \quad (3.3)$$

where u is a universal, but complicated, function. Insofar as the five initial assumptions are fulfilled, Eq. (3.3) is valid for gas, liquid, and solid. This is the general expression of the principle of corresponding states.

I shall now consider various deductions from formula (3.3) and examine to what extent they are borne out by experimental data for neon, argon, krypton, xenon, nitrogen, oxygen, carbon monoxide, and methane.

4. CRITICAL STATE

In Table I, the first row gives the molecular weights, the next four rows[2] the critical tempera-

[2] The sources of experimental data are collected in an appendix.

TABLE I.

	Formula	Ideal	Ne	A	Kr	Xe	N_2	O_2	CO	CH_4
1.	M		20.18	39.94	83.7	131.3	28.02	32.00	28.00	16.03
2.	T_c/°K		44.8	150.7	209.4	289.8	126.0	154.3	133.0	190.3
3.	ρ_c/g cm^{-3}		0.484	0.5308	0.9085	1.155	0.311	0.430	0.301	0.162
4.	v_c/cm^3 mole^{-1}		41.7	75.3	92.1	113.7	90.2	74.5	93.2	98.8
5.	P_c/atmos.		26.9	48.0	54.1	58.2	33.5	49.7	34.5	45.7
6.	$P_c v_c/RT_c$	0.292	0.305	0.292	0.290	0.278	0.292	0.292	0.294	0.289
7.	T_B/°K		121	411.5			327		~345	491
8.	T_B/T_c	2.7	2.70	2.73			2.59		2.6	2.58
9.	T_b/°K (P = 1 atmos.)		27.2	87.3	120.9	165.1	77.3	90.1	81.6	112.5
10.	T_s/°K ($P = P_c/50$)		25.2	86.9	122.0	167.9	74.1	90.1	78.9	110.5
11.	T_b/T_c		0.608	0.580	0.577	0.570	0.614	0.583	0.613	0.591
12.	T_s/T_c	0.58	0.563	0.577	0.582	0.580	0.588	0.583	0.593	0.581
13.	L_e/R°K		224*	785	1086	1520	671	820	727	1023
14.	L_e/RT_b		8.25	8.98	8.98	9.19	8.68	9.11	8.91	9.10
15.	L_e/RT_s	9.05	8.9	9.04	8.91	9.06	9.06	9.11	9.22	9.26
16.	$10^3\alpha$			4.54			5.88	3.85	4.91	
17.	αT_c	0.68		0.68			0.74	0.59	0.65	
18.	T_t/°K		24.6	83.8	116.0	161.3	63.1	54.4	68.1	90.6
19.	T_t/T_c	0.555	0.549	0.557	0.553	0.557	0.501	0.352	0.512	0.476
20.	L_f/R°K		40.3	141.3	196.2	276	86.8	53.5	100.7	113
21.	L_f/RT_t	1.69	1.64	1.69	1.69	1.71	1.37	0.98	1.48	1.25
22.	P_t/atmos.		0.425	0.682	0.721	0.810				
23.	$100P_t/P_c$	1.37	1.58	1.42	1.33	1.39				
24.	v_l/cm^3			28.14	34.13	42.68				
25.	v_g/cm^3			24.61	29.65	37.09				
26.	v_l/v_g			1.144	1.151	1.151				
27.	C_P/R at T_s	5.4		5.5	5.35	5.35	6.75	6.45	7.25	

* This is not a measured value, but an approximate value based on vapor pressure measurements.

ture T_c, the critical density ρ_c, the critical molar volume v_c, and the critical pressure P_c. The sixth row gives the values of $P_c v_c/RT_c$, which van der Waals pointed out should have a universal value. It will be seen that the values for argon, krypton, nitrogen, oxygen, carbon monoxide, and methane all lie within 1.5 percent of 0.292. The value for neon is somewhat higher and that for xenon somewhat lower. In the former case the deviation is presumably a quantal effect; in the latter there is no obvious explanation.

5. VIRIAL COEFFICIENT AND BOYLE POINT

The equation of state of a gas up to moderate pressures may be expressed in the form

$$Pv/RT = 1 + B(T)/v, \qquad (5.1)$$

where $B(T)$ is called the second virial coefficient, which is positive at high and negative at low temperatures. The temperature T_B at which $B(T)$ changes sign is called the Boyle point. In the seventh row of Table I are given values of T_B and in the eighth values of T_B/T_c. The values for neon, argon, and oxygen agree within 1 percent; those for nitrogen, carbon monoxide, and methane are 5 to 8 percent lower.

The dependence of $B(T)$ on T is not expressible by any particularly simple relation. $B(T)/v_c$ is plotted against T/T_c in Fig. 1 for neon, argon, nitrogen, and oxygen, for which alone of the substances being considered accurate data at low pressures are available. It will be seen that the three substances are roughly fitted by the same curve.

