Mathematical Methods In Theoretical Physics

Edited by

Wesley E. Brittin
Department of Physics and Astrophysics
University of Colorado

Colorado Associated University Press
Boulder, Colorado

Colorado Associated University Press
1424 Fifteenth Street
Boulder, Colorado 80302

Library of Congress Catalog Number 59-13034
ISBN 0-87081-047-2

THE LECTURERS

Eytan Barouch
Massachusetts Institute of Technology, Cambridge, Mass.

Wesley E. Brittin
University of Colorado, Boulder, Colorado

Joseph Dreitlein
University of Colorado, Boulder, Colorado

J. G. Garrison
University of California, Livermore, California

M. D. Girardeau
University of Oregon, Eugene, Oregon

Rudolf Haag
Hamburg University, Hamburg, Germany

N. M. Hugenholtz
University of Groningen, Groningen, The Netherlands

Herbert Jehle
George Washington University, Washington, D. C.

John R. Klauder
Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

Elliott H. Lieb
Massachusetts Institute of Technology, Cambridge, Mass.

Joel L. Lebowitz
Yeshiva University, New York, New York

Harry Morrison
University of California, Berkeley, California

David Ruelle
Institute for Advanced Study, Princeton, New Jersey

Arthur Y. Sakakura
Colorado School of Mines, Golden, Colorado

G. L. Sewell
Queen Mary College, London, England

L. Streit
Syracuse University, Syracuse, New York

F. Strocchi
Princeton University, Princeton, New Jersey

J. Wong
University of California, Livermore, California

FOREWORD

This volume contains material based on lectures given during the second part of the Colorado Summer Institute for Theoretical Physics held during July and August, 1971.

The portion of the Institute represented here is devoted to mathematical physics, with topics ranging over statistical mechanics, dissipative systems, composite particles, algebraic methods and field theory.

Volumes XIV represent the last of the current series of the Boulder Lectures in Theoretical Physics since necessary support is no longer available. It is hoped that the Institutes have served a useful purpose by stimulating young scientists as well as old to work in some of the fascinating fields which have been covered. The Institutes have certainly played an important role in physics at the University of Colorado and at this time I wish to thank all who have participated over the years.

The Institute was sponsored by the National Science Foundation.

I wish to thank the lecturers and the participants for their effort for a lively Institute and to the secretary, Mrs. Charlotte Walker for her invaluable contribution to the organization of the Institute. I would also like to extend my appreciation to Mrs. Walker for the typing of the manuscript.

Boulder, August 1973

Wesley E. Brittin

CONTENTS

STATISTICAL MECHANICS OF THE XY-MODEL

Eytan Barouch
Department of Mathematics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Introduction

Spin systems play a major role within the general framework of statistical mechanics and many body theory, since they provide us with explicit examples for which a great deal of exact results are known.

A spin system is specified by a Hamiltonian whose components are spin-spin interactions, and spin-external field interactions. Because of mathematical limitations we restrict ourselves to nearest-neighbor interactions only.

A famous example of a spin system is the Ising model (Ising 1925) introduced by Ising who computed its partition function in one dimension, and Onsager (1944) computed the partition function for the two dimensional lattice with no magnetic field. The Hamiltonian that Onsager studied is given by

$$H_I = -J_1 \sum \sigma^x_{n,m} \sigma^x_{n+1,m} - J_2 \sum \sigma^x_{n,m} \sigma^x_{n,m+1} \qquad (1.1)$$

Onsager found an explicit phase transition as a logarithmic singularity of the specific heat at a critical temperature $T_c \neq 0$, and Yang (1952) computed the spontaneous magnetization by a perturbation theory for $T < T_c$ with the famous $|T-T_c|^{1/8}$ result.

The Ising model represents a "classical" system, since every term in (1.1) commutes with each other. An

obvious generalization of (1.1) to quantum systems with nearest neighbor interactions is the generalized Heisenberg model (in one dimension)

$$H_{xyz} = -\sum_{j=1}^{N} \{A\sigma_j^x\sigma_{j+1}^x + B\sigma_j^y\sigma_{j+1}^y + C\sigma_j^z\sigma_{j+1}^z\} . \qquad (1.2)$$

The complexity of H_{xyz} vs. H_I is very clear since even the ground state of (1.2) is highly nontrivial. Particular cases of (1.2) were studied for a long time. Bethe (1931) found the ground state eigenvector when $A = B = C$, and Hulthen found the ground state eigenvalue. Lieb, Schultz and Mattis (1961) and Katsura (1962) studied the XY-model for which $C = 0$, Yang and Yang (1966) studied the cases $A = B$.

The general case (1.2) was not understood until very recently. Baxter (1971) computed exactly the ground state of (1.2) and his complicated results contain all the other cases as particular limiting cases!

Two-dimensional "classical" lattices can be studied by constructing a "transfer matrix" introduced by Kramers and Wannier (1941), and the log of its largest eigenvalue gives the free energy per site, in the thermodynamic limit. Commutation relations of V with Hamiltonians of quantum lattices suggest that the mathematical tools developed for one are very handy for the other.

McCoy and Wu (1967) demonstrated that a linear Hamiltonian commutes with V of the general six vertex ferroelectrics, Sutherland (1970) demonstrated that the transfer matrix of the eight vertex ferroelectrics commutes with (1.2) for a special choice of A, B, C. Baxter derived the ground state energy of (1.2) using the brilliant method he developed for the 8-vertex problem.

Another example is the relation of the XY-model with a transverse field to the Ising model. Suzuki (1971) shows that the XY-Hamiltonian (Lieb, Schultz and Mattis 1961, Katsura 1962) given by

$$H = \sum_j \{(1+\gamma)\sigma_j^x\sigma_{j+1}^x + (1-\gamma)\sigma_j^y\sigma_{j+1}^y - h\sigma_j^z\} \qquad (1.3)$$

commutes with the transfer matrix of the Ising model with the identification

$$K_i = J_i/kT = \beta J_i$$

$$\tanh K_1^* = e^{-2K_1}, \quad \cosh 2K_1^* = \gamma^{-1}, \text{ and } \tanh 2K_2 = (1-\gamma^2)^{1/2}/h \qquad (1.4)$$

We devote these lectures to the physical properties of (1.3), and because of the commutation relations of Suzuki it is natural to expect an extensive use of the mathematics developed for the Ising model.

There are four major topics we are going to discuss.

(a) Ground state energy and thermodynamics of (1.3).

It is well known that a one-dimensional system with finite nearest neighbor interactions does not exhibit a phase transition at any finite temperature. However, it is not at all clear, that the ground state energy, and the thermodynamic functions at T = 0 are analytic functions of the coupling constants. We wish to study the effect of symmetry or lack of symmetry on the analytical properties of the macroscopic averages. Some symmetry breaking points are apparent ($\gamma=0$), some are not ($h=1$, $h^2+\gamma^2=1$).

The symmetry properties of (1.3) manifest themselves in the behavior of the correlation functions ρ_{vu}. We find a long range order in the x direction of the ground state, namely

$$\lim_{R\to\infty} \rho_{xx} = \begin{cases} \frac{1}{4}(-1)^R \frac{2}{1+\gamma}\left[\gamma^2(1-h)^2\right]^{1/4} & h < 1 \\ 0 & h > 1 \end{cases} . \qquad (1.5)$$

(b) Dynamical properties of many particle systems very near thermal equilibrium are most commonly studied in terms of the time delayed correlations

$$\rho_{vu}(R,t) = \langle S_o^v(0) S_R^u(t) \rangle \ . \tag{1.6}$$

Neimeijer (1967) was able to compute $\rho_{zz}(R,t)$ exactly, and found that all contributions come from two-particle excitations, with t^{-1} approach to the infinite time limit, instead of the commonly believed exponential approach. This result led Mazur (1969) to develop his criteria for nonergodicity of a system, and he demonstrated that the system is not ergodic. Later on, McCoy, Barouch and Abraham (1971) and Johnson and McCoy (1971) studied the rest of the $\rho_{uv}(R,t)$, and found distinction between $h > 1$ or $h < 1$. For $h > 1$, contributions to $\rho_{xx}(R,t)$ come from 1,2,3,... excitations, where for $h < 1$, we have only even number of excitations contributing to the asymptotic series. The only other exact result known is $\rho_{xz}(R,t)$, computed by Johnson and McCoy.

(c) In 1968, McCoy and Wu presented a detailed analysis of the Ising model with random exchange energies. They found that the logarithmic singularity of the specific heat rounds off, infinitely differentiable but non-analytic. Smith (1970) introduced these ideas to the isotropic (γ=0) XY model, and was able to study the influence of these random impurities exactly, using the pioneering work of Dyson (1953) on a random chain of harmonic oscillators. He finds that the singularities of the ground state functions become infinitely differentiable.

(d) Our last topic is non-equilibrium phenomenae, introduced to the XY-model by Niemeijer (1967). Let h in (1.3) be given by

$$h(t) = \begin{cases} a & t \leq 0 \\ h_1(t) & t > 0 \end{cases} \tag{1.6}$$

The system is assumed in thermal equilibrium, and at a specified time t = 0 we turn on a time dependent field. The natural question is whether the thermodynamic functions will approach equilibrium, and what is their asymptotic behavior for large t. Niemeijer (1967) and Barouch, McCoy and Dresden (1970) studied the z-direction magnetization for a step function h(t) namely $h_1(t) = b$. It was found that the infinite time limit is a non-equilibrium limit. If b = 0, the "zero field" magnetization does not vanish. Furthermore, we find a division into regions in the long time behavior of $m_z(t)$. If $h > 1-\gamma^2$, the long time behavior of $m_z(t)$ is $O(t^{-3/2})$ with two oscillating frequencies, independent of γ. If $h < 1-\gamma^2$, $m_z(t)$ is $O(t^{-1/2})$ with one frequency dependent on γ. On the boundary, $m_z(t)$ is $O(t^{-3/4})$.

Our approach is exact solution of the Liouville equation for the density matrix $\rho(t)$

$$i \frac{d}{dt} \rho(t) = [H(t), \rho(t)]. \tag{1.7}$$

We reduce (1.7) to a second order differential equation of the form

$$V''(t) + [\Lambda^2 + \psi(t)]V(t) = 0 \tag{1.8}$$

and express $\rho(t)$ in terms of the solution of (1.8). Doing so, we find that no matter how slowly h(t) varies with time, $\lim_{t\to\infty} m_z(t)$ exists, but this is not an equilibrium limit. Another expression of the nonergodicity of the system is total destruction of the long range order of ρ_{xx}.

It was felt, however, that a <u>local</u> spin would thermalize, and it was found (Abraham, Barouch, Gallavotti, and Martin-Löf, 1970) that an internal spin thermalizes like t^{-1} where a boundary spin (Tjon 1970) thermalizes like t^{-3}.

Ground State and Thermodynamics

The XY Hamiltonian was defined last time by

$$H = \sum_j [(1+\gamma)S_j^x S_{j+1}^x + (1-\gamma)S_j^y S_{j+1}^y - hS_j^z] \qquad (2.1)$$

In order for H to be uniquely defined we choose cyclic boundary condition, namely $S_{N+1}^\alpha \equiv S_1^\alpha$ where $\alpha = x,y,z$, and N is the number of spins in the chain.

We diagonalize H in four steps, following LSM and Katsura.

(i) Express S_j^x, S_j^y, S_j^z in terms of creation and destruction operators.

$$S_j^x = \tfrac{1}{2}(b_j^\dagger + b_j), \quad S_j^y = \frac{1}{2i}(b_j^\dagger - b_j), \quad S_j^z = b_j^\dagger b_j - \tfrac{1}{2}. \qquad (2.2)$$

The operators b_i satisfy a mixed set of commutation relations

$$[b_i^\dagger, b_j] = [b_i^\dagger, b_j^\dagger] = [b_i, b_j] = 0 \qquad i \neq j \qquad (2.3)$$

and the anticommutation relations

$$\{b_i, b_i^\dagger\} = 1 \qquad b_i^2 = (b_i^\dagger)^2 = 0. \qquad (2.4)$$

(ii) Jordan-Wigner transformation. We express the operators $b_j, b_j^\dagger$ in terms of Fermi operators. Let

$$c_\ell = \exp[\pi i \sum_{j=1}^{\ell-1} b_j^\dagger b_j] b_\ell$$
$$c_\ell^\dagger = b_\ell^\dagger \exp[-\pi i \sum_{j=1}^{\ell-1} b_j^\dagger b_j] \qquad (2.5)$$

It is an easy matter to verify the identities

$$b_j^\dagger b_{j+1} = c_j^\dagger c_{j+1}, \quad b_j^\dagger b_{j+1}^\dagger = c_j^\dagger c_{j+1}^\dagger \tag{2.6}$$

$$b_N^\dagger b_1 = -c_N^\dagger c_1 \exp(i\pi \sum_j c_j^\dagger c_j)$$

$$b_N^\dagger b_1^\dagger = -c_N^\dagger c_1^\dagger \exp(i\pi \sum_j c_j^\dagger c_j). \tag{2.7}$$

Inserting (2.6), (2.7), and (2.2) into (2.1), we obtain

$$H = \tfrac{1}{2} \{ \sum_{j=1}^{N} [c_j^\dagger c_{j+1} + c_{j+1}^\dagger c_j + \gamma(c_j^\dagger c_{j+1}^\dagger + c_{j+1} c_j)] -$$

$$- 2h \sum_{j=1}^{N} (c_j^\dagger c_j - \tfrac{1}{2}) - [c_N^\dagger c_1 + c_1^\dagger c_N + \gamma(c_N^\dagger c_1^\dagger + c_1 c_N)] \times$$

$$\times [1 + \exp(i\pi \sum c_j^\dagger c_j)] \} . \tag{2.8}$$

The last term in (2.8) is the only term that is not quadratic in the Fermi operators $c_j, c_j^\dagger$, and it comes from the imposed cyclic boundary conditions.

LSM observed that in most of the thermodynamics averages, the last term can be dropped and called the "modified boundary condition" c-cyclic. This is indeed correct in the thermodynamic limit $N \to \infty$, and we will adopt the c-cyclic condition for most of the discussions. However, for clarity and completeness we outline here the treatment of (2.8) (Katsura 1962). This treatment is needed to demonstrate the difficulties that rise in computing the transverse time-delayed correlations.

Let

$$P_{\pm} = \tfrac{1}{2}[1 \pm \exp(i\pi \sum_{j=1}^{N} c_j^{\dagger} c_j)] \tag{2.9}$$

The operator $P_+(P_-)$ is the projection operator for states with an even (odd) number of c_j excitations. The Hamiltonian (2.8) then decomposes into

$$H = H_+ P_+ + H_- P_- \tag{2.10}$$

where we assume N to be even, and $H_{\pm}$ are given by

$$H_{\pm} = \tfrac{1}{2} \{ \sum_{j=1}^{N-1} [c_j^{\dagger} c_{j+1} + c_{j+1}^{\dagger} c_j + \gamma(c_j^{\dagger} c_{j+1}^{\dagger} + c_{j+1} c_j) - 2h \sum_{j=1}^{N} (c_j^{\dagger} c_j - \tfrac{1}{2}) \pm [c_N^{\dagger} c_1 + c_1^{\dagger} c_N + \gamma(c_N^{\dagger} c_1^{\dagger} + c_1 c_N)]] \}. \tag{2.11}$$

The decomposition (2.10) is expected, since H commutes with the parity operator $P_+ - P_-$. Therefore, when acting on a state with even (odd) number of c_j excitations, H may be replaced by $H_+(H_-)$. Therefore, the c-cyclic condition means replacing H by H_+ which is permitted for calculation of expectation values of even operators, like S_j^z, $S_j^x S_{j+R}^x$, but not for odd operators like S_j^x, $S_j^x S_k^y S_\ell^x$ and so forth. Those who are interested in utmost rigour are referred to Katsura's paper (1962) who gave a very thorough treatment of this recondite point.

Using the c-cyclic condition we write H as (with the understanding that N is very large)

$$H = \tfrac{1}{2} \sum_{j=1}^{N} [(c_j^{\dagger} c_{j+1} + \gamma c_j^{\dagger} c_{j+1}^{\dagger}) - 2h c_j^{\dagger} c_j] + \tfrac{1}{2} N h. \tag{2.12}$$

(iii) Fourier decomposition

Let

$$c_j^\dagger = N^{-\frac{1}{2}} \sum_{p=-N/2}^{N/2} \exp(ij\varphi_p) a_p^\dagger \qquad (2.13)$$

$$c_j = N^{-\frac{1}{2}} \sum_{p=-N/2}^{N/2} \exp(-ij\varphi_p) a_p \qquad \varphi_p = 2\pi p/N$$

By direct substitution of (2.31) into (2.12) we obtain

$$H = \sum_{p=1}^{N/2} H_p \qquad (2.14)$$

where

$$H_p = (\cos \varphi_p - h)\ (a_p^\dagger a_p + a_{-p}^\dagger a_{-p}) - \frac{i}{2} \gamma \sin \varphi_p (a_p^\dagger a_{-p}^\dagger + a_p a_{-p}) + h \qquad (2.15)$$

Since one obtains

$$[H_p, H_q] = 0, \qquad (2.16)$$

all H_p can be diagonalized simultaneously.

(iv) Bogoliubov-Valatin transformation.

To diagonalize H_p-h, we change the phase of a_p, a_{-p} by $e^{i\pi/4}$, and write linear combinations of the form

$$e^{i\pi/4} a_p = \cos \theta_p \eta_p + \sin \theta_p \eta_{-p}^\dagger$$
$$e^{i\pi/4} a_{-p} = \cos \theta_p \eta_{-p} - \sin \theta_p \eta_p^\dagger \qquad (2.17)$$

where

$$\tan 2\theta_p = -\frac{\gamma \sin \varphi_p}{\cos \varphi_p - h}, \quad 0 \le \theta \le \frac{\pi}{2} \tag{2.18}$$

and the transformation is cannonical since $\theta_p = -\theta_{-p}$. By direct substitution of (2.17) and (2.18) into (2.14) we obtain

$$H = \sum_p \Lambda(\varphi_p)(\eta_p^\dagger \eta_p - \tfrac{1}{2}) + \tfrac{1}{2}Nh \tag{2.19}$$

where $\eta_p^\dagger \eta_p$ is a Fermi number operator, and

$$\Lambda(\varphi) = [\gamma^2 \sin^2\varphi + (\cos \varphi - h)^2]^{\frac{1}{2}} \tag{2.20}$$

We finally obtain the ground state per particle to be

$$E = \tfrac{1}{2}h - \frac{1}{2N} \sum_p \Lambda(\varphi_p) \tag{2.21}$$

which in the thermodynamic limit $N \to \infty$ becomes

$$E = \tfrac{1}{2}h - \frac{1}{2\pi} \int_0^\pi d\varphi [\gamma^2 \sin^2\varphi + (\cos \varphi - h)^2]^{\frac{1}{2}}. \tag{2.22}$$

The ground state (2.22) is definitely not an analytic function of h, γ for all h, γ. For instance let h = 0. Then

$$E = -\frac{1}{\pi} \int_0^{\pi/2} d\varphi \, [1 - (1-\gamma^2) \sin^2\varphi]^{\frac{1}{2}} = -\frac{1}{\pi}\mathcal{E}[(1-\gamma^2)^{\frac{1}{2}}] \tag{2.23}$$

where $\mathcal{E}(k)$ is the complete elliptic integral of the second kind (GR 8.112), with a singularity proportional to $\gamma^2 \log\gamma$ for $\gamma \sim 0$. (See appendix).

The function $\Lambda(\varphi)$ is the one particle excitation, and it has a gap that disappears at h = 1. So one might expect a symmetry breaking point at this value. This point will be further discussed later on.

We now turn to discuss the thermodynamics of the system. Since H in (2.19) is expressed in terms of non-interacting Fermions with "kinetic energy" $\Lambda(\varphi)$, the free energy per site is then given by

$$\beta f(h,\gamma,\beta) = -\frac{1}{\pi}\int_0^{\pi} \ln\{2\cosh[\tfrac{1}{2}\beta\Lambda(\varphi)]\}\,d\varphi \qquad (2.24)$$

The rest of the thermodynamics is then straightforward.

The magnetization in the z-direction is given by

$$m_z = \frac{1}{2\pi}\int_0^{\pi} d\varphi(h - \cos\varphi)\tanh[\tfrac{1}{2}\beta\Lambda(\varphi)][\Lambda(\varphi)]^{-1} \qquad (2.25)$$

the internal energy U is given by

$$U = -\beta^{-1}\frac{1}{2\pi}\int_0^{\pi}\Lambda(\varphi)\tanh[\tfrac{1}{2}\beta\Lambda(\varphi)]\,d\varphi \qquad (2.26)$$

the specific heat c

$$c = \frac{\partial U}{\partial T} = \{-\frac{k}{2\pi}\int_0^{\pi}\Lambda(\varphi)\tanh[\tfrac{1}{2}\beta\Lambda(\varphi)]\,d\varphi +$$

$$+\beta^{-3}\frac{k}{2\pi}\int_0^{\pi}\Lambda^2(\varphi)\{1+\tanh^2[\tfrac{1}{2}\beta\Lambda(\varphi)]\}d\varphi\} \qquad (2.27)$$

and the susceptibility χ_z is given by

$$\chi_z = \frac{\partial m_z}{\partial h} \qquad (2.28)$$

It is very clear that there is no phase transition at any finite temperature, since $m_z(h) = 0$ when $h = 0$.

At the ground state [T=0 or $\beta=\infty$] we have a different situation. The interesting functions are $m_z(h)$, $\chi_z(h)$, and their singular behavior was studied by Niemeijer.

For $\gamma = 0$ the magnetization behaves in a non-analytic way, namely

$$m_z(h) = \begin{cases} \frac{1}{2} & h \geq 1 \\ \frac{1}{2} - \frac{1}{\pi} \text{ arc cosh} & 0 \leq h \leq 1 \end{cases} \qquad (2.29)$$

For $\gamma \neq 0$, Niemeijer evaluated the magnetization numerically, and found a continuous non-analytic behavior of $m_z(h)$ at $h = 1$.

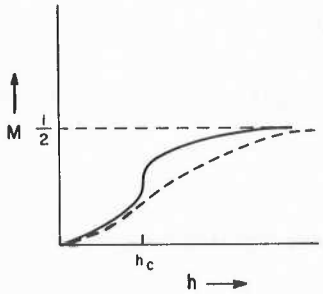


Fig.1a. $\langle M^z \rangle$ as function of h for T = 0 (solid line) and T > 0 (dashed line).

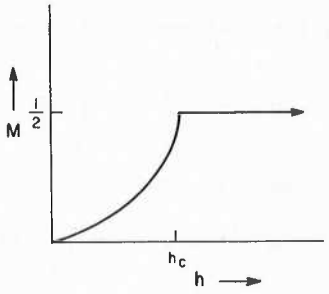


Fig.1b. Magnetization of the ground state as function of h for $\gamma = 0$.

(Th. Niemeijer, Physica 36, 377, 1967)

The ground state susceptibility χ_z exhibits a logarithmic singularity near h = 1. To see that rewrite 2.28 for $\beta = \infty$:

$$\chi_z(h) = \frac{\partial}{\partial h}\frac{1}{2\pi}\int_0^{\pi}\frac{(h-\cos\varphi)\,d\varphi}{[(h-\cos\varphi)^2+\gamma^2\sin^2\varphi]}$$

$$= \frac{1}{2\pi}\int_0^{\pi}\frac{d\varphi}{[(h-\cos\varphi)^2+\gamma^2\sin^2\varphi]^{1/2}}$$

$$-\frac{1}{2\pi}\int_0^{\pi}\frac{(h-\cos\varphi)^2\,d\varphi}{[(h-\cos\varphi)^2+\gamma^2\sin^2\varphi]^{3/2}}$$

$$= \frac{1}{2\pi}\int_0^{\pi}\frac{\gamma^2\sin^2\varphi}{[(h-\cos\varphi)^2+\gamma^2\sin^2\varphi]^{3/2}}\,d\varphi \qquad (2.30)$$

To evaluate $\chi_z(h)$ write $\int_0^{\pi} = \int_0^{\delta} + \int_{\delta}^{\pi}$ where δ is small but finite. The susceptibility is given by

$$\chi_z(h) = I_1(\delta) + I_2(\delta) \qquad (2.31)$$

where $I_2(\delta)$ is a smooth function of h so we need to study only I_1, for $\varphi \sim \delta$, namely

$$I_1 \simeq \frac{1}{2\pi}\int_0^{\delta}\frac{\gamma^2\varphi^2\,d\varphi}{[\gamma^2\varphi^2+(h-1)^2+\varphi^2(h-1)]^{3/2}} \qquad (2.32)$$

$$I_1(\delta,h,\gamma) \sim \frac{1}{2\pi}\gamma^2 p^{-3/2}\{-\delta[\delta^2+p(h-1)^2]^{-1/2}$$

$$+\log[\delta+(\delta^2+(h-1)^2p)^{\frac{1}{2}}] - \tfrac{1}{2}\log[p(h-1)^2]\} \qquad (2.33)$$

and p is given by $p = [\gamma^2+(h-1)^2]^{-1}$.

The dominant behavior near h = 1 comes from the last term of (2.33), so $\chi_z(h,\gamma)$ can be written as

$$\chi_z(h,\gamma) \simeq -\frac{1}{2\pi}\gamma^2[\gamma^2+(h-1)^2]^{-3/2}\log|h-1| + f(h) \qquad (2.34)$$

where f(h) is bounded and continous in h. Graphically, $\chi_z(h)$ behaves as

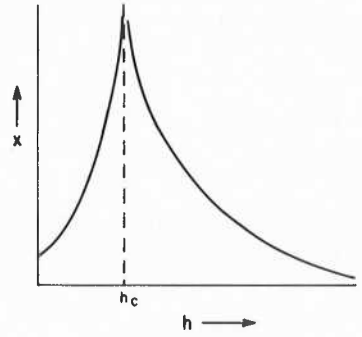


Fig. 2a. $\chi^z(h)$ as function of h for T = 0.

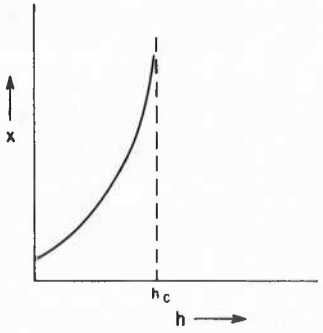


Fig.2b. Susceptibility of the ground state for $\gamma = 0$ as function of h.

(Th. Niemeijer, Physica 36, 377, 1967)

Appendix

Using BMP Vol. 1, (318), and Vol. 2 (110, form 12) we obtain [$0 < k = 1 - \gamma^2$]

$$\mathcal{E}(k) = \tfrac{1}{2}\pi \; {}_2F_1(-\tfrac{1}{2},\tfrac{1}{2},1,k) =$$

$$= \frac{(1-k)}{2\pi} \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{3}{2})\Gamma(\frac{1}{2}+n)}{n!\;(n+1)!} [h_n - \log(1-k)][1-k]^n$$

$$= \frac{\gamma^2}{2\pi} \sum_{n=0}^{\infty} \frac{\Gamma(n+\frac{3}{2})\Gamma(\frac{1}{2}+n)}{n!\;(n+1)!} [h_n - 2\,\lg\,\gamma]\gamma^{2n}$$

where h_n is given by

$$h_n = \psi(n+1) + \psi(n+2) - \psi(n+\tfrac{1}{2}) - \psi(n+\tfrac{3}{2})$$

and the leading singularity for $\gamma \sim 0$ is proportional to

$\gamma^2 \lg \gamma$.

$$m_z(h,\ \gamma=0) = \frac{1}{2\pi} \int_0^{\pi} \frac{h - \cos\varphi}{|h - \cos\varphi|}\, d\varphi = \tfrac{1}{2} \text{ if } h \geq 1$$

$0 \leq h \leq 1$:

$$m_z(h,\ \gamma=0) = \frac{1}{2\pi} \int_0^{\arccos(h)} d\varphi + \frac{1}{2\pi} \int_{\arccos(h)}^{\pi} d\varphi$$

$$= -\frac{1}{2\pi} \arccos(h) + \frac{1}{2\pi} [\pi - \arccos(h)\ ]$$

$$= \tfrac{1}{2} - \frac{1}{\pi} \arccos(h)$$

Spin Correlation functions

Spin correlation functions are very important in lattice statistics, since they contain information about a possible long range order.

The equilibrium, equal time correlation functions are defined by

$$\rho_{vv}(\ell,m) = \langle S_\ell^v S_m^v \rangle \qquad v = x,y,z \tag{3.1}$$

LSM write these correlation functions in terms of the operators b_j, $b_j^\dagger$ as

$$\begin{aligned}
\rho_{zz}(\ell,m) &= \langle (b_\ell^\dagger b_\ell - \tfrac{1}{2})\ (b_m^\dagger b_m - \tfrac{1}{2}) \rangle \\
\rho_{xx}(\ell,m) &= \tfrac{1}{4} \langle (b_\ell^\dagger + b_\ell)\ (b_m^\dagger + b_m) \rangle \\
\rho_{yy}(\ell,m) &= \tfrac{1}{4} \langle (b_\ell^\dagger - b_\ell)\ (b_m^\dagger - b_m) \rangle
\end{aligned} \tag{3.2}$$

Define new operators

$$A_i = c_i^\dagger + c_i, \qquad B_i = c_i^\dagger - c_i \tag{3.3}$$

and observe the identity

$$e^{i\pi c_i^\dagger c_i} = A_i B_i = -B_i A_i \tag{3.4}$$

We wish to express ρ_{vv} in terms of the $c_j^\dagger$, c_j.

$$\rho_{xx}(\ell,m) = \tfrac{1}{4}\langle (c_\ell^\dagger - c_\ell)\ [\exp(i\pi \sum_{\ell+1}^{m-1} c_j^\dagger c_j)\ ](c_m^\dagger + c_m)\rangle$$

$$= \tfrac{1}{4}\langle B_\ell A_{\ell+1} B_{\ell+1} \cdots A_{m-1} B_{m-1} A_m \rangle \tag{3.5a}$$

$$\rho_{yy}(\ell,m) = \tfrac{1}{4}(-1)^{m-\ell} \langle A_\ell B_{\ell+1} A_{\ell+1} \cdots B_{m-1} A_{m-1} B_m \rangle \tag{3.5b}$$

$$\rho_{zz}(\ell,m) = \tfrac{1}{4}\langle A_\ell B_\ell A_m B_m \rangle \tag{3.5c}$$

Fubini and Caianello (1952) show by the use of Wick's theorem (1950) that expectation values of the type (3.5) are given in terms of Pfaffians. In particular we have

$$\rho_{xx}(m-\ell) =$$

$$\tfrac{1}{4}\mathrm{pf}\begin{vmatrix} S_{\ell,\ell+1} & S_{\ell,\ell+2} \cdots S_{\ell,m-1} & G_{\ell,\ell+1} \cdots & G_{\ell,m} \\ & \vdots & \vdots & \\ & S_{m-2,m-1} & G_{m-2,\ell+1} \cdots & G_{m-2,m} \\ & & G_{m-1,\ell+1} & G_{m-1,\ell+2} \cdots G_{m-1,m} \\ & & & Q_{\ell+1,\ell+2} \cdots Q_{\ell+1,m} \\ & & & \vdots \\ & & & Q_{m-1,m} \end{vmatrix} \tag{3.6}$$

where

$$S_{\ell,m} = \langle B_\ell B_m \rangle = S(m-\ell)$$

$$Q_{\ell,m} = \langle A_\ell A_m \rangle = Q(m-\ell) \qquad (3.7)$$

$$G_{\ell,m} = G(m-\ell) = \langle B_\ell A_m \rangle \qquad (3.8)$$

An important simplification occurs in (3.6) for the equilibrium case, namely $S_{\ell,m} = Q_{\ell,m} = \delta_{\ell,m}$. Then the Pfaffian is equal to the determinant formed from its non-zero entries $G(m-\ell)$.

The three spin-spin correlation functions are given in terms of $G(R)$ as

$$\rho_{xx} = \tfrac{1}{4} \begin{vmatrix} G_{-1} & G_{-2} & \cdots & G_{-R} \\ G_0 & G_{-1} & \cdots & G_{-R+1} \\ \vdots & & & \\ G_{R-2} & G_{R-3} & \cdots & G_{-1} \end{vmatrix} \qquad (3.9a)$$

$$\rho_{yy} = \tfrac{1}{4} \begin{vmatrix} G_1 & G_0 & G_{-1} \cdots & G_{-R+2} \\ G_2 & G_1 & \cdots & G_{-R+3} \\ \vdots & \vdots & & \vdots \\ G_R & G_{R-1} & & G_1 \end{vmatrix} \qquad (3.9b)$$

$$\rho_{zz} = m_z^2 - \tfrac{1}{4} G_R G_{-R} \qquad (3.9c)$$

It is clear that ρ_{zz} is the easiest to deal with, since it involves only simple products of G_R. However,

ρ_{xx} and ρ_{yy} are much harder to evaluate, since one deals with large Toeplitz determinants [Barouch and McCoy, 1971].

To evaluate G_R use (2.13) and the translation invariance to obtain

$$G_R = \frac{1}{N^2} \sum_{p,q,j} \langle \exp(\tfrac{2\pi i}{N}[j(p+q) + Rq]) a_p^\dagger a_q^\dagger + \exp(\tfrac{2\pi i}{N}[j(-p+q) + Rq]) a_p a_q^\dagger - \exp(\tfrac{2\pi i}{N}[j(p-q) - Rq]) a_p^\dagger a_q - \exp(\tfrac{2\pi i}{N}[j(-p-q) - Rq]) a_p a_q \rangle \tag{3.10}$$

Performing the sum over j and taking $N \to \infty$ yield

$$-G_R = \bar{G}_R = = -\frac{1}{\pi}\int_0^\pi d\varphi \frac{\tanh[\frac{1}{2}\beta\Lambda(\varphi)]}{\Lambda(\varphi)}[-\cos\varphi R(\cos\varphi - a) + \sin\varphi R \sin\varphi] = = -\frac{1}{2\pi}\int_{-\pi}^{\pi} (\tfrac{1}{2}\beta) e^{-i\varphi R} T[\tfrac{1}{2}\beta\Lambda(\varphi)](-\cos\varphi + a + i\gamma\sin\varphi)\, d\varphi \tag{3.11}$$

with $T(x) = \frac{\tanh x}{x}$.