6. DENSITIES OF COEXISTENT LIQUID AND VAPOR

If ρ_l denotes the density of the liquid and ρ_g that of the vapor in mutual equilibrium at the

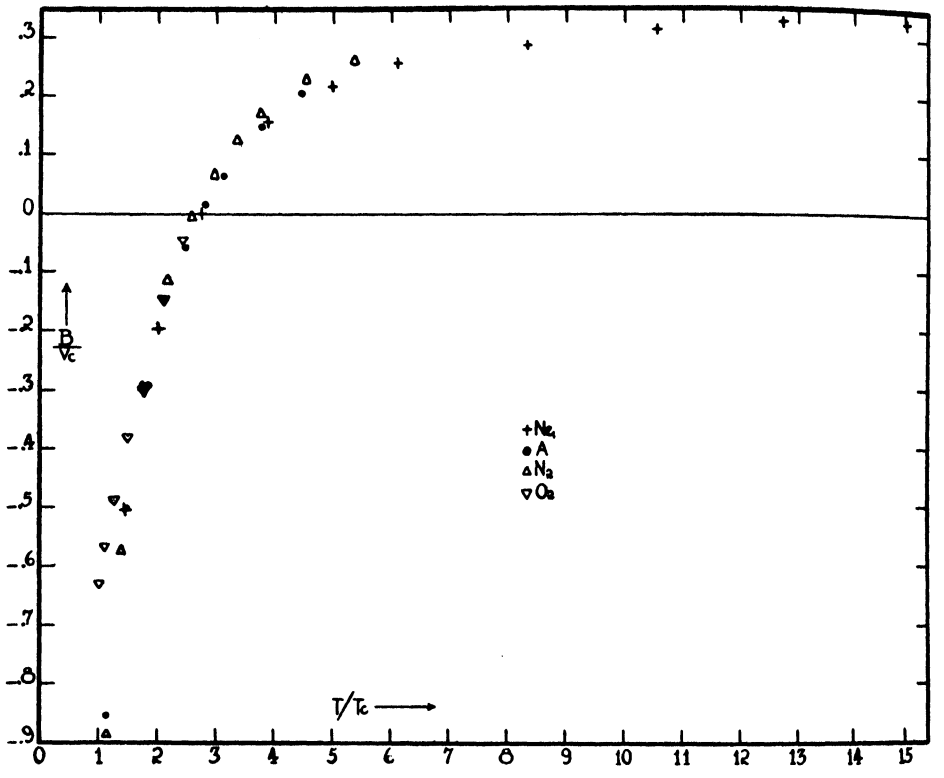


FIG. 1.

temperature T, while ρ_c denotes the critical density, then according to the principle of corresponding states one should expect ρ_l/ρ_c and ρ_g/ρ_c to be universal functions of T/T_c. In Fig. 2 are plotted all the experimental data for the substances under consideration. It will be seen that except for carbon monoxide and methane most of the points lie on or near a single curve.

The curve drawn through the points for argon is represented by the empirical formulae

$$\frac{\rho_l}{\rho_c}=1+\frac{3}{4}\left(1-\frac{T}{T_c}\right)+\frac{7}{4}\left(1-\frac{T}{T_c}\right)^{\frac{1}{3}}, \qquad (6.1)$$

$$\frac{\rho_g}{\rho_c}=1+\frac{3}{4}\left(1-\frac{T}{T_c}\right)-\frac{7}{4}\left(1-\frac{T}{T_c}\right)^{\frac{1}{3}} \qquad (6.2)$$

By addition and subtraction of (6.1) and (6.2) one obtains

$$\frac{\rho_l+\rho_g}{2\rho_c}=1+\frac{3}{4}\left(1-\frac{T}{T_c}\right), \qquad (6.3)$$

$$\frac{\rho_l-\rho_g}{\rho_c}=\frac{7}{2}\left(1-\frac{T}{T_c}\right)^{\frac{1}{3}} \qquad (6.4)$$

Formula (6.3) expresses the well-known law of the rectilinear diameter. Formula (6.4) is new. The inaccuracy of these formulae as applied to argon is generally only one or two parts per thousand of ρ_c or of ρ_l. But the percentage inaccuracy in ρ_g increases with decrease of temperature and becomes serious below $T\simeq 0.65T_c$. In fact according to the formulae, ρ_g would become zero at $T=0.565T_c$, but this is below the triple point $T_t=0.58T_c$. It is therefore not recommended that

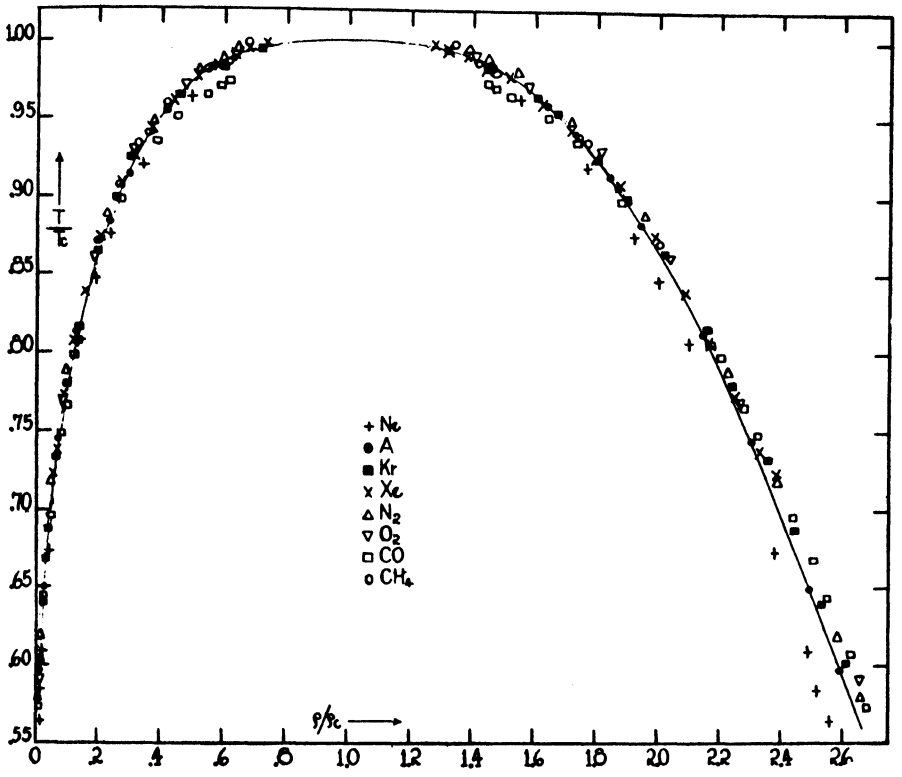


FIG. 2.

these formulae should be used for computing values of ρ_g. There are however occasions when one requires relatively accurate values not of ρ_g itself but of $(\rho_l-\rho_g)/\rho_c$; on such occasions formula (6.4) in view of its extreme simplicity and surprisingly high accuracy has much to recommend it. An example of its use will occur in Section 16.

7. VAPOR PRESSURE

At temperatures considerably below the critical temperature, say $T<0.65T_c$, when formula (6.2) for ρ_g becomes inaccurate it is convenient to consider the equilibrium vapor pressure P rather than ρ_g. According to the principle of corresponding states one should expect P/P_c to be a universal function of T/T_c. In particular the temperatures T_s at which the equilibrium pressure P is one-fiftieth of the critical pressure should be corresponding temperatures for different substances and the ratio of T_s to T_c should have a universal value. On the other hand T_b the boiling points at a pressure of one atmosphere are not corresponding temperatures for different substances. In rows 9 and 10 of Table I are given T_b the boiling point at a pressure of one atmosphere, and T_s the boiling point at a pressure one-fiftieth the critical pressure. In rows 11 and 12 are given the ratios T_b/T_c and T_s/T_c. It will be seen that the values of the latter are, as expected, more nearly the same than the values of the former.

8. ENTROPY OF EVAPORATION

According to Trouton's rule the molar entropies of evaporation for different substances have

roughly the same value at their boiling-points under a pressure of one atmosphere. Hildebrand[3] has found that the molar entropies of evaporation are more nearly the same at temperatures at which ρ_σ has the same value. Neither Trouton's rule nor Hildebrand's accords with the principle of corresponding states, which requires rather that the molar entropies of evaporation should be equal at corresponding temperatures. In particular they should be equal at a standard temperature $T_s=0.58T_c$ at which the vapor pressure is one-fiftieth the critical pressure. In row 13 of Table I are given the molar heats of evaporation L_s divided by the gas constant R. In rows 14 and 15 are given the molar entropies of evaporation divided by the gas constant R at the temperatures T_b and T_s respectively. It will be seen that the values of L_s/RT_s are, as they should be, more nearly alike than the values of L_s/RT_b.

9. COEFFICIENT OF THERMAL EXPANSION

An extremely sensitive test of the accuracy of formula (6.1) for the density ρ_l of the liquid is obtained by using this formula to calculate α the coefficient of thermal expansion. According to this formula α will vary with the temperature, but only slowly at low temperatures. At the temperature $T_s=0.58T_c$, the calculated value of $\alpha(T_c-T_s)$ is 0.287 and that of αT_c is 0.68. The only experimental values of α are given in row 16 of Table I and the corresponding values of αT_c in row 17. The agreement for nitrogen, oxygen, and carbon monoxide is as good as one could expect. In the case of argon the agreement is exact. This is not an independent test of the principle of corresponding states, but is a demonstration of the high accuracy of formula (6.1) as applied to argon.

10. TRIPLE-POINT TEMPERATURE

As pointed out in Section 2, it is only for monatomic molecules that the principle of corresponding states may be expected to apply to the solid state, because the rotational motion of diatomic and polyatomic molecules can hardly be unrestricted in a crystal. In row 18 of Table I are given the triple-point temperatures T_t and in row 19 the ratios of T_t to T_c. There is good agreement between the monatomic molecules of argon, krypton, xenon, and even neon. For the diatomic molecules and methane the agreement is as expected appreciably less good. The fact that oxygen deviates considerably more than nitrogen and methane is interesting. This may possibly have some connection with the fact that the gaseous molecule O_2 is in a triplet state. The molar entropy consequently contains an electronic contribution $R\log 3$, which presumably persists in the liquid, but has disappeared[4] at the lowest temperatures at which thermal measurements have been made on crystalline oxygen. Whether it disappears by formation of O_4 molecules or otherwise is not known, nor in what temperature range it disappears, but it may be expected to exert a marked effect on the equilibrium properties of the solid phase.

11. ENTROPY OF FUSION

The molar heats of fusion L_f divided by the gas constant R are given in row 20 of Table I and the molar entropies of fusion divided by R in row 21. The agreement between the values for the monatomic molecules is good.

12. TRIPLE-POINT PRESSURES

The triple-point pressures P_t of the monatomic species are given in row 22 of Table I and the ratios $100P_t/P_c$ in row 23. The values for argon, krypton, and xenon, but not neon, agree within ± 4 percent. Since an increase of 4 percent in the equilibrium vapor pressure of the crystal, or of the liquid, corresponds to an increase in absolute temperature of only 0.4 percent, this is a much more sensitive test than that of Section 10.

13. EXPANSION ON FUSION

In rows 24 and 25 of Table I are given the molar volumes v_l and v_s of liquid and solid, respectively, at the triple point for the monatomic species. In row 26 are given the ratios of v_l to v_s. As previously pointed out by Clusius and Weigand[5] the uniformity of these ratios is remarkable.

[3] J. Hildebrand, J. Am. Chem. Soc. **37**, 970 (1915); **40**, 45 (1918).

[4] W. Giauque and H. Johnston, J. Am. Chem. Soc. **51**, 2300 (1929).

[5] K. Clusius and K. Weigand, Zeits. f. Physik. Chemie **B46**, 1 (1940).