Asymptotic results of ρ_{zz} for large R, at the ground state can be now readily obtained as:

(a) $\gamma = 0$

$$\rho_{zz}(R) = \begin{cases} m_z^2 - \left(\frac{\sin[R \operatorname{arccosh}]}{\pi R}\right)^2 & h < 1 \\ \frac{1}{4} & h \geq 1 \end{cases} \tag{3.12a}$$

(b) $\gamma \neq 0, \quad h = 1$

$$\rho_{zz}(R) \sim m_z^2 - \tfrac{1}{4}(\pi R)^{-2}[1 + \tfrac{1}{4}(\gamma R)^{-2} + O(R^{-3})] \quad (3.12b)$$

(c) $h^2 = 1-\gamma^2 \qquad \gamma \neq 0$

$$\rho_{zz}(R) = m_z^2 \quad (3.12c)$$

(d) $\gamma \neq 0 \qquad 0 \leq h^2 < 1-\gamma^2$

$$\rho_{zz}(R) \sim m_z^2 - \alpha^2 R R^{-2} \pi^{-1} \mathrm{Re}\{e^{i\psi(R+1)}[\frac{(1-e^{2i\psi})}{(1-\alpha^2 e^{-2i\psi})}]^{1/2}\}$$

$$\times \mathrm{Re}\{e^{i\psi(R-1)}[\frac{(1-\alpha^2 e^{-2i\psi})}{(1-e^{2i\psi})}]^{1/2}\} \quad (3.12d)$$

where $\cos\psi = h(1-\gamma^2)^{-\frac{1}{2}}$ and $\alpha = \frac{1-\gamma}{1+\gamma} < 1$

(e) $1 - \gamma^2 < h^2 < 1$

$$\rho_{zz}(R) \sim m_z^2 - \tfrac{1}{8}\lambda_2^{-2R} R^{-2} \pi^{-1}\{1 + O(R^{-1})\} \quad (3.12e)$$

(f) $h > 1$

$$\rho_{zz}(R) \sim \tfrac{1}{4} - \tfrac{1}{8}\lambda_2^{2R} R^{-2} \pi^{-1}\{1 + O(R^{-1})\} \quad (3.12f)$$

and λ_2 is given by

$$\lambda_2 = \{h - [h^2 - (1-\gamma^2)]^{1/2}\}/(1-\gamma).$$

This correlation function ρ_{zz} reveals more structure than expected intuitively. We see the boundary $\gamma^2 + h^2 = 1$ in which ρ_{zz} is R independent. We also see that the approach to the Limit $R \to \infty$ is exponential everywhere except

on the boundaries, where it is algebraic. In particular, one would like to interpret the $h > 1$ region as an "easy axis" region. However, in order to be able to say that, we need information about vanishing of possible long range orders at $h = 1$.

It is interesting to note that for $0 \le h^2 < 1 - \gamma^2$ the approach of ρ_{zz} to its limit is oscillatory. This is also the region for which the equivalence to the Ising model does not hold. Suzuki (1971) calls it the "quantum region", and the outside of the unit circle $h^2 + \gamma^2 = 1$ the "classical region", with this circle acting like a natural boundary.

The asymptotic expansions for finite temperature can be obtained in the same fashion, and are given in Barouch and McCoy (1971), eq. (6.1)-(6.11).

We turn our attention to the transverse correlations $\rho_{yy}(R)$, $\rho_{xx}(R)$ at the ground state $\beta = \infty$. We make an extensive use of Szëgo's theorem about the asymptotic properties of Toeplitz determinants, and refer the reader to the paper by Hartwig and Fisher (1969) for a detailed exposition of this topic, motivated by the analysis presented by T. T. Wu (1966).

Szëgo's theorem: Let C_R be the $R \times R$ Toeplitz determinant

$$C_R = \begin{vmatrix} C_0 & C_{-1} & \cdots & C_{-R+1} \\ C_1 & C_0 & & C_{-R+2} \\ \vdots & \vdots & & \vdots \\ C_{R-1} & C_{R-2} & \cdots & C_0 \end{vmatrix} \tag{3.13}$$

where c_n is given by

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-in\varphi} \; c(e^{i\varphi}) \; d\varphi \tag{3.14}$$

If

(i) $\sum_{n=-\infty}^{\infty} |c_n| < \infty$

(ii) $\sum_{n=-\infty}^{\infty} |n||c_n|^2 < \infty$

(iii) $c(e^{i\varphi}) \neq 0$ on the unit circle

(iv) $\ln c(e^{i\varphi})$ is a periodic function of φ with a period 2π (winding index zero). Then the asymptotic value of C_R is given by

$$C_R \doteqdot e^{k_o R} \exp\left(\sum_{n=1}^{\infty} n k_n k_{-n} \right) \tag{3.15}$$

where k_n are given by

$$\ln c(e^{i\varphi}) = \sum_{n=-\infty}^{\infty} k_n e^{in\varphi} \tag{3.16}$$

Conditions (iii) and (iv) are very delicate, and have to be tested rigorously.

These conditions are obeyed for ρ_{xx} when $h < 1$. Condition (iv) is violated for ρ_{yy} and ρ_{xx} for $h > 1$. This is not too serious, since T.T. Wu designed a method that bypasses this difficulty. However, at $h = 1$ condition (iii) is violated, and there we have only partial answers.

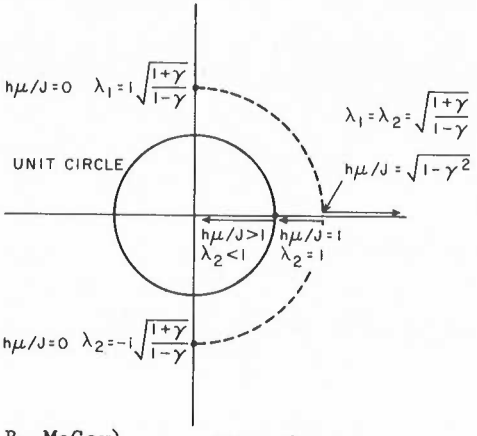
Define $\bar{G}_R = C_{R+1}$. Then ρ_{xx} is given by (3.13) with $T = 0$ as

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i\varphi n} \left[\frac{(1-\lambda_1^{-1}e^{i\varphi})(1-\lambda_2^{-1}e^{i\varphi})}{(1-\lambda_1^{-1}e^{-i\varphi})(1-\lambda_2^{-1}e^{-i\varphi})} \right]^{\frac{1}{2}} d\varphi \tag{3.17}$$

with

$$\lambda_{1,2} = \{h \pm [h^2 - (1-\gamma^2)]^{\frac{1}{2}}\}/(1-\gamma) \qquad (3.18)$$

It is instructive to study the motion of $\lambda_1, \lambda_2, \lambda_1^{-1}$, λ_2^{-1} in the complex $e^{i\varphi}$ plane for fixed γ



(E.Barouch & B. McCoy)
(Phys. Rev. A3, 786, 1971)

Fig. 3

Let us tabulate the values of Ind $c(e^{i\varphi})$:

Case	Function	Index
(a)	ρ_{xx} $h < 1$	0
(b)	ρ_{xx} $h > 1$	+1
(c)	ρ_{yy} $h > 1$	-1
(d)	ρ_{yy} $h < 1$	-2

(a) ρ_{xx} for $h < 1$

We can apply Szëgo's theorem directly

$$\tfrac{1}{2} \ln\left[\frac{(1 - \lambda_1^{-1}e^{i\varphi})(1 - \lambda_2^{-1}e^{i\varphi})}{(1 - \lambda_1^{-1}e^{-i\varphi})(1 - \lambda_2^{-1}e^{-i\varphi})}\right] = \sum_{n=-\infty}^{+\infty} k_n e^{i\varphi n} \quad (3.19)$$

$$k_o = 0$$

$$k_n = (-1)^n \tfrac{1}{2}(\lambda_1^{-n} + \lambda_2^{-n}) \quad (3.20)$$

$$k_{-n} = -(-1)^n \tfrac{1}{2}(\lambda_1^{-n} + \lambda_2^{-n})$$

So

$$\sum_{n=1}^{\infty} n k_n k_{-n} = -\tfrac{1}{4} \sum_{n} \frac{1}{n}\{(\lambda_1^{-2})^n + (\lambda_2^{-2})^n + 2[(\lambda_1\lambda_2)^{-1}]^n\}$$

$$= \tfrac{1}{4} \log[(1-\lambda_1^{-2})(1-\lambda_2^{-2})(1-\lambda_1^{-1}\lambda_2^{-1})^2] \quad (3.21)$$

Substitution of (3.21) in (3.15) yields for $h < 1$

$$\lim_{R\to\infty} C_R = [(1-\lambda_1^{-2})(1-\lambda_2^{-2})(1-\lambda_1^{-1}\lambda_2^{-1})^2]^{\frac{1}{4}} =$$

$$= 2(1+\gamma)^{-1}[\gamma^2(1-h^2)]^{\frac{1}{4}} \quad (3.22)$$

and the first term in the asymptotic evaluation of $\rho_{xx}(R)$ is given by

$$\rho_{xx}(R) \doteq (-1)^R[2(1+\gamma)]^{-1}[\gamma^2(1-h^2)]^{\frac{1}{4}}. \quad (3.23)$$

At the boundary value $h^2 = 1-\gamma^2$ we are able to calculate ρ_{xx} exactly, and find that ρ_{xx} is R independent

$$\rho_{xx} = \tfrac{1}{4}(-1)^R \, 2\gamma/(1+\gamma) .$$

By the same direct method we find that for $\gamma = 0$ $h \geq 1$ $\rho_{xx} = \rho_{yy} = 0$.

The case $h = 1$ $\gamma \neq 0$ is closely related to Wu's $T = T_c$ in the Ising model. The result is

$$\rho_{xx} \doteq \tfrac{1}{4}(-1)^R [2\gamma/(1+\gamma)](\gamma R)^{-1/4} e^{1/4} 2^{1/12} A^{-3} [1+0(R^{-2})] \tag{3.24}$$

where $A = 1.282\ 427\ 130$ is the Glaisher's constant.

The case $\gamma = 0$ $h < 1$ has both λ on the unit circle. We are unable to evaluate ρ_{xx} on this line. The only result available is $\gamma = h = 0$ due to McCoy.

$$\rho_{xx} = \rho_{yy} \doteq \tfrac{1}{4}(-1)^R \, e^{1/2} 2^{1/6} A^{-6} \, R^{-1/2}(1-(-1)^R \, \tfrac{1}{8}R^{-2}+\ldots) \tag{3.25}$$

(b) ρ_{xx} for $h > 1$

Consider the transpose of (3.13), then the index of the resulting generating function is -1. We may write $\rho_{xx} \doteq \tfrac{1}{4}(-1)^R \, B_R$

$$B_R = \begin{vmatrix} b_{-1} & b_{-2} & \cdots & b_{-R} \\ b_0 & b_{-1} & & b_{-R+1} \\ b_{R-2} & b_{R-3} & & b_{-1} \end{vmatrix} \tag{3.26}$$

with

$$b_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i\varphi n}\left[\frac{(1-\lambda_1^{-1}e^{-i\varphi})(1-\lambda_2 e^{-i\varphi})}{(1-\lambda_1^{-1}e^{i\varphi})(1-\lambda_2 e^{i\varphi})}\right]^{\frac{1}{2}} d\varphi$$

In order to evaluate B_R we follow Wu almost word for word. Consider the determinant D_{R+1} defined by

$$D_{R+1} = \begin{vmatrix} b_0 & & \\ \vdots & & B_R \\ b_{R-1} & b_{R-2}\ \cdots & b_0 \end{vmatrix} \tag{3.27}$$

Then

$$B_R = (-1)^R D_{R+1} X_R \tag{3.28}$$

D_{R+1} is a "good" Toeplitz matrix, whose limit as $R \to \infty$ is given by

$$\lim_{R\to\infty} (-1)^R D_{R+1} = [(1-\lambda_2^2)(1-\lambda_1^{-2})(1-\lambda_1^{-1}\lambda_2)^2]^{\frac{1}{4}}$$

and X_R is the corner element of the inverse of D_{R+1}, and is determined by a finite Wienner-Hopf sum equation. Let $d(\xi)$ be given by ($d(\xi)$ has index = 0)

$$d(\xi) = \left[\frac{(1-\lambda_1^{-1}\xi^{-1})(1-\lambda_2\xi^{-1})}{(1-\lambda_1^{-1}\xi)(1-\lambda_2\xi)}\right]^{\frac{1}{2}} \tag{3.29}$$

Then $d(\xi)$ has a unique factorization

$$[d(\xi)]^{-1} = P(\xi)Q(\xi^{-1}) \tag{3.30}$$

for $|\xi| = 1$ such that $P(\xi)$ and $Q(\xi)$ are analytic for $|\xi| < 1$ and continuous and nonzero for $|\xi| \le 1$. Explicitly, $P(\xi)$ and $Q(\xi)$ are given by

$$P(\xi) = \left[\frac{1-\lambda_1^{-1}\xi}{1-\lambda_2\xi}\right] = [Q(\xi)]^{-1} \tag{3.31}$$

Wu shows that X_R is given asymptotically by

$$X_R \doteq \frac{1}{2\pi i} \oint_{|\xi|=1} d\xi\, \xi^{R-1} P(\xi^{-1})\, [Q(\xi)]^{-1}$$

$$= \frac{1}{2\pi i} \oint_{|\xi|=1} d\xi\, \xi^{R-1} \left[\frac{(1-\lambda_1^{-1}\xi^{-1})(1-\lambda_1^{-1}\xi)}{(1-\lambda_2\xi)(1-\lambda_2\xi^{-1})}\right]^{\frac{1}{2}} \tag{3.32}$$

Performing the tedious asymptotic expansion of (3.32) we find for large R

$$\rho_{xx} \doteq (-1)^R \tfrac{1}{4} \pi^{-\frac{1}{2}} \lambda_2^R R^{-\frac{1}{2}} [(1-\lambda_1^{-2})(1-\lambda_2^2)^{-1}(1-\lambda_1^{-1}\lambda_2^{-1})^2]^{\frac{1}{4}}$$

$$\times \{1 + O(R^{-1})\}. \tag{3.33}$$

(c) The study of ρ_{yy} with $h > 1$ is very similar to ρ_{xx} with $h > 1$, and is discussed in II.

(d) In the case of ρ_{yy} with $h < 1$ we have a generating function with index $= -2$. We add two rows and two columns, and proceed in the same fashion as before, where X_R is replaced by

$$\begin{vmatrix} y_R & y_{R+1} \\ y_{R-1} & y_R \end{vmatrix}$$

according to Theorem 4 of Hartwig and Fisher. Details and results of these considerations are available in II. We have also computed the next order terms for $h < 1$, and found a monotonic approach for $h^2 > 1-\gamma^2$ and oscillatory approach for $h^2 < 1-\gamma^2$.

ρ_{yy} for h = 1.

Pfeuty (1970) has shown that for $\gamma = 1$ $\rho_{yy} = -(4R^2-1)\rho_{xx}$. Combining this with (5.31) of Wu we find for $\gamma \neq 0,1$

$$\rho_{yy}(R) \doteq -\tfrac{1}{4}(-1)^R \gamma(1+\gamma) \tfrac{1}{8}(\gamma R)^{-9/4} e^{7A} 2^{1/12} A^{-3}[1+O(R^{-2})]. \tag{3.34}$$

Time delayed correlations.

Define the time delayed correlations

$$\rho_{v,u}(R,t) = \langle S_j^v(0)\ S_{j+R}^u(t)\rangle. \tag{4.1}$$

Dynamical properties of many-body systems very near thermal equilibrium are almost uniquely studied in terms of time delayed correlations of the type (4.1).

The importance of (4.1) and its relation to experiments (like scattering, NMR, etc.) led theoreticians to look for a nontrivial model, for which (4.1) can be computed exactly. Niemeijer observed that $\rho_{zz}(R,t)$ can be computed exactly for the XY-model we are studying. The function $\rho_{zz}(R,t)$ is conceptually simple to obtain since it is the only one which does not contain S_j^x or S_j^y.

Later on, we (McCoy, Barouch and Abraham 1971) studied the ground state properties of $\rho_{xx}(R,t)$ and $\rho_{yy}(R,t)$, and Johnson and McCoy (1971) completed the study for $u \neq v$.

(a) $\rho_{zz}(R,t)$

We wish to compute

$$\rho_{zz}(R,t) = \frac{1}{N}\sum_{\ell} \frac{T_r[e^{-\beta H} S_\ell^z e^{iHt} S_{\ell+R}^z e^{-iHt}]}{T_r[e^{-\beta H}]}. \tag{4.2}$$

In (4.2) we are evolving an even operator, therefore, one can use the c-cyclic condition. Rewriting H and S_j^z in terms of the operators $\eta_j^\dagger$, η_j we have (using (2.13), and (2.17))

$$H = \sum_p \Lambda_p \eta_p^\dagger \eta_p + \text{const.} \tag{4.3}$$

$$S_j^z = \frac{1}{N} \sum_{p,q} \{\exp[ij(\varphi_p - \varphi_q)][\cos\theta_p \eta_p^\dagger + \sin\theta_p \eta_{-p}]$$

$$[\cos\theta_q \eta_q + \sin\theta_q \eta_{-q}^\dagger]\}-\tfrac{1}{2} \tag{4.4}$$

and we use

$$\exp[it\Lambda_q \eta_q^\dagger \eta_q]\eta_q \exp[-it\Lambda_q \eta_q^\dagger \eta_q] = e^{-it\Lambda_q} \eta_q. \tag{4.5}$$

By substitution of (4.5), (4.4), (4.3) in (4.2) one obtains Neimeijer's result for T = 0 as

$$\rho_{zz}(R,t) = m_z^2 + \left[\frac{1}{4\pi}\int_{-\pi}^{\pi} \exp[i(R\varphi + t\Lambda(\varphi))]d\varphi\right]^2$$

$$-\left[\frac{1}{4\pi}\int_{-\pi}^{\pi} \exp[i(R\varphi + t\Lambda(\varphi))]\,\frac{(\cos\varphi - h)}{\Lambda(\varphi)}d\varphi\right]^2$$

$$-\left[\frac{1}{4\pi}\int_{-\pi}^{\pi} \exp[i(R\varphi + t\Lambda(\varphi))]\,\frac{\gamma \sin\varphi}{\Lambda(\varphi)}\right]^2 \tag{4.6}$$

It is interesting to note that as $t \to \infty$ the approach to the limit is $\sim t^{-1}$ and not exponential as several approximation schemes predict. The usefulness of (4.6) as an exact result manifests itself in Mazur's approach to ergodic theory (1969) who proved that m_z is not an ergodic variable.

We now turn our attention to the rest of the correlations $\rho_{uv}(R,t)$, and would like to demonstrate the inapplicability of the above method. Consider $\rho_{vu}(R,t)$ as

$$\rho_{vu}(R,t) = \frac{1}{Z}\,\mathrm{tr}[e^{-\beta H} S_o^v e^{iHt} S_R^u e^{-iHt}] \tag{4.7}$$

where $Z = \mathrm{tr}[e^{-\beta H}]$ is the partition function, and consider an expansion in terms of eigenvectors of H, namely

$$\rho_{vu}(R,t) = Z^{-1} \sum_{m,n} e^{-\beta E_m} \langle E_m | S_o^v | E_n \rangle e^{it(E_m - E_n)} \langle E_n | S_R^u | E_m \rangle \tag{4.8}$$

Let u = x, v = x. Since S_j^x is a product of odd number of Fermi operators, the only non-zero matrix elements are between eigenstates of H^+ and H^- defined by (2.11). One might add that the difficulty in obtaining these matrix elements is similar to Yang's (1952) study of the spontaneous magnetization in the Ising model. We bypass this difficulty by considering 4-spin correlations $C_{xx}(R,t)$ defined by

$$C_{xx}(N,R,t) = \langle S^x_{1+\frac{N}{2}}(t)\, S^x_{1-R+N}(t)\, S^x_{1-R+\frac{N}{2}}(0) \rangle \tag{4.9}$$

where we keep the number of sites large but finite. C_{xx} may be evaluated in terms of matrix elements of even operators only, and by the use of the cluster property we have

$$\lim_{N\to\infty} C_{xx}(N,R,t) = [\rho_{xx}(R,t)]^2 \tag{4.10}$$

Admittedly this method is a poor man's way of obtaining the results, but this is the only one we know.

To evaluate $C_{xx}(N,R,t)$ defined by (4.9) we apply Wick's theorem once more, and obtain a block Toeplitz determinant

$$C^2_{xx}(N,R,t) = \begin{vmatrix} 0 & S_x & T_x & U_x \\ -\tilde{S}_x & 0 & -U_x & V_x \\ -T_x & U_x & 0 & -S_x \\ -U_x & -V_x & \tilde{S}_x & 0 \end{vmatrix} \tag{4.11}$$

$$C^2_{xx}(N,R,t) = \begin{vmatrix} 0 & S_x & T_x & U_x \\ -\tilde{S}_x & 0 & -U_x & V_x \\ -T_x & U_x & 0 & -S_x \\ -U_x & -V_x & \tilde{S}_x & 0 \end{vmatrix} \tag{4.11}$$

where superscript $\sim$ means transpose. Each entry is $(\frac{N}{2} - R) \times (\frac{N}{2} - R)$ matrix whose elements for $T = 0$ are given by

$$(S_x)_{m,n} = \frac{1}{N} \sum_{\varphi} e^{-i(m-n-1)\varphi} G(\varphi) \tag{4.12a}$$

$$(T_x)_{m,n} = \frac{1}{N} \sum_{\varphi} e^{-i(m+n+R)\varphi} e^{-i\Lambda(\varphi)t} G(\varphi) \tag{4.12b}$$

$$(U_x)_{m,n} = -\frac{1}{N} \sum_{\varphi} e^{i(m+n+R+1)\varphi} e^{-i\Lambda(\varphi)t} \tag{4.12c}$$

$$(V_x)_{m,n} = -\frac{1}{N} \sum_{\varphi} e^{i(m+n+R+2)\varphi} e^{-i\Lambda(\varphi)t} \tag{4.12d}$$

and $G(\varphi)$ is given by

$$G(\varphi) = e^{-i\varphi} \left[\frac{(1-\lambda_1^{-1} e^{i\varphi})(1-\lambda_2^{-1} e^{i\varphi})}{(1-\lambda_1^{-1} e^{-i\varphi})(1-\lambda_2^{-1} e^{-i\varphi})} \right]^{\frac{1}{2}} \tag{4.13}$$

To evaluate (4.11) asymptotically we used the scheme developed by Cheng and Wu for $\langle S_{oo} S_{M,N} \rangle$ in the Ising model. This derivation is long, and we do not wish to present it here. Details are available in paper IV (McCoy Barouch and Abraham).

The results are

(a) $h < 1$

where superscript $\sim$ means transpose. Each entry is $(\frac{N}{2} - R) \times (\frac{N}{2} - R)$ matrix whose elements for T = 0 are given by

$$(S_x)_{m,n} = \frac{1}{N} \sum_{\varphi} e^{-i(m-n-1)\varphi} G(\varphi) \tag{4.12a}$$

$$(T_x)_{m,n} = \frac{1}{N} \sum_{\varphi} e^{-i(m+n+R)\varphi} e^{-i\Lambda(\varphi)t} G(\varphi) \tag{4.12b}$$

$$(U_x)_{m,n} = -\frac{1}{N} \sum_{\varphi} e^{i(m+n+R+1)\varphi} e^{-i\Lambda(\varphi)t} \tag{4.12c}$$

$$(V_x)_{m,n} = -\frac{1}{N} \sum_{\varphi} e^{i(m+n+R+2)\varphi} e^{-i\Lambda(\varphi)t} \tag{4.12d}$$

and $G(\varphi)$ is given by

$$G(\varphi) = e^{-i\varphi} \left[\frac{(1-\lambda_1^{-1}e^{i\varphi})(1-\lambda_2^{-1}e^{i\varphi})}{(1-\lambda_1^{-1}e^{-i\varphi})(1-\lambda_2^{-1}e^{-i\varphi})} \right]^{\frac{1}{2}} \tag{4.13}$$

To evaluate (4.11) asymptotically we used the scheme developed by Cheng and Wu for $\langle S_{oo} S_{M,N} \rangle$ in the Ising model. This derivation is long, and we do not wish to present it here. Details are available in paper IV (McCoy Barouch and Abraham).

The results are

(a) $h < 1$

$$\rho_{xx}(R,t) \doteq \rho_{xx}(\infty) \{1 + (2\pi)^{-1} \oint d\xi \oint d\eta\, \xi^R \eta^{-R} (\xi-\eta)^{-2} e^{-it[\Lambda(\xi)+\Lambda(\eta)]} \tfrac{1}{2}[M(\xi,\eta) + M(\eta,\xi) - 1]\} \tag{4.14}$$

where the contours are the unit circles and the η contour is indented outward at $\xi = \eta$, and

$$M(\xi,\eta) = \left[\frac{(1-\lambda_1^{-1}\xi^{-1})(1-\lambda_2^{-1}\xi^{-1})(1-\lambda_1^{-1}\xi)(1-\lambda_2^{-1}\xi)}{(1-\lambda_1^{-1}\eta^{-1})(1-\lambda_2^{-1}\eta^{-1})(1-\lambda_1^{-1}\eta)(1-\lambda_2^{-1}\eta)} \right]^{\frac{1}{2}} \tag{4.15}$$

(b) $h > 1$

$$\rho_{xx}(R,t) \doteq (-1)^R \tfrac{1}{4}[(1-\lambda_2^2)(1-\lambda_1^{-2})(1-\lambda_1^{-1}\lambda_2)^2]^{\frac{1}{4}}$$

$$\frac{1}{2\pi i}\oint d\xi\ \xi^{R-1}\ e^{-it\Lambda(\xi)}\left[\frac{(1-\lambda_1^{-1}\xi)(1-\lambda_1^{-1}\xi^{-1})}{(1-\lambda_2\xi)(1-\lambda_2\xi^{-1})}\right]^{\frac{1}{2}} \qquad (4.16)$$

and the square root is defined postive at $\xi = -1$.

(c) $h = 1$

We simply don't know what to do.

These results should be compared with $\rho_{zz}(R,t)$ where all the contributions are from two particle excitations. Here, for $h > 1$ we get contributions from 1,2,3,... excitations and for $h < 1$ from 2,4,6,... excitations, since (4.15) and (4.16) are the first terms in the expansions of $\rho_{xx}(R,t)$.

The only other exact result is $\rho_{xz}(R,t)$ for $h < 1$, derived by Johnson and McCoy (1971) and is given by

$$\rho_{xz}(R,t) = \tfrac{1}{4}(-1)^R\left(\frac{2}{1+\gamma^2}\right)^{1/2}[\gamma^2(1-h^2)]^{1/8}\ (S\text{-}A) \qquad (4.17)$$

where

$$S = \frac{1}{2\pi i}\oint_{|z|=1} dz\ z^{-2}\left[\frac{(1-\lambda_1^{-1}z)(1-\lambda_2^{-1}z)}{(1-\lambda_1^{-1}z^{-1})(1-\lambda_2^{-1}z^{-1})}\right]^{\frac{1}{2}} \qquad (4.18)$$

$$A = W_1 - W_2 \qquad (4.19)$$

$$W_1 = \frac{1}{2\pi i}\oint dz\ e^{i\Lambda t}\ z^{-R}[(1-\lambda_1^{-1}z)(1-\lambda_2^{-1}z)]^{-\frac{1}{2}}$$

$$\times\ \frac{1}{2\pi i}\oint dz'(z'z-1)^{-1}\ e^{i\Lambda t}\ (z')^{-(R+1)}[(1-\lambda_1^{-1}z')(1-\lambda_2^{-1}z')]^{\frac{1}{2}} \qquad (4.20)$$

$$W_2 = \frac{1}{2\pi i} \oint dz\, e^{i\Lambda t}\, z^R [(1-\lambda_1^{-1}z)(1-\lambda_2^{-1}z)]^{-\frac{1}{2}}$$

$$\times \frac{1}{2\pi i} \oint dz'(1-zz')\, e^{i\Lambda t}\, (z')^{R-1} [1-\lambda_1^{-1}z')(1-\lambda_2^{-1}z')]^{\frac{1}{2}}$$

(4.21)

Random Impurities

It is well known that no physical material is 100% pure. There are several kinds of impurities in solids like foreign ions, some rare or common isotopes and so forth. We will address ourselves to the problem of "frozen in" impurities. The impurities, randomly distributed with some normalized distribution P(z(m)) have the interesting effect that singularities associated with phase transitions tend to round off, infinitely differentiable, but non-analytic. This statement is primarily based on the work of McCoy and Wu (1968) for the Ising model, who found rounding of the specific heat near T_c instead of the famous $\log|T-T_c|$ divergence derived by Onsager (1944). McCoy and Wu's paper partially motivated the work of Smith which we are about to discuss.

In this lecture we wish to discuss the thermodynamics of an isotropic XY chain with $\gamma = 0$, but with random coupling constants, and study their effect on the singularities of the thermodynamic functions discussed earlier. Smith introduces the Hamiltonian H

$$H = \sum_{m=1}^{N-1} \{J(m)(S_m^x S_{m+1}^x + S_m^y S_{m+1}^y) - h \sum_{m=1}^{N} S_m^z\} \qquad (5.1)$$

where J(m) are independent random variables.

Step (i) and (ii) can be carried over, namely

$$H = \tfrac{1}{2}Nh + \sum_{m,n} c_m^\dagger A_{mn} c_n \qquad (5.2)$$

with

$$A_{mm} = -h \qquad A_{m,m+1} = A_{m+1,m} = J(m)$$
$$A_{mn} = 0 \quad \text{otherwise.} \tag{5.3}$$

Since (5.2) is quadradic, the transformation

$$c_m = \sum_\alpha f_\alpha(m)\mu_\alpha, \qquad c_m^\dagger = \sum_\alpha f_\alpha^*(m)\mu_\alpha^\dagger \tag{5.4}$$

yields

$$H = \text{Const.} + \sum_\alpha \lambda_\alpha \mu_\alpha^\dagger \mu_\alpha \tag{5.5}$$

with free energy

$$\beta f = 2h\beta - \frac{1}{N} \sum_\alpha \ln[1 + e^{\beta\lambda_\alpha}] \tag{5.6}$$

Define $A = B - hI$ where I is the unit matrix, and an eigenvector of B with eingenvalue δ is clearly an eigenvector of A with eingenvalue $\lambda = \delta - h$. Our task is to find the distribution of the eigenvalues of B for $N \to \infty$, and this was done by Dyson (1953) in his brilliant analysis of the random chain of harmonic oscillators.

Smith, following Dyson, defines

$$\Omega(x) = \lim_{N\to\infty} \frac{1}{N} \sum_{\alpha=1}^{N} \ln\,(1+x\delta_\alpha) = \int_{-\infty}^{\infty} \ln(1+x\delta)\, dM\,(\delta) \tag{5.7}$$

(the branch of the log is taken in $(-\pi,\pi)$), where $M(\delta)$ is the limiting distribution function of the eigenvalues of B, obtained by the relation

$$\lim_{\epsilon\to 0} \mathrm{Re}\left[\frac{1}{i\pi}\, \Omega(-x+i\epsilon)\right] = \int_{1/x}^{\infty} dM(\delta) = 1 - M\left(\frac{1}{x}\right) \tag{5.8}$$

Let $D(x) = M'(x)$ be the density of states, then (5.6) becomes

$$\beta f = 2h\beta - \int_{-\infty}^{\infty} \ln[1 + e^{-\beta(x-h)}]D(x)dx \qquad (5.9)$$

So to determine the thermodynamic behavior of the random system one needs to compute $\Omega(x)$.

Expanding the log in (5.7) before taking the limit, together with $\mathrm{tr}\, B^{2n+1} = 0$ yield

$$\sum_{\alpha=1}^{N} \ln(1+x\delta_\alpha) = -\tfrac{1}{2}\sum_{n=1}^{\infty} \frac{1}{n} x^{2n}\mathrm{Tr}B^{2n} = \sum_{m=1}^{N} \ln\{1-\sigma(m,x)\} \qquad (5.10)$$

and $\sigma(m,x)$ is a continued fraction with recurrence relation

$$\sigma(m,x) = \frac{x^2J^2(m)}{[1-\sigma(m+1,x)]} \qquad (5.11)$$

Setting $\rho(m,y) = -\sigma(m,iy)$ we have

$$\rho(m,y) = \frac{y^2J^2(m)}{[1+\rho(m+1,y)]} \qquad (5.12)$$

and $\Omega(iy)$ is given by

$$\Omega(iy) = \lim_{N\to\infty} \frac{1}{N} \sum_{m=1}^{N} \ln[1+\rho(m,y)] \qquad (5.13)$$

In the limit of large N, $\rho(m,y)$ tends to the limiting distribution $f(\rho)$ obtained by

$$f(\rho) = \int d\rho' \int f(\rho')\delta\left(\rho - \frac{y^2z}{1+\rho'}\right)P(z)\, dz \qquad (5.14)$$

where $P(z)$ is the distribution of the random variables $\frac{1}{4}J^2(m)$. Carrying out the z integration we obtain an integral equation for $f(\rho)$

$$f(\rho) = \int_0^\infty f(\rho') \frac{(1+\rho')}{y^2} P\left[\frac{(1+\rho')}{y^2}\right] d\rho' \qquad (5.15)$$

and the explicit P(z) we choose is given by

$$P(z) = \frac{n^n}{(n-1)!} z^{n-1} e^{-nz} \qquad (5.16)$$

Dyson solved (5.15) exactly!!

$$f(\rho) = [K_n(y^2)]^{-1} \rho^{n-1}(1+\rho)^{-n} \exp(-\frac{n\rho}{y^2}) \qquad (5.17)$$

with the normalization K_n given by

$$K_n(y^2) = \int_0^\infty \rho^{n-1}(1+\rho)^{-n} \exp(-\frac{n\rho}{y^2})\, d\rho \qquad (5.18)$$

and Ω is found to be

$$\Omega_n(iy) = [K_n(y^2)]^{-1} \int_0^\infty \rho^{n-1}(1+\rho)^{-n} \ln(1+\rho) \exp(-\frac{n\rho}{y^2})\, d\rho \qquad (5.19)$$

In order to compute M_n and hence D_n for large n one needs to analytically continue $\Omega_n(iy)$, compute the discontinuity from the negative real axis, and study the result asymptotically. Smith has done that and his asymptotic results are (for large finite n)

$$D_n(y) \sim \frac{1}{2\pi}\{(1-y^2)^{-1/2} + \frac{1}{4n}(1-y^2)^{-3/2} + O(n^{-2})\} \qquad y < 1$$

$$D_n(y) \sim \frac{y}{\pi \sinh\theta}\{2n\theta(\cosh\theta - 1) + \theta - 1\}\exp[-\theta - 2n(\sinh\theta - \theta)] \qquad y > 1$$

with $\theta = \operatorname{arc\,cosh}(y^2-1)$

$$D_n(y) \sim n^{1/3}\{a_1 - a_2[n^{2/3}(y^2-1)]^2 + O([n^{2/3}(y^2-1)]^3\}$$

$$|n^{2/3}(y^2-1)| < 1 \qquad y \sim 1$$

and

$$a_1 = \frac{2^{1/3}3^{7/6}\Gamma^3(\frac{2}{3})}{\pi\Gamma^2(\frac{1}{3})} \sim .18 \qquad a_2 \sim .53 \tag{5.20}$$

In the limit $n \to \infty$, the Poisson distribution becomes a δ function, and the density of states $D_\infty(y)$ becomes

$$D_\infty(y) = \begin{cases} \frac{1}{2\pi}(1-y^2)^{-1/2} & y < 1 \\ 0 & y > 1 \end{cases} \tag{5.21}$$

and is shown in Fig. 4

Once $D_n(y)$ is obtained, we can study the thermodynamic functions. We find that m_z tends to smooth up at the ground state

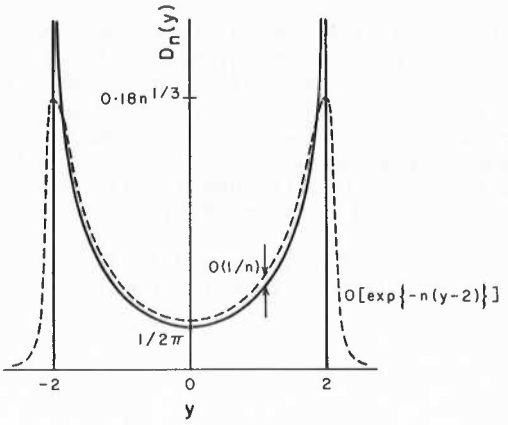


Fig. 4. Sketch of density of states $D_n(y)$ for $n\to\infty$ and for n very large. Full curve, $n\to\infty$; Broken curve, n very large. (E.R. Smith, J.Phys. C.: S.S. Phys., Vol. 3, 1419, 1970.)