14. HEAT CAPACITY OF LIQUID

According to the principle of corresponding states the molar heat capacity at corresponding temperatures should be the same for monatomic molecules. That this is at least roughly true is shown by the values given in row 27 of Table I. The application of the principle to diatomic molecules is less simple. One of the assumptions used to deduce the principle is that the molecules should rotate freely in the liquid. If this assumption is accurately fulfilled, then the rotational degrees of freedom should make precisely the same contribution to each of the thermodynamic functions in the liquid as in the gas. In particular the rotational contribution to the molar heat capacity of the diatomic species should be just R. Hence the total molar heat capacity of the diatomic species should exceed that of the monatomic species by just R. Actually the excess, as shown by the values in row 27 of Table I, is greater than R. This is understandable if there is a small restriction of the rotation increasing with decreasing temperature.

15. EXTENSION OF PRINCIPLE TO SURFACES

In order to extend the principle of corresponding states to surfaces, one needs to revise assumption III as follows: "The intramolecular degrees of freedom are assumed to be completely independent of the volume per molecule and of the distance of the molecule from the surface." From this assumption it follows that the excess free energy F_s due to an interface of area A containing N_s molecules is of the form

$$-\frac{F_s}{N_s kT}=\log \psi_s\left(\frac{kT}{\epsilon_0}\,\frac{A}{N_s R_0^2}\right), \tag{15.1}$$

where ψ_s is a complicated but universal function of kT/ϵ_0 and $A/N_sR_0^2$. It is also evident that at equilibrium between liquid and vapor $A/N_sR_0^2$ is like v_l/R_0^3 a universal function of T/T_c consequently $F_sR_0^2/kTA$ is a universal function of T/T_c. But for a single component system the surface tension γ is equal to F_s/A. Hence $\gamma R_0^2/kT$ is a universal function of T/T_c. Finally, using the fact that v_c/R_0^3 has a universal value, one concludes that $\gamma v_c^{\frac{2}{3}}/kT$ and so also $\gamma v_c^{\frac{2}{3}}/kT_c$ are universal functions of T/T_c. How well this prediction is borne out by the experimental data will be examined in the two succeeding sections.

16. TEMPERATURE DEPENDENCE OF SURFACE TENSION

Since the surface tension of a liquid decreases with increasing temperature and vanishes at the critical point, the simplest possible form of empirical relation between γ and T is

$$\gamma=\gamma_0\left(1-\frac{T}{T_c}\right)^{1+r}, \tag{16.1}$$

where r is a constant. I shall show that this relation with $r=2/9$ in fact adequately represents the experimental data, but for the moment I leave the value of r unspecified. Van der Waals suggested a formula such as (16.1) with $r=0.234$ and Ferguson[6] has proposed $r=0.21\pm0.015$.

One of the earliest and best known empirical relations relating to the temperature dependence of γ is that of Eötvos, namely

$$\gamma v_l^{\frac{2}{3}} \propto (1-T/T_c). \tag{16.2}$$

This relation was however found to be inexact and various modifications of it have been proposed, of which the most satisfactory is that of Katayama,[7] namely

$$\gamma y^{-\frac{2}{3}} \propto (1-T/T_c), \tag{16.3}$$

where y is defined by

$$y=1/v_l-1/v_c, \tag{16.4}$$

or

$$yv_c=(\rho_l-\rho_g)/\rho_c. \tag{16.5}$$

Katayama showed that the experimental data for various organic substances were in excellent accord with his relation (16.3). How accurately this relation fits the data for argon, nitrogen, etc. can only be confirmed by trial, but anticipating such an examination I shall tentatively and provisionally assume that the relation (16.3) holds for these substances.

By substitution of (16.5) into formula (6.4), one sees that this can be expressed in the form

$$y \propto (1-T/T_c)^{\frac{2}{3}}. \tag{16.6}$$

---

[6] A. Ferguson, Trans. Faraday Soc. 19, 407 (1923); Proc. Phys. Soc. 52, 759 (1940).

[7] Katayama, Science Reports Tôhoku Imperial University 4, 373 (1916).

TABLE II. $\gamma=\gamma_0(1-T/T_c)^{11/9}$.

Ne			A			N_2			O_2		
$T_c=44.8°K$			$T_c=150.7°K$			$T_c=126.0°K$			$T_c=154.3°K$		
$\gamma_0=15.1$ dyne/cm			$\gamma_0=36.31$ dyne/cm			$\gamma_0=28.4$ dyne/cm			$\gamma_0=38.4$ dyne/cm		
	γ dyne/cm			γ dyne/cm			γ dyne/cm			γ dyne/cm	
T °K	calc.	obs.	T °K	calc.	obs.	T °K	calc.	obs.	T °K	calc.	obs.
24.8	5.64	5.61	85.0	13.16	13.19	70.0	10.54	10.53	70.0	18.34	18.35
25.7	5.33	5.33	87.0	12.67	12.68	75.0	9.40	9.39	75.0	17.02	17.0
26.6	5.02	4.99	90.0	11.95	11.91	80.0	8.29	8.27	80.0	15.72	15.73
27.4	4.75	4.69				85.0	7.20	7.20	85.0	14.44	14.5
28.3	4.45	4.44				90.0	6.14	6.16	90.0	13.17	13.23
$V_c=41.7$ cm³/mole			$V_c=75.3$ cm³/mole			$V_c=90.2$ cm³/mole			$V_c=74.5$ cm³/mole		
$\gamma_0 V_c^{\frac{1}{3}}T_c^{-1}$			$\gamma_0 V_c^{\frac{1}{3}}T_c^{-1}$			$\gamma_0 V_c^{\frac{1}{3}}T_c^{-1}$			$\gamma_0 V_c^{\frac{1}{3}}T_c^{-1}$		
$=4.05$ erg deg.$^{-1}$ mole$^{-\frac{1}{3}}$			$=4.3$ erg deg.$^{-1}$ mole$^{-\frac{1}{3}}$			$=4.5$ erg deg.$^{-1}$ mole$^{-\frac{1}{3}}$			$=4.4$ erg deg.$^{-1}$ mole$^{-\frac{1}{3}}$		