To compute the susceptibility for T = 0 we have

$$\chi_n(h) = \lim_{\beta\to\infty} \beta \int_{-\infty}^{\infty} d\mu(\mu-h)^2 \{1+e^{-\beta(\mu-h)}\}^{-2} e^{-\beta(\mu-h)} D_n(\mu)$$

$$= D_n(h) \qquad (5.21)$$

and $D_n(h)$ is infinitely differentiable.

A recent study is in the process of being completed. We study the Hamiltonians with random magnetic moments where the resulting integral equation is too hard to solve exactly, and we study the smoothing of the transition near h = 1.

Non-equilibrium.

Most studies of nonequilibrium phenomenae start from the Liouville equation for the density matrix $\rho(t)$

$$i \frac{d}{dt} \rho(t) = [H^S(t), \rho(t)] \qquad (6.1)$$

It is not at all clear which approximations, if any, are appropriate in given circumstances. Existence of non-trivial examples for which (6.1) is exactly solved, enables us to evaluate the effectiveness and legitimacy of such approximations. In the following lecture we solve (6.1) exactly for the XY-model, where we allow the magnetic field to depend on time explicitly. Below we follow I (Barouch, McCoy and Dresden 1970).

Since stages (i), (ii), and (iii) are independent of the field, we may start from the Hamiltonian

$$H = \sum_p H_p$$

$$= \tfrac{1}{2} \sum_{p>0}^{N/2} \{\alpha_p(t)[a_p^\dagger a_p + a_{-p}^\dagger a_{-p}] + \tfrac{1}{2} i\delta_p [a_p^\dagger a_p^\dagger + a_p a_p] + 2h(t)\} \qquad (6.2)$$

with

$$\alpha_p(t) = 2[\cos \varphi_p - h(t)], \qquad \delta_p = -2\gamma \sin \varphi_p \tag{6.3}$$

As we saw before, each H_p in (6.2) commutes with each other, which means that the space upon which H acts, decomposes into non-interacting subspaces.

Let $(|0\rangle;\ a_p^\dagger a_{-p}^\dagger |0\rangle;\ a_{-p}^\dagger |0\rangle)$ be the basis for the pth subspace in the Heisenberg picture. The Hamiltonian then becomes a matrix $\bar{H}(t)$

$$\bar{H}(t) = \sum_p [I \otimes I \otimes \ldots \otimes \bar{H}_p(t) \otimes I \ldots \otimes I] \tag{6.4}$$

where $\otimes$ is the direct product, I the 4×4 unit matrix and

$$\bar{H}_p(t) = \begin{vmatrix} h(t) & \frac{1}{2}i\delta_p & 0 & 0 \\ -\frac{1}{2}i\delta_p & 2\cos\varphi_p - h(t) & 0 & 0 \\ 0 & 0 & \cos\varphi_p & 0 \\ 0 & 0 & 0 & \cos\varphi_p \end{vmatrix} \tag{6.5}$$

Let $U_p(t)$ be the time evolution matrix in the pth subspace given by

$$i\frac{d}{dt} U_p(t) = U_p(t)\bar{H}_p(t) \qquad U_p(0) = I \tag{6.6}$$

The Hamiltonian $H^S(t)$ is then given by

$$H^S(t) = \sum_p [I \otimes I \otimes \ldots \otimes H_p^S(t) \otimes \ldots \otimes I] \tag{6.7}$$

where

$$H_p^S(t) = U_p(t)\bar{H}_p(t)U_p^{\dagger}(t) \tag{6.8}$$

Because of (6.7) or (6.4), the density matrix at $t = 0$ is given by

$$\rho(0) = e^{-\beta\bar{H}_1(0)} \otimes e^{-\beta\bar{H}_2(0)} \otimes \ldots e^{-\beta\bar{H}_{N/2}(0)} \tag{6.9}$$

This particular algebraic form, together with (6.7) suggests that a solution of (6.1) with boundary condition (6.9) would have a similar form

$$\rho(t) = \rho_1(t) \otimes \rho_2(t) \otimes \ldots \otimes \rho_{N/2}(t) \tag{6.10}$$

This is indeed true, if

$$\begin{aligned} i\frac{d}{dt}\rho_p(t) &= [H_p^S(t), \rho_p(t)], \\ \rho_p(0) &= e^{-\beta\bar{H}_p(0)} = e^{-\beta H_p^S(0)} \end{aligned} \tag{6.11}$$

In other words, all we have to do in order to obtain $\rho(t)$ is to solve (6.6). The only nontrivial part of (6.6) is the upper left block, and U_{11}, U_{12}, U_{21}, U_{22} can be easily determined if one of them is known.

Let

$$h = b + h_1(t) \text{ with } \lim_{t\to\infty} h_1(t) = 0 \tag{6.12}$$

and

$$U_{11}(t) = V(t)e^{-it\cos} \tag{6.13}$$

The same algebra yields

$$V''(t) + [\Lambda^2(b) + \psi(t)]V(t) = 0 \tag{6.14}$$

with b.c. $V(0) = 1$, $V'(0) = i[\cos\varphi - h(0)]$, with two independent solutions $W_1(t)$, $W_2(t)$, and $\psi(t)$ is given by

$$\psi(t) = h_1^2(t) - 2(\cos\varphi - b)h_1(t) + ih_1'(t) \tag{6.15}$$

Example 1.

$$h(t) = \begin{cases} a & t \le 0 \\ b & t > 0 \end{cases} \tag{6.16}$$

$$V(t) = i\,\frac{\cos\varphi - b}{\Lambda(b)}\,\sin[t\Lambda(b)] + \cos[t\Lambda(b)] \tag{6.17}$$

Example 2.

$$h(t) = \begin{cases} a & t \le 0 \\ b + (a-b)e^{-Kt} \end{cases} \tag{6.18}$$

$$V(t) = A_1W_1(t) + A_2W_2(t) \tag{6.19a}$$

where

$$W_1(t) = \exp\left[i\Lambda(b)t + i\frac{(a-b)}{K}\,e^{-Kt}\right]$$

$${}_1F_1\left[\frac{i}{K}[\Lambda(b)+b-\cos\phi];\,1+\frac{2i\Lambda(b)}{K};\,-\frac{2i(a-b)}{K}\,e^{-Kt}\right] \tag{6.19b}$$

$$W_2(t) = \exp\left[-i\Lambda(b)t + i\frac{(a-b)}{K}\,e^{-Kt}\right]$$

$${}_1F_1\left[\frac{i}{K}[\Lambda(b)+b-\cos\varphi];\,1+\frac{2i\Lambda(b)}{K};\,-\frac{2i(a-b)}{K}\,e^{-Kt}\right] \tag{6.19c}$$

and the constants A_1, A_2 are given by

$$A_1 = \frac{W_2'(0) - i(\cos\varphi - a)W_2(0)}{W_1(0)W_2'(0) - W_2(0)W_1'(0)} \tag{6.19d}$$

$$A_2 = \frac{iW_1(0)(\cos\varphi - a) - W_1'(0)}{W_1(0)W_2'(0) - W_2(0)W_1'(0)} \tag{6.19e}$$

The first question we want to study is the approach of $m_z(t)$ to its infinite time limit.

To compute $m_z(t)$, we observe that $\frac{1}{N}\sum_j S_j^z$ can be written as

$$N^{-1}\sum_p M_p = N^{-1}\sum_p [a_p^\dagger a_p + a_{-p}^\dagger a_{-p} - 1]$$

and

$$m_z(t) = \frac{1}{N}\sum \frac{\mathrm{Tr}[M_p U_p(t)\rho_p(0)U_p^\dagger(t)}{\mathrm{Tr}\ \rho_p(0)} \tag{6.20}$$

Using (6.19a), $m_z(t)$ can be expressed in terms of $W_1(t)$ and $W_2(t)$ as

$$m_z(t) = N^{-1}\sum_p \frac{\tanh[\frac{1}{2}\beta\Lambda(h(0))]}{\Lambda(h(0))} F[W_1(t), W_2(t)] \tag{6.21}$$

and $F[W_1(t), W_2(t)]$ is given explicitly in (4.7) of I. Example 1 yields for the step function (6.16)

$$m_z(t) = \frac{1}{N}\sum_p \frac{\tanh[\frac{1}{2}\beta\Lambda(a)]}{\Lambda(a)\Lambda^2(b)} \{\cos[2\Lambda(b)t]\gamma^2(a-b)\sin^2\varphi_p$$

$$- (\cos\varphi_p - b)[(\cos\varphi_p - a)(\cos\varphi_p - b) + \gamma^2\sin^2\varphi]\} \tag{6.22}$$

It is clear from (6.22) that if N is finite and large and $t \to \infty$, the limit does not exist, and one may

wish to compute the Poincare cycle as an explicit function of N.

In the thermodynamic limit $N \to \infty$, the sum becomes an integral

$$m_z(t) = \frac{1}{2\pi}\int_0^{\pi} d\varphi \frac{\tanh[\frac{1}{2}\beta\Lambda(a)]}{\Lambda(a)\Lambda^2(b)} \{\cos[2\Lambda(b)t]\gamma^2(a-b)\sin^2\varphi -$$

$$- (\cos\varphi - b)[(\cos\varphi - a)(\cos\varphi - b) + \gamma^2\sin^2\varphi]\} \qquad (6.23)$$

This result was derived earlier by Niemeijer. There are several interesting limits to check.

(i) $t = 0$: $m_z(0)$ becomes the equilibrium result given by (2.)

(ii) $\gamma = 0$: Since $[\sum_j S_j^z, \sum_j S_j^x S_{j+1}^x + S_j^y S_{j+1}^y] = 0$ one expects no time dependence of $\langle \sum_j S_j^z \rangle$, and in (6.23) the time dependent term is proportional to γ^2.

(iii) $a = b$: No jump, and again $m_z(t) = m_z(0)$.

(iv) $t \to \infty$:

$$m_z(\infty) = \frac{1}{2\pi}\int_0^{\pi} d\varphi \frac{\tanh[\frac{1}{2}\beta\Lambda(a)](a - \cos\varphi)}{\Lambda(a)}$$

$$\times \left\{ 1 + \frac{(b-a)\gamma^2\sin^2\varphi}{(a-\cos\varphi)[(b-\cos\varphi)^2 + \gamma^2\sin^2\varphi]} \right\} \qquad (6.24)$$

This is not the equilibrium magnetization, since $b = 0$ does not yield $m_z(\infty) = 0$. The system, even after infinite time, remembers that it was subjected to an external field a, through the nonzero function in the curly brackets of (6.24). This is an explicit expression of the non-ergodicity of the system.

One might raise the suspicion that the nonergodic behavior of $m_z(t)$ is due to the special discontinuous case (6.16). In other words, one might hope that a continuous very slow change of $h(t)$ would yield an equilibrium limit. To demonstrate this hope to be false we went through the pain of playing the same game with example (6.18), and computed $m_z(K,t)$. We found that taking the limit $\lim_{K\to\infty}\lim_{t\to\infty} m_z(K,t)$ gives back (6.24) and $\lim_{K\to\infty}\lim_{t\to\infty} m_z(K,t)$ gives a complicated expression [I (6.14)], that shares with (6.24) the unpleasant feature of failing to vanish at $b = 0$. We can safely conclude that a global change of the magnetic field results in a nonergodic magnetization $m_z(t)$.

It is interesting to study the long time behavior of $m_z(t)$ (6.23). We find (I) one more division into regions.

(i) $h > 1-\gamma^2$: $m_z(t)$ approaches its limit like $t^{-3/2}$, and oscillates with two interfering frequencies, exchange type and larmor type.

(ii) $h < 1-\gamma^2$: The leading term of $m_z(t)$ decays like $t^{-\frac{1}{2}}$ with a single frequency $2\gamma[1-b^2(1-\gamma^2)^{-1}]^{1/2}$. In the next term all three frequencies are present.

(iii) $h = 1-\gamma^2$: Boundary case for which $m_z(t)\sim t^{-3/4}$.

This division to regions rises from the number of extremal points of the one particle spectrum Λ. If we consider $\cos\varphi = y$, Λ is given by

$$\Lambda(y,b) = [\gamma^2(1-y^2) + (b-y)^2]^{1/2} \qquad -1 \le y \le 1$$

In the case (i) Λ is monotonic and has two extremal points at the boundaries $y = \pm 1$, in case (ii) Λ has 3 extremal points at the boundaries and at $y_0 = b/(1-\gamma^2)$, and in the boundary case (iii), one of its endpoints coincides with y_0.

In figure (5) we show a numerical and asymptotic study of $m_z(t)$ for case (i), where the interference of the two frequencies is quite clear.

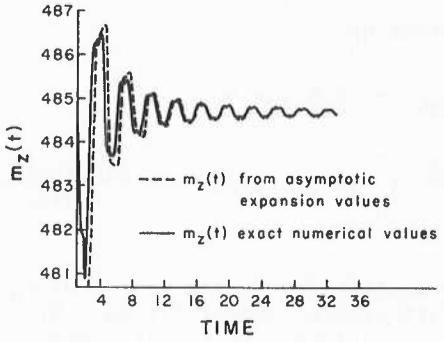


Fig. 5. $m_z(t)$ exact (numerical) and asymptotic for large t.
$a = 10$, $b = 2$
$\gamma = \frac{1}{2}$, $\beta = 1$

(E. Barouch, B. McCoy, M. Dresden)
(Phys. Rev. A2, 1075, 1970)

We wish to investigate more of the nonergodic features of the system by examining the spin correlation functions.

By a similar method of evaluating $m_z(t)$ we find for the step function case (6.16)

$$G_R = -\frac{1}{\pi}\int_0^{\pi} d\varphi \cos \varphi R \left(\frac{\tanh[\frac{1}{2}\beta\Lambda(a)]}{\Lambda(a)\Lambda^2(b)}\right)\{[\gamma^2\sin^2\varphi + $$

$$+(\cos\varphi - a)(\cos\varphi - b)](\cos\varphi - b) - (a-b)\gamma^2\sin^2\varphi$$

$$\cos[2\Lambda(b)t]\} + \frac{\gamma}{\pi}\int_0^{\pi} d\varphi \sin\varphi R \sin\varphi \left(\frac{\tanh[\frac{1}{2}\beta\Lambda(a)]}{\Lambda(a)\Lambda^2(b)}\right)$$

$$\times \{[\gamma^2 \sin 2\varphi + (\cos\varphi - a)(\cos\varphi - b)] +$$

$$+ (a-b)(\cos\varphi - b)\cos[2\Lambda(b)t]\} \qquad (6.25)$$

and S_R is given by

$$S_R = \langle A_i A_{i+R}\rangle = \langle B_i B_{i+R}\rangle =$$

$$= \frac{\gamma(a-b)}{\pi}\int_0^{\pi} d\varphi \sin\varphi \sin\varphi R \frac{\sin[2t\Lambda(b)]}{\Lambda(a)\Lambda(b)} \qquad (6.26)$$

In the equilibrium ($a = b$ or $t = 0$) S_R vanishes identically. Furthermore, when $t \to \infty$ $S_R \to 0$. In these two cases the correlations are Toeplitz determinants. However, for finite t, we have a full Pfaffian, which forces us to try to evaluate a block Toeplitz determinant. This we are unable to do (III), and we can just estimate the most dominant term, up to an unknown multiplicative constant.

Since the analysis is conceptually simple and quite tedious, let me summarize our conclusions, and refer you for details to III.

(1) $\lim_{t\to\infty} \rho_{zz}(R,t) \neq$ Equilibrium $\rho_{zz}(R)$

(2) At the ground state $\lim_{R\to\infty}\lim_{t\to\infty} \rho_{xx}(R,t) = 0$ namely there is a destruction of the long range order.

(3) For finite long time, the correlation functions approach their nonergodic limits with the same power laws and same frequencies as $m_z(t)$.

After the conclusions that fundamental thermodynamic functions like magnetization and correlations are nonergodic, one might believe that all thermodynamic averages are nonergodic, and do not tend to their equilibrium values. This question was studied by Girardeau (1969). He considered the Fourier component $M_q = \Sigma_j S_j^z \cos q_i$,

and proved that $\langle M_q(t)\rangle \to 0$ as $t \to \infty$ when $h = o$ which is the equilibrium values if $0 < q < \pi$, where $\langle M_q(t)\rangle$ is given by

$$\langle M_q(t)\rangle = \mathrm{Tr}[\rho(0)e^{itH}M_q e^{-itH}] \tag{6.27}$$

$$\rho(0) = \exp\{-[\beta H + \lambda M_q]\}/\mathrm{Tr}\{-[\beta H + \lambda M_q]\} \tag{6.28}$$

and the parameter λ measures the prescribed initial values, and its deviation from an equilibrium state. The cases $q = 0, \pi$ are different, since their limit $\neq 0$. Note that $q = 0$ corresponds to our example (6.16) with $b = 0$, $a = \lambda$, and $q = \pi$ corresponds to the staggered case.

When one convinces himself that the system is nonergodic, the natural question to ask is why. Is it because of the decomposition of the system into noninteracting subspaces? Is it because of the low dimensions of the system? Is it because the system is isolated and is not coupled to a heat bath?

At this stage it is not too wise to point at a specific "reason" and claim its responsibility for the nonergodic behavior of the system. So to gain some insight into the meaning of these questions we studied the time behavior of a single magnetic impurity at the boundary (Tjon 1970) using the weak coupling approximation, and exactly inside the chain (Abraham, Barouch, Gallavotti and Martin-Löf 1970). In both cases thermalization was obtained, namely $m_z(t)$ approached its correct limit, but as a power law t^{-1} (for internal spins) or t^{-3} (for a boundary spin).

We present now the analysis of ABGM for the isotropic case $\gamma = 0$ and $h = 0$. The Hamiltonian is given by

$$H = \tfrac{1}{4}\sum_{j=1}^{N}(\sigma_j^x\sigma_{j+1}^x + \sigma_j^y\sigma_{j+1}^y) + h(t)\sigma_m^z \equiv H_o + h(t)\sigma_m^z \tag{6.29}$$

with

$$h(t) = \begin{cases} h & t \leq 0 \\ 0 & t > 0 \end{cases} \tag{6.30}$$

Thermalization occurs if

$$\lim_{t\to\infty} \lim_{N\to\infty} \langle \sigma_n^z(t) \rangle = 0 \tag{6.31}$$

The difficulty involved in this problem is breaking of the translation invariance, thus Fourier decomposition does not yield decomposition to noninteracting subspaces. Furthermore, we can look at the rest of the chain acting on the single spin like a heat bath.

We proceed with the standard stages (i), (ii), (iii) and obtain

$$H_o = E_o + \sum_q \cos q \; a_q^\dagger a_q \tag{6.32a}$$

$$1 + \sigma_n^z = \frac{2}{N} \sum_{qq'} e^{in(q'-q)} \; a_q^\dagger a_{q'} \tag{6.32b}$$

and q are the solutions of

$$e^{iqN} = 1 \tag{6.32c}$$

We study the expectation value

$$\langle 1 + \sigma_n^z(t) \rangle = [\mathrm{Tr}\, e^{-\beta H}]^{-1} \, \mathrm{Tr}[e^{-\beta H} e^{iH_o t} (1+\sigma_n^z) e^{-iH_o t}] \tag{6.33}$$

Using (6.32) one obtains

$$e^{iH_o t}(1+\sigma_n^z)e^{-iH_o t} = \frac{2}{N}\sum_{qq'} \exp\{i[n(q'-q) + t(\cos q - \cos q')]\}a_q^\dagger a_{q'}, \tag{6.34}$$

Since H is quadratic in the Fermi operators $a_q, a_q^\dagger$, it may be written as

$$H = E_1 + \sum_j \lambda_j \alpha_j^\dagger \alpha_j \tag{6.35}$$

where α_j are Fermi operators related to a_q by the unitary transformation

$$\alpha_j = \sum_q U_{jq} a_q \tag{6.36}$$

Combining (6.36), (6.34) and (6.33) we obtain

$$\langle 1 + \sigma_n^z(t)\rangle = \\ = \frac{2}{N}\sum_{qq'} \exp\{[n(q'-q)+t(\cos q - \cos q')]\} \sum_j U_{jq}^* U_{jq'} \langle a_j^\dagger a_j\rangle \tag{6.37}$$

where $\langle a_j^\dagger a_j\rangle$ is the Fermi occupation number given by $[1 + \exp(\beta\lambda_j)]^{-1}$.

The coefficients U_{jq} are determined from the eigenvalue problem

$$(\lambda_j - \cos q)U_{jq} = \frac{2h}{M}\sum_{q'} u_{jq'} e^{i(q-q')m} \tag{6.38}$$

and there are two possibilities

(i) $\lambda_j = \cos q_o$ for some q_o that solves (6.32c)

$$U_{jq} = 2^{-\frac{1}{2}}(\delta_{q',q_o} e^{iqm} - \delta_{q',-q_o} e^{-iqm}) \tag{6.39}$$

(ii) $\lambda_j \neq \cos q$

$$U_{jq} = e^{iqm}/N(\lambda_j)(\lambda_j - \cos q) \tag{6.40}$$

where λ_j are the zeros of

$$F_N(\lambda) = 1 - \frac{2h}{N} \sum_q (\lambda - \cos q)^{-1} \tag{6.41}$$

and the normalization is given by

$$|N(\lambda_j)|^2 = \sum_q (\lambda_j - \cos q)^{-2} = \frac{N}{2h} \frac{\partial F}{\partial \lambda}\Big|_{\lambda=\lambda_j} \tag{6.42}$$

Combining (6.39) - (6.42) and taking the thermodynamic limit we finally obtain

$$\langle \sigma_n^z(t) \rangle = \frac{2h}{\pi i} \oint_c \frac{G(\lambda, n-m, t)\, G(\lambda, m-n, -t)}{(1 + e^{\beta\lambda})F(\lambda)}\, d\lambda \tag{6.43}$$

where the contour c in the complex λ plane avoids the zeros of $1 + e^{\beta\lambda}$, enclosing the zeros of $F_N(\lambda)$, and the functions G, F are given by

$$G(\lambda,1,t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\exp\{i[-1q + t \cos q]\}}{\lambda - \cos q}\, dq \tag{6.44a}$$

$$F(\lambda) = 1 - 2h(\lambda^2-1)^{-\frac{1}{2}} \tag{6.44b}$$

and asymptotic study of (6.43) for large t shows approach to 0 like t^{-1}. This power law also governs the approach of the correlation functions to their equilibrium nonzero limit. We have also studied $\gamma \neq 0$, $h \neq 0$, and obtained similar results.

The last question we mention is that if we turn on a field, the magnetization does not approach equilibrium (ABGM) no matter how slowly the field rises with time from its initial zero value. The result is given by

$$\langle\sigma_n^z(t) = \mathrm{Re}\Big[\int_{-\pi}^{\pi} dk(1 + e^{\beta \cos k})^{-1} \times \int_{-\pi}^{\pi} dp \exp\{i[t(\cos k - \cos p)]\Big] + (k-p)(n-m)\mu(k,p,t) + 2\int_{-\pi}^{\pi} dk(1+e^{\beta \cos k})^{-1} \times \Big|\int_{-\pi}^{\pi} dp \exp\{i[t \cos p - p(n-m)\mu(k,p,t)\Big|^2 \tag{6.45}$$

with

$$\mu(k,p,t) = 2 \int_0^t e^{it' \cos p} h(t') X_{p'}(t') dt' \tag{6.46}$$

and $x_p(t)$ is to be determined from the Volterra equation

$$x_p(t) = e^{-it \cos p} - i \int_0^t dt' J_0(t-t') h(t') X_p(t') \tag{6.47}$$

which can be solved exactly for $h(t') = h \neq 0$.

ACKNOWLEDGMENT

The author wishes to thank his colleague, Prof. S. Orszag, for his interest in the present work, and for partial financial support through his A.P. Sloan Research Fellowship.

REFERENCES

D. Abraham, E. Barouch, G. Gallavotti, A. Martin-Löf, Phys. Rev. Lett. 25, 1449, (1970).

Stut. of App. Math. L, 121 (1971), ibid (to be published).

E. Barouch and M. Dresden, Phys. Rev. Lett. 23, 114, (1969).

E. Barouch, B. McCoy and M. Dresden, Phys. Rev. A2, 1075, (1970). This paper is referred to as I.

E. Barouch, B. McCoy, Phys. Rev. A3, 786, (1971) (referred to as II), ibid A3, 2137, (1971) (referred to as III).

R. Baxter, Phys. Rev. Lett., 26, 832, (1971), ibid 26, 834, (1971). Preprint (2) 1971.

H. Bethe Zeit Physik, 71, 205, (1931).

E. R. Caianello and S. Fubini, Nuovo Cimento, 9, 1218, (1952).

H. Cheng and T. T. Wu, Phys. Rev. 164, 7/9, (1967).

F. J. Dyson, Phys. Rev. 92, 1331, (1953).

M. D. Girardeau, Phys. Lett. 30A, 442 (1969).

G. Gallavotti, Stud. of Appl. Math, L, 89, (1971).

R. Hartwig and M. Fisher, Arch. Rat. Mech. 32, 190, (1969).

E. Ising, Z. Physik 31, 253, (1925).

L. Hulthen, Arkiv. Mat. Astron. Fysik 26A, Na11, (1938).

J. D. Johnson and B. McCoy, Phys. Rev. 4A, 2314, (1971).

H. Kramers and G. Wannier, Phys. Rev. 60, 252, (1941).

S. Katsura, Phys. Rev. 127, 1508, (1962).

S. Katsura, T. Horiguchi and M. Suzuki, Physica 45, 67 (1970).

E. Lieb, T. Schultz and D. Mattis, Ann. of Phys. 16, 941, (1961). (referred to as LSM)

P. Mazur, Physica 43, 533, (1969).

B. McCoy, Phys. Rev. 173, 531, (1968).

B. McCoy, E. Barouch and D. Abraham, Phys. Rev. 4A 2331, (1971).

B. McCoy and T. T. Wu, Nuovo Cummento 56, 311, (1960). Phys. Rev. 176, 631 (1968); 188, 982 (1969). The two dimensional Ising Model, Harvard University Press, (1971).

Th. Niemeijer, Physica, 36, 377, (1967), 39, 313, (1968).

Y. Nambu, Prog. Theor. Phys. 5, 1 (1950).

L. Onsager, Phys. Rev. 65, 117, (1944).

P. Pfeuty, Ann. of Phys. 57, 79, (1970).

S. Pikin and V. Tsukenrik, Soviet Phys. JETP 23, 914, (1966).

M. Suzuki, J. Phys. Soc. Japan 21, 2140, (1966).
Prog. Theor. Phys. 43, 882, (1970).
Phys. Lett. 34A, 94, (1971).

E. R. Smith, J. Phys. C, Solid St. Phys. 3, 1419, (1970). Preprint (1971).

B. Sutherland, J.M.P. 11, 3183, (1970).

J. Tjon, Phys. Rev. 2B, 2711, (1970).

H. vanVianen and J. Tjon, Physica 48, 497 (1970).

T. T. Wu, Phys. Rev. 149, 380 (1966).

C. N. Yang, Phys. Rev. 85, 809, (1952).

C. N. Yang and C. P. Yang, Phys. Rev. 147, 303 (1966); 150, 321 (1966); 150, 327 (1966); 151, 258 (1966).

COMPOSITE PARTICLES IN MANY-BODY SYSTEMS
METHOD OF STOLT AND BRITTIN

Wesley E. Brittin
University of Colorado
Boulder, Colorado

I. Introduction

Many situations exist in astrophysics, plasma physics, chemical physics, etc. where it would be meaningful and useful to have theories dealing with equilibrium and transport properties relating to various bound or composite particles (atoms, molecules, ions) as well as to "free" or "unbound" particles (nuclei, electrons) and to electromagnetic radiation (photons). A variety of special techniques has been invented to treat some specific problems in the above categories. These treatments range from very crude empirical and theoretical guesses to very sophisticated field-theoretic procedures. From the standpoint of basic physics these problems pose certain difficulties and there exists at present no unified approach to them. It would be very desirable to have a theory that would be as complete for these problems as is the kinetic theory of gases for a tenuous system composed of classical stable molecules.

An important first step in the desired direction was taken in 1963 by M. Girardeau[1], who showed that

> "... a second-quantization representation for many-atom systems can be developed in which the atomic annihilation and creation operators satisfy elementary boson or fermion commutation relations, i.e., the atoms behave like point particles. In this representation the Hamiltonian, expressed as a function of the local atomic field operators, takes the familiar form of a sum of a quadratic part representing independent-particle (here independent atom) energies and a quartic part representing two-body interactions."

In order to take account of the fact that composites cannot be actual bosons or fermions, a subsidiary condition imposed on the overall state vector enforced correct symmetry for the elementary particles which built up the composites. In the expression for Girardeau's elementary destruction and creations operators A_α and A_α^* the index α labels the atomic states and includes the center of mass motion as well as the internal motion of the particles making up the atom. In addition, the atomic states include all continuum (ionized) states of the atom as well as bound states. The question of how to introduce only bound state composites was not treated at that time.

In the spring of 1971 R. H. Stolt and W. E. Brittin found a way of introducing bound composites for relatively simple systems [2]. Subsequently A. Y. Sakakura[3] and M. Girardeau[4] found other methods. Although there are many unresolved (even in principle) problems, the subject has reached a certain maturity. My lectures have only to do with the methods discovered by Stolt and Brittin, since Sakakura[5] and Girardeau[6] give accounts of their important work elsewhere in these lecture proceedings.

II. Preliminary

We consider a system containing N protons and N electrons* which interact through Coulomb forces. Let $\psi = \psi(x_1 \ldots x_N; y_1 \ldots y_N)$ be an arbitrary square integrable function (wave function) of the proton and electron coordinates x_k, y_l (x,y includes positions and spins). The set of all such functions forms a Hilbert space which we label $\mathcal{H}$. The space $\mathcal{H}$ contains functions belonging to all symmetry classes and not only those which represent physical states. The physical states are represented by functions ψ_A which are completely antisymmetric with respect to permutations of proton (electron) coordinates. The subspace $A \subset \mathcal{H}$ of physical states is obtained from $\mathcal{H}$ by projection with $\hat{A}$, the total antisymmetrizer given by

$$\hat{A} = \frac{1}{N!^2} \sum \epsilon_P \hat{U}_P \sum \epsilon_{P'} \hat{V}_{P'} \qquad (2.1)$$

* It is straightforward to generalize these considerations to situations where there are different numbers N_p, N_e of protons and electrons.

where P (P′) goes over all proton (electron) coordinate permutations and

$$(\hat{U}_P\psi)(x_1\ldots x_N;y_1\ldots y_N) \equiv \psi(x_{\alpha_1}\ldots x_{\alpha_N};y_1\ldots y_N) \qquad (2.2)$$

where $P(1\ldots N) = (\alpha_1\ldots\alpha_N)$, and

$$(\hat{V}_{P'}\psi)(x_1\ldots x_N;y_1\ldots y_N) \equiv \psi(x_1\ldots x_N;y_{\beta_1}\ldots y_{\beta_N}) \qquad (2.3)$$

where $P'(1\ldots N) = (\beta_1\ldots\beta_N)$

Thus

$$A = \hat{A}\,\mathcal{K} \qquad (2.4)$$

where $\hat{A}$ is the projector* defined by expression 2.1. The condition that ψ_A represent a physical state is that

$$\hat{A}\psi_A = \psi_A \qquad (2.5)$$

We wish to describe situations in which some of the electrons and protons have combined to form bound hydrogen atoms. Naturally in the general situation, where the system might be in a highly condensed phase, for example, it may not be useful to ask for a description of the system in terms of atoms. Perhaps other "clusters" or "composites" may be more useful in that case. However as an example of our method, let us think of physical situations where it is meaningful to speak of the system as "having" "bound" atoms and "free" electrons and protons. I wish to emphasize at this point that our treatment has no approximations in it in so far as the description of physical states is concerned, although its utility may depend on whether or not the physical situation approximates the description we choose to use.

Let the system be placed in a box of volume V where V is large but finite. Then we may introduce a complete orthonormal set of one-proton states $\varphi_i(x)$, $i = 1,2,3,\ldots$ and a complete orthonormal set of one-electron states $\psi_j(y)$, $j = 1,2,3,\ldots$ so that any $\psi(x_1\ldots x_N;y_1\ldots y_N) \in \mathcal{K}$

* The conditions that an operator $\hat{A}$ be a projector are that $\hat{A}^2 = \hat{A} = \hat{A}^*$. These can be verified for $\hat{A}$ defined above.

may be expanded as

$$\psi(x_1 \ldots x_N; y_1 \ldots y_N) \equiv \psi(x_N y_N) =$$

$$= \sum_{\substack{i_1 \ldots i_N \\ j_1 \ldots j_N}} C_{i_1 \ldots i_N j_1 \ldots j_N} \varphi_{i_1}(x_1) \ldots \varphi_{i_N}(x_N) \psi_{j_1}(y_1) \ldots \psi_{j_N}(y_N) \tag{2.6}$$

In fact, because of the orthogonality of the φs and ψs,

$$C_{i_1 \ldots i_N j_1 \ldots j_N} \equiv C_{ij} =$$

$$= \int dx_1 \ldots dx_N dy_1 \ldots dy_N \overline{\varphi}_{i_1}(x_1) \ldots \overline{\varphi}_{i_N}(x_N) \overline{\psi}_{j_1}(y_1) \ldots \overline{\psi}_{j_N}(y_N) \psi(x_N y_N). \tag{2.7}$$

If $\psi(x_N y_N)$ is a physical state $C_{i_1 \ldots i_N j_1 \ldots j_N}$ is completely antisymmetric in $i_1 \ldots i_N$ rsp. $j_1 \ldots j_N$, and conversely. If ψ, ψ' are any two functions in $\mathcal{H}$ $(\psi, \psi') = \int dx(N) dy(N) \overline{\psi}(x(N), y(N)) \psi'(x(N), y(N)) = \sum \overline{C}_{ij} C'_{ij}$.

Let us now introduce two-particle bound states $\varphi_\alpha(x,y)$. These states are to represent isolated <u>bound</u> electron-proton states (including the center of mass motion). They are taken to be orthonormal

$$\int \overline{\varphi}_{\alpha_1}(xy) \varphi_{\alpha_2}(xy) dxdy = \delta_{\alpha_1 \alpha_2}, \tag{2.8}$$

but not complete. We have

$$\sum_\alpha \varphi_\alpha(xy) \overline{\varphi}_\alpha(x'y') = P_B(xy;x'y') \tag{2.9}$$

where $P_B(xy;x'y')$ is the coordinate space representation of $\hat{P}_B$ the 2-particle bound state projector. That is

$$(\hat{P}_B \psi)(xy) = \int P_B(xy;x'y') \psi(x'y') dx'dy' \tag{2.10}$$

Further

$$\| \hat{P}_B \psi \| < \| \psi \| \tag{2.11}$$

where $\| \psi \|^2 = \int \overline{\psi}(xy) \psi(xy) dxdy$, which expresses the fact

that the projector onto bound states is less than the unit operator.

Let us introduce "bound" many particle functions through the 2-particle states φ_α by defining the functions

$$\psi^M_{\alpha_1..\alpha_M;i_{M+1}..i_N;j_{M+1}..j_N}(x_1..x_N;y_1..y_N) \equiv \psi^M_{\alpha ij}(x(N)y(N))$$

$$= \varphi_{\alpha_1}(x_1y_1)\varphi_{\alpha_2}(x_2y_2)..\varphi_{\alpha_M}(x_My_M)\varphi_{i_{M+1}}(x_{M+1})..\varphi_{i_N}(x_N)$$

$$\times \psi_{j_{M+1}}(y_{M+1})..\psi_{j_N}(y_N). \qquad (2.12)$$

These functions are orthonormal for a given M

$$(\psi^M_{\alpha ij},\psi^M_{\alpha' i' j'}) = \int dx(N)dy(N)\bar{\psi}^M_{\alpha ij}(x(N)y(N))\psi^M_{\alpha' i' j'}(x(N)y(N))$$

$$= \delta_{\alpha\alpha'}\delta_{ii'}\delta_{jj'} \equiv \delta_{\alpha_1\alpha_1'}\delta_{\alpha_2\alpha_2'}\ldots\delta_{\alpha_M\alpha_M'}\delta_{i_{M+1}i_{M+1}'}$$

$$\ldots\delta_{i_Ni_N'}\delta_{j_{M+1}j_{M+1}'}\ldots\delta_{j_Nj_N'}, \qquad (2.13)$$

where the integral $\int dx(N)dy(N)\ldots$ is extended over all $x_1\ldots x_N;y_1\ldots y_N$, including summation over spin variables. These functions span a subspace P_M of $\mathcal{H}$ defined by

$$P_M = \{\psi;\psi = \sum_{\alpha ij} C^M_{\alpha ij}\ \psi^M_{\alpha ij};\ \sum_{\alpha ij}|C^M_{\alpha ij}|^2 \text{ finite}\}. \qquad (2.14)$$

P_M may be considered to correspond to those states having M (or more) bound atoms, although P_M contains functions that are not completely antisymmetric in electron and proton coordinates. The subspace P_M contains P_{M+1} since $\varphi_{\alpha_{M+1}}(x_{M+1}y_{M+1})$ can be expanded in terms of $\varphi_{i_{M+1}}(x_{M+1})$ and $\psi_{j_{M+1}}(y_{M+1})$. We express this in the customary manner, $P_M \supset P_{M+1}$. In particular P_o is the entire space $\mathcal{H}$ since $\varphi_i\ \psi_i$; are complete one-particle states, thus

$$\mathcal{H} = P_o \supset P_1 \supset \ldots \supset P_M \supset \ldots \supset P_N \qquad (2.15)$$

The projector for $\hat{P}_M$ is expressed as

$$\hat{P}_M = \sum_{\alpha ij} |M\alpha ij\rangle \langle M\alpha ij| \tag{2.16}$$

which means that

$$(\hat{P}_M\psi)(x_1y_1\ldots x_Ny_N) = \sum_{\alpha ij} \psi^M_{\alpha ij}(x_1\ldots y_N)\langle M\alpha ij|\psi\rangle \tag{2.17}$$

with

$$\langle M\alpha ij|\psi\rangle \equiv \int dx_N dy_N \overline{\psi}^M_{\alpha ij}(x(N)y(N))\psi(x(N)y(N)). \tag{2.18}$$

The relationship (2.15) is reexpressed in terms of the projectors $\hat{P}_M$ as

$$\hat{1} = \hat{P}_0 > \hat{P}_1 > \hat{P}_2 \ldots > \hat{P}_M > \ldots > \hat{P}_N \tag{2.19}$$

If $P(1) \equiv \Sigma|\alpha\rangle\langle\alpha|$, is the single atom bound state projector, i.e. cf. (2.9) $(\hat{P}(1)\psi)(x_1y_1) = \sum_\alpha \varphi_\alpha(x_1y_1)\int\overline{\varphi}_\alpha(x_1'y_1')$ $\times\ \psi(x_1'y_1')dx_1'dy_1'$, we may express $\hat{P}_M$ as

$$\hat{P}_M = \hat{P}(1) \otimes \hat{P}(2) \otimes \hat{P}(3) \otimes \ldots \otimes \hat{P}(M) \otimes \hat{1}_{N-M} \tag{2.20}$$

where $\hat{1}_{N-M}$ is the unit operator for functions of the variables $x_{M+1}, y_{M+1} \ldots x_N, y_N$. Hence P_M maybe written,

$$\hat{P}_M = \hat{P}(1)\otimes\hat{P}(2)\otimes..\otimes\hat{P}(M)\otimes\hat{1}(M+1)\otimes\hat{1}(M+2)\otimes..\otimes\hat{1}(N) \tag{2.21}$$

where $\hat{1}(R)$ refers to the unit operator for functions of the variables x_R, y_R. The subspaces P_M although not physical, or at least not entirely so, do somehow correspond to states having M or more bound atoms. Indeed if the "real" atoms in the system are far enough apart the functions in P_M may represent physical states very closely. However, this is not what we seek. We would like to find functions corresponding to precisely M bound atoms. This is done rather simply. Since $\hat{P}_M > \hat{P}_{M+1}$

$$\hat{P}_{M+1}\hat{P}_M = \hat{P}_M\hat{P}_{M+1} = \hat{P}_{M+1}, \tag{2.22}$$

so

$$\hat{R}_M \equiv \hat{P}_M - \hat{P}_{M+1} = \hat{P}_M - \hat{P}_M\hat{P}_{M+1} = \hat{P}_M(\hat{1}-\hat{P}_{M+1}) = \hat{P}_M\hat{P}^{\perp}_{M+1} = \hat{P}^{\perp}_{M+1}\hat{P}_M \tag{2.23}$$

has the property that functions lying in $R_M \equiv \hat{R}_M \mathcal{H}$ are in P_M but orthogonal to P_{M+1}. Hence R_M "has" M or more bound atoms but not M+1 or more bound atoms — i.e. R_M "has" precisely M bound atom states or more precisely is the subspace for functions corresponding to M "bound" atoms, N-M "free" electrons, and N-M "free" protons. The entire function space $\mathcal{H}$ may be decomposed directly into orthogonal subspaces according to the scheme

$$\mathcal{H} = R_0 \oplus R_1 \oplus \ldots \oplus R_M \oplus \ldots \oplus R_N, \tag{2.24}$$

which corresponds to the identity

$$\begin{aligned}\hat{1} = \hat{P}_0 &= (\hat{P}_0-\hat{P}_1)+(\hat{P}_1-\hat{P}_2)+\ldots+(\hat{P}_M-\hat{P}_{M+1})+\ldots+\hat{P}_N \\ &= \hat{P}_0\hat{P}^{\perp}_1+\hat{P}_1\hat{P}^{\perp}_2+\ldots+\hat{P}_M\hat{P}^{\perp}_{M+1}+\ldots+\hat{P}_N \\ &= \hat{R}_0+\hat{R}_1+\ldots\hat{R}_M+\ldots+\hat{R}_N. \end{aligned} \tag{2.25}$$

It is easy to see that $\hat{R}_M\,\hat{R}_{M'} = \delta_{MN'}\,\hat{R}_M$, e.g. if $\psi_0 \in \hat{R}_0 = \hat{P}^{\perp}_1$ then ψ_0 is orthogonal to P_1 and hence to $P_2, P_3, \ldots$ which are contained in P_1, etc. Therefore we may decompose any $\psi \in \mathcal{H}$ into orthogonal components $\psi_M \equiv \hat{R}_M\psi$. If

$$\psi = \sum_{M=0}^{N} \psi_M$$

and

$$\psi' = \sum_{M=0}^{N} \psi'_M ,$$

then the orthogonality of R_M implies

$$(\psi,\psi') = \sum_{M=0}^{N} (\psi_M,\psi'_M) \tag{2.26}$$

Explicit expressions for $\hat{R}_M$ are obtained directly if we use the decomposition of $\hat{P}_M$ given in equation 21:

$$\hat{R}_M = \hat{P}_M - \hat{P}_{M+1} = \hat{P}(1)\otimes\hat{P}(2)\otimes\ldots\otimes\hat{P}(M)\otimes\hat{P}(M+1)^{\perp}\otimes\hat{1}(M+2)\otimes\ldots\otimes\hat{1}(N) \tag{2.27}$$

where $\hat{P}(M+1)^{\perp} \equiv \hat{1}(M+1)-\hat{P}(M+1)$ is the projector onto the

single atom unbound states;

$$(P(M+1)^{\perp}\psi)(M+1) = \psi(M+1) - \sum_{\alpha}\varphi_{\alpha}(M+1)\langle\alpha|\psi\rangle \equiv \psi^{F}(M+1)$$

(recall (M+1) here refers to the variables x_{M+1}, y_{M+1}). Since $R_M \subset P_M$ we may write

$$\psi_M = \sum_{\alpha ij} C^M_{\alpha ij}\, \psi^M_{\alpha ij}. \tag{2.28}$$

Then

$$\begin{aligned} C^M_{\alpha ij} &= (\psi^M_{\alpha ij}, \psi_M) = (\psi^M_{\alpha ij}\hat{R}_M\psi) = (\hat{R}_M\psi^M_{\alpha ij}\psi) \\ &= (\chi^M_{\alpha ij}, \psi) \end{aligned} \tag{2.29}$$

with

$$\chi^M_{\alpha ij} \equiv \hat{R}_M\psi^M_{\alpha ij} = (\hat{P}_M - \hat{P}_{M+1})\psi^M_{\alpha ij} = \psi^M_{\alpha ij} - \hat{P}_{M+1}\psi^M_{\alpha ij}\,, \tag{2.30}$$

which can also be written as

$$\chi^M_{\alpha ij} = \varphi_{\alpha_1}(1)\varphi_{\alpha_2}(2)\ldots\varphi_{\alpha_M}(M)\varphi^F_{i_{M+1}j_{M+1}}(M+1)\varphi_{M+2}(x_{M+2})$$

$$\ldots\varphi_{i_N}(x_N)\psi_{j_{M+2}}(y_{M+2})\ldots\psi_{j_N}(y_N) \tag{2.31}$$

where, as above,

$$\varphi^F_{ij}(k) \equiv \varphi_i(x_k)\psi_j(y_k) - [\hat{P}_B(k)(\varphi_i\psi_j)](k). \tag{2.32}$$

Unfortunately the coefficients $C^M_{\alpha ij}$ in the expansion of ψ_M do not have all the properties required for a description of the system in terms of "bound" atoms and "free" particles. Such coefficients for physical states ψ_A should be completely symmetric in $\alpha_1\ldots\alpha_M$ and completely antisymmetric in $i_{M+1}\ldots i_N$ and in $j_{M+1}\ldots j_N$. For physical states, $C^M_{\alpha ij}$ are <u>symmetric</u> in $\alpha_1\ldots\alpha_N$ but the presence of $\varphi^F_{i_M j_M}$ in the M+1 th entry spoils complete antisymmetry for $C^M_{\alpha ij}$ in the indices $i_{M+1}\ldots i_N$ and $j_{M+1}\ldots j_N$. (The $C^M_{\alpha ij}$s are completely antisymmetric in i_{M+2}, $i_{M+3}\ldots$, i_N and in j_{M+2}, $j_{M+3},\ldots,j_N$ however).

The decomposition 2.28 for physical states may be useful in chemical problems because of its simplicity.

However, see appendix B which gives additional reasons for not using this representation. We may at this point mention a similar decomposition first used by A.Y. Sakakura[3]. From the expansion for $\hat{R}_M$ (Eq. 2.27) we note that the M+1 st entry is $\hat{P}(M+1)^{\perp}$, so if $\varphi_\beta(x,y)$ form a complete set of states for the "unbound" proton-electron system, we may expand any ψ using the orthogonal basis

$$\Phi^M_{\alpha_1\cdots\alpha_M\beta i_{M+2}\cdots i_N j_{M+2}\cdots j_N} \equiv \Phi^M_{\alpha\beta ij} \equiv \varphi_{\alpha_1}(1)\varphi_{\alpha_2}(2)\cdots\varphi_{\alpha_M}(M)$$

$$\varphi_\beta(M+1)\varphi_{i_{M+2}}(x_{M+2})\cdots\varphi_{i_N}(x_N)\psi_{j_{M+2}}(y_{M+2})\cdots\psi_{j_N}(y_N) \quad (2.33)$$

which for fixed M spans the subspace R_M. In fact

$$\hat{R}_M = \sum_{\substack{\alpha_1\cdots\alpha_M,\beta \\ i_{M+2}\cdots i_N \\ j_{M+2}\cdots j_N}} |\Phi^M_{\alpha\beta ij}\rangle\langle\Phi^M_{\alpha\beta ij}| \quad (2.34)$$

For physical states the coefficients $C^M_{\alpha_1\cdots\alpha_M\beta i_{M+2},\cdots j_N}$ are completely symmetric with respect to interchanges of $\alpha_1\cdots\alpha_M$ and β and completely antisymmetric with respect to interchanges of $i_{M+2}\cdots j_N$. In physical terms the system may be thought of in terms of a given number M of "bound" atoms, one "unbound" atom, and N-(M+1) "free" electrons and N-(M+1) "free" protons. In later work Sakakura has managed to eliminate this "crazy" β-boson.

In conclusion to this section it is to be noted that we have achieved an orthogonal expansion 2.28 similar to but distinct from that used by Girardeau[6], i.e.*

$$\psi = \Sigma\psi_M = \sum_{\alpha_1\cdots\alpha_M} f^M_{\alpha_1\cdots\alpha_M}(x_{M+1}\cdots x_N; y_{M+1}\cdots y_N)\varphi_{\alpha_1}(1)\varphi_{\alpha_2}(2)\cdots\varphi_{\alpha_M}(M) \quad (2.35)$$

---

* $$f^M_{\alpha_1\cdots\alpha_M}(x_{M+2},\ldots,y_N) \equiv \sum_{\substack{i_{M+1}\cdots i_N \\ j_{M+1}\cdots j_N}} C^M_{\alpha_1\cdots\alpha_M i_{M+1}\cdots i_N j_{M+1}\cdots j_N}\varphi_{i_{M+1}}(x_{M+1})\cdots\psi_{j_N}(y_N)$$

which for a physical state $\psi = \hat{A}\psi$ has coefficient functions $f^M_{\alpha_1\ldots\alpha_M}$ which are completely symmetric in $\alpha_1\ldots\alpha_M$, but which are not completely antisymmetric in $x_{i_{M+1}}\ldots x_{i_N}$ and in $y_{j_{M+1}}\ldots y_{j_N}$. Of course, since for physical states, $\psi = \hat{A}\psi$,

$$\hat{A}\psi = \psi = \Sigma\, \hat{A}\psi_M = \Sigma\, \psi_M, \tag{2.36}$$

and in $\hat{A}\psi_M$ the coefficient functions f^M_α are antisymmetrized. However projection with $\hat{A}$ spoils the orthogonality: $\hat{A}\psi_M \neq \psi_M$ and $(\hat{A}\psi_M, \hat{A}\psi_{M'}) \neq 0$ for $M \neq M'$.

III. The Stolt-Brittin Method

We have mentioned that projection in general spoils orthogonality, that is if $(\psi_1,\psi_2) = 0$ and $\hat{P}$ is some projector, then $(\hat{P}\psi_1,\hat{P}\psi_2) = (\hat{P}\psi_1,\psi_2) \neq 0$, in general. On the other hand if P_2 is a closed subspace of $\mathcal{H}$ which is contained in the closed subspace P_1 of $\mathcal{H}$, and if $\hat{A}$ is a projector, then $\hat{A}P_2$ is contained in $\hat{A}P_1$, e.g. $P_1 \supset P_2 \Rightarrow \hat{A}P_1 \supset \hat{A}P_2$.* The subspaces P_M introduced in II have the property $\mathcal{H} = P_0 \supset P_1 \supset\ldots\supset P_M \supset\ldots\supset P_N$, therefore

$$A = \hat{A}\mathcal{H} = \hat{A}P_0 \supset \hat{A}P_1 \supset\ldots\hat{A}P_M \supset\ldots\supset \hat{A}P_N \tag{3.1}$$

which means that we may decompose A, the subspace of physical wave functions into an orthogonal set of physical subspaces A_M,

$$A = A_0 \oplus A_1 \oplus\ldots\oplus A_M \oplus\ldots\oplus A_N \tag{3.2}$$

where

$$A_M = \hat{A}P_M \quad \hat{A}P_{M+1} \tag{3.3}$$

is the direct difference of the subspaces $\hat{A}P_M$ and $\hat{A}P_{M+1}$. The subspace A_M consists of those states ψ^A_M which are in $\hat{A}P_M$ but which are orthogonal to all elements of $\hat{A}P_{M+1}$.

* We really are talking about the closed subspaces $\overline{\hat{A}P_1}, \overline{\hat{A}P_2}$ obtained by forming the closures of $\hat{A}P_1$ and $\hat{A}P_2$, but to keep the notation simple we simply write $\hat{A}P$ for $\overline{\hat{A}P}$.

We define A_M to be the physical subspace corresponding precisely to M "bound" atoms and N-M "free" protons and N-M "free" electrons. Thus any $\psi = \hat{A}\psi$ may be expanded into orthogonal components ψ_M

$$\psi = \hat{A}\psi = \sum_{M=0}^{N} \psi_M, \tag{3.4}$$

and $(\psi_M, \psi_{M'}) = p_M \delta_{MM'}$ where $p_M = (\psi_M, \psi_M)$ may be regarded as the probability that an observation of the system will result in finding precisely M "bound" atoms present. The average number $\langle M \rangle$ of bound atoms present is just

$$\langle M \rangle = \sum_{M'} M' p_{M'} \tag{3.5}$$

which, of course, in general changes with time. The projectors $\hat{A}_M$ for the subspaces A_M are not nearly so easy to compute as, for example, $\hat{R}_M$. In fact we know of no ways, except those requiring infinite processes, of actually computing ψ_M, given ψ. Let us, however, proceed with the problem, since in practice we will use approximation procedures in any case.

The result of projecting the subspace P_M with $\hat{A}$ results in a subspace $\hat{A}P_M$ for which we would like to find the projector (which we denote by $\hat{A} \circ \hat{P}_M$). It is so designated because we anticipate that it may be compounded in some fashion from the projectors $\hat{A}$ and $\hat{P}_M$. We can form $\hat{A}P_M$ by taking the closure of $\hat{A}\hat{P}_M^{\infty}$. We now observe that

$$(\hat{A} \circ \hat{P}_M)\hat{A}\hat{P}_M = \hat{A}P_M \tag{3.6}$$

i.e. $\hat{A} \circ \hat{P}_M$ is a left projector of $\hat{A}\hat{P}_M$. In fact it is the left projector of $\hat{A}\hat{P}_M$, which means that

$$\hat{A} \circ \hat{P}_M = \inf \{\hat{Q};\ \hat{Q} \text{ a projector};\ \hat{Q}\hat{A}\hat{P}_M = \hat{A}P_M\}. \tag{3.7}$$

Knowing what $\hat{A} \circ \hat{P}_M$ is, however, does not offer much guidance for computing it. In order to find an expression for $\hat{A} \circ \hat{P}_M$ let us split A into two orthogonal subspaces

$$A = (\hat{A}P_M) \oplus (\hat{A}P_M)_A^{\perp} \tag{3.8}$$

where $(\hat{A}P_M)_A^{\perp}$ is the orthogonal compliment of $\hat{A}P_M$ in A

which means $(\hat{A}\hat{P}_M)^{\perp}_{\hat{A}} = A - (\hat{A}\hat{P}_M)$. If ψ is contained in $(AP_M)^{\perp}_{\hat{A}}$ it is orthogonal to all vectors of the form $\hat{A}\hat{P}_M x$, $x \in \mathcal{H}$, in addition to being in A. Therefore

$$(\psi, \hat{A}\hat{P}_M x) = (\hat{A}\psi, \hat{P}_M x) = (\psi, \hat{P}_M x) = (\hat{P}_M \psi, x) = 0 \qquad (3.9)$$

which shows that $\psi \in P_M{}^{\perp}$ and $\psi \in A$, hence $\psi \in A \wedge P_M{}^{\perp}$, the subspace common to A and $P_M{}^{\perp}$. Therefore

$$\hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge (\hat{1} - \hat{P}_M) = \hat{A} - \hat{A} \wedge \hat{P}_M{}^{\perp}. (*) \qquad (3.10)$$

The intersection $\hat{A} \wedge \hat{P}_M{}^{\perp}$ of the projectors $\hat{A}$ and $\hat{P}_M{}^{\perp}$ can be expressed (see Appendix A) as

$$\hat{A} \wedge \hat{P}_M{}^{\perp} = \lim_{n\to\infty}\text{-s}(\hat{A}\hat{P}_M{}^{\perp})^n = \lim_{n\to\infty}\text{-s}(\hat{A}\hat{P}_M{}^{\perp}\hat{A})^n = \lim_{n\to\infty}\text{-s}(\hat{A} - \hat{A}\hat{P}_M\hat{A})^n \qquad (3.11)$$

where lim-s is meant limit in the strong sense: i.e. $\lim_{mn\to\infty} \|(\hat{A}\hat{P}_M{}^{\perp})^m\psi - (\hat{A}\hat{P}_M{}^{\perp})^n\psi\| \to 0$ for all $\psi \in \mathcal{H}$. The subspaces A_M may now be constructed from the projectors $\hat{A} \circ \hat{P}$;

$$\hat{A}_M = \hat{A} \circ \hat{P}_M - \hat{A} \circ \hat{P}_{M+1} = \hat{A} \wedge \hat{P}_{M+1}{}^{\perp} - \hat{A} \wedge \hat{P}_M{}^{\perp}. \qquad (3.12)$$

We now have the necessary tools to treat situations involving changing numbers of bound atoms. From the orthogonal decomposition

$$\psi = \hat{A}\psi = \sum_{M=0}^{N} \hat{A}_M \psi = \sum_{M=0}^{N} \psi_M, \qquad (3.13)$$

and the Schrödinger equation $i\hbar \frac{\partial \psi}{\partial t} = \hat{H}\psi$ for the N-proton, N-electron system

$$i\hbar\dot{\psi} = \sum_{M=0}^{N} i\hbar\dot{\psi}_M = \hat{H}\psi = \sum_{M=0}^{N} (\hat{H}\psi)_M \qquad (3.14)$$

where $(\hat{H}\psi)_M \equiv \hat{A}_M \hat{H}\psi$. Therefore

$$i\hbar \frac{\partial}{\partial t} \psi_M = \hat{A}_M \hat{H}\psi = \sum_{M'} \hat{A}_M \hat{H} \hat{A}_{M'} \psi = \sum_{M'} \hat{A}_M \hat{H} \hat{A}_{M'} \psi_{M'} = \sum_{M'} \hat{H}_{MM'} \psi_{M'} \qquad (3.15)$$

---

* 3.10 can also be written $\hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge \hat{P}_M{}^{\perp} = \hat{A} \wedge (\hat{1} - \hat{A} \wedge \hat{P}_M{}^{\perp}) = \hat{A} \wedge \{\hat{A} \wedge \hat{P}_M{}^{\perp}\}^{\perp} = \hat{A} \wedge (\hat{A}^{\perp} \vee \hat{P}_M)$ where $\hat{P}_1 \vee \hat{P}_2 \equiv (\hat{P}_1{}^{\perp} \wedge \hat{P}_2{}^{\perp})^{\perp}$ for two projectors $\hat{P}_1, \hat{P}_2$. We may also express $\hat{A}_M$ as $\hat{A} \circ \hat{P}_M \wedge \hat{P}_{M+1}{}^{\perp}$ (see appendix C).

where the Hamiltonian operator $\hat{H}_{MM'} \equiv \hat{A}_M \hat{H} \hat{A}_{M'}$ maps the M′-bound atom subspace into the M-atom subspace. We note that $\hat{H}^*_{MM'} = (\hat{A}_M \hat{H} \hat{A}_{M'})^* = \hat{A}_{M'} \hat{H} A_M = \hat{H}_{M'M}$, so that on A, $\hat{H}$ has the correct Hermitian character

$$\hat{H} \to \hat{A}\hat{H}\hat{A} = \sum_{MM'} \hat{H}_{MM'} = \left(\sum_{MM'} \hat{H}_{MM'}\right)^* = \sum_{MM'} \hat{H}_{M'M}. \qquad (3.16)$$

For some problems, it may be useful to use the formalism as developed to this point. One can easily introduce "atomic" and "free" particle observables and carry out a quantum mechanical treatment for them. However, we have lost much of the simplicity which we had when we dealt with the functions $\psi^M_{\alpha ij}$. We would like to transform our theory back to the simpler subspaces P_M. How can this be done? Well, we notice that A_M is a subspace of $\hat{A}P_M$ and a non zero vector $\psi \in A_M$ is of the form $\psi = \hat{A}P_M\chi$ for some χ, $\chi \in \mathcal{H}$. Let us look at those $\chi_0 \in P_M$ which are mapped into zero by $\hat{A}\hat{P}_M$: $\hat{A}\hat{P}_M\chi_0 = 0 \Rightarrow (\psi, \hat{A}\hat{P}_M\chi_0) = 0$, $\psi \in \mathcal{H}$ so $0 = (P_M \hat{A}\psi, \chi_0)$, and hence χ_0 is in the orthogonal compliment of $P_M \circ A$ in P_M. Thus it follows that non zero vectors in $P_M \circ A$ are mapped by $\hat{A}\hat{P}_M$ into non-zero vectors in $A \circ P_M$, and non zero vectors in $A \circ P_M$ are mapped by $\hat{P}_M\hat{A} = (\hat{A}\hat{P}_M)^*$ into non zero vectors in $P_M \circ A$. The following simple diagram illustrates our result:*

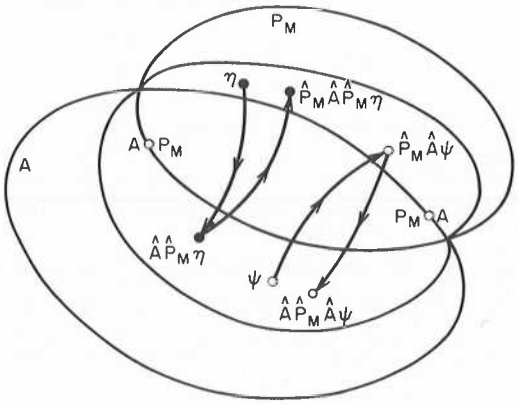


Figure I

* This result is well-known to some mathematicians (e.g. those who know it well).

The mapping $\hat{A}\hat{P}_M\hat{A} = \hat{A}\hat{P}_M(\hat{A}\hat{P}_M)^*$ is 1-1 on A^oP_M and $\hat{P}_M\hat{A}\hat{P}_M = (\hat{A}P_M)^*(\hat{A}P_M)$ is 1-1 on $P_M{}^oA$ (see Fig I). Hence on A^oP_M we may normalize $\hat{P}_M\hat{A}$ (which maps A^oP_M onto $P_M{}^oA$) and introduce*

$$\hat{W}_M \equiv (\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}\,\hat{P}_M\hat{A}. \tag{3.17}$$

$\hat{W}_M$ maps A^oP_M onto $P_M{}^oA$ and

$$\hat{W}_M\hat{W}_M{}^* = (\hat{P}_M\hat{A}P_M)^{-\frac{1}{2}}\,\hat{P}_M\hat{A}\;\hat{A}\hat{P}_M(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}$$

$$= \begin{cases} 1 \text{ (on } P_M{}^oA) \\ 0 \text{ on } (P_M{}^oA)^{\perp} \end{cases}$$

$$= \hat{P}_M{}^o\hat{A} \tag{3.18}$$

Similarly

$$\hat{W}_M{}^*\hat{W}_M = \hat{A}^o\hat{P}_M \tag{3.19}$$

thus $\hat{W}_M, (\hat{W}_M{}^*)$ is a (partial) isometric mapping of A^oP_M onto $P_M{}^oA$ ($P_M{}^oA$ onto A^oP_M).

More technically, $\hat{W}_M$ is the essentially unique factor appearing in the polar decomposition of $\hat{P}_M\hat{A}$

$$\hat{P}_M\hat{A} = \sqrt{\hat{P}_M\hat{A}\hat{P}_M}\;\hat{W}_M = \hat{W}_M\sqrt{\hat{A}\hat{P}_M\hat{A}} \tag{3.20}$$

Similarly**

$$\hat{A}\hat{P}_M = \sqrt{\hat{A}\hat{P}_M\hat{A}}\;\hat{W}_M{}^* = \hat{W}_M{}^*\sqrt{\hat{P}_M\hat{A}\hat{P}_M} \tag{3.21}$$

---

* To be more meaningful, it is perhaps better to recognize that $\hat{W}_M = (\hat{P}_M{}^o\hat{A})(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}(\hat{P}_M{}^o\hat{A})\hat{P}_M\hat{A}(\hat{A}^o\hat{P}_M)$ since $(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}$ has meaning only on $P_M{}^oA$.

** On A^oP_M, $\hat{P}_M\hat{A}$ has an inverse $(\hat{P}_M\hat{A})^{-1}$ acting on $P_M{}^oA$, so

$$\hat{P}_M\hat{A}\hat{P}_M = \hat{P}_M\hat{A}\;\hat{A}\hat{P}_M\cdot\hat{P}_M\hat{A}\cdot(\hat{P}_M\hat{A})^{-1},\quad (\hat{P}_M^2=\hat{P}_M,\ \hat{A}^2=\hat{A})$$

$$= \hat{P}_M\hat{A}\sqrt{\hat{A}\hat{P}_M\hat{A}}(P_MA)^{-1}(P_MA)\sqrt{\hat{A}\hat{P}_M\hat{A}}(\hat{P}_M\hat{A})^{-1}$$

$$= [\hat{P}_M\hat{A}\sqrt{AP_MA}\,(\hat{P}_M\hat{A})^{-1}]2$$

thus $\sqrt{\hat{P}_M\hat{A}\hat{P}_M} = \hat{P}_M\hat{A}\sqrt{\hat{A}\hat{P}_MA}\,(\hat{P}_M\hat{A})^{-1}$ and $\sqrt{\hat{P}_M\hat{A}\hat{P}_M}\,\hat{P}_M\hat{A}=\hat{P}_M\hat{A}\sqrt{\hat{A}\hat{P}_M\hat{A}}$ or $\hat{P}_M\hat{A}(\hat{A}\hat{P}_M\hat{A})^{-\frac{1}{2}} = (\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}\,\hat{P}_M\hat{A} = \hat{W}_M$.

(See Appendix A for a more technical presentation.)

By "essentially" unique, I mean that $\hat{W}_M$ is unique in the sense that any $\hat{W}_M'$ satisfying 3.20 having the same domain and range as $\hat{W}_M$, is equal to $\hat{W}_M$. The (partial) isometry $\hat{W}_M$ maps A^oP_M onto $P_M{}^oA$ it therefore maps the subspace A_M of A^oP_M onto a subspace C_M of $P_M{}^oA \subset P_M$. By our definition

$$C_M \equiv \hat{W}_M A_M \tag{3.22}$$

$$A_M = \hat{W}_M{}^* C_M \tag{3.23}$$

The projector $\hat{C}_M$ onto C_M can be written

$$\hat{C}_M = \hat{W}_M \hat{A}_M \hat{W}_M{}^*. \tag{3.24}$$

Clearly $\hat{C}_M$ as defined above is self-adjoint. Further $\hat{C}_M{}^2 = \hat{W}_M\hat{A}_M\hat{W}_M{}^*\hat{W}_M\hat{A}_M\hat{W}_M{}^* = \hat{W}_M\hat{A}_M\hat{A}^o\hat{P}_M\hat{A}_M\hat{W}_M = \hat{W}_M\hat{A}_M{}^2\hat{W}_M{}^* = \hat{C}_M$, since $\hat{A}_M < \hat{A}^o\hat{P}_M$. Thus $\hat{C}_M$ is a projector. We now show that C_M is stable under $\hat{C}_M$:

$$\begin{aligned}\hat{C}_M C_M = \hat{C}_M\hat{W}_M A_M &= \hat{W}_M\hat{A}_M\hat{W}_M{}^*\hat{W}_M A_M \\ &= \hat{W}_M\hat{A}_M\hat{A}^o\hat{P}_M A_M \\ &= \hat{W}_M\hat{A}_M A_M = \hat{W}_M A_M = C_M.\end{aligned}$$ *

Eq. 2.24 can be inverted to yield $\hat{A}_M = \hat{W}_M{}^*\hat{C}_M\hat{W}_M$.