Now eliminating y between (16.3) and (16.6), one obtains

$$y \propto (1-T/T_c)^{11/9}, \tag{16.7}$$

which can also be written

$$\gamma=\gamma_0(1-T/T_c)^{11/9}, \tag{16.8}$$

in agreement with (16.1) with $r=2/9$.

To recapitulate, for argon, nitrogen, and other substances under consideration the validity of formula (16.6) has already been established in Section 6. Taking account of this, formula (16.8) and Katayama's formula (16.3) are mutually inter-dependent. Verification of either implies the accuracy of the other and conversely. I shall now show that formula (16.8) holds satisfactorily for the substances under discussion for which experimental data exist, namely neon, argon, nitrogen, and oxygen.

In Table II at the top, are given the experimental value of the critical temperature T_c and the arbitrarily assumed value of γ_0 for each substance. Below this the values of γ calculated according to (16.8) are compared with the observed values over a range of temperatures. The agreement is good throughout and could hardly be improved by amending the value 2/9 assumed for r.

If now instead of eliminating y between (16.3) and (16.6) to obtain (16.7), one eliminates T/T_c one obtains

$$\gamma \propto y^s, \quad s=3(1+r)=3\tfrac{2}{3}. \tag{16.9}$$

This is in definite disagreement with the better known relation of McLeod[8] according to which s is 4. Actually McLeod based his relation on half a dozen organic compounds, but a mere glance at his figures is sufficient to show that γ in fact varies as some power of y less than 4.

It is outside the scope of the present paper to discuss the best value of s to represent the over-all average behavior of organic compounds. I merely wish to emphasize that for the substances having the simplest and most symmetrical molecules the value of s is not 4 but $3\frac{2}{3}$. Any molecular theory[9] or model of the simplest molecules leading to $s=4$ is therefore spurious.

17. ABSOLUTE VALUE OF SURFACE TENSION

In Section 15 it was shown that according to the principle of corresponding states $\gamma v_c^{\frac{1}{3}}T_c^{-1}$ should be a universal function of T/T_c. Seeing that for the substances under consideration the dependence of γ on the temperature is represented by formula (16.8), it follows that $\gamma_0 v_c^{\frac{1}{3}}T_c^{-1}$ should have a universal value. How nearly this is the case is shown at the bottom of Table II. The agreement between argon, nitrogen, and oxygen is remarkable; that of neon is less good, presumably owing to quantal effects.

18. DEVIATIONS FROM THE PRINCIPLE

In the column of Table I preceding neon and labelled "Ideal," are given values of those quantities which should be common to all sub-

[8] J. McLeod, Trans. Faraday Soc. **19**, 38 (1923).
[9] Fowler and Guggenheim, *Statistical Thermodynamics* (Cambridge University Press New York, 1939), p. 451.

stances conforming to the principle of corresponding states. It has been shown that argon, krypton, and xenon conform with a high degree of accuracy to the principle in all their properties; that nitrogen, oxygen, carbon monoxide, and methane also conform with fair accuracy in the gaseous and liquid states but not in the solid state. The behavior corresponding to argon, krypton, and xenon should be considered as an ideal one to which other substances may be expected to conform more or less according to the extent to which the assumptions of Section 2 are or are not satisfied. Departures of a particular substance from this ideal behavior must not be regarded as unsatisfactory flaws in the principle, but rather as giving interesting information concerning fundamental differences between the molecules of the substance in question and the ideal molecule visualized in Section 2. The varied effects of such deviations from the ideal have been discussed by others, notably by Hildebrand.[10] by Bernal,[11] and by Pitzer.[1]

APPENDIX

The following sources of experimental data have been used:

1. Critical data, densities of coexistent liquid and vapor.
   Neon: Mathias, Crommelin, and Onnes, Comptes rendus **175**, 933 (1922).
   Mathias and Crommelin, Comm. Leiden Supp. 52 (1924).
   Argon: Onnes and Crommelin, Comm. Leiden No. 131a (1912).
   Krypton: Mathias, Crommelin, and Meihuizen, Comptes rendus **204**, 630 (1937).
   Xenon: Patterson, Cripps, and Gray, Proc. Roy. Soc. **86**, 579 (1912).
   Nitrogen: Mathias, Onnes, and Crommelin, Comm. Leiden No. 145c (1914).
   Oxygen: Mathias and Onnes, Comm. Leiden No. 117 (1911).
   Carbon monoxide: Crommelin, Bijleveld, and Grigg, Comm. Leiden **20** No. 221b (1932).
   Methane: Cardoso, J. Chim. Phys. **13**, 312 (1915).

2. Boyle points.
   Neon, argon, nitrogen: Holborn and Otto, Zeits f. Physik **33**, 9 (1925).
   Carbon monoxide: rough interpolation of data of Bartlett, J. Am. Chem. Soc. **52**, 1374 (1930).
   Methane: Keyes and Burkes, J. Am. Chem. Soc. **49**, 1403 (1924).