Any physical state ψ may now be decomposed into orthogonal M-atom states:

$$\psi = \hat{A}\psi = \sum_{M=0}^{N} \hat{A}_M\psi = \sum_{M=0}^{N} \psi_M, \tag{3.25}$$

and the ψ_M in A_M may be related to ξ_M in C_M by

$$\xi_M = \hat{W}_M\psi_M, \quad \psi_M = \hat{W}_M{}^* \xi_M. \tag{3.26}$$

Since $\hat{W}_M{}^*$ is a partial isometry $(\psi_M,\psi_M) = (\xi_M,\xi_M)$, and

$$(\psi,\psi) = \|\psi\|^2 = \sum_{M=0}^{N} \|\psi_M\|^2 = \sum_{M=0}^{N} \|\xi_M\|^2. \tag{2.27}$$

* Note that if $\hat{C}_M\psi = 0$, $\hat{W}_M\hat{A}_M\hat{W}_M{}^*\psi = 0$, or for any $x \in \mathcal{H}$, $0 = (x,\hat{W}_M\hat{A}_M\hat{W}_M{}^*\psi) = (\hat{W}_M\hat{A}_M\hat{W}_M{}^*x,\psi) = 0$, so if x is an arbitrary element of C_M, $0 = (x,\psi) \Rightarrow \psi \in C_M{}^\perp$ which establishes $\hat{C}_M$ as the projector for C_M.

Further if ψ' is another physical state,

$$(\psi,\psi') = \sum_{M=0}^{N}(\psi_M,\psi_M') = \sum_{M=0}^{N}(\xi_M,\xi_M') \tag{3.28}$$

Since $\xi_M \in C_M \subset P_M{}^oA \subset P_M$, we may expand ξ_M in terms of $\psi^M_{\alpha ij}$,

$$\xi_M = \sum_{\alpha ij} C^M_{\alpha ij}\ \psi^M_{\alpha ij}, \tag{3.29}$$

and since $\xi_M \in P_M{}^oA$, $\xi_M = \hat{P}_M\hat{A}\psi$ for some $\psi \in \mathcal{H}$, which means that $C^M_{\alpha ij}=(\psi^M_{\alpha ij}\ \xi_M)=(\psi^M_{\alpha ij}\ \hat{P}_M\hat{A}\psi)=(\hat{A}\hat{P}_M\psi^M_{\alpha ij},\psi)=(\hat{A}\psi^M_{\alpha ij},\psi)$. This means that the $C^M_{\alpha ij}$ are completely symmetric in $\alpha_1\ldots\alpha_M$, completely antisymmetric in $i_{M+1},\ldots i_N$ and completely antisymmetric in $j_{M+1},\ldots j_N$. The scalar product 3.28 becomes

$$(\psi,\psi') = \sum_{M=0}^{N}\ \sum_{\alpha ij} \overline{C}^M_{\alpha ij}\ C'^M_{\alpha ij} \tag{3.30}$$

We have now demonstrated a correspondence between physical states ψ and a collection $\{C^M_{\alpha ij}\}$ of coefficients

$$\psi \rightarrow \{C^M_{\alpha ij}\} \tag{3.31}$$

Further the coefficients have the symmetry corresponding to "bound" bose atoms and "free" fermi protons and electrons. Not every set $\{C^M_{\alpha ij}\}$ having the correct symmetry corresponds to a physical state however. Only those $C^M_{\alpha ij}$ which through 3.29 give rise to a $\xi_M \in C_M$ correspond to a physical state. The restriction on $C^M_{\alpha ij}$ that they represent physical states is that $\xi_M=\hat{C}_M\xi_M$, or

$$C^M_{\alpha ij} = (\psi^M_{\alpha ij},\xi_M) = (\psi^M_{\alpha ij},\hat{C}_M\xi_M) = (\hat{C}_M\psi^M_{\alpha ij},\xi_M). \tag{3.32}$$

We may write

$$\hat{C}_M\psi^M_{\alpha ij} = \sum_{\alpha' i' j'} \psi^M_{\alpha' i' j'}\ \langle M\alpha' i' j'|\hat{C}_M|M\alpha ij\rangle$$

Hence

$$C^M_{\alpha ij} = (\hat{C}_M C)^M_{\alpha ij} \equiv \sum_{\alpha' i' j'}\langle M\alpha ij|\hat{C}_M|M\alpha' i' j'\rangle C^M_{\alpha' i' j'} \tag{3.33}$$

is the additional required restriction on the $C^M_{\alpha ij}$.

If we use 3.17 for $\hat{W}_M$, the expression 3.24 for $\hat{C}_M$ becomes

$$\hat{C}_M = (\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}\,\hat{P}_M\hat{A}\,\hat{A}_M\,\hat{A}\hat{P}_M(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}$$

$$= (\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}\,\hat{A}_M(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}} \qquad (3.34)$$

Since $\hat{A}_M\hat{P}_M = \{\hat{A} \wedge \hat{P}_{M+1}{}^{\perp} - \hat{A} \wedge \hat{P}_M{}^{\perp}\}\hat{P}_M = (\hat{A} \wedge \hat{P}_{M+1}{}^{\perp})\hat{P}_M$, the above expression becomes

$$\hat{C}_M = (\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}\,(\hat{A} \wedge \hat{P}_{M+1}{}^{\perp})(\hat{P}_M\hat{A}\hat{P}_M)^{-\frac{1}{2}}. \qquad (3.35)$$

The coefficients $C^M_{\alpha_1\ldots\alpha_M\, i_{M+1}\ldots i_N\, j_{M+1}\ldots j_N}$ are completely symmetric in $\alpha_1\ldots\alpha_M$, and completely antisymmetric in $i_{M+1}\ldots i_N$ and in $j_{M+1}\ldots j_N$. We may introduce more general state vectors $C^{M_aN_pN_e}_{\alpha_1\ldots\alpha_{M_a} i_1\ldots i_{N_p}\, j_1\ldots j_{N_e}}$ and define operators a_α, $a_\alpha{}^*$, a_i, $a_i{}^*$, b_j, $b_j{}^*$ through expressions of the type:

$$(a_\alpha C)^{M_aN_pN_e}_{\alpha_1\ldots\alpha_{M_a} i_1\ldots i_{N_p} j_1\ldots j_{N_e}} = \sqrt{M_a+1}\; C^{M_a+1,N_pN_e}_{\alpha_1\ldots\alpha_{M_a}\alpha, i_1\ldots i_{N_p} j_1\ldots j_{N_e}} \qquad (3.36)$$

$$(a_\alpha{}^* C)^{M_aN_pN_e}_{\alpha_1\ldots\alpha_{M_a} i_1\ldots i_{N_p} j_1\ldots j_{N_e}} =$$

$$\sqrt{M_a}\; S_{\alpha_1\ldots\alpha_M}\; C^{M_a-1,N_pN_e}_{\alpha_1\ldots\alpha_{M_a-1} i_1\ldots i_{N_p} j_1\ldots j_{N_e}}\,\delta_{\alpha_M\alpha}, \qquad (3.37)$$

where $S_{\alpha_1\ldots\alpha_M}$ symmetrizes the indices $\alpha_1\ldots\alpha_M$, i.e. $S_{\alpha_1\ldots\alpha_{M_a}} = \frac{1}{M_a!}\sum_{(\alpha)} P(\alpha)$ where $P(\alpha)$ permutes $(\alpha_1\ldots\alpha_M)$ and the sum goes over all $M_a!$ permutions of $\alpha_1\ldots\alpha_{M_a}$. The operators a_i, b_j etc. are defined similarly except that in $a^*{}_i$,

$\hat{b}^*_j$ the symmetrizer is replaced by an antisymmetrizer.* The operators $\hat{a}_\alpha$, $\hat{a}^*_\alpha$ are adjoint to each other and satisfy the usual Bose commutation relations

$$[\hat{a}_\alpha,\hat{a}_\beta] = 0 \quad [\hat{a}_\alpha,\hat{a}^*_\beta] = \delta_{\alpha\beta} \tag{3.38}$$

Similarly $\hat{a}_i,\hat{a}^*_j$ and $\hat{b}_j,\hat{b}^*_j$ satisfy the usual Fermi anticommutation relation

$$\{\hat{a}_i,\hat{a}_j\} = \{\hat{b}_i,\hat{b}_j\} = 0 \tag{3.39}$$

$$\{\hat{a}_i,\hat{a}^*_j\} = \{\hat{b}_i,\hat{b}^*_j\} = \delta_{ij} \tag{3.40}$$

$$\{\hat{a}_i,\hat{b}_j\} = \{\hat{a}^*_i,\hat{b}^*_j\} = 0 \tag{3.41}$$

With the aid of the destruction operators $\hat{a}_\alpha$, $\hat{a}_i$, $\hat{b}_j$ we may define the field operators:

$$\hat{\chi}(x,y) = \sum_\alpha \hat{a}_\alpha\varphi_\alpha(x,y), \quad \text{"Bound" atomic field,} \tag{3.42}$$

$$\hat{\varphi}(x) = \sum_i \hat{a}_i\varphi_i(x), \quad \text{"Free" proton field,} \tag{3.43}$$

and

$$\hat{\psi}(y) = \sum_i \hat{b}_j\psi_j(y), \quad \text{"Free" electron field.} \tag{3.44}$$

These operators obey the commutation (anticommutation) relations

$$[\hat{\chi}(xy),\hat{\chi}(x'y')] = 0, \; [\hat{\chi}(x,y),\hat{\chi}^*(x'y')] = \langle xy|\hat{P}_B|x'y'\rangle \tag{3.45}$$

$$\{\hat{\varphi}(x),\hat{\varphi}(x')\} = \{\hat{\varphi}(x),\hat{\psi}(y)\} = \{\hat{\psi}(y),\hat{\psi}(y')\} = 0, \text{etc.}, \tag{3.46}$$

$$\{\hat{\varphi}(x),\hat{\varphi}^*(x')\} = \delta(x-x'), \tag{3.47}$$

$$\{\hat{\psi}(y),\hat{\psi}^*(y')\} = \delta(y-y'). \tag{3.48}$$

* The definitions of a_i, a_i^* have an additional factor $(-1)^{N_e}$ so as to require that a_i, b_j have the standard anticommutation relations 3.41 for distinct fermions. i.e.

$$(a_i C)^{M_a N_p N_e}_{(\alpha) i_1 \ldots i_{N_p}(j)} = (-1)^{N_e} \sqrt{N_p+1}\, C^{M_a N_p+1,\, N_e}_{(\alpha) i_1 \ldots i_{N_p} i(j)}$$

and with no factor $(-1)^{N_p}$ in the definition of b_j.

We are dealing now with a Fock space $\mathfrak{F}$ whose vectors Ψ are sequences of functions $\{\psi^{M_a N_p N_e}(z_1 \ldots z_{M_a}; x_1 \ldots x_{N_p}; y_1 \ldots y_{N_e})\}$, $z_\alpha = x_\alpha y_\alpha$, $M_\alpha, N_p, N_e = 0,1,2..$ with the scalar product (Ψ,Ψ') of two elements of $\mathfrak{F}$ expressed as

$$(\Psi,\Psi') = \sum_{M_a N_p N_e} \int \psi^{*M_a N_p N_e}(z_1 \ldots z_{M_a}; x_1 \ldots x_{i_{N_p}}; y_1 \ldots y_{N_e})$$
$$\psi'^{M_a N_p N_e}(z_1 \ldots z_{M_a}; x_1 \ldots x_{i_{N_p}}; y_1 \ldots y_{N_e})$$
$$dz_1 \ldots dz_{M_a}\, dx_1 \ldots dx_{i_{N_p}}\, dy_1 \ldots dy_{N_e} \qquad (3.49)$$

The state Ψ may be generated from the vacuum $|0)$ through

$$\Psi = \sum_{M_a, N_p, N_e} \frac{1}{\sqrt{M_a! N_p! N_e!}} \int dx_1 \ldots dx_{N_p}\, dy_1 \ldots dy_{N_e}\, dz_1 \ldots dz_{M_a}$$
$$\psi^{M_a N_p N_e}(z_1 \ldots y_{N_e}) \chi^*(z_{M_a}) \ldots \chi^*(z_1) \varphi^*(x_{N_p}) \ldots \varphi^*(x_1)$$
$$\times\ \psi^*(y_{N_e}) \ldots \psi^*(y_1) |\, 0) \qquad (3.50)$$

The results up to this point have been rigorous, with no approximations of any kind. We observe that in order to transform the usual quantum mechanical basis to our new "composite particle" basis we must be able to compute expressions of the form $\hat{P}_1 \wedge \hat{P}_2$ which involve limiting processes $(\hat{P}_1 \wedge \hat{P}_2)\psi = \lim_{N\to\infty} (\hat{P}_1 \hat{P}_2)^N \psi$. We would like to find an expression for H, the Hamiltonian, in the new basis. We first expand the physical wave function ψ:

$$\psi = \hat{A}\psi = \sum_M \hat{A}_M \psi = \sum_M \psi_M = \sum_M \hat{W}_M{}^* \xi_M \qquad (3.51)$$

where $\xi_M \equiv \hat{W}_M \psi_M = \hat{W}_M \hat{A}_M \psi$. Similarly

$$\hat{H}\psi = \hat{A}\hat{H}\psi = \sum_M \hat{A}_M \hat{H}\psi = \sum_M (\hat{H}\psi)_M = \sum_M \hat{W}_M{}^* (\tilde{H}\xi)_M$$

where

$$(\tilde{H}\xi)_M \equiv \hat{W}_M(\hat{H}\psi)_M = \hat{W}_M\hat{A}_M\hat{H}\psi = \sum_{M'} \hat{W}_M\hat{A}_M\hat{H}\hat{A}_{M'}\hat{W}_{M'}{}^*\xi_{M'}, \quad (3.52)$$

or

$$(\tilde{H}\xi)_M = \sum_{M'} \tilde{H}_{MM'}\xi_{M'}, \text{ with } \tilde{H}_{MM'} = \hat{W}_M\hat{A}_M\hat{H}\hat{A}_{M'}\hat{W}_{M'}{}^* = \hat{W}_M\hat{H}\hat{W}_{M'}{}^*.$$

The terms $H_{MM'}$, $M \neq M'$, in the Hamiltonian correspond to transitions $\xi_{M'} \rightarrow \xi_M$ in which the number of bound atoms changes. Therefore, if we can find reasonable expressions for the $\tilde{H}_{MM'}$, we will be able to look at questions relating to chemical reactions, rates of ionization, etc., from a many-body point of view. The existence of the composite particle basis has now been established, at least for the rather simple system "bound Hs", "free protons" and "free electrons". The theory is complete in that it does not distinguish between tenuous ionized hydrogen at high temperature and dense solid hydrogen at low temperature. It is to be expected, therefore, that the general formalism be very complex. However, it was not designed to be useful in the general case. (It would not be very useful to describe solid H in terms of "bound" atoms and "free" electrons and protons, even though it is in principle possible). On the other hand, if it is sensible to think of the system as being composed of composites and free particles, our method should be useful, once we are able to introduce appropriate approximations to our general formulae. For example suppose that the system is sufficiently tenuous that only those electron-proton pairs which have formed bound atoms contribute to the bound state component. For such states

$$\sqrt{\hat{P}_M\hat{A}\hat{P}_M}\,\xi_M = \lambda_M\xi_M \quad (3.53)$$

where $\lambda_M{}^2 = M!((N-M)!)^2/(N!)^2$

Further, in this approximation

$$(\chi^M_{\alpha ij}|\tilde{H}_{MM'}|\chi^{M'}_{\alpha' i' j'}) = (\chi^M_{\alpha ij}|\hat{A}\hat{H}\hat{A}|\chi^{M'}_{\alpha' i' j'})/\lambda_M\lambda_{M'}$$

The Hamiltonian can now be expressed in second quantized form* as

* First established by R. S. Stolt

$$
\begin{aligned}
\tilde{H} = {} & \int \hat{\chi}^*(x,y)(T_p+T_e+V_{ep})\hat{\chi}(x,y)\,dxdy + \int \hat{\varphi}^*(x)T_p\hat{\varphi}(x)\,dx \\
& + \int \hat{\psi}^*(y)T_e\hat{\psi}(y)\,dy + \int \hat{\psi}^*(y)\hat{\varphi}^*(x)V_{ep}\hat{\varphi}(x)\hat{\psi}(y)\,dxdy \\
& + \int \hat{\chi}^*(x,y)\hat{\psi}^*(y')(V_{ee}+V_{ep})\hat{\psi}(y')\hat{\chi}(x,y)\,dxdydy' \\
& + \int \hat{\chi}^*(x,y)\hat{\varphi}^*(x')(V_{pp}+V_{pe})\hat{\varphi}(x')\hat{\chi}(x,y)\,dx'dxdy \\
& + \tfrac{1}{2}\int \hat{\chi}^*(x,y)\hat{\chi}^*(x'y')(V_{ee}+V_{pp}+V_{pe}+V_{ep})\hat{\chi}(x'y')\hat{\chi}(x,y)\,dxdx'dydy' \\
& + \tfrac{1}{2}\int \hat{\varphi}^*(x)\hat{\varphi}(x')V_{pp}\hat{\varphi}(x')\hat{\varphi}(x)\,dxdx' \\
& + \tfrac{1}{2}\int \hat{\psi}^*(y)\hat{\psi}^*(y')V_{ee}\hat{\psi}(y')\hat{\psi}(y)\,dydy' \\
& - \int \hat{\chi}^*(x,y)\hat{\varphi}^*(x')I_p h_{epp}\hat{\varphi}(x')\hat{\chi}(x,y)\,dx'dxdy \\
& - \int \hat{\chi}^*(x,y)\hat{\psi}^*(y')I_e h_{eep}\hat{\psi}(y')\hat{\chi}(x,y)\,dxdydy' \\
& - \tfrac{1}{2}\int \hat{\chi}^*(x,y)\hat{\chi}^*(x'y')I_e h_{2e,2p}\hat{\chi}(x'y')\hat{\chi}(x,y)\,dxdx'dydy' \\
& + \int \hat{\varphi}^*(x)\hat{\psi}^*(y)(T_p+T_e+V_{ep})\hat{\chi}(x,y)\,dxdy \\
& + \int \hat{\chi}^*(x,y)(T_p+T_e+V_{ep})\hat{\psi}(y)\hat{\varphi}(x)\,dxdy \\
& + \int \hat{\chi}^*(x,y)\hat{\varphi}^*(x')\hat{\psi}^*(y')(V_{aa}-I_e h_{2e,2p})\hat{\chi}(x'y')\hat{\chi}(x,y)\,dxdx'dydy' \\
& + \int \hat{\chi}^*(x,y)\hat{\chi}(x'y')(V_{aa}-I_e h_{2e,2p})\hat{\psi}(y')\hat{\varphi}(x')\hat{\chi}(xy)\,dxdx'dydy' \\
& + \int \hat{\varphi}^*(x)\hat{\psi}^*(y)\hat{\psi}^*(y')V_{ea}\hat{\psi}(y')\hat{\chi}(x,y)\,dxdydy' \\
& + \int \hat{\chi}^*(x,y)\hat{\psi}^*(y')V_{ae}\hat{\psi}(y')\hat{\psi}(y)\hat{\varphi}(x)\,dxdydy' \\
& + \int \hat{\varphi}^*(x)\hat{\varphi}^*(x')\hat{\psi}^*(y)V_{pa}\hat{\varphi}(x')\hat{\chi}(x,y)\,dxdx'dy \\
& + \int \hat{\chi}^*(x,y)\hat{\varphi}^*(x')V_{ap}\hat{\psi}(y)\hat{\varphi}(x')\hat{\varphi}(x)\,dxdx'dy\ .
\end{aligned}
\tag{3.54}
$$

In the above expression V_{ep} is the single electron-proton interaction, V_{ap} the single atom-proton interaction, I_e the electron pair exchange operator, $h_{2e,2p}$ the 2-electron 2-proton interaction, etc. This approximate Hamiltonian is probably sufficient for many problems. The terms all have direct physical significance. The first three terms represent free bound atoms, free protons, and free electrons. The next term ($\hat{\Psi}^*\hat{\varphi}^* V_{ep} \hat{\Psi}\varphi$) represents the unbound electron proton interaction. Continuing we have the electron-atom interaction, the proton atom interaction, etc. It may be instructive to write the approximate Hamiltonian 3.54 as

$$\begin{aligned}\tilde{H} = {} & \tilde{T}_a + \tilde{T}_p + \tilde{T}_e + \tilde{V}_{ep} + \tilde{V}_{ea} + \tilde{V}_{pa} + \tilde{V}_{aa} + \tilde{V}_{ee} + \tilde{V}_{pp} \\ & + \tilde{E}_{ea} + \tilde{E}_{pa} + \tilde{E}_{aa} + \tilde{V}(ep \leftarrow a) + \tilde{V}(a \leftarrow ep) \\ & + \tilde{V}(epa \leftarrow aa) + \tilde{V}(aa \leftarrow epa) + \tilde{V}(eep \leftarrow ea) \\ & + \tilde{V}(ea \leftarrow eep) + \tilde{V}(epp \leftarrow pa) + \tilde{V}(pa \leftarrow epp) , \qquad (3.55)\end{aligned}$$

where we list below expressions for the various terms: (using $\hat{a}_\alpha, \hat{a}_i, \hat{b}_j$ etc. instead of the fields $\hat{\chi}, \hat{\varphi}, \hat{\Psi}$)

$$\tilde{T}_a = \sum_{\alpha\alpha'} \hat{a}_\alpha^+ \langle \alpha | T_p + T_e + V_{ep} | \alpha' \rangle \hat{a}_{\alpha'};$$

$$\tilde{T}_p = \sum_{ii'} \hat{a}_i^+ \langle i | T_p | i' \rangle \hat{a}_i;$$

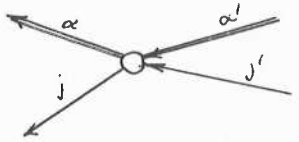
$$\tilde{T}_e = \sum_{jj'} \hat{b}_j^+ \langle j | T_e | j' \rangle \hat{a}_{j'}$$

$$\tilde{V}_{ep} = \sum_{iji'j'} \hat{a}_i^* \hat{b}_j^*$$

$$\langle ij | V_{ep} | i'j' \rangle \hat{b}_{j'} \hat{a}_{i'};$$

$$\tilde{V}_{ea} = \sum_{\alpha_1\alpha' jj'} \hat{a}_\alpha^* \hat{b}_j^*$$

$$\langle \alpha j | V_{ee} + V_{ep} | \alpha' j' \rangle \hat{b}_{j'} \hat{a}_{\alpha'};$$

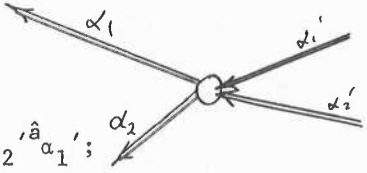


$$\tilde{V}_{pa} = \sum_{\alpha\alpha' ii'} \hat{a}_\alpha^* \hat{a}_i^*$$

$$\langle \alpha i | V_{pe} + V_{pp} | \alpha' i' \rangle \hat{a}_{i'} \hat{a}_{\alpha'};$$

$$\tilde{V}_{aa} = \tfrac{1}{2} \sum_{\alpha_1\alpha_2\alpha_1'\alpha_2'} \hat{a}_{\alpha_1}^* \hat{a}_{\alpha_2}^*$$

$$\langle \alpha_1\alpha_2 | V_{ee} + V_{pp} + V_{ep} + V_{pe} | \alpha_1' a_2' \rangle \hat{a}_{\alpha_2'} \hat{a}_{\alpha_1'};$$



$$\tilde{V}_{ee} = \tfrac{1}{2} \sum_{j_1 j_2 j_1' j_2'} \hat{b}_{j_1}^* \hat{b}_{j_2}^*$$

$$\langle j_1 j_2 | V_{ee} | j_1' j_2' \rangle \hat{b}_{j_2'} \hat{b}_{j_1'};$$

$$\tilde{V}_{pp} = \tfrac{1}{2} \sum_{i_1 i_2 i_1' i_2'} \hat{a}_{i_1}^* \hat{a}_{i_2}^*$$

$$\langle i_1 i_2 | V_{pp} | i_1' i_2' \rangle \hat{a}_{i_2'} \hat{a}_{i_1'};$$

$$\tilde{E}_{ea} = -\sum_{\alpha\alpha' jj'} \hat{a}_\alpha^* \hat{b}_j$$

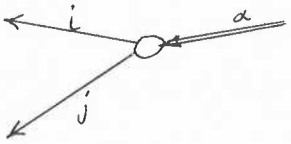
$$\langle \alpha j | I^e h_{eep} | \alpha' j' \rangle \hat{b}_{j'} \hat{a}_{\alpha'}$$

$$\tilde{E}_{pa} = -\sum_{\alpha\alpha ii'} \hat{a}_\alpha^* \hat{a}_i^*$$

$$\langle \alpha i | I^p h_{epp} | \alpha' i' \rangle \hat{a}_{i'} \hat{a}_{\alpha'};$$

$$\tilde{E}_{aa} = -\tfrac{1}{2}\sum_{\alpha_1\alpha_2\alpha_1'\alpha_2'} \hat{a}_{\alpha_1}^{*}\hat{a}_{\alpha_2}^{*}$$

$$\langle\alpha_1\alpha_2|I^{e}h_{ee\text{-}pp}|\alpha_1'\alpha_2'\rangle\hat{a}_{\alpha_2'}\hat{a}_{\alpha_1'};$$

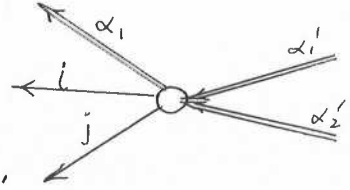


$$\tilde{V}(ep\leftarrow a) = \sum_{\alpha ij} \hat{a}_i^{*}\hat{b}_j^{*}$$

$$\langle ij|T_p+T_e+V_{ep}|\alpha\rangle\hat{a}_{\alpha};$$

$$\tilde{V}(a\leftarrow ep) = \tilde{V}(ep\leftarrow a)^{*} = \sum_{\alpha ij} \hat{a}_{\alpha}^{*}$$

$$\langle\alpha|T_p+T_e+V_{ep}|ij\rangle\hat{b}_j\hat{a}_j;$$



$$\tilde{V}(aep\leftarrow aa) = \sum_{\alpha_1\alpha_1'\alpha_2' ij} \hat{a}_{\alpha_1}^{*}\hat{a}_i^{*}\hat{b}_j^{*}$$

$$\langle\alpha_1 ij|V_{aa}-I^{e}h_{ee\text{-}pp}|\alpha_1'\alpha_2'\rangle\hat{a}_{\alpha_2'}\hat{a}_{\alpha_1'}$$

$$\tilde{V}(aa\leftarrow aep) = V(aep\leftarrow aa)^{*} =$$

$$\sum_{\alpha_1\alpha_1'\alpha_2' ij} \hat{a}_{\alpha_1'}^{*}\hat{a}_{\alpha_2'}^{*}$$

$$\langle\alpha_1'\alpha_2'|V_{aa}-I^{e}h_{ee\text{-}pp}|\alpha_1 ij\rangle$$

$$\hat{b}_j\hat{a}_i\hat{a}_{\alpha_1};$$

$$\tilde{V}(epp\leftarrow pa) = \sum_{i_1 i_2 j \alpha i} \hat{a}_{i_1}^{*} \hat{a}_{i_2}^{*} \hat{b}_{j}^{*}$$

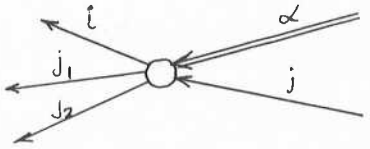
$$\langle i_1 i_2 j | V_{pa} | \alpha i \rangle \, \hat{a}_i \hat{a}_\alpha$$

$$= (\tilde{V}(pa\leftarrow ppe))^{*}$$

$$V(eep\leftarrow ea) = \sum_{i j_1 j_2 \alpha j} \hat{a}_i^{*} \hat{b}_{j_1}^{*} \hat{b}_{j_2}^{*}$$

$$\langle i j_1 j_2 | V_{ea} | \alpha j \rangle \hat{b}_j \hat{a}_\alpha$$

$$= (\tilde{V}(ea\leftarrow epe))^{*}$$



It must be emphasized that the diagrams shown above correspond to terms in the Hamiltonian and are not representative of any perturbative scheme. The last seventeen terms represent the basic vertices representing interaction and exchange in this approximation. It is hoped that even this simplified Hamiltonian will prove to be a useful aid toward the treatment of elementary chemical kinetics etc. from the standpoint of basic quantum theory.

APPENDIX A

Projectors, Partial Isometries, Polar Decomposition

1. Definition of Projectors and Elementary Properties

It is sufficient for our purposes to consider projectors defined on a Hilbert space $\mathcal{H}$. We assume known that if $P_1 \subset \mathcal{H}$ is a closed vector subspace of $\mathcal{H}$, then $P_1^{\perp}$, the set of all vectors in $\mathcal{H}$ which are orthogonal to P_1, is also a closed linear subspace of $\mathcal{H}$. Further, the direct sum $P_1 \oplus P_1^{\perp}$ is $\mathcal{H}$ itself;

$$\mathcal{H} = P_1 \oplus P_1^{\perp} \qquad (A.1)$$

which means that for every $\chi \in \mathcal{H}$ we may write uniquely

$$\chi = \chi_1 + \chi_1^{\perp}$$

with $\chi_1 \in P_1$, $\chi_2 \in P_1^{\perp}$ and

$$\|\chi\|^2 = (\chi,\chi) = \|\chi_1\|^2 + \|\chi_1^{\perp}\|^2.$$

We define projection operators or projectors $\hat{P}_1$, $\hat{P}_1^{\perp}$ by

$$\hat{P}_1\chi = \chi_1, \quad \hat{P}_1^{\perp}\chi = \chi_1^{\perp} = \chi - \chi_1$$

It then follows[7] that $\hat{P}_1, \hat{P}_1^{\perp}$ are bounded linear operators acting on $\mathcal{H}$ having norm one and having the properties

$$\hat{P}_1\hat{P}_1^{*} = \hat{P}_1, \qquad \hat{P}_1^{\perp}\hat{P}_1^{\perp *} = \hat{P}_1^{\perp}, \qquad \hat{P}_1\hat{P}_1^{\perp} = 0,$$

and

$$\hat{1} = \hat{P}_1 + \hat{P}_1^{\perp}$$

is a decomposition of the unit operator on $\mathcal{H}$. Conversely if $\hat{P}$ is any bounded operator on $\mathcal{H}$ satisfying $\hat{P}\hat{P}^* = \hat{P}$ (or equivalently $\hat{P} = \hat{P}^*$, $\hat{P}^2 = \hat{P}$), then $\hat{P}\mathcal{H} \equiv P$ is a closed linear subspace of $\mathcal{H}$ and $\hat{P}$ projects vectors of $\mathcal{H}$ onto P. Further $\hat{P}_{\perp} = \hat{1} - \hat{P}$ is the projector onto $P^{\perp} \equiv \mathcal{H} - P$

We now consider some elementary properties of projectors related to various combinations of projectors. We state these properties as theorems with proofs.

Theorem 1: If $\hat{P}_1$ and $\hat{P}_2$ are projectors on $\mathcal{H}$, the product $\hat{P}_1\hat{P}_2$ is a projector iff $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$.

Proof: $(\hat{P}_1\hat{P}_2)^* = \hat{P}_2^*\hat{P}_1^* = \hat{P}_2\hat{P}_1$, so the condition is necessary. It is also sufficient for

$$(\hat{P}_1\hat{P}_2)^2 = \hat{P}_1(\hat{P}_2\hat{P}_1)\hat{P}_2 = \hat{P}_1^2\hat{P}_2^2 = \hat{P}_1\hat{P}_2.$$

An important special case obtains when $\hat{P}_1\hat{P}_2 = 0$. In this case the subspaces P_1 and P_2 are orthogonal. For let $\chi_1 \in P_1$ and $\chi_2 \in P_2$.

Then

$$(\chi_1,\chi_2) = (\hat{P}_1\chi_1,\hat{P}_2\chi_2) = (\chi_1,\hat{P}_1\hat{P}_2\chi_2) = 0$$

The closed subspace corresponding to the projector $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$ is that which is common to P_1 and P_2 * and which we denote by $P_1 \wedge P_2$. That is

$$P_1 \wedge P_2 = \{\chi\in\mathfrak{K};\ \hat{P}_1\chi = \chi,\ \hat{P}_2\chi = \chi\}.$$

We denote the projector onto $P_1 \wedge P_2$ by $\hat{P}_1 \wedge \hat{P}_2$. We note that the projector $\hat{P}_1 \wedge \hat{P}_2$ is defined independently whether $\hat{P}_1$ and $\hat{P}_2$ commute. However only if $\hat{P}_1,\hat{P}_2$ commute do we have

$$\hat{P}_1 \wedge \hat{P}_2 = \hat{P}_1\hat{P}_2,\ (\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1).$$

Theorem 2: The sum $\hat{P}_1 + \hat{P}_2$ of two projectors $\hat{P}_1$ and $\hat{P}_2$ is a projector iff $\hat{P}_1\hat{P}_2 = 0$.

Proof: The sum is Hermitian since each term is. Therefore we need only to have

$$(\hat{P}_1 + \hat{P}_2)^2 = \hat{P}_1 + \hat{P}_2 + \hat{P}_1\hat{P}_2 + \hat{P}_2\hat{P}_1 = \hat{P}_1 + \hat{P}_2,$$

or

$$\hat{P}_1\hat{P}_2 + \hat{P}_2\hat{P}_1 = 0$$

This latter expression when multiplied on the left and right by $\hat{P}_1$ yields

$$\hat{P}_1\hat{P}_2 + \hat{P}_1\hat{P}_2\hat{P}_1 = 0$$

and

$$\hat{P}_1\hat{P}_2\hat{P}_1 + \hat{P}_2\hat{P}_1 = 0$$

Hence $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$, and by the first line $2\hat{P}_1\hat{P}_2 = 0$ or $\hat{P}_1\hat{P}_2 = 0$.

Theorem 3: The difference $\hat{P}_1-\hat{P}_2$ is a projector iff $\hat{P}_1\hat{P}_2 = \hat{P}_2$.

---

* $\hat{P}_1\chi = \hat{P}_1\hat{P}_1\hat{P}_2\chi = \hat{P}_1\hat{P}_2\chi = \chi$
$\hat{P}_2\chi = \hat{P}_2\hat{P}_1\hat{P}_2\chi = \hat{P}_1\hat{P}_2\ \chi = \hat{P}_1\hat{P}_2\chi = \chi$
conversely if $\chi\in P_1$ and $\chi\in P_2$, $\hat{P}_1\hat{P}_2\chi=\hat{P}_1\chi=\chi$, $\hat{P}_2\hat{P}_1\chi=\hat{P}_2\chi$.

Proof:

$$(\hat{P}_1 - \hat{P}_2)^2 = \hat{P}_1 + \hat{P}_2 - \hat{P}_1\hat{P}_2 - \hat{P}_2\hat{P}_1 = \hat{P}_1 - \hat{P}_2$$

Hence

$$2\hat{P}_2 = \hat{P}_1\hat{P}_2 + \hat{P}_2\hat{P}_1 .$$

If we multiply this equation on left (right) by $\hat{P}_1$, we find

$$2\hat{P}_1\hat{P}_2 = \hat{P}_1\hat{P}_2 + \hat{P}_1\hat{P}_2\hat{P}_1$$

and

$$2\hat{P}_2\hat{P}_1 = \hat{P}_1\hat{P}_2\hat{P}_1 + \hat{P}_2\hat{P}_1$$

Hence $\hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$ and $2\hat{P}_2 = 2\hat{P}_1\hat{P}_2$ or $\hat{P}_2 = \hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$. If $\hat{P}_1-\hat{P}_2$ is a projector we have $P_1 \supseteq P_2$, since $\chi\in P_2 \Rightarrow \hat{P}_2\chi=\chi$ and $\hat{P}_1(\hat{P}_2\chi)=\hat{P}_1\chi=\hat{P}_1\hat{P}_2\chi=\hat{P}_2\chi=\chi$, which shows that χ is in P_1. Conversely if $P_1 \supseteq P_2$, $(\hat{P}_1-\hat{P}_2)\chi=\chi_1-\chi_2\in P_1$ further $\hat{P}_2(\chi_1-\chi_2)=\hat{P}_2\chi_1-\hat{P}_2\chi_1=\hat{P}_2\chi_1-\chi_2=\chi_2-\chi_2=0$

Whenever $P_1 \supseteq P_2$ or equivalently whenever $\hat{P}_2=\hat{P}_1\hat{P}_2$, we write $\hat{P}_1 \geq \hat{P}_2$, which introduces a <u>partial ordering</u> "$\geq$" into the set of all projectors. That is

(a) $\hat{P} \geq \hat{P}$,

(b) if $\hat{P}_1 \geq \hat{P}_2$ and $\hat{P}_2 \geq \hat{P}_1$, then $\hat{P}_1 = \hat{P}_2$

(c) if $\hat{P}_1 \geq \hat{P}_2$ and $\hat{P}_2 \geq \hat{P}_3$, then $\hat{P}_1 \geq \hat{P}_3$.

We note that $\hat{0} \leq \hat{P} \leq \hat{1}$ for all projectors $\hat{P}$. We can now see the consequences of commutativity for two projectors $\hat{P}_1,\hat{P}_2$, for given $\hat{P}_1\hat{P}_2=\hat{P}_2\hat{P}_1$, then the three quantities $\hat{P}_1-\hat{P}_1\hat{P}_2=\hat{P}_1(\hat{1}-\hat{P}_2)=\hat{P}_1\hat{P}_2^{\perp}$, $\hat{P}_2-\hat{P}_1\hat{P}_2=\hat{P}_2(1-\hat{P}_1)=\hat{P}_2\hat{P}_1^{\perp}$, and $\hat{P}_1\hat{P}_2$ are mutually orthogonal projectors. Hence their sum $\hat{P}_1\hat{P}_2^{\perp}+\hat{P}_2\hat{P}_1^{\perp}+\hat{P}_1\hat{P}_2=\hat{P}_1+\hat{P}_2-\hat{P}_1\hat{P}_2$ is a projector. Further this projector corresponds to the smallest closed subspace $P_1 \vee P_2$ containing P_1 and P_2. We write, therefore,

$$\hat{P}_1 \vee \hat{P}_2 = \hat{P}_1 + \hat{P}_2 - \hat{P}_1\hat{P}_2, \text{ if } \hat{P}_1\hat{P}_2 = \hat{P}_2\hat{P}_1$$

However $P_1 \vee P_2$ is defined for <u>any</u> two closed subspaces P_1,P_2, but $\hat{P}_1 \vee \hat{P}_2$ is given by the above if and only if $\hat{P}_1\hat{P}_2=\hat{P}_2\hat{P}_1$.

In general we may define

$$\hat{P}_1 \wedge \hat{P}_2 = g.l.b\ \{\hat{P};\ \hat{P} \le \hat{P}_1,\ \hat{P} \le \hat{P}_2\}$$

and

$$\hat{P}_1 \vee \hat{P}_2 = l.u.b\ \{\hat{P};\ \hat{P} \ge \hat{P}_1,\ \hat{P} \ge \hat{P}_2\}$$

The existence of $\hat{P}_1 \wedge \hat{P}_2$, and $\hat{P}_1 \vee \hat{P}_2$ for every pair $\hat{P}_1, \hat{P}_2$ of projectors means that the set of all projectors with the partial ordering as defined above forms a lattice, the lattice of all projectors on $\mathcal{H}$ or equivalently the lattice of all subspaces of $\mathcal{H}$. We would like to find the projectors $\hat{P}_1 \wedge \hat{P}_2$ and $\hat{P}_1 \vee \hat{P}_2$ for general projectors $\hat{P}_1, \hat{P}_2$. However, we need only an expression for $\hat{P}_1 \wedge \hat{P}_2$, since we shall show below that

$$\hat{P}_1 \vee \hat{P}_2 = (\hat{P}_1^{\perp} \wedge \hat{P}_2^{\perp})^{\perp}.$$

We now find an expression for $\hat{P}_1 \wedge \hat{P}_2$. Note that if $\chi \in \hat{P}_1 \wedge \hat{P}_2$, then $\hat{P}_1\chi = \chi$ and $\hat{P}_2\chi = \chi$ so that $(\hat{P}_1\hat{P}_2)^N\chi = \chi$. This suggests that we explore the properties of $(\hat{P}_1\hat{P}_2)^N$. Let * $Z_N \equiv (\hat{P}_1\hat{P}_2)^N\chi$ for any $\chi \in \mathcal{H}$. Then

$$\|P_2 Z_N\|^2 = \|(\hat{P}_1 + \hat{1} - \hat{P}_1)\hat{P}_2 Z_N\|^2$$

$$= \|\hat{P}_1\hat{P}_2 Z_N\|^2 + \|(\hat{1} - \hat{P}_1)\hat{P}_2 Z_N\|^2$$

$$= \|Z_{N+1}\|^2 + \|(\hat{1} - \hat{P}_1)\hat{P}_2 Z_N\|^2$$

Similarly

$$\|Z_N\|^2 = \|\hat{P}_2 Z_N\|^2 + \|(\hat{1} - \hat{P}_2)Z_N\|^2 ,$$

so

$$\|Z_{N+1}\|^2 = \|\hat{P}_2 Z_N\|^2 - \|(\hat{1} - \hat{P}_1)\hat{P}_2 Z_N\|^2$$

and

$$\|\hat{P}_2 Z_N\|^2 = \|Z_N\|^2 - \|(1 - \hat{P}_2)Z_N\|^2.$$

---

* I am indebted to Peter Breitenlohner for suggesting this argument, which will show the existence of $\lim_{N\to\infty} \|Z_N\| = \lim_{N\to\infty} \|P_2 Z_N\|$.

Therefore

$$\|Z_N\| \geq \|\hat{P}_2 Z_N\| \geq \|Z_{N+1}\|$$

this shows that $\lim_{N\to\infty} \|Z_N\| = a_o$ exists, and also that $\lim_{N\to\infty} \|P_2 Z_N\|$ exists and is equal to a_o. Similarly $\lim_{N\to\infty} \|(1 - \hat{P}_1)\hat{P}_2 Z_N\| = 0$. To prove that $\lim\text{-}S(\hat{P}_1\hat{P}_2)^N$ exists, we must now establish that $\lim \|Z_M - Z_N\|$, as M and N tend independently to infinity, exists and is zero. This proof is not quite as simple as we might like*, although it is straightforward. It is patterned after the proof of von Neuman[7], who established the existence of $\lim\text{-}S(\hat{P}_1\hat{P}_2\hat{P}_1)^N$ and $\lim\text{-}S(\hat{P}_2\hat{P}_1\hat{P}_2)^N$, and showed that $\hat{P}_1 \wedge \hat{P}_2 = \lim_{N\to\infty}\text{-}S(\hat{P}_1\hat{P}_2\hat{P}_2)^N = \lim_{N\to\infty}\text{-}S(\hat{P}_2\hat{P}_1\hat{P}_2)^N$. We have

$$\|Z_M - Z_N\| = (Z_M - Z_N, Z_M - Z_N) = (Z_M, Z_M) + (Z_N, Z_N) - (Z_M, Z_N) - (Z_N, Z_M),$$

so if we can show that each of the four scalar products has a common limit, the existence of $\lim_{N\to\infty} Z_N$ is established. Consider

$$g_{MN} \equiv (Z_M, Z_N) = ((\hat{P}_1\hat{P}_2)^M \chi, (\hat{P}_1\hat{P}_2)^N \chi)$$

$$\equiv ((\hat{P}_2\hat{P}_1)^N (\hat{P}_1\hat{P}_2)^M \chi, \chi) = (\hat{P}_2(\hat{P}_1\hat{P}_2)^{M+N-1}\chi, \chi)$$

$$\equiv g_{M+N-1}$$

So

$$\|Z_M - Z_N\| = g_{2M} + g_{2N} - 2g_{M+N}.$$

If we can now show that $\lim_{K\to\infty} g_K$ exists, the $\lim\text{-}S(\hat{P}_1\hat{P}_2)^N$ exists. The quantity $g_K \equiv (\hat{P}_2(\hat{P}_1\hat{P}_2)^K\chi, \chi) = ((\hat{P}_2\hat{P}_1\hat{P}_2)^K\chi, \chi)$ is real and non negative. This follows from the fact that $\hat{P}_2\hat{P}_1\hat{P}_2$ is a non-negative self-adjoint operator. Hence

---

* Prof. B. Misra constructed a proof based on the spectral resolutions of the self-adjoint operators $\hat{P}_2\hat{P}_1\hat{P}_2$ and $\hat{P}_1\hat{P}_2\hat{P}_1$. However, I prefer a more elementary proof which I give here.

$(\hat{P}_2\hat{P}_1\hat{P}_2)^{\frac{1}{2}}$ exists, and

$$g_K = ((\hat{P}_2\hat{P}_1\hat{P}_2)^{K/2}\chi,(\hat{P}_2\hat{P}_1\hat{P}_2)^{K/2}\chi) = \|(\hat{P}_2\hat{P}_1\hat{P}_2)^{K/2}\chi\|^2 \geq 0$$

Now $g_{K+1} = ((\hat{P}_2\hat{P}_1\hat{P}_2)y,y)$ where $y \equiv (\hat{P}_2\hat{P}_1\hat{P}_2)^{K/2}\chi$,

so

$$g_{K+1} = (\hat{P}_1\hat{P}_2y,P_2y) = (\hat{P}_1\hat{P}_2y,\hat{P}_1\hat{P}_2y),$$

or

$$g_{K+1} = \|(\hat{P}_1\hat{P}_2)y\|^2 \leq \|\hat{P}_2y\|^2 \leq \|y\|^2 = g_K$$

so g_K is a non-increasing sequence of non-negative real numbers, and hence the sequence has a limit, $\lim_{K\to\infty} g_K = g_o$. Therefore $\|Z_M-Z_N\| \to g_o + g_o - 2g_o = 0$. It is now simple to establish that the limiting vector $Z_o = \lim_{N\to\infty} (\hat{P}_1\hat{P}_2)^N\chi$ has the properties $\hat{P}_1Z_o = Z_o$, $\hat{P}_2Z_o = Z_o$,,and that Z_o is obtained from χ by a projector, and that this projector is $\hat{P}_1\wedge\hat{P}_2$. Thus

$$\hat{P}_1 \wedge \hat{P}_2 = \lim_{N\to\infty}\text{-}S(\hat{P}_1\hat{P}_2)^N$$

$$= \lim_{N\to\infty}\text{-}S(\hat{P}_1\hat{P}_2\hat{P}_1)^N$$

$$= \lim_{N\to\infty}\text{-}S(\hat{P}_2\hat{P}_1\hat{P}_2)^N.$$

By the subspace $P_1 \vee P_2$ is meant, the closed subspace obtained by forming all linear combination of $\chi_1\in P_1$ and $\chi_2\in P_2$ together with limits of sequences of such linear combinations. Any vector y in $(P_1 \vee P_2)^\perp$ has the property that

$$(y,Z_1) = 0, \qquad Z_1\in P_1$$

$$(y,Z_2) = 0, \qquad Z_2\in P_2$$

that is $y\in P_1^\perp$ and $y\in P_2^\perp$, hence $y\in P_1^\perp \wedge P_2^\perp$. Therefore

$$P_1 \vee P_2 = \mathcal{H} - (P_1^\perp \wedge P_2^\perp),$$

and

$$\hat{P}_1 \vee \hat{P}_2 = (\hat{P}_1^\perp \wedge \hat{P}_2^\perp)^\perp,$$

or

$$\hat{P}_1 \vee \hat{P}_2 = \hat{I} - (\hat{I}-\hat{P}_1) \wedge (\hat{I}-\hat{P}_2).$$

If $[\hat{P}_1,\hat{P}_2] = 0$,

$$\hat{P}_1 \vee \hat{P}_2 = \hat{1} - (\hat{1}-\hat{P}_1)(\hat{1}-\hat{P}_2)$$

$$= \hat{P}_1 + \hat{P}_2 - \hat{P}_1\hat{P}_2, \text{ as before.}$$

The above result may be generalized to

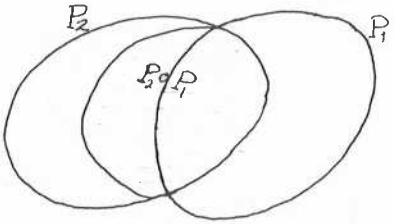
$$\bigvee_\alpha \hat{P}_\alpha = (\bigwedge_\alpha \hat{P}_\alpha^{\perp})^{\perp}$$

for any collection $\{\hat{P}_\alpha\}$ of projectors.

We may apply a projector $\hat{P}_2$ to each vector in the subspace $P_1 = \hat{P}_1\mathcal{H}$:

$$\hat{P}_2\hat{P}_1\mathcal{H} = \hat{P}_2P_1$$

The result is in general not a closed subspace. However, we may close it $\overline{\hat{P}_2\hat{P}_1\mathcal{H}} \equiv P_2 \circ P_1$ and ask what is the projector $\hat{P}_2 \circ \hat{P}_1$ for this subspace. In order to find $P_2 \circ \hat{P}_1$ we consider the symbolic diagram:



Let χ be in the complement of $P_2 \circ P_1$ in P_2. Then

$$(\chi,\hat{P}_2\hat{P}_1Z) = 0 \qquad \forall Z\in\mathcal{H}$$

or

$$(\hat{P}_1\hat{P}_2\chi,Z) = (\hat{P}_1\chi,Z) \qquad \forall Z\in\mathcal{H}.$$

Hence

$$\hat{P}_1\chi = 0, \text{ so}$$

$\chi\in P_1^{\perp}$ and $\chi\in P_1^{\perp} \wedge P_2$. Thus the complement of $P_2 \circ P_1$ in P_2 is $P_1^{\perp} \wedge P_2$.

Therefore

$$P_2 \circ P_1 = P_2 - P_2 \wedge P_1^{\perp}, \text{ or}$$

$$\hat{P}_2 \circ \hat{P}_1 = \hat{P}_2 - \hat{P}_2 \wedge \hat{P}_1^{\perp}$$

Support: Closed Range (cf. Dixmier[8])*

Consider now a bounded operator $\hat{A}$. We define the null space $N(\hat{A})$ of $\hat{A}$ by

$$N(\hat{A}) = \{\chi, \chi \in \mathcal{H}, \hat{A}\chi = 0\}$$

The support of $\hat{A}$ is defined to be the orthogonal complement of $N(\hat{A})$

$$S(\hat{A}) = N(\hat{A})^{\perp}$$

with projector $\hat{S}(\hat{A})$. Clearly

$$\hat{A} = \hat{A}\,\hat{S}(\hat{A})$$

Further

$$\hat{S}(\hat{A}) = \inf \{\hat{S}; (\hat{S}\hat{S}^* = \hat{S}), \hat{A} = \hat{A}\hat{S}\}.$$

$\hat{S}(\hat{A})$ is therefore called the right projector of $\hat{A}$. $\hat{S}(\hat{A}) \equiv RP(\hat{A})$ If $\chi \neq 0$, $\chi \in S(\hat{A})$ then $\hat{A}\chi \neq 0$ for otherwise $\hat{A}\chi = 0 \Rightarrow \chi \in N(\hat{A})$ which is orthogonal to $S(\hat{A})$.

We define $\hat{A}\mathcal{H} \equiv R(\hat{A})$, the closed range of $\hat{A}$ with projector $\hat{R}(\hat{A})$

e.g. $$\hat{R}(\hat{A})\mathcal{H} = R(\hat{A}) = \hat{A}\mathcal{H}$$

clearly $$\hat{R}(\hat{A})\hat{A} = \hat{A}$$

and in fact

$$\hat{R}(\hat{A}) = \inf \{\hat{R}, \hat{R}\hat{R}^* = \hat{R}, \hat{R}\hat{A} = \hat{A}\}$$

So $\hat{R}(\hat{A}) \equiv LP(\hat{A})$, the left projector of $\hat{A}$.

Theorem: $R(\hat{A}) = S(\hat{A}^*)$, or in terms of projectors

* The remainder of material presented in this appendix follows Dixmier[8] very closely. It is included because we need the results to more firmly establish our own.

$$\hat{R}(\hat{A}) = \hat{S}(\hat{A}^*)$$

Proof:

$$\hat{A} = \hat{R}(\hat{A})\hat{A}\hat{S}(\hat{A})$$

Therefore

$$\hat{A}^* = \hat{S}(\hat{A})\hat{A}^*\hat{R}(\hat{A})$$

and hence

$$\hat{S}(A) \geq \hat{R}(\hat{A}^*),$$

and

$$\hat{R}(\hat{A}) \geq \hat{S}(\hat{A}^*)$$

Therefore

$$\hat{R}(\hat{A}) \geq \hat{S}(A^*) \geq \hat{R}(A),$$

and

$$\hat{S}(\hat{A}^*) = \hat{R}(\hat{A})$$

Partially Isometric Operators (Cf Dixmier [8])

Let $\hat{u}$ be a bounded operator on $\mathcal{H}$, $\hat{S}(\hat{u})$ its support, $\hat{u}$ is said to be partially isometric if $\hat{u}$ is isometric on $S(\hat{u}) = \hat{S}(\hat{u})\mathcal{H}$. Then $R(\hat{u}) = \hat{u}\mathcal{H} = \hat{u}S(\hat{u})$. The range of $\hat{u}$ is a closed subspace of $\mathcal{H}$ and $\hat{u}$ maps $S(\hat{u})$ isometrically on $R(\hat{u})$.

$\hat{S}(\hat{u})$ is called the initial projector of $\hat{u}$ and $S(\hat{u})$ the initial subspace of $\hat{u}$.

$\hat{R}(\hat{u})$ is called the final projector of $\hat{u}$ and $R(\hat{u})$ the final subspace of $\hat{u}$. Let $\chi \in S(\hat{u})$, then $y = \hat{u}\chi \in R(\hat{u})$ and for all $Z \in \mathcal{H}$, we have

$$(\chi, Z) = (\hat{S}(\hat{u})\chi, Z) = (\chi, \hat{S}(\hat{u})Z),$$

and since both χ, and $\hat{S}(\hat{u})Z$ are in $S(\hat{u})$,

$$(\chi, Z) = (\hat{u}\chi, \hat{u}\hat{S}(\hat{u})Z) = (\hat{u}\chi, \hat{u}Z)$$

$$= (y, \hat{u}Z) = (\hat{u}^*y, Z) \qquad \forall Z \in \mathcal{H}$$

$$\therefore \chi = \hat{u}^*y$$

Now from $(\chi, Z) = (\hat{u}^*\hat{u}\chi, Z)$

follows $\hat{u}^*\hat{u}\chi = \chi \qquad \forall \chi \in S(\hat{u})$

and therefore since $\hat{u}y = 0$ implies $y \in S(\hat{u})^{\perp}$,

$$\hat{u}^*\hat{u} = \hat{S}(\hat{u})$$

Now $S(\hat{u}^*) = R(\hat{u})$, $R(\hat{u}^*) = S(\hat{u})$, and if $y \in R(\hat{u}) = S(\hat{u}^*)$,

$$\hat{u}^*y = \chi, \ \hat{u} \text{ is } \underline{\text{isometric}}:$$

$$y \in R(\hat{u}) = S(\hat{u}^*) \text{ since } y = \hat{u}\chi \text{ is isometric.}$$

Again

$$(y,Z) = (\hat{R}(\hat{u})y,Z) = (y,\hat{R}(\hat{u})Z)$$

$$= (\hat{u}^*y,\hat{u}^*\hat{R}(\hat{u})Z) = (\hat{u}^*y,\hat{u}^*Z) \quad (\hat{u}^*S(\hat{u}^*) = \hat{u}^*)$$

$$= (\hat{u}\hat{u}^*y,Z) \qquad \forall Z$$

$$uu^*y = y \qquad \forall y \in R(\hat{u}),$$

Hence $\hat{u}\hat{u}^* = \hat{R}(\hat{u})$.

Conversely:

Suppose $\hat{W}^*\hat{W} = \hat{S}$ is a <u>projector</u>

$$\|\hat{W}\chi\|^2 = (\hat{W}\chi,\hat{W}\chi) = (\hat{W}^*\hat{W}\chi,\chi)$$

$$= (\hat{S}\chi,\chi) = (\hat{S}\chi,\hat{S}\chi) = \|\hat{S}\chi\|^2$$

Thus $\hat{W}$ is isometric on $S = \hat{S}\mathcal{H}$ and zero <u>elsewhere</u>
Further

$$\hat{W}\hat{W}^* = \hat{R}(\hat{W}).$$

If $\hat{V}\hat{V}^* = \hat{R}$ is a projector then $\hat{V}^*$ is partially isometric on $R(\hat{V}) = S(\hat{V}^*)$ and $\hat{V}$ is partially isometric on $\hat{S}(\hat{V})$.

<u>Polar Decomposition</u>:

Let A be a bounded operator on $\mathcal{H}$, and $\hat{S}(\hat{A})$ be the projector on $S(\hat{A})$, and $\hat{R}(\hat{A})$ be the projector on $R(\hat{A})$. Let us define the operator $|\hat{A}| \equiv (\hat{A}^*\hat{A})^{\frac{1}{2}}$. Then for $\forall Z \in \mathcal{H}$

$$\|\hat{A}Z\|^2 = (\hat{A}Z,\hat{A}Z) = (\hat{A}^*\hat{A}Z,Z)$$

$$= (|\hat{A}|^2Z,Z) = (|\hat{A}|Z,|\hat{A}|Z) = \| \ |\hat{A}|Z \ \|^2$$

Hence $S(\hat{A}) = S(|\hat{A}|)$ and $R(|\hat{A}|) = S(|\hat{A}|)$, since $|\hat{A}|^* = |\hat{A}|$. Further the correspondence

$$|\hat{A}|\chi \to \hat{A}\chi$$

is a linear and isometric mapping $\hat{W}'$ from S(A) to R(A):

$$\hat{A}\chi = \hat{W}'|\hat{A}|\chi$$

We extend it by closure to all $R(\hat{A})$:

$$\hat{A} = \hat{W}|\hat{A}|.$$

$$\hat{S}(\hat{W}) = S(\hat{A}),$$

$$\hat{R}(\hat{W}) = \hat{R}(\hat{A}).$$

The expression $\hat{A} = \hat{W}|\hat{A}|$ is called polar decomposition of $\hat{A}$. We state the important results:

$$\hat{W}^*\hat{W} = \hat{S}(\hat{A})$$

$$\hat{W}\hat{W}^* = \hat{R}(\hat{A})$$

We now adopt Dixmier's notation to fit our needs. $\hat{P}_M$, $\hat{A}$ are projectors with subspaces P_M and A.
We have seen that

$$\hat{R}(\hat{A}\hat{P}_M) = \hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge \hat{P}_M^{\perp}$$

$$= \hat{S}((\hat{A}\hat{P}_M)^*) = \hat{S}(\hat{P}_M\hat{A})$$

$$\hat{R}(\hat{A}\hat{P}_M) = \hat{S}(\hat{P}_M\hat{A}) = \hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge \hat{P}_M^{\perp}$$

$$\hat{R}(\hat{P}_M\hat{A}) = \hat{S}(\hat{A}\hat{P}_M) = \hat{P}_M \circ \hat{A} = \hat{P}_M - \hat{P}_M \wedge \hat{A}^{\perp}$$

We know therefore that there exists an isometric mapping $\hat{W}_M$ from $S(P_M\hat{A}) = \overline{\hat{A}P_M}$ to $R(\hat{P}_M\hat{A}) = \overline{\hat{P}_M A}$

$$\hat{W}_M : \hat{A}P_M \to \hat{P}_M A$$

$$\hat{W}_M : \overline{\hat{A}P_M} \ni \psi_M \to \xi_M = \hat{W}_M\psi_M \in \overline{\hat{P}_M A}$$

$$\hat{W}_M^*\hat{W}_M = \hat{S}(\hat{P}_M\hat{A}) = \hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge \hat{P}_M^{\perp}$$

Further

$$\hat{P}_M\hat{A} = \hat{W}|\hat{P}_M\hat{A}| = |\hat{A}\hat{P}_M|\hat{W}$$

$$= (\hat{P}_M\hat{A}\hat{P}_M)^{\frac{1}{2}}\hat{W}$$

$$\hat{W}_M\hat{W}_M^* = \hat{R}(\hat{P}_M\hat{A}) = \hat{P}_M \circ \hat{A} = \hat{P}_M - \hat{P}_M \wedge \hat{A}^{\perp}$$

Thus $\psi_M \in \hat{A}P_M$ is mapped into

$$\xi_M = \hat{W}_M\psi_M, \qquad \xi_M \in \overline{\hat{P}_M A} = P_M \circ A$$

and

$$\hat{W}_M^*\xi_M = \hat{W}_M^*\hat{W}_M\psi_M = \hat{S}(\hat{P}_M \circ \hat{A})\psi_M$$

$$= \hat{R}(\hat{A}\hat{P}_M)\psi_M = (\hat{A} \circ \hat{P}_M)\psi_M = \psi_M$$

Similarly

$$\xi_M \in \overline{\hat{P}_M A} = P_M \circ A$$

is mapped into

$$\psi_M = W_M^*\xi_M,$$

and

$$\hat{W}_M\psi_M = \hat{W}_M\hat{W}_M^*\xi_M$$

$$= \hat{R}(\hat{P}_M\hat{A})\xi_M$$

$$= (\hat{P}_M \circ \hat{A})\xi_M = \xi_M$$

APPENDIX B

Additional Comments on R_M

Let us introduce the unitary permutation operator $\hat{u}(P,Q)$ defined by

$$(\hat{u}_{(P,Q)}\psi)(x_1 \cdots x_N; y_1 \cdots y_N) = \psi(x_{\alpha_1} \cdots x_{\alpha_N}; y_{\beta_1} \cdots y_{\beta_N})$$

where P is the permutation $(\alpha_1 \ldots \alpha_N)$ of $(1 \ldots N)$ and Q is the permutation $(\beta_1 \ldots \beta_N)$ of $(1 \ldots N)$. With this definition

$$\hat{A} = \frac{1}{N!^2} \sum_P \sum_Q \epsilon_P \epsilon_Q \hat{u}_{(P,Q)}$$

where the sums extend over all possible permutations P,Q, and ϵ_P, ϵ_Q are the signatures of P and Q respectively. We now write $\hat{P}_M$ in the form

$$\hat{P}_M = \hat{P}(1) \otimes \hat{P}(2) \otimes \ldots \hat{P}(M) \otimes \hat{I}(M+1) \ldots \otimes \hat{I}(N)$$

$$\hat{P}_M = \hat{P}(1,1) \otimes \hat{P}(2,2) \otimes \hat{P}(3,3) \ldots \otimes \hat{P}(M,M)$$

(supressing the unit operators on the remaining N-M electron and proton variables) in which the first number K of K,K in $\hat{P}(K,K)$ refers to proton coordinates and the second to electron coordinates. With this notation

$$\hat{R}_M = \hat{P}_M \otimes \hat{P}(M+1,M+1)^{\perp} = \hat{P}_M \hat{P}(M+1,M+1)^{\perp}$$

where $\hat{P}(M+1,M+1)$ etc. includes $\hat{1} \times \hat{1} \times \ldots \times \hat{1} \times \hat{1}$ except at the (M+1)st position.

A first objection to the use of $\psi^M_{\alpha ij}$ might be that, since the electrons and protons are identical and $\psi^M_{\alpha ij}$ assigns the first M electrons and first M proton to the bound atoms, $\psi^M_{\alpha ij}$ could not be a good basis. This objection is not valid however, for suppose we use a new basis in which a permuted order for the electron and proton is used:

$$\psi^M_{(P,Q)\alpha ij} \equiv \hat{u}_{(P,Q)} \psi^M_{\alpha ij} .$$

Then for a physical state ψ, $\hat{u}_{(P,Q)}\psi = \xi_P \xi_Q \psi$, $u^*_{(P,Q)}\psi = \epsilon_P \epsilon_Q \psi$,

$$\hat{P}^M_{(P,Q)} \equiv \sum_{\alpha ij} |\psi^M_{(PQ)\alpha ij}\rangle\langle\psi^M_{(P,Q)\alpha ij}| = \hat{u}_{(P,Q)} \hat{P}^M \hat{u}^*_{(P,Q)}$$

$$\hat{P}^{(M+1)}_{(P,Q)} \equiv \hat{u}_{(P,Q)} \hat{P}^{M+1} \hat{u}^*_{(P,Q)}, \text{ and } \hat{R}^M_{(P,Q)} = \hat{u}_{(P,Q)} \hat{R}^M \hat{u}^*_{(P,Q)}$$

In the new basis $\psi = \sum_M \psi^M_{(P,Q)}$, where

$$\psi^M_{(P,Q)} = \hat{u}_{(P,Q)}\hat{R}^M\hat{u}^*_{(PQ)}\psi = \epsilon_P\epsilon_Q\hat{u}_{(PQ)}\hat{R}^M\psi = \epsilon_P\epsilon_Q\hat{u}_{(P,Q)}\psi_M$$

thus the coefficients $C^M_{(P,Q)\alpha ij}$ in the new basis have the form

$$\begin{aligned} C^M_{(P,Q)\alpha ij} &= (\psi^M_{(PQ)\alpha ij}, \psi^M_{(PQ)}) \\ &= (\hat{u}_{(p,Q)}\psi^M_{\alpha ij}, \epsilon_P\epsilon_Q\hat{u}_{(PQ)}\psi_M) \\ &= \epsilon_P\epsilon_Q(\hat{u}_{(P,Q)}\psi^M_{\alpha ij}\hat{u}_{(P,Q)}\psi_M) \\ &= \epsilon_P\epsilon_Q(\psi^M_{\alpha ij}, \psi_M) = \epsilon_P\epsilon_Q C^M_{\alpha ij}. \end{aligned}$$

Therefore for a physical state the new coefficients differ from the old by a single overall factor $\epsilon_P\epsilon_Q \pm 1$, so the description in terms of $C^M_{\alpha ij}$ for the orthogonal subspace R_M is essentially independent of which electrons or protons we label $1,\ldots,N$.

A further objection might be that although we may expand a physical state ψ into orthogonal (non physical) states ψ_M

$$\psi = \sum_{M=0}^N \psi_M \qquad \psi_M = \hat{R}_M\psi$$

and although

$$\psi = \hat{A}\psi = \sum_M^N \hat{A}\psi_M$$

the $\hat{A}\psi_M$ are not orthogonal and therefore do not correspond (directly) to physical M-atom states.

This disadvantage can be further elucidated by considering that

$$\hat{R}_M < \hat{P}(M,M)$$

and $\hat{R}_M A$ cannot be considered directly as the physical M-atom subspace since if we examine $\hat{R}_M \wedge \hat{A}$ we note that

$$\hat{P}(M,M)\hat{R}_M \wedge \hat{A} = \hat{R}_M \wedge \hat{A}.$$

Further since

$$\hat{u}_{(P,Q)}\hat{P}(M,M) = \hat{P}(PM,QM)\hat{u}_{(P,Q)}$$

and A is invariant under $\hat{u}(P,Q)$, we have

$$\hat{P}(PM,QM)\hat{R}_M \wedge \hat{A} = \hat{R}_M \wedge \hat{A}$$

But we know that $\hat{P}(M+1,M+1)R_M = 0$, and hence

$$\hat{R}_M \wedge \hat{A} = 0,$$

so R_M does not contain any physical states.

This result is not as disturbing as it may seem. We are using a decomposition of a physical states ψ into non-physical components $R_M\psi$, which, although strange in appearance, is not incorrect. It becomes somewhat a matter of taste whether or not to use the simple $R_M\psi$'s with $C^M_{\alpha ij}$'s not having the symmetry we would like.

Finally we mention that not all $C^M_{\alpha ij}$ having the correct symmetry* represent physical states. Only those satisfying a subsidiary condition can be states. The subsidiary condition is determined from

$$C^M_{\alpha ij} = (\chi^M_{\alpha ij},\psi) = (\chi^M_{\alpha ij},\hat{A}\psi) = (\hat{A}\chi^M_{\alpha ij},\psi).$$

However we may expand $\hat{A}\chi^M_{\alpha ij} = \Sigma_{M',\alpha',i',j'}\chi^{M'}_{\alpha' i' j'}(\chi^{M'}_{\alpha' i' j'},\hat{A}\chi^M_{\alpha ij})$ to obtain

$$C^M_{\alpha ij} = \Sigma_{M',\alpha',i',j'}(\chi^M_{\alpha ij},A\chi^{M'}_{\alpha' i' j'})C^{M'}_{\alpha' i' j'}$$

* Correct symmetry here means that $C^M_{\alpha_1\cdots\alpha_M i_{M+1}\cdots i_N, j_{M+1}\cdots j_N}$ are completely symmetric in $\alpha_1\cdots\alpha_M$, completely antisymmetric in $i_{M+2},\cdots i_N$, and completely antisymmetric in $j_{M+2},\cdots j_N$.