3. Virial coefficients.
   Holborn and Otto, Zeits f. Physik **33**, 5 (1925).

4. Vapor pressures.
   Landolt-Börnstein tables.

5. Heats of evaporation and of fusion, triple-points, heat capacities of liquids.
   Neon: Clusius, Zeits. f. Physik. Chemie **B4**, 1 (1929).
   Clusius, and Riccoboni, Zeits. f. Physik. Chemie **B38**, 81 (1937).
   Argon: Eucken and Hauck, Zeits. f. Physik. Chemie **134**, 161 (1928).
   Clusius, Zeits. f. Physik. Chemie **31**, 459 (1936).
   Frank and Clusius, Zeits. f. Physik. Chemie **B42**, 395 (1939).
   Krypton: Clusius, Kruis, and Konnertz, Ann. d. Physik **33**, 642 (1938).
   Xenon: Clusius and Riccoboni, Zeits. f. Physik. Chemie **B38**, 81 (1937).
   Nitrogen: Giauque and Clayton, J. Am. Chem. Soc. **55**, 4875 (1933).
   Oxygen: Giauque and Johnston, J. Am. Chem. Soc. **51**, 2300 (1929).
   Carbon monoxide: Clayton and Giauque, J. Am. Chem. Soc. **54**, 2610 (1932).
   Methane: Frank and Clusius, Zeits. f. Physik. Chemie **B36**, 291 (1937).

6. Coefficients of thermal expansion.
   Baly and Donnan, J. Chem. Soc. **81**, 907 (1902).

7. Densities of liquid and crystal at triple-point.
   Clusius and Weigand, Zeits. f. Physik. Chemie **B46** (1940).

8. Surface tensions.
   Neon: van Urk, Keesom, and Nijhoff, Comm. Leiden No. 182b (1926).
   Argon: Baly and Donnan, corrected by Rudorf, Ann. d. Physik [4] **29**, 764 (1909).
   Nitrogen, oxygen: Baly and Donnan, Trans. Chem. Soc. **81**, 907 (1902).

[10] J. Hildebrand, J. Chem. Phys. **7**, 233 (1939).
[11] J. Bernal, Address to Chemical Society, February 3, (1938).

BIBLIOGRAPHY

The entries in this bibliography of monographs and review articles on the classical equilibrium theory of fluids are identified as to principal topical content by the letters in parenthesis following each entry. The meaning of these letters are:

(C) Statistical mechanical theories of capillarity

(E) Statistical theory of mechanical (equilibrium) properties of fluids, e.g., elasticity of real fluids

(EPF) Electromagnetic properties of fluids

(P) Phenomenological description of fluid properties (extensive summaries)

(PF) Foundations and principles of statistical mechanics

(RSF) Radiation and particle scattering by fluids

(SE) Statistical mechanics of solutions of electrolytes

(SMF) Statistical mechanics and thermodynamics of fluids

(SN) Statistical mechanics of solutions of nonelectrolytes

(T) Thermodynamics texts of special bearing to fluid theory

1. Barker, J. A., "Lattice Theories of the Liquid State," Macmillan, New York, 1963. (SMF),(NS)

2. Bauer, E., et al., "Changements de Phases," Société de Chimie Physique, Presses Universitaires de France, Paris, 1952. (SMF)(T),(SN),(P)

Bibliography (Cont'd.) - 2

3. Beattie, J. A. and Stockmayer, W. H., "A Treatise on Physical Chemistry," edited by H. S. Taylor, Van Nostrand, New York, 1951. (SMF),(T)

4. Becker, R., "Theorie der Wärme," Springer, Berlin, 1955. (PF)

5. Bergmann, L., "Der Ultraschall," 6th Edition, Hirzel, Stuttgart, 1954. (P),(E),(T)

6. Bogoliubov, N. N., "Problems of a Dynamical Theory in Statistical Physics," translated from Russian by E. K. Gora in "Studies in Statistical Mechanics," Vol. 1, edited by J. De Boer and G. E. Uhlenbeck, North Holland, Amsterdam, 1962. (SMF)

7. Boltzmann, L., "Vorlesungen über der kinetische Theorie der Gase," Leipzig, 1920, 3rd Edition. (SMF)

8. Born, M. and Green, H. S., "A General Kinetic Theory of Liquids," Cambridge, London, 1949. (SMF)

9. Brown, G. H. and Shaw, W. G., "Liquid Crystals," Chem. Rev. 57, 1049, (1957). (T),(P)

10. Brown, W. B., Discussion, "Thermodynamic and Transport Properties of Fluids," Int. Union Pure and Appl. Chem. and Inst. Mech. Eng., London, 1958. (SMF)

11. Buff, F. B. and Kirkwood, J. G., "Solutions of Nonelectrolytes," Ann. Rev. Phys. Chem. 2, 51, (1951). (SN)

12. Callen, H. B., "Thermodynamics," Wiley, New York, 1960. (T)

13. Cole, G. H. A., "The kinetic theory of monotomic liquids at ordinary temperatures," Rep. Prog. Phys. 19, 1, (1956). (SMF)

Bibliography (Cont'd.) - 3

14. Davies, R. O. and Jones, G. O., "Thermodynamic and kinetic properties of glass," Advance in Phys. 2: 370-410, (1953). (T)

15. De Boer, J., "Molecular distribution and equation of state of gases," Repts. Progr. Phys. 12, 305, (1949). (SMF)

16. De Boer, J., "Theory of monatomic liquids," Nuovo Cimento 10 Suppl. No. 3, 225-50, (1953). (SMF)

17. Dobbs, E. R. and Jones, G. D., "Theory and properties of solid argon," Rep. Progr. Phys. 20, 516, (1957). (SMF)

18. Epstein, P. S., "Textbook of Thermodynamics," Wiley, New York, 1937. (T)

19. Falkenhagen, H., "Elektrolyte," S. Hirzel, Leipzig, 1953. (SE)

20. Fisher, I. Z., "Statistical Theory of Liquids," (in Russian) State Publishing House, Physico - Math Literature, Moscow, 1961. (SMF)