In the above result $\chi^M_{\alpha ij}$ are just those defined by Eq. 31:

$$\chi^M_{\alpha ij} = \hat{R}_M \psi^M_{\alpha ij} = \hat{P}_M \psi^M_{\alpha ij} - \hat{P}_{M+1} \psi^M_{\alpha ij} = \psi^M_{\alpha ij} - \hat{P}_{M+1} \psi^M_{\alpha ij}$$

$$= \varphi_{\alpha_1}(x_1 y_1) .. \varphi_{\alpha_M}(x_M y_M) \varphi^F_{i_{M+1} j_{M+1}} \varphi_{i_{M+2}} .. \varphi_{i_N} \psi_{j_{M+2}} .. \psi_{j_N}(y_N).$$

APPENDIX C

Proof that

$$\hat{A}_M = (\hat{A} \circ \hat{P}_M) \wedge P_{M+1}^{\perp}$$

We start with expression 3.12 for $\hat{A}_M$

$$\hat{A}_M = \hat{A} \wedge \hat{P}_{M+1}^{\perp} - \hat{A} \wedge \hat{P}_M^{\perp}$$

$$= \hat{A} \wedge \hat{P}_{M+1}^{\perp} (\hat{A} \wedge \hat{P}_M^{\perp})^{\perp},$$

since $\hat{A} \wedge \hat{P}_M^{\perp} < \hat{A} \wedge \hat{P}_{M+1}^{\perp}$

$$= \hat{P}_{M+1}^{\perp} \wedge \hat{A} \wedge (\hat{A} \wedge \hat{P}_M^{\perp})^{\perp}$$

since $(\hat{A} \wedge \hat{B}) \wedge \hat{C} = \hat{A} \wedge (\hat{B} \wedge \hat{C})$

$$= \hat{P}_{M+1}^{\perp} \wedge \{\hat{A} \wedge \hat{A} \circ \hat{P}_M\}.$$

But

$$\hat{A} \circ \hat{P}_M = \hat{A} - \hat{A} \wedge \hat{P}_M^{\perp} = \hat{A} \wedge (\hat{A} \wedge P_M^{\perp})^{\perp} \text{ and}$$

$\hat{A} \wedge (\hat{A} \circ \hat{P}_M) = \hat{A} \circ \hat{P}_M$ since $\hat{A} \circ \hat{P}_M < \hat{A}$, and therefore

$$\hat{A}_M = \hat{P}_{M+1}^{\perp} \wedge (\hat{A} \circ \hat{P}_M) = (\hat{A} \circ \hat{P}_M) \wedge \hat{P}_{M+1}^{\perp},$$

(or)

$$\hat{A}_M = \{\hat{A} - \hat{A} \wedge \hat{P}_M^{\perp}\} \wedge \hat{P}_{M+1}^{\perp},$$

(or)

$$\hat{A}_M = \{\hat{A} - \hat{A} \wedge (\hat{1} - \hat{P}_M)\} \wedge (\hat{1} - \hat{P}_{M+1}.$$

REFERENCES

1. Girardeau, M., J. Math. Phys. 4 1096 (1963).

2. Stolt, R. H. and Brittin, W. E., Phys. Rev. Lett. 27, 616 (1971).

3. Sakakura, A. Y., Phys. Rev. Lett. 27, 822 (1971).

4. Girardeau, M., Phys. Rev. Lett. 27, 1416 (1971).

5. Sakakura, A. Y., This volume, 483.

6. Girardeau, M., This volume, 147.

7. von Neumann, J., Functional Operators, Vol. II pp 51-59, Princeton Univ. Press, Princeton, New Jersey (1950).

8. Dixmier, J., Les Algèbres d'operateurs dans l'espace Hilbertien (Algèbres de von Neumann) Gauthier-Villars, Paris (1969) 2nd Ed.

SOME HEURISTIC REMARKS ON ALGEBRAIC TECHNIQUES IN STATISTICAL PHYSICS[1]

Joseph Dreitlein*
University of Colorado
Boulder, Colorado

Introduction

The introduction of sophisticated algebraic techniques into statistical physics[2] has divided physicists into two non-communicating groups - those who categorically denounce them and those who view the techniques as a door opening to a view which could lead to the formulation and even solution of some of the outstanding problems of statistical physics. The barrier which separates the two groups of physicists is the language, arcane to the average physicist and Shakespearian to the devotee. The main purpose of these lectures is to attempt to bridge the gap between the two groups of physicists by wording familiar situations in both the secular and the professional mathematical language.

Physicists have been faced more than once before with mathematical refinements. Even Einstein reacted somewhat adversely[3] to the mathematical reformulation of special relativity with the statements "Since the mathematicians have attacked the relativity theory, I myself no longer understand it any more" and "The people in Göttingen sometimes strike me, not as if they wanted to help formulate something clearly, but as if they wanted only to show us physicists how much brighter they are than we". Nevertheless, Einstein went on to develop the general theory of relativity using the highly mathematical theory which sprang from Minkowski's formulation.

Briefly stated, the algebraic approach is designed to handle the peculiarities which arise when a system is considered infinite in the sense of having an infinite number of particles in an infinite volume but with finite density. The peculiarities encountered include non-equivalent

*This work was supported by the National Science Foundation under Grant number NSF GP-19479

representations of the commutation relations and, most spectacular of all, a breakdown of the symmetry[4] which the Hamiltonian of the finite system possesses. This symmetry breakdown is the tocsin of a phase transition.

Two issues cognate to the subject are worth noting. The first is that not all statistical phenomena are explained by taking the thermodynamic limit ($N\to\infty$, $V\to\infty$, $\frac{N}{V}\to\rho$) for which the algebraic method is specifically designed. For example, thermal conductivity is volume dependent for highly purified solids, such as solid He^4, at low temperatures. There are also conjectures[5] that near the critical point of fluids, certain observables may depend upon the number of particles N in the system.

The second observation is that the algebraic techniques used in statistical physics are also appropriate to the quantized relativistic field problem. Indeed, the only exact analysis to date of such fields, the Φ_2^4 superrenormalizable model, uses heavily the algebraic and analytic apparatus of the genre discussed here. The status of such investigations is nicely reviewed by Jaffe[6]. Questions such as uniqueness of the vacuum (called cyclic state below) and symmetry breaking are shared by both relativistic and non-relativistic theories.

Compared to the problems of formulating and understanding quantum electrodynamics, the non-relativistic many body problems appear almost insignificant. Yet a hope is that a thorough understanding of the latter case will serve as a springboard for taking off into relativistic domains. For some, this may be reason enough for formulating statistical physics algebraically.

I. Algebraic Description of Discrete Finite Systems

For quantum systems described by a finite dimensional Hilbert space (C^N), the measurement of a dynamical variable associated with the self-adjoint operator A involves repeated measurements on states concocted each in an identical manner. The measurements yield up a number $\langle A\rangle$ which can be theoretically calculated from the expression

$$\langle A\rangle = \mathrm{tr}\ \rho\ A$$

The density matrix ρ is hermitian, of trace unity and non-negative. It describes a mixed state unless $tr\rho = tr\rho^2$. In none of the considerations in section I need ρ be considered to describe an equilibrium situation.

It is useful to redescribe this simple picture in a more formidable algeraic language with the goal of formulating a scheme which carries over to description of systems with an arbitrarily large number of particles (idealized to systems with infinite number of particles). With an appropriate algebraic language, one can hope at least to state clearly such important physical questions as what is meant by multiple phases in the thermodynamic limit, by symmetry breaking and by the development of irreversibility from microscopically reversible finite systems.

1. The Space of Observables and States

Certainly in the set of observables $\mathfrak{A}$ all hermitian operators should be included. Furthermore, if it be realized that any operator on C^N can be written as the linear combination of two hermitian operators (viz, $A = (\frac{A+A^+}{2}) + i(\frac{A-A^+}{2i})$), there is little point in not extending the set of operators under consideration to include all N x N matrices. The restriction of the algebra of observables to purely hermitian operators has, however, been investigated by Jordan, von Neumann and Wigner[7]. The question of observables becomes more acute for a spatially infinite system because of the finiteness of measuring instruments.

The set of all N x N complex matrices forms the complete matrix algebra $\mathfrak{A}$ on C^N. It is an N^2 dimensional vector space and closed under matrix multiplication. This means, if $A\in\mathfrak{A}$ and $B\in\mathfrak{A}$ then

$$(\lambda_1 A + \lambda_2 B)\in\mathfrak{A} \quad \text{(vector space property)}$$
$$AB\in\mathfrak{A} \tag{1}$$

where λ_1 and λ_2 are complex numbers. The algebra possesses an involution-hermitian conjugation- and can be normed by setting

$$\|A\| = \sup_{\psi \in C^N} \frac{\|A\psi\|}{\|\psi\|} \tag{2}$$

where $\|\psi\|$ is the norm of the vector ψ in C^N
The norm has the defining properties

$$\begin{aligned}
&(N.1) \quad \|A\| \geq 0; \quad \|A\| = 0 \quad \text{iff } A = 0 \qquad (3)\\
&(N.2) \quad \|A+B\| \leq \|A\| + \|B\|\\
&(N.3) \quad \|\lambda A\| = |\lambda| \, \|A\| \qquad \lambda \in C\\
&(N.4) \quad \|AB\| \leq \|A\| \, \|B\|\\
&(N.5) \quad \|I\| = 1 \qquad (I=\text{identity})\\
&(N.6) \quad \|A^+\| = \|A\|
\end{aligned}$$

The normed algebra with these properties is called a C*-algebra. It is closed with respect to this norm. Further mathematical details can be found in reference 8.

A state on the algebra determines the results of any physical measurement on an ensemble prepared in a way described by the state. It is an assignment of a complex number (real for hermitian operators) $\omega(A)$ to every element $A \in \mathfrak{A}$. These numbers should agree with the rules of interpreting quantum theory and, of course, describe the experiment. To be a state, the functional ω must be linear (superposition principle) and non-negative for positive observables. Require therefore

$$\begin{aligned}
&(S.1) \quad \omega(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \omega(A_1) + \lambda_2 \omega(A_2) \qquad (4)\\
&(S.2) \quad \omega(A^+A) \geq 0\\
&(S.3) \quad \omega(I) = 1
\end{aligned}$$

The last requirement (S.3) means that if the identity I be resolved into mutually orthogonal projectors (physical alternatives); $I = \Sigma \rho_n$, $\rho_n \rho_m = \delta_{nm} \rho_n$, then one of the mutually exclusive possibilities associated with these projectors is sure to occur.

What are the possible states and how can they be classified? To answer this question, a geometric picture of state space $\mathcal{S}$ may be heuristic. The set of all (bounded) linear functionals on $\mathcal{U}$ itself forms a space, here of dimension N^2 over the complex numbers. This follows because the combination $\lambda_1 f_1 + \lambda_2 f_2$ of two linear functionals is itself a linear functional because it satisfies S.1. In this so called space dual to $\mathcal{U}$, the set of non-negative functionals form a convex subset which is geometrically a cone. A set $\mathcal{S}$ is convex if $f_1 f_2 \in \mathcal{S}$ implies that $\lambda_1 f_1 + (1-\lambda) f_2 \in \mathcal{S}$ for λ real and between 0 and 1. The set of all states satisfying S.1, S.2 and S.3 is a cross section of this cone and itself a convex set.

For the finite dimensional case under consideration, the linear functionals representing states have the representation

$$\omega(A) = \mathrm{tr}(\rho A) \tag{5}$$

where the density matrix ρ is any positive hermitian matrix normalized to $\mathrm{tr}\ \rho = 1$ (Hermiticity of ρ follows automatically from linearity and non-negativity of ω). States realized with the density matrix construction are called normal states. Every state of a finite discrete system being considered in this chapter is normal.

2. Symmetry and the Classification of States

The set of all physical states $\mathcal{S}$ ordinarily does not come into consideration when studying a physical system. The reason is that the admissible states are restricted by the prearranged experimental conditions. For example, the system studied may be in equilibrium or it may be spatially homogeneous. Both these conditions are brought about by allowing the system to interact with itself and the environment for a sufficiently long period of time. In addition then to specifying the observables and the possible states, physical systems have imposed upon them various symmetry properties.

To describe a symmetry of a system, the automorphisms of the algebra of observables $\mathfrak{A}$ are used. An automorphism of $\mathfrak{A}$ is a mapping of the elements of $\mathfrak{A}$ into themselves

$$\alpha: \quad A \to \alpha(A) \tag{6}$$

which preserves the algebraic properties

$$(A.1) \quad \alpha(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \alpha(A_1) + \lambda_2 \alpha(A_2) \tag{7}$$

$$(A.2) \quad \alpha(A_1 A_2) = \alpha(A_1)\alpha(A_2)$$

$$(A.3) \quad \|\alpha(A)\| = \|A\|$$

$$(A.4) \quad \alpha(A^+) = [\alpha(A)]^+$$

In particular, the set of automorphisms considered may form a group. Let $\{g_1 g, \ldots\}$ be elements of the group. Then the group property is

$$\alpha_{g_1}[\alpha_{g_2}(A)] = \alpha_{g_1 g_2}(A) \tag{8}$$

An example of a symmetry group of a system is the group of time displacements $\alpha_t(A)$

$$\alpha_{t_1}[\alpha_{t_2}(A)] = \alpha_{t_1+t_2}(A) \tag{9}$$

By definition, a state is invariant under the group G if

$$\omega[\alpha_g(A)] = \omega(A) \text{ for all } A \in \mathfrak{A} \text{ and } g \in G. \tag{10}$$

States invariant under the group of time displacements are called equilibrium states. The states of interest may be restricted by demanding that ω belong to the set of states invariant under some symmetry.

Suppose that the set of admissible states is restricted to some subset $\mathfrak{R}$ of $\mathfrak{S}$ which is convex. Every convex set possesses a set of extreme points $E(\mathfrak{R})$ with the defining property

(E) If $\omega \in E(\mathfrak{R})$ and $\omega = \lambda\omega_1 + (1-\lambda)\omega_2$, then $\omega_1 = \omega_2 = \omega$

for real λ between 0 and 1. (11)

The states in $E(\mathfrak{R})$ are the pure states. All other states are mixed states (For a finite system, a pure state in the set of all possible states $\mathfrak{S}$ entails $\mathrm{tr}\,\rho^2 = 1$). Remember that this terminology is relative to the set $\mathfrak{R}$ being considered.

When the states of a system are restricted to those invariant under a group G, the subset of states will be called $\mathfrak{S}_G$, the set of G-invariant states. States which cannot be decomposed into a mixture of two differing states invariant under G are called extremal invariant with respect to G and will be denoted by $E(\mathfrak{S}_G)$.

For the mixed states of the system, it is always possible to decompose the state ω as a linear combination of other states in an infinite variety of ways. For example, let the set $\mathfrak{S}_T$ be the states invariant under the time displacement group T of some Hamiltonian. Let the spectral decomposition of the density matrix ρ be $\sum_\alpha \rho_\alpha p_\alpha$ then,

$$\omega(A) = \Sigma \rho_\alpha \omega_\alpha(A) \tag{12}$$

$$\omega_\alpha(A) = \mathrm{tr}\, p_\alpha A$$

Even this decomposition is not unique if any two of the ρ_α are equal (Try $\rho = 1$!). In this case the state ω has been written as a linear combination of extremal states. In order that all points of a convex set in N dimensions be a unique combination of extremal points, the convex set must be a simplex σ^{N-1}. (A simplex σ^{N-1} is the Nth element of the sequence: point, interval, triangle, tetrahedron ...). The physical import of these considerations on extreme points will become clear only when infinite systems with multiple coexisting phases are considered.

It is clear that the symmetry properties of a system do not restrict the set of admissible states to a degree sufficient enough to specify the state.

Besides homogeneity and stationarity, ordinarily the temperature is set and suffices to determine the state. While this restriction cannot be imposed as a symmetry, it can be imposed as a condition on the state. Consider

$$\omega(AB) = \frac{\mathrm{tr}\ e^{-\beta\mathcal{H}}(AB)}{\mathrm{tr}\ e^{-\beta\mathcal{H}}} = \frac{\mathrm{tr}\ e^{-\beta\mathcal{H}}B\ e^{-\beta\mathcal{H}}A\ e^{\beta\mathcal{H}}}{\mathrm{tr}\ e^{-\beta\mathcal{H}}} \quad (13)$$

Define

$$\alpha_\beta(A) = e^{-\beta\mathcal{H}}\ A\ e^{\beta\mathcal{H}} \quad (14)$$

The above development motivates the imposition of the KMS condition[9] on the state

$$\omega(AB) = \omega(B\alpha_\beta(A)) \quad \text{(for all A and B in } \mathfrak{A}) \quad (15)$$

This condition suffices to prescribe the state of any finite discrete system uniqely. The proof goes as follows:

(i) $\mathrm{tr}\ \rho(AB) = \mathrm{tr}\ \rho\ Be^{-\beta\mathcal{H}}Ae^{\beta\mathcal{H}} = \mathrm{tr}(e^{-\beta\mathcal{H}}Ae^{\beta\mathcal{H}}\rho)B$ (16)
so $\mathrm{tr}(\rho A - e^{-\beta\mathcal{H}}Ae^{\beta\mathcal{H}}\rho)B = 0$

(ii) Since B is arbitrary, choose it equal to the hermitian adjoint of the expression in parentheses and conclude $\rho A = e^{-\beta\mathcal{H}}\ Ae^{\beta\mathcal{H}}\rho$

(iii) Premultiply by $e^{\beta\mathcal{H}}$ and observe $e^{\beta\mathcal{H}}\rho$ commutes with all A, hence is a multiple of the identity. Normalization yields $\rho = e^{-\beta\mathcal{H}}/\mathrm{tr}\ e^{-\beta\mathcal{H}}$.

For infinite systems, while $\lim_{N\to\infty} e^{-\beta\mathcal{H}_N}$ will not exist ($\mathcal{H}_N$=N particel Hamiltonian), it is often true that $\lim_{N\to\infty} e^{-\beta\mathcal{H}}Ae^{\beta\mathcal{H}_N}$ will exist and offer a means of compu-

3. The GNS Representation

The most familiar realization of the linear functionals ω and the algebra of observables has already been cited, viz the specification of the state by a density matrix and the representation of the algebra by the N x N matrices with $\omega(A) = \mathrm{tr}(\rho A)$.

Another possibility is to represent the algebra of observables by matrices acting in a Hilbert space $C^M (M > N)$ and obtain the state functional ω as a single matrix element of the observables represented. The GNS (Gelfand, Naimark, Segal) construction[10] leads to such a representation both for finite and for infinite systems.

To see the main feature of the GNS representation, consider a pure density matrix which is necessarily of the operator form

$$\rho = |\Phi\rangle\langle\Phi| \qquad \langle\Phi|\Phi\rangle = 1 \tag{17}$$

The observable A records as the number

$$\omega(A) \equiv \langle A\rangle = \langle\Phi|A|\Phi\rangle \tag{18}$$

One might wonder if there is always a representation of any ω of this type even for mixed states to be precise, does there exist a representation of observable operators $\mathfrak{D}_\omega(A)$ acting in a Hilbert space $\mathfrak{H}_\omega$ such that

$$\omega(A) = (\Phi, \mathfrak{D}_\omega(A)\Phi) \tag{19}$$

The GNS construction leads to such a representation. Furthermore, the representation of this form is rendered unique if we demand that $|\Phi\rangle$ be <u>cyclic</u>. A vector Φ is cyclic relative to the algebra $\mathfrak{A}$ if the set $\{\mathfrak{D}_\omega(A)\Phi \,|\, A\in\mathfrak{A}\}$ is dense in the space $\mathfrak{H}_\omega$, that is, for any vector ψ in the space $\mathfrak{H}_\omega$ $||\psi - \mathfrak{D}_\omega(A)\Phi||$ can be made as small as desired by a suitable choice of A. In other words, the space $\mathfrak{H}_\omega$ has a set of operators $\mathfrak{D}_\omega(\mathfrak{A})$ which are not too sparse. The cyclic vector Φ is often called the vacuum.

Recall first the defining properties of a representation

$$(R_1)\ \mathfrak{D}_\omega(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1 \mathfrak{D}_\omega(A_1) + \lambda_2 \mathfrak{D}_\omega(A_2) \tag{20}$$

$$(R_2)\ \mathfrak{D}_\omega(A_1 A_2) = \mathfrak{D}_\omega(A_1)\mathfrak{D}_\omega(A_2)$$

$$(R_3)\ \mathfrak{D}_\omega(A^+) = [\mathfrak{D}_\omega(A)]^+$$

To see what is involved in the construction, the details will be worked out first for systems whose defining representation operates in C^2. The algebra of observables $\mathfrak{A}$ in the space (physically realized as the space of a spin $\frac{1}{2}$ particle or of photon polarizations) consists of elements σ of the form

$$\sigma = a_o\sigma_o + a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 \tag{21}$$

with σ_o as the identity and the multiplication law

$$\sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_h \qquad i,j,h = 1,2,3. \tag{22}$$

With no loss of generality, the density matrix may be chosen as

$$\rho = \frac{e^{\gamma\sigma_3}}{tr(e^{\gamma\sigma_3})} \tag{23}$$

The linear functional thereby defined has the action

$$\omega(\sigma_o) = 1 \; ; \quad \omega(\sigma_1) = \omega(\sigma_2) = 0$$
$$\omega(\sigma_3) = \tanh\gamma \tag{24}$$

Seek a representation of the form

$$\omega(\sigma) = \langle\Phi|\mathfrak{D}_\omega(\sigma)|\Phi\rangle \tag{25}$$

Clearly $\mathfrak{D}_\omega(\sigma)$ cannot be an irreducible representation of the algebra for $\gamma \neq \infty$. (There is only one irreducible representation of the complete matrix algebras.) The next simplest representation is the 4-dimensional one

$$\mathfrak{D}(\sigma) = \begin{bmatrix} \sigma & 0 \\ 0 & \sigma \end{bmatrix} \tag{26}$$

Setting $\Phi = (\Phi_1\Phi_2\Phi_3\Phi_4)^T$ (T is the transposition to column form), seek Φ such that

$$\langle\Phi|\mathfrak{D}(\sigma_o)|\Phi\rangle = |\Phi_1|^2+|\Phi_2|^2+|\Phi_3|^2+|\Phi|^2 = 1 \tag{27}$$

$$\langle\Phi|\mathfrak{D}(\sigma_1)|\Phi\rangle = (\Phi_1{}^*\Phi_2+\Phi_2{}^*\Phi_1)+(\Phi_3{}^*\Phi_4+\Phi_4{}^*\Phi_3) = 0$$

$$\langle\Phi|\mathfrak{D}(\sigma_2)|\Phi\rangle = i(\Phi_1{}^*\Phi_2-\Phi_2{}^*\Phi_1)+i(\Phi_3{}^*\Phi_4-\Phi_4{}^*\Phi_3) = 0$$

$$\langle\Phi|\mathfrak{D}(\sigma_3)|\Phi\rangle = |\Phi_1|^2-|\Phi_2|^2+|\Phi_3|^2-|\Phi_4|^2 = \tanh\gamma$$

The solution is

$$|\Phi_1| = \sqrt{\frac{1+\tanh\gamma}{2}}\ ,\quad |\Phi_4| = \sqrt{\frac{1-\tanh\gamma}{2}} \tag{28}$$

$$|\Phi_2| = |\Phi_3| = 0$$

It is convenient to take instead a related representation

$$\mathfrak{D}(\sigma) = \begin{bmatrix}\sigma & 0\\ 0 & \hat{\sigma}\end{bmatrix} \qquad \hat{\sigma} = \sigma_1\sigma\sigma_1 \tag{29}$$

for then we have (with a phase choice) the product form

$$\Phi = \begin{bmatrix}\sqrt{\frac{1+\tanh\gamma}{2}}\\ -\sqrt{\frac{1-\tanh\gamma}{2}}\end{bmatrix} \otimes \binom{1}{0}$$

$$\mathfrak{D}(\sigma) = \left[\frac{1+\sigma_3}{2}\right]\otimes\sigma + \left[\frac{1-\sigma_3}{2}\right]\otimes\sigma \tag{30}$$

Recall that if $\Phi\in C^N$ and $\chi\in C^{N'}$ then $\Phi\otimes\chi$, the direct product is a vector in $C^{N+N'}$ with components $\varphi_i\chi_j$.

While the representation is reducible, Φ is cyclic (if $\gamma \neq \infty$) since

$$\mathfrak{D}\left[\sigma\left[\frac{1+\sigma_3}{2}\right]\right]\Phi = \begin{bmatrix}\sqrt{\frac{1+\tanh\varphi}{2}}\\ 0\end{bmatrix}\otimes\sigma\binom{1}{0}$$

$$\mathfrak{D}\left[\sigma\left[\frac{1-\sigma_3}{2}\right]\right]\Phi = \begin{bmatrix}0\\ \frac{1-\tanh\varphi}{2}\end{bmatrix}\otimes\hat{\sigma}\binom{1}{0} \tag{31}$$

clearly spans the space C^4 since $\sigma\binom{1}{0}$ generates the space C^2.

To summarize, the GNS construction provides us with a representation of the algebra of observables $\mathfrak{A}$ such that the state ω is represented by the matrix element of a cyclic vector

$$\omega(A) = (\Phi, \Omega_\infty(A)\Phi) \tag{32}$$

Formally, this insures that all calculations at finite temperature "look like" calculations at zero temperature.

Given a linear positive functional (a state) ω and a C^*-algebra of observables, GNS have proved the existence of such a representation with a cyclic vector Φ. The representation is irreducible if and only if ω is a pure state. Note that in the representation constructed above, the state is pure at $\gamma=\infty$ (zero temperature) and that Φ is no longer cyclic. The GNS construction then leads to the two dimensional representation

$$\omega(\sigma) = (\Phi, \sigma\Phi) \; ; \quad \Phi = \binom{1}{0} \tag{33}$$

4. The Representation of Finite Dimensional Algebras

The explicit construction of the GNS representation for the finite dimensional discrete systems provides a concrete case illustrating the structure and properties of such representations.

With no loss of generality, the density matrix specifying an N level system may be assumed diagonal with eigenvalues $\rho_1\rho_2\ldots\rho_N$. Let E^{nm} be the matrix with entry 1 in the n^{th} row, m^{th} column and zero elsewhere. Then

$$\mathrm{tr}\ \rho\ E^{nm} = \rho_n \delta_{nm} \tag{34}$$

The GNS representation then must have

$$\langle\Phi|\mathfrak{A}_{\omega}(E^{nm})|\Phi\rangle = \rho_n \delta_{nm} \tag{35}$$

All elements of the algebra are obtainable from E^{nm} since

$$A = \sum_{rs} a_{rs} E^{rs} \tag{36}$$

where of course the bases E^{rs} are dictated by the state ρ. The multiplication law in $\mathfrak{A}$ is determined by

$$E^{mn}E^{rs} = \delta_{nr}E^{ms} \tag{37}$$

Introduce the transposition operators P_N operating on C^N. These have the defining property

$$P_n(C_1C_2\ldots C_n\ldots C_N)^T=(C_nC_2\ldots C_1\ldots C_N)^T \tag{38}$$

$$P_n=P_n^{+} \qquad P_n^{2}=1$$

Write

$$\mathfrak{A}(E^{rs}) = \sum E^{mn} \otimes P_nE^{rs}P_n \tag{39}$$

Clearly,

$$\mathfrak{A}(E^{mn})\mathfrak{A}(E^{rs}) = \delta_{nr}\mathfrak{A}(E^{ms}).$$

$$\mathfrak{A}([E^{mn}]^{+})=[\mathfrak{A}(E^{mn})]^{+} \tag{40}$$

Set $\Phi = (A_1A_2\ldots A_n)^T \otimes (100\ldots0)^T$. Then since

$$(100\ldots0)P_nE^{rs}P_n(100\ldots0)^T= \begin{cases} 1 & n=r=s \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

only the terms with r=s in $\langle\Phi|\mathfrak{A}(E^{rs})|\Phi\rangle$ are non-zero But

$$\langle\Phi|\mathfrak{A}(E^{rs})|\Phi\rangle = A_r^{*}A_r \tag{42}$$

Selection of $A_r^{*}A_r=\rho_r$ yields a representation. The phase can be selected arbitrarily, say

$$A_r = \sqrt{\rho}_r \tag{43}$$

Is Φ cyclic? If no $\rho_s=0$, simple considerations show that Φ is cyclic. In this non-degenerate case, cyclicity follows immediately from the expression

$$\mathfrak{O}(A)\Phi=\sum_s (0\ldots A_s\ldots)^T \otimes (A_{5s}A_{2s}A_{3s}\ldots A_{1s}\ldots A_{Ns})^T \tag{44}$$

The fact that $A_s=\sqrt{\rho}_s\neq 0$ for all s and the entries A_{rs} $A=\sum_{rs} A_{rs}E^{rs}$ insures that $\mathfrak{O}(\mathfrak{A})\Phi$ spans the space. Of course, the representation is reducible in the general case since $(E^{mm}\otimes 1)\Phi$ is a space invariant under $\mathfrak{O}(\mathfrak{A})$

If $\rho_N=\rho_{N-1}=\cdots=\rho_{R+1}=0$, the GNS representation is in a space of NR dimensions. Write in the degenerate case

$$\mathfrak{O}(E^{rs}) = \sum_{n=1}^{R} E^{nn} \otimes P_n E^{rs} P_n$$

$$\Phi = (A_1\ldots A_R)^T \otimes (10\ldots 0)_N^T$$

$$A_s = \sqrt{\rho}_s \tag{45}$$

The vector Φ is cyclic. In particular for a pure state

$$\mathfrak{O}(E^{rs}) = E^{rs}.$$

As expected and the representation is N dimensional.

The reduction of the dimension of the Hilbert space C^{N^2} as the "mixed state becomes purer" can be described more formally. A left ideal $\mathfrak{I}_L$ of an algebra $\mathfrak{A}$ is a subspace of the algebra with the property

$$AB\in\mathfrak{I}_L \quad \text{for all } A\ \mathfrak{A},\ B\ \mathfrak{I}_L \tag{46}$$

One particular way of constructing a left ideal is to consider the set of all elements B such that $\operatorname{tr}\rho(B^+B)=0$. The set so formed is a left ideal as a simple calculation shows. To reduce the dimension of the representing space from C^{N^2} to C^{NR}, find a projection operator π such that

$$\pi\mathfrak{A}(B^{+}B)\pi = 0 \qquad \text{for } B\in\mathfrak{I}_L$$

$$\omega(\pi\mathfrak{A}(A)\pi) = \omega(A) \qquad \text{for } A\notin\mathfrak{I}_L \tag{47}$$

The GNS construction then leads to the Hilbert space C^{NR} spanned by the vectors in $\pi\psi$ where $\psi\in C^{N^2}$. This is what was done in the previous paragraph.

To conclude the discussion of the algebras acting in C^N, it is interesting to note that although the GNS representation is generically reducible for finite T, the algebraic representation is economical in the following sense. Let $\mathfrak{A}(\mathfrak{U})'$ be the set of operators in $\mathcal{H}_\omega$ which commute with all $\mathfrak{A}(\mathfrak{U})$. Then

$$\mathfrak{A}(\mathfrak{U})' \cap \mathfrak{A}(\mathfrak{U}) = \lambda\mathfrak{A}(E) \tag{48}$$

$\mathfrak{A}(E)$ = representation of the identity

Such a representation is called <u>primary</u>. In addition, the property

$$\mathfrak{A}(\mathfrak{U})'' \equiv [\mathfrak{A}(A)']' = \mathfrak{A}(\mathfrak{U}) \tag{49}$$

can be verified. This last property defines the representation algebra as a <u>von Neumann algebra</u>. A von Neumann algebra is a factor if it satisfies the condition of Eq. (48). To prove these statements it suffices to exhibit the elements of $\mathfrak{A}(\mathfrak{U})'$ explicitly. B belongs to $\mathfrak{A}(\mathfrak{U})'$ if it is of the form

$$\sum_{nm}^{R} b_{nm}\,(E^{nm}\otimes P_n P_m) \tag{50}$$

for arbitrary complex b_{nm}.

II. <u>Algebraic Treatment of Infinite Lattice Systems</u>

The remaining presentation will be concerned with infinite systems of a special type - lattice systems. A lattice Z^d is the set of all elements indexed by a d-tuplet of integers $(n_1,n_2\ldots n_d) \equiv n$. Physical lattices occur naturally in crystals where d=3, of course and n locates an elementary atomic grouping from which the crystal is made. A subset Λ of the lattice is defined to be any subset of the lattice points.

To each point $n \in Z^d$ of the lattice, associate an s-dimensional Hilbert sapce C_n^s. For example, if the lattice is the idealization of a crystal with spin $\frac{1}{2}$ particles on each site then the space C_n^s is the two dimensional complex space describing the spin at site n on the d=3 dimensional lattice. Direct products (also called tensor products) of these spaces

$$\bigotimes_{n\in\Lambda} C_n^s \tag{51}$$

are the spaces of interest here. From now on, s is taken to be 2. For Λ a finite subset, the Hilbert space of the direct product is finite dimensional and the analysis of operators in this space is straight forward. On the other hand, when Λ is taken to be Z^d itself, the infinite tensor product has novel properties[11]. It leads to an inseparable Hilbert space.

1. Infinite Tensor Products

Nothing topologically or algebraically eventfully happens if a set of N spins is described by the tensor product of their Hilbert spaces. The tensor product $\bigotimes_{n}^{N} C_n^2$ is defined to be the 2^N dimensional Hilbert space obtained by forming all linear combinations of the N tuplet

$$(\varphi_1,\varphi_2,\ldots\varphi_N) = \varphi \text{ with } \varphi_r \in C_r^2 \tag{52}$$

with the properties

$$(\varphi_1,\varphi_2,\ldots\lambda_1\varphi_{r_1} + \lambda_2\varphi_{r_2}\ldots\varphi_N) = \lambda_1(\varphi_1\varphi_3\ldots\varphi_{r_1}\ldots\varphi_N) \tag{53}$$

$$+ \lambda_2(\varphi_1\varphi_2\ldots\varphi_{r_2}\ldots\varphi_N) \qquad (\varphi,\chi) = (\varphi_1,\chi_1)\ldots(\varphi_N,\chi_N)$$

Upon this space operate all the bounded operators $B(\bigotimes_{n}^{N} C_n^2)$ which are generated by taking sums and products of the basic operators

$$\begin{aligned} &1\otimes 1\otimes\ldots\otimes\sigma^r\otimes 1\otimes\ldots\otimes 1 \quad \text{(abbreviated } \sigma^r\text{)} \\ &\sigma^r = a_o^{(r)} + a_1^{(r)}\sigma_1^{(r)} + a_2^{(r)}\sigma_2^{(r)} + a_3^{(r)}\sigma_3^r \end{aligned} \tag{54}$$

The action of $\sigma^{(r)}$ is defined by its action on a φ-vector

$$\sigma^{(r)}\varphi \equiv (\varphi_1\varphi_2\ldots\sigma^{(r)}\varphi_r\ldots\varphi_N)$$
$$\sigma^r(\lambda_1\varphi + \lambda_2\chi) = \lambda_1\sigma^r\varphi + \lambda_2\sigma^r\chi \tag{55}$$

In statistical physics, the simplified thermodynamical properties of systems are expected to be exhibited only when $N \to \infty$. Something mathematically eventful does happen when the number of spins is assumed infinite.