21. Fournet, G. in Handbuch der Physik, (edited by S. Flügge), Vol. XXXII, Springer, Berlin, 1957, p. 238, "Etude de la structure des fluids et des substance amorphes au moyen de la diffusion des rayons α." (RSF)

22. Fowler, R. H., "Statistical Mechanics," Cambridge, London, 1936. (PF)

23. Fowler, R. H. and Guggenheim, E. A., "Statistical Thermodynamics," Cambridge, London, 1939. (PF),(SMF),(C),(SN), (SE),(EPF)

Bibliography (Cont'd.) - 4

24. Frenkel, J., "Kinetic Theory of Liquids," Oxford, London, 1946. (SMF),(E),(C),(SN)
25. Friedman, H. L., "Ionic Solution Theory," Interscience, New York, 1962. (SE)
26. Frisch, H. L., "The Equation of State of the Classical Hard Sphere Fluid," Adv. Chem. Phys. 6, 229 (1964).
27. Gibbs, J. W., "The Collected Works of J. W. Gibbs," (Yale, New Haven, 1948), Vol. II. (PF)
28. Gibbs, J. W., "The Collected Works of J. W. Gibbs," (Yale, New Haven, 1948), Vol. I. (T)
29. Gingrich, N. S., "The Diffraction of X-rays by Liquid Elements," Rev. Mod. Phys. 15, 90 (1943). (RSF)
30. Green, H. S., "Molecular Theory of Fluids," North Holland, Amsterdam, 1952. (SMF),(C),(E),(RSF)
31. Guggenheim, E. A., "Thermodynamics," North Holland, Amsterdam, 1949. (T)
32. Guggenheim, E. A., "Mixtures," Oxford, London, 1952, (SN)
33. Harasima, A., "Theory of Liquids," (in Japanese), Iwanami-Shoten, Tokyo, 1954. (SMF),(C)
34. Harasima, A., "Molecular Theory of Surface Tension," Adv. Chem. Phys. 1, 203, (1958). (C)
35. Harned, H. S. and Owen, B., "The Physical Chemistry of Electrolyte Solutions," 3rd Edition, Reinholt, New York, 1958. (SE)
36. Hildebrand, J. H. and Scott, R. L., "The Solubility of Nonelectrolytes," Reinhold, New York, 1950. (SN)

Bibliography (Cont'd.) - 5

37. Hill, T. L., "Statistical Mechanics," McGraw-Hill, New York, 1956. (PF),(SMF),(SN)

38. Hill, T. L., "Introduction to Statistical Thermodynamics," Addison-Wesley, Reading, 1960. (SMF),(SN),(SE)

39. Hirschfelder, J. O., Curtiss, C. F. and Bird, R. B., "Molecular Theory of Gases and Liquids," Wiley, New York, 1954. (SMF),(RSF),(C),(EPF)

40. Huang, K., "Statistical Mechanics," Wiley, New York, 1963. (PF),(SMF)

41. Jones, G. O., "Glass," Methuen, London, 1956. (T),(SMF),(P)

42. Kahn, B., Dissertation, "Statistical Mechanics of Fluids," Utrecht, 1938. (SMF)

43. Kauzmann, W., "Nature of the glassy state and the behavior of liquids at low temperatures," Chem. Rev. 43, 219, (1948). (T),(SMP),(P)

44. Khinchin, A. I., "Mathematical Foundations of Statistical Mechanics," Dover, New York, 1949. (PF)

45. Kihara, T., "Intermolecular Forces and Equation of State of Gases," Adv. Chem. Phys. 1, 267, (1958). (SMF)

46. Kimball, G. E., "A Treatise on Physical Chemistry," edited by H. S. Taylor and S. Glasstone, Van Nostrand, New York, 1951. (SMF)

47. Kuenen, J. P., "Theorie der Verdampfung und Verflüssigung von Gemischen," Barth, Leipzig, 1906. (T),(SN),(P)

48. Landau, L. D. and Lifshitz, E. M., "Statistical Physics," Pergamon, London, 1958. (PF),(SMF),(SN),(SE),(C)

Bibliography (Cont'd) - 6

49. Landsberg, P. T., "Thermodynamics," Interscience, 1961. (T)

50. Lewis, G. N. and Randall, M., "Thermodynamics," McGraw-Hill, New York, 1923. (T)

51. Margenau, H., "Van der Waals forces," Rev. Mod. Phys. 11, 1, (1939). (SMF)

52. Massignon, D., "Mécanique Statistique des Fluides," Dunod, Paris, (1957). (SMF)

53. Mayer, J. E. and Mayer, M. G., "Statistical Mechanics," Wiley, New York, 1940. (SMF),(PF)

54. Mayer, J. E., "Theory of Real Gases," in Handbuch der Physik, (edited by S. Flügge), Vol. XII, Springer, Berlin, 1958. (SMF)

55. Mazur, P., "On Statistical Mechanics and Electromagnetic Properties of Matter," Adv. Chem. Phys. 1, 309, (1958). (EPF)

56. Miller, A. R., "The Theory of Solutions of High Polymers," Oxford, London, 1948. (SN)

57. Münster, A., "Statistische Thermodynamik," Springer, Berlin, 1956. (PF),(SMF),(RSF)

58. Münster, A., "Statistische Thermodynamik Kondensierten Phasen," in Handbuch der Physik, (edited by S. Flügge), Vol. XIII, Springer, Berlin, 1962. (SNF),(EPF),(RSF), (SN),(SE)