The <u>complete tensor product</u> $\overset{\infty}{\otimes} C_n^2$ is defined to be the Hilbert space constructed by taking all finite linear combinations of vectors of the form

$$\varphi = (\varphi_1\varphi_2\varphi_3\ldots) \qquad \varphi_r \in C_s^3 \tag{56}$$

and completing the space with the help of an inner product defined by

$$(\varphi,\chi) = \prod_i^{\infty} (\varphi_i,\chi_i) \tag{57}$$

The value of this inner product is defined to be finite only when the product converges. It is zero otherwise by definition. The Hilbert space so constructed $\overset{\infty}{\otimes} C_n^2$ is non-separable since uncountably many vectors can be made mutually orthogonal. Such a space is not useful to represent the physics of quantum systems.

To obtain a physically relevant Hilbert space for the infinite dimensional system, select a vector φ of the form given above. A vector $\chi = (\chi_1\chi_2\chi_3\ldots)$ is defined to be equivalent to φ if only a finite number of χ_i differ from the corresponding φ_i. The usual inner product is formed and is well defined for all finite linear combinations of vectors of the form φ. A separable Hilbert space $\mathcal{H}(\varphi)$ results by taking all finite linear combinations of vectors in the same equivalence class and completing the space with the help of the norm formed from the inner product. The Hilbert space $\mathcal{H}(\varphi)$ is called the <u>incomplete tensor product</u>.

As a specific example, let $\Phi^{(+)}$ represent a vector with all spins up. All vectors with a finite number of spins down belong to $\mathcal{H}(\Phi^{+})$. Another equivalence class is generated by $\Phi^{(-)}$ representing all spins down. In $\mathcal{H}(\Phi^{-})$ are all vectors with a finite number of spins up. One notable peculiarity of infinite systems is discernible here - every vector of $\mathcal{H}(\Phi^{+})$ is orthogonal to every vector in $\mathcal{H}(\Phi^{-})$.

The representation of algebras in $\mathcal{H}(\Phi^{+})$ and $\mathcal{H}(\Phi^{-})$ can be unitarily inequivalent representations as the following example shows. Suppose that in the algebra under discussion, there is the bounded operator

$$A \equiv \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \sigma_3^n \tag{58}$$

On $\mathcal{H}(\varphi^{+})$, this operator is represented by 1, on $\mathcal{H}(\varphi^{-})$ it is represented by -1. Clearly there is no unitary transformation such that

$$\begin{aligned} &VA^{(+)}V^{-1} = A^{(-)} \\ &A^{(+)} = 1,\ A^{(-)} = -1 \end{aligned} \tag{59}$$

There is a class of Hilbert spaces which should be considered as physical equivalent to $\mathcal{H}(\varphi^{+})$. All such spaces are generated from vectors χ which are called <u>weakly equivalent</u> to φ^{+}. The definition of weakly equivalent vectors is motivated as follows. Consider a transformation V acting on $\Phi^{(+)}$, which has the action

$$\begin{aligned} V\Phi^{(+)} &= (e^{i\Phi_1}\alpha_1,\ e^{i\Phi_2}\alpha_2,\ e^{i\Phi_3},\ldots) \\ \Phi^{+} &= (\alpha_1,\alpha_2,\alpha_3,..) \end{aligned} \tag{60}$$

Physically, the choice of the phases should make no difference in the state. V may be written formally as

$$V = \bigotimes_{n}^{\infty} e^{i\sigma_3^n\varphi_n} \tag{61}$$

but the matrix element

$$(\Phi^+, V\Phi^+) = \prod_{n}^{\infty} e^{i\varphi_n} \tag{62}$$

may be ill defined if $\Sigma\varphi_n$ is unbounded. Instead of V, the "renormalized" form may be used

$$V' = \bigotimes_{n}^{\infty} e^{i(\sigma_3^n - 1)\varphi_n} \tag{63}$$

Then $V'\Phi^{(+)} = \Phi^+$. However, the effect of V' on operator is not completely trivial. For example,

$$(V')^+ \sigma_+^{(n)} V' = e^{i\varphi_n} \sigma_+^{(n)} \tag{64}$$

Clearly, V' is a unitary operator in the separable space $\mathcal{H}(\varphi^+)$ and yields an equivalent representation of the algebra.

The criterion for weak equivalence of two vectors φ and χ is that

$$\sum_n |1 - |(\varphi_n, \chi_n)|| < \infty . \tag{65}$$

For vectors in the same equivalence class, the analogous criterion is

$$\sum_n |1 - (\varphi_n, \chi_n)| < \infty . \tag{66}$$

One advantage of working with the algebra of operators rather than with the vectors in a Hilbert space follows from the above discussion. Rays, not vectors, are uniquely associated with physical states. For infinite systems, the phase arbitrariness leads to the necessity of classifying vectors in weak equivalence classes. The algebraic formulation of the structure of infinite systems is, in this sense, strongly motivated.

2. Algebra of Observables on Lattices

For the ideal infinite system, the algebra of observables should be somewhat restricted to accord with the finiteness of the apparati of observation. A second practical demand is that the infinite system be mathematically tractable. Among the other requirements, surely one needs a simple device for constructing and specifying states of an infinite dimensional system. In order to carry over most of the ideas developed in section I for finite systems, it suffices to assemble the observables into a C^*-algebraic structure called a quasi-local algebra.

Consider any finite subset Λ of the lattice. On the subset, the algebra of observables $\mathfrak{A}(\Lambda)$ is taken to be the algebra of all bounded operators. For lattice systems of spin $\frac{1}{2}$ particles, $\mathfrak{A}(\Lambda)$ is simply the algebra of all operators in a finite dimensional Hilbert space $\bigotimes^{\Lambda} C_2^n$ of dimension $2^{N(\Lambda)}$ where $N(\Lambda)$ is the number of points in Λ. The set of all such algebra

$$\{\mathfrak{A}(\Lambda): \Lambda \text{ finite}, \Lambda \in Z^d\} \tag{67}$$

provides plenty of observables. The super-algebra formed by the union of all these algebras $\bigcup_{\Lambda} \mathfrak{A}(\Lambda)$ is the algebra of local observables. In the algebra $\bigcup_{\Lambda} \mathfrak{A}(\Lambda)$ a particular operator $A(\Lambda_1)$ acts as the unit operator on the sublattice which is the complement of Λ_1 and has, of course, its original action on the Hilbert space $\mathfrak{K}(\Lambda_1)$.

The set of observables is conveniently made slightly larger by completing the algebra in the norm topology. The resultant algebra $\overline{\bigcup_{\Lambda} \mathfrak{A}(\Lambda)}^{N}$ is called the quasi-local algebra of observables. The reason for completing the algebra is two-fold. One is physical. For most Hamiltonians driving the operator $A(\Lambda)$, the time evolute of $A(\Lambda)$ will be contained in the quasi-local algebra but not in the local algebra. The explanation for this behavior is essentially the reason why a strictly localized wave packet in ordinary quantum-mechanical Hilbert space becomes non-localized as soon as the time evolution operator works on it even though its norm is

preserved. The second reason for completing the local algebra is to obtain a C^*-algebra. For C^*-algebras with a state defined on the system, the GNS construction yields a Hilbert space with a cyclic vector.

3. States of Infinite Lattice Systems

Given the algebra $\mathfrak{A}$ of quasi-local observables, the states of a physical system can be defined just as was done for the finite system. A state ω is a positive normalized linear form (or functional) satisfying s.1, s.2, s.3 of section I.

To construct a state by the canonical procedure, first take any observable in the local algebra, say, $A \in \mathfrak{A}(\Lambda_1)$. Form the linear functional $\omega_\Lambda(A)$ on the algebra $\mathfrak{A}(\Lambda)$ where $A > \Lambda_1$. The limit is now taken as $\Lambda \to Z^d$. Define.

$$\omega(A) = \lim_{\Lambda \to Z^d} \omega_\Lambda(A) \tag{68}$$

The question of the existence of this limit involves the nature of the Hamiltonian of the system. Extensive investigations have recently shown that such limits exist. In particular, Lebowitz and Lieb [12] have shown the existence of the thermodynamic limit of the free energy for quantum systems with Coulomb interactions. Therefore, assume $\omega(A)$ exists for all A of the local algebra. To define the state on the quasi-local algebra extend the functional to the algebra completed in the norm. Such an extension exists and is unique.

The specification of the C^*-algebra and the state ω then provides the basic ingredients for the GNS construction.

4. Symmetries and the Classification of States

The realization of a symmetry on the algebra $\mathfrak{A}$ is given by specifying the automorphisms of the algebra induced by the symmetry group. It is assumed that the state is invariant under the automorphism

$$\omega\,[\alpha(A)] = \omega(A) \tag{69}$$

which means that the physical system admits the symmetry. When the algebra is realized as a set of bounded operators acting on the Hilbert space obtained from the GNS construction, it turns out that the automorphism is implemented by a unitary operator which leaves the cyclic vector invariant.

States invariant under the symmetry group G are called G-invariant states.

Instead of introducing more definitions and classification schemes at this point, specific examples of infinite lattice systems will now be used to clarify the properties of infinite systems.

III. Simple Models of Infinite Lattice Systems

To illustrate the rather abstract formulation of section II, two specific models of infinite lattice systems will be described in some detail. The first model is the one dimensional Ising model; the second the Ising-Weiss model: the latter exhibits a phase transition at finite temperature.

A. One-Dimensional Ising Model

The one dimensional Ising model consists of a set of spin $\frac{1}{2}$ particles on a one dimensional lattice interacting via nearest neighbor interaction. The characteristic feature is that the interaction depends only upon the third component of the spin on each lattice site, a limiting case of the Heisenberg one dimensional spin system where the interaction becomes highly anisotropic.

1. Hamiltonian of the Finite System

For an N-spin system, the Hamiltonian is

$$\mathcal{H}_N = -J \sum_{n=1}^{N} \sigma_3^{(n)} \sigma_3^{(n+1)}\,;\quad \sigma_3^{(N+1)} \equiv \sigma_3^{(1)} \tag{70}$$

Cyclic boundary conditions have been chosen.

The Hamiltonian is invariant under the cyclic group whose action is to translate the spins on the sites in a cycle. Another symmetry admitted by the Hamiltonian is the symmetry group of a disc $\mathfrak{D}_\infty$. All elements are generated by rotations about the z axis through an arbitrary angle ϕ and rotation of all spins through π radians about any axis in the 1-2 plane. In the infinite limit $N \to \infty$, the group $\mathfrak{D}_\infty$ will be broken at $T = 0$.

The partition function for the finite Hamiltonian $\mathcal{H}_N$ is evaluated by introducing the transfer matrix

$$T_{ij} = e^{\beta J_{ij}} \tag{71}$$

which takes on the form

$$T = e^{\gamma}\,\tau_o + e^{-\gamma}\,\tau_1 \tag{72}$$

where τ_i are Pauli matrices in the transfer space and τ_o is the unit matrix. Then, a simple observation yields

$$Z_N = \sum_{\sigma_3^{(n)} = \mp 1} e^{\beta J \sum_{n=1}^{N} \sigma_3^{(n+1)}} \underset{(T.S)}{=} \operatorname{tr} T^N = \operatorname{tr} \begin{pmatrix} e^{\gamma} & e^{-\gamma} \\ e^{-\gamma} & e^{\gamma} \end{pmatrix}^N \tag{73}$$

$$\gamma = \beta J$$

(T.S.) = transfer space

2. States of the System for Arbitrary T

To determine the state of the system, the limiting functional

$$\omega(A) = \lim_{N \to \infty} \omega_N(A) \tag{74}$$

will be computed. It is sufficient to compute

$$\omega_N\left(\sigma^{(n_1)}\,\sigma^{(n_2)} \cdots \sigma^{(n_r)}\right) \tag{75}$$

where $\sigma^{(n_i)}$ is an element of the complete matrix algebra at site n_i. It even proves sufficient to compute an even more specialized form

$$\omega_N\left(\sigma_3^{(n_1)}\sigma^{(n_2)}\cdots\sigma_3^{(n_r)}\right) \quad (N\geq n_r>n_{r-1}>\cdots>n_1\geq 1) \tag{76}$$

since these alone are non-zero as a symmetry argument shows.

The basic observation needed to evaluate the correlation function

$$C_N(m_1m_2\cdots m_r) \equiv \omega_N\left(\sigma_3^{n_1}\sigma_3^{n_1+m_1}\sigma_3^{(n_1+m_1+m_2)}\cdots \sigma^{(n_1+m_1+m_2\cdots+m_r)}\right) \tag{77}$$

is that

$$C_N(m_1m_2\cdots m_r) = \frac{\operatorname{tr}_{(T.S)} T^{n_1}\tau_3T^{m_1}\tau_3T^{m_2}\tau_3\cdots T^{N-n_1-m_1-m_2\cdots}}{\operatorname{tr}_{(T.S)} T^N} \tag{78}$$

where τ_3 is a matrix in the transfer space (T.S). It is an easy matter to evaluate the expression by utilizing the identity

$$T = (2\sinh^2\gamma)^{\frac{1}{2}}\, e^{\tau_1\zeta}$$
$$\tanh\zeta = e^{-2\gamma}\,; \quad \zeta = (\tanh\gamma)^{-\frac{1}{2}} \tag{79}$$

and the state functional becomes

$$\omega\left(\sigma_3^{(n_1)}\sigma_3^{(n_1+m_1)}\sigma_3^{(n_1+m_1+m_2)}\ldots\sigma_3^{(n_1+m_1+m_3+\ldots+m_r)}\right)$$

$$= \begin{cases} 0 & (r \text{ even}) \\ (\tanh\gamma)^{m_1+m_3+\ldots+m_r} & (r \text{ odd}) \end{cases} \tag{80}$$

Observe that the correlation functions have the factorization property

$$\omega\left(\sigma_3^{(n_1)}\sigma_3^{(n_1+m_1)}\sigma_3^{(n_1+m_1+m_2)}\ldots\right) = \tag{81}$$

$$= \omega\left(\sigma_3^{(n_1)}\sigma^{(n_1+m_1)}\right)\omega\left(\sigma_3^{(n_1+m_1+m_2)}\sigma_3^{(n_1+m_1+m_2+m_3)}\right)\ldots$$

so that all that is needed is

$$\omega\left(\sigma_3^{(n_1)}\sigma_3^{(n_1+m_1)}\right) = (\tanh\gamma)^{m_1} \tag{82}$$

Note that the Griffith's inequalities are clearly satisfied.

To evaluate the expectation value of any element of the algebra of quasi-local observables, the state functional is extended to the elements of the norm closure of the generating elements whose state functional has been evaluated above.

As an example, the energy per particle e is found to be

$$e = \omega\left(\lim_{N\to\infty} -\frac{J}{N}\sum_{n=1}^{N}\sigma_3^{(n)}\sigma_3^{n+1}\right)$$

$$= \omega\left(\lim_{N\to\infty} -J\,\omega\left(\sigma_3^{(1)}\sigma_3^{(2)}\right)\right) = -J\omega\left(\sigma_3^{(1)}\sigma_3^{(2)}\right) \tag{83}$$

$$= -J\tanh\beta J$$

3. Cluster Properties

The cluster properties of a system provide a means of studying properties of the state ω. To define the cluster property, start with a group G under which the state is invariant $\omega[\alpha_g(A)] = \omega(A)$. For the one dimensional Ising model, the groups of interest are F the direct product of the translation group T and the group D_∞ and the time displacement group U.

A state is strongly clustering if

$$\lim_{g\to\infty} \omega(\alpha_g(A)\, B) = \omega(A)\,\omega(B) \quad \text{(for all A and B)} \tag{84}$$

It suffices to check the clustering properties for the model for $\omega(\sigma_3^n\sigma_3^m)$. Consider the translation group

$$\lim_{a\to\infty} \omega\left(\alpha_a(\sigma_3^n)\sigma_3^m\right) = \lim_{a\to\infty} (\tanh\gamma)^{(m-n+a)} \quad (m>n)$$
$$\alpha_a(\sigma_3^n) = \sigma_3^{n+a} \tag{85}$$

For non-zero temperature

$$\lim_{a\to\infty} \omega\left(\alpha_a(\sigma_3^n)\sigma_3^m\right) = 0 = \omega(\sigma_3^n)\,\omega(\sigma_3^m) \tag{86}$$

On the other hand, for zero temperature, $\tanh\gamma = 1$ but $\omega(\sigma_3^n)\,\omega(\sigma_3^m) = 0$.

For zero temperature, the system fails to cluster strongly, an indication of long range order.

The extremal invariant states now possess the strongly clustering properties on account of the trivial nature of the long range order properties.

4. The KMS Condition

Finally, the KMS condition may be applied to the determination of the state of the system. Demand

$$\omega(AB) = \omega(B\,\alpha_\beta(A))$$
$$\lim_{N\to\infty} \alpha_\beta(A) = e^{-\beta\mathcal{H}_N} A\, e^{-\beta\mathcal{H}_N} \tag{87}$$

Then

$$\alpha_\beta(\sigma_3^n) = \sigma_3^n$$
$$\alpha_\beta(\sigma_+^{(n)}) = e^{2\beta J[\sigma_3^{(n+1)}+\sigma_3^{(n-1)}]}\,\sigma_+^{(n)}\ ;\ \sigma_+ = \sigma_1+i\sigma_2 \tag{88}$$
$$\alpha_\beta(\sigma_-^{(n)}) = e^{2\beta J[\sigma_3^{n+1}+\sigma_3^{(n-1)}]}\,\sigma_-^{(n)}\ ;\ \sigma_- = \sigma_1-i\sigma_2$$

Use now relations such as:

$$\omega(\sigma_+^{(n)}\sigma_-^{(n)}) = \omega(\sigma_-^n \alpha_\beta(\sigma_+^n))$$
$$\omega(\sigma_-^{(n)}\sigma_+^{(n)}) = \omega(\sigma_+^n \alpha_\beta(\sigma_-^{(n)})) \tag{89}$$
$$\omega\left[\left\{\sigma_+^{(n)},\sigma_-^{(n)}\right\}\right] = 4\ ,\ \omega\left[\left\{\sigma_+^{(n)},\sigma_-^{(n)}\right\}\right] = 4\omega(\sigma_3^{(n)})$$

and translational invariance to obtain

$$2\omega(\sigma_3^n\sigma_3^{n+1})\cosh 2\gamma - \omega(\sigma_3^n\sigma_3^{n+2})\sinh 2\gamma = \sinh 2\gamma \tag{90}$$

The validity of this relationship can be checked against known results. This illustrates how the KMS condition can be used as a calculational tool.

B. The Ising-Weiss Model

In a sense the Ising-Weiss model is the opposite extreme from the nearest neighbor interaction just considered. Each spin now couples equally but weakly with all spins with interaction Hamiltonian

$$\mathcal{H}_N = -\frac{J}{N}\left(\sum_{n=1}^{N} \sigma_3^{(n)}\right)^2 \tag{91}$$

Clearly the space dimension of the lattice is irrelevant.

1. The Canonical Construction of the State

To evaluate the partition function, it is easiest to use a basis in which $\Sigma\sigma_3^n$ is diagonal, for then

$$Z_N = \mathrm{tr}\, e^{-\beta\mathcal{H}_N} = \sum_M \Pi(M)\, e^{\beta\frac{J}{N}M^2} \tag{92}$$

$$M = N, N-2, \cdots, -N \ .$$

$\Pi(M)$ is the multiplicity of orthogonal states in the space of fixed M. A simple calculation yields

$$\Pi(M) = N! \Big/ \left(\frac{N-M}{2}\right)! \left(\frac{N+M}{2}\right)! \tag{93}$$

As $N\to\infty$, $\Pi(M)$ sharply peaks around $M = o$ because of the voluminous phase space. Indeed, an asymptotic evaluation of $\Pi(M)$ yields

$$\Pi(M) \to \frac{1}{\sqrt{2\pi N}}\, e^{\frac{-M^2}{2N}} \tag{94}$$

On the other hand, in the partition function the energy favors large M^2 again exponentially. The conflict between the demand of large phase space and low energy is resolved at high temperatures by the domination of the phase space. To see this, use Eq. (94) tentatively in Eq. (92). Then

$$Z \to \sum_M \frac{1}{\sqrt{2\pi N}}\, e^{\left(\frac{2\beta J-1}{2}\right)\frac{M^2}{N}} \tag{95}$$

If $2\beta J<1$, then indeed the approximation (Eq. 94) is self-consistent since $\left\langle \left(\frac{M}{N} \right)^2 \right\rangle \to 0$.

The situation changes at the critical temperature

$$2\beta_c J = 1$$
$$T_c = \frac{2J}{k} \tag{96}$$

for now Eq. (94) can no longer be used self-consistently. Write

$$Z_N = \sum_M e^{\ell n \Pi(M)} e^{\frac{\beta J}{N} M^2} \tag{97}$$

and seek the maximum of the function

$$\ell n \Pi(M) + \frac{\beta J}{N} M^2 \tag{98}$$

Stirling's approximation can be used for $\Pi(M)$ for all cases except T=o. The condition for an extreme value is then

$$-\frac{1}{2} \ell n \left(\frac{1+m}{1-m} \right) + 2\beta J m = 0$$
$$m = \lim_{N\to\infty} \left(\frac{M}{N} \right) \tag{99}$$

That this is the Weiss expression for magnetization follows from the identity $\tanh^{-1}\chi = \frac{1}{2} \ell n \left(\frac{1+\chi}{1-\chi} \right)$.

Below T_c the m=o solution must be excluded as shown above. Above T_c, m=o.

The state of the system is now easily determined. Above T_c, the state is independent of T and is given by

$$\omega(A) = \omega(\sigma^{(n_1)}\sigma^{(n_2)}\cdots\sigma^{(n_r)}) = o \qquad (A\neq 1)$$
$$\omega(A) = \omega(1) = 1 \tag{100}$$

Below T_c, again all observables not of the form of linear combinations (and their limits) of $\sigma_3^{(n_1)}\sigma_3^{(n_2)}\cdots$ have $\omega(A) = o$. Otherwise, the state is determined by

$$\omega(\sigma_3^{(n_1)}\sigma_3^{(n_2)}\cdots\sigma_3^{(n_r)}) = o \qquad (r \text{ odd})$$
$$\omega(\sigma_3^{(n_1)}\sigma_3^{(n_2)}\cdots\sigma_3^{(n_r)}) = m^r \qquad (r \text{ even}) \tag{101}$$

where it is assumed that no two of the indices n_i are equal.

The ground state needs separate investigation but again leads to the above result with m=1.

2. Symmetry Breaking

Below T_c, the state of the system can be written as the linear combination of two states, each extremal invariant under lattice translations. In each of the two states the spins are all aligned in the same direction with

$$\omega_1(\sigma_3^{(n)}) = |m|$$
$$\omega_2(\sigma_3^{(n)}) = -|m| \tag{102}$$

Further considerations run along the same line indicated in part A of this section.

3. The Cyclic Vector of the GNS Construction

For T=o, the GNS construction leads to the cyclic vector

$$\phi = \frac{1}{\sqrt{2}}\binom{1}{1} \otimes \left[\prod_{\otimes}\binom{1}{o}\right] \tag{103}$$

and the observables are represented by

$$\mathcal{A}_\omega\left(\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}\right) = \left(\frac{1+\sigma_3}{2}\right)\otimes\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}$$

$$+\left(\frac{1-\sigma_3}{2}\right)\otimes\hat{\sigma}^{(n_1)}\hat{\sigma}^{(n_2)}\ldots\hat{\sigma}^{(n_r)} \qquad (104)$$

The only non-obvious fact is the cyclicity of ϕ. Consider

$$\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}\left(\frac{1+\sigma_3}{2}^{(n_1)}\right)\left(\frac{1+\sigma_3}{2}^{(n_2)}\right)\ldots\left(\frac{1+\sigma_3}{2}^{(n_r)}\right). \qquad (105)$$

Applied to ϕ, this yields

$$\frac{1}{\sqrt{2}}\binom{1}{0}\otimes\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}\prod_{\otimes}\binom{1}{0}$$

Similarly,

$$\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}\left(\frac{1-\sigma_3}{2}^{(n_1)}\right)\left(\frac{1-\sigma_3}{2}^{(n_2)}\right)\ldots\left(\frac{1-\sigma_3}{2}^{(n_r)}\right) \qquad (106)$$

applied to ϕ yields

$$\frac{1}{\sqrt{2}}\binom{0}{1}\otimes\sigma^{(n_1)}\sigma^{(n_r)}\ldots\sigma^{(n_r)}\prod_{\otimes}\binom{1}{0} \qquad (107)$$

From these relations, it is clear that the algebra generates a dense set when applied to ϕ.

The extremal invariant states can be obtained by writing

$$\phi = \left\{\left(\frac{1+\sigma_3}{2}\right)\otimes\prod_{\otimes}1\right\}\phi + \left\{\left(\frac{1-\sigma_3}{2}\right)\otimes\prod_{\otimes}1\right\}\phi \equiv \frac{\phi_1}{\sqrt{2}} + \frac{\phi_2}{\sqrt{2}} \qquad (108)$$

The state ω then decomposes

$$\omega = \frac{\omega_1}{2} + \frac{\omega_2}{2} \qquad (109)$$

For $o < T < T_c$, the GNS representation is different and is, of course, reducible since the state is not pure. However, the explicit construction is again easy to carry out (mainly because the system is separable). Suppose the state is extremal invariant with respect to translations. Then let

$$\omega(\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}) = \bigotimes^{\pi}\left[\frac{1+\sigma_3^{(n_i)}}{2}\otimes\sigma^{(n_i)} + \left(\frac{1-\sigma_3^{(n_i)}}{2}\right)\otimes\hat{\sigma}^{(n_i)}\right] \tag{110}$$

The representation of the algebra yields a vector representation of the state with the cyclic vector

$$\phi = \bigotimes^{\pi}\left[\begin{pmatrix}\sqrt{\frac{1+m}{2}}\\ \sqrt{\frac{1-m}{2}}\end{pmatrix}\otimes\begin{pmatrix}1\\ o\end{pmatrix}\right] \tag{111}$$

where m is the solution of Eq. (99).

Above T_c, the same GNS construction is realized but with $m=o$.

Finally, the remaining case to be considered is the representation of the state obtained by the thermodynamic limiting procedure for $o < T < T_c$. At first sight, it might seem that all that it is necessary to do is to combine the $+|m|$ and $-|m|$ representation in the following manner :

$$\omega(\sigma^{(n_1)}\sigma^{(n_2)}\ldots\sigma^{(n_r)}) =$$
$$-\left(\frac{1+\sigma_3}{2}\right)^{(n_i)}\otimes\bigotimes^{\pi}\left[\frac{1+\sigma_3}{2}^{(n_i)}\otimes\sigma^{(n_i)} + \left(\frac{1-\sigma_3}{2}\right)^{(n_i)}\otimes\hat{\sigma}^{(n_i)}\right] \tag{112}$$
$$+\left(\frac{1-\sigma_3}{2}\right)^{n_i}\otimes\bigotimes^{\pi}\left[\left(\frac{1+\sigma_3}{2}\right)^{n_i}\otimes\hat{\sigma}^{(n_i)} + \left(\frac{1-\sigma_3}{2}\right)^{(n_i)}\otimes\sigma^{(n_i)}\right]$$

with the cyclic vector given by

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \bigotimes^{\pi} \left[\begin{pmatrix} \sqrt{\frac{1+m}{2}} \\ \sqrt{\frac{1-m}{@}} \end{pmatrix} \otimes \begin{pmatrix} 1 \\ o \end{pmatrix} \right] \tag{113}$$

This construction yields a representation but ϕ is not cyclic. Thus the representation space is "too big". It can be reduced in half by observing that if

$$I = 1 \otimes \bigotimes^{\pi} 1^{(n_i)} \otimes \bigotimes^{\pi} \sigma_1^{(n_i)} \tag{114}$$

then $I\mathfrak{D}_{\omega}(A)\ I$ is also a representation with the cyclic vector

$$\phi' = \frac{\left(\frac{1+I}{2}\right)\phi}{\left(\phi, \frac{1+I}{2}\phi\right)^{\frac{1}{2}}} \tag{115}$$

This is the correct GNS representation for the case considered.

Conclusions and Outlook

The purpose of these simple considerations was to put the algebraic formulation of statistical physics in concrete terms. Of course, the whole power of the algebraic approach is the general conclusions which can be drawn from abstract generalities. There is a wealth of literature and reviews now on the abstract structure of the algebraic properties of infinite system.[13] What has been attempted here is the most elementary treatment of the principles involved in the hope of broadening the class of physicists who might find these techniques useful and constructive.

REFERENCES

1. Heuristic means that no attempt at mathematical rigor is intended. The constructive examples given herein are designed to motivate the reader to study and utilize the more abstract algebraic and analytic tools.

2. One of the early papers stressing the algebraic structure of statistical systems was that of R. Haag (The Mathematical Structure of the Bardeen-Cooper-Schrieffer Model) Nuovo Cimento 25, 287 (1962).

3. These quotations are given in the biography of Einstein by Ronald W. Clark (World Publishing Company, 1971).

4. For a review of the role of broken symmetries in both relativistic and non-relativistic theory, see G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, "Broken Symmetries and the Goldstone Theorem" in "Advances in Particle Physics" Vol. 2 (Interscience Publishers, 1968) edited by R. L. Cool and R. E. Marshak.

5. See, for example, M. Fisher "The Nature of Critical Points" in Lectures in Theoretical Physics, Vol. VII C (1964), University of Colorado Press.

6. Arthur M. Jaffe, Rev. Mod. Phys. 41, 675 (1969).

7. P. Jordan, J. von Neumann and E. Wigner, "On the Algebraic Generalization of the Quantum Mechanical Formalism", Ann. Math. 35, 29 (1934).

8. J. P. Eckmann, M. Guenin, "Methodes Algébrique en Mécanique Statistique", Springer Verlag (1969).

9. KMS stands for Kubo, Martin and Schwinger who used this condition in a more conventional fashion. The algebraization of the condition is due to Haag, Hugenholtz and Winninh, Commun. Math. Phys. 5, 215 (1967).

References (cont'd)

10. The GNS construction is given in Ref. 8. For the use of the GNS construction in quantum field theory see Michael Reed "The GNS Construction - A Pedagogical Example" Brandeis Lectures Vol. 2 (1970).

11. These novel properties were first analyzed in the mathematical literature by von Neumann, J. Comp . Math 6, 1 (1939).

12. J. L. Lebowitz and Elliot H. Lieb, Phys. Rev. Lett. 22, 631 (1969).

13. See, for example, other review articles in this volume, and the article by N. M Hugenholtz in "Fundamental Problems in Statistical Mechanics" Vol II (North Holland Publishing Company, 1968).

BROKEN SYMMETRY IN RESTRICTED GEOMETRIES†

J. C. Garrison and J. Wong
Lawrence Livermore Laboratory, University of California
Livermore, California 84550
and
Harry L. Morrison
Dept. of Physics, University of California
Berkeley, California 94720
and
Lawrence Livermore Laboratory, University of California
Livermore, California 94550

I. INTRODUCTION

The subject of these talks will be an application of C^*-algebra techniques to the problem of broken symmetry and long-range order in systems with restricted geometries. We will try to make the discussion more concrete by talking about thin-film systems; that is, physical systems that are constrained to lie between two infinitely extended parallel planes with separation L. This problem has been studied by a number of people[1-6] using the methods of conventional, "finite-volume," statistical mechanics. There are, however, drawbacks to these proofs. In the first place, the arguments only apply to the small number of order parameters that have been considered so far; and, in the second place, the proofs use the method of Bogoliubov quasi-averages, which has not been established within the algebraic approach. The purpose of these talks is to present a new proof that avoids the use of quasi-averages and does not require the specification of an order parameter.

Since Professors Haag and Hugenholtz are giving a series of lectures on the foundations of the algebraic

† This work was performed under the auspices of the U.S. Atomic Energy Commission.

approach, we shall mention only a few points that are of importance for this application. We shall take the concrete view that the algebra of observables $\mathfrak{A}$ is realized by bounded operators on the canonical Fock space $\mathfrak{h}_F(\Gamma)$ corresponding to the configuration space Γ relevant to the problem. In the present case, Γ is the space between the two bounding planes. The groups of physical symmetry operations (e.g., time translations, space translations, rotations, etc.) will be represented by automorphisms on $\mathfrak{A}$; we assume, in addition, the following local structure: Each continuous one-parameter automorphism group $F = \{\alpha_\lambda\}$ is <u>locally generated</u>; that is, for each finite volume $V \subset \Gamma$ there is a self-adjoint operator $Q(V)$ acting on $\mathfrak{h}_F(\Gamma)$ and

$$\alpha_\lambda A = \lim_{n\to\infty} \exp[i\lambda Q(V_n)]\, A\, \exp[-i\lambda Q(V_n)], \forall A \in \mathfrak{A}, \alpha_\lambda \in F,$$

where $V_{n+1} \supset V_n$ and $\bigcup_n V_n = \Gamma$. We will assume that physically admissable states ω on $\mathfrak{A}$ can be extended to act on the unbounded operators $Q(V)$. More precisely, we require that the local algebra $\mathfrak{A}(V)$ includes the spectral projections of $Q(V)$ and that the state ω is <u>locally normal</u>; that is, the restriction of ω to $\mathfrak{A}(V)$ is given by a density matrix. The precise assumption is then that the local density matrix gives well-defined average values for unbounded operators like $Q(V)$. This is a reasonable assumption since ω is supposed to be the thermodynamic limit of local Gibbs density matrices.

The groups of particular interest are those that leave the Hamiltonian invariant; their infinitesimal generators are usually defined by a density $q(r)$ which satisfies a continuity equation:

$$Q(V) = \int_V d^3r\, q(r),$$

$$\frac{\partial q(r,t)}{\partial t} + \nabla \cdot I(r,t) = 0,$$

where I is the current associated with q.

Finally, we will need the Kubo-Martin-Schwinger (KMS) condition. A state ω is said to be a <u>KMS state if</u>

$$\int dt\, f(t - i\beta)\omega(BA(t)) = \int dt\, f(t)\omega(A(t)B) \quad \forall f \in \mathfrak{D},\ A,\ B \in \mathfrak{A},$$

where $A \to A(t)$ denotes the time-translation authomorphism and $\mathfrak{D}$ is the space of C^∞ functions with compact support.

With this background in mind, we can go on to construct the necessary machinery for the study of long-range order and broken symmetry in thin films.

II. LONG-RANGE ORDER AND BROKEN SYMMETRY

We have so far discussed the action of symmetry operations