59. Münster, A., "Prinzipien der statistischen Mechanik," in Handbuch der Physik, edited by S. Flügge, Vol. III/2, Springer, Berlin, 1959. (PF)

Bibliography (Cont'd.) - 7

60. Partington, J. R., "An Advanced Treatise on Physical Chemistry," Vol. 1, Properties of Gases, (1949), Vol. 2, Properties of Liquids, (1951), Longmans, London, 1949-1951. (P)

61. Pesce, B., ed., "Electrolytes," Pergamon, London, 1962. (SE)

62. Pippard, A. B., "The Elements of Classical Thermodynamics," Cambridge, London, 1960. (T)

63. Pitzer, K. S. and Curl, R. F., The Thermodynamic properties of normal fluids, p. 1, in "Thermodynamic and Transport Properties of Fluids," Int. Union Pure and Appl. Chem. and Inst. Mech. Eng., London, 1958, Joint conference on thermodynamic and transport properties of fluids, Proc., London. (SMF)

64. Prager, S., "The One-Dimensional Plasma," Adv. Chem. Phys. 4, 201, (1962). (PF)

65. Prigogine, I., "The Molecular Theory of Solutions," North Holland, Amsterdam, 1957. (SN),(SMF),(T)

66. Prigogine, I. and Defay, R., "Chemical Thermodynamics," Vol. 1, (translated by D. H. Everett), Longmans, London, 1954. (T),(SN),(SE)

67. Reiss, H., "The Scaled Particle Theory of Fluids," Adv. Chem. Phys. (in press)

68. Rice, O. K., "Critical phenomena in binary liquid systems," Chem. Rev. 44, 69, (1949). (SMF),(P)

69. Robinson, R. and Stokes, R., "Electrolyte Solutions," Butterworth, London, 1955. (SE)

Bibliography (Cont'd.) - 8

70. Rowlinson, J. S., "Intermolecular forces and the properties of matter," Quart. Rev. Chem. Soc., London 8, 168, (1954). (SMF)

71. Rowlinson, J. S., Sutton, J. R., and Weston, J. F., "Thermodynamic and Transport Properties of Fluids," Int. Union Pure Appl. Chem. and Inst. Mech. Eng., London, 1958. (T),(P)

72. Rowlinson, J. S., "Liquids and Liquid Mixtures," Academic, New York, 1959. (SMF),(SN),(T)

73. Rushbrooke, G. S., "Introduction to Statistical Mechanics," Oxford, London, 1949. (PF),(SN)

74. Salsburg, Z. W., "Solutions of Nonelectrolytes," Ann. Rev. Phys. Chem. 8, 55, (1957). (SN)

75. Scatchard, G., Chem. Revs. 8, 321, (1931); 44, 7, (1949). (SE)

76. Schneider, W. G., "Changements des Phases," Soc. Chim. Phys., Paris, 1952. (SMF),(T),(P)

77. Schrödinger, E., "Statistical Thermodynamics," Cambridge, London, 1948. (PF)

78. Temperley, H. N. V., "Changes of State," Cleaver-Hume, London, 1956. (SMF),(T)

79. Ter Haar, D., "Elements of Statistical Mechanics," Rinehart, New York, 1954. (PF)

80. Timmermans, J., "Les solutions concentrées," Masson, Paris, 1936. (SN),(T)

81. Tisza, L., "On the General Theory of Phase Transitions," in "Phase Transformations in Solids," edited by R. Smoluchowski, J. Mayer and Weyl, Wiley, New York, 1951. (T)

Bibliography (Cont'd.) - 9

82. Tolman, R. C., "The Principles of Statistical Mechanics," Oxford, London, 1938. (PF)

83. Tompa, H., "Polymer Solutions," Butterworths Sci. Publ., London, 1956. (SN)

84. Uhlenbeck, G. E. and Ford, G. W., "The Theory of Linear Graphs with Applications to the Theory of the Virial Development of the Properties of Gases," in "Studies in Statistical Mechanics," Vol. 1, edited by J. De Boer and G. E. Uhlenbeck, North Holland, Amsterdam, 1962. (SMF)

85. Van der Waals, J. D., "Die Kontinuität des gasförmingen und flüssigen Zustandes," Vol. 1 (1899), Vol. 2 (1900), Barth, Leipzig. (T),(P)

86. Van der Waals, J. D. and Kohnstamm, P., "Lehrbuch der Thermodynamik Barth, Leipzig, (1908). (T),(SN)

87. Van Laar, J. J., "Die Thermodynamik Einheitlicher Stoffe und Binärer Gemische," Noordhoff, Groningen, 1936. (T),(SN)

88. Wilson, A. H., "Thermodynamics and Statistical Mechanics," Cambridge, 1957. (T),(PF)

89. Yvon, J., "La théorie statistique des fluides et l'équation d'état," Actualités scientifique et industrielles 203, Hermann, Paris, 1935. (SMF)

90. Yvon, J., "Fluctuations en densité," Actualités scientifique et industrielles 542, Hermann, Paris, 1937. (SMF)

91. Yvon, J., "La Propagation et la diffusion de la lumière," Actualités 543, Hermann, Paris, 1937. (RSF),(SMF)

92. Zernike, F., "L'Opalescence critique," Dissertation, Amsterdam, 1915. (SMF),(RSF)

93. Zernike, J., "Chemical Phase Theory," Kluwer, Antwerp, 1956. (T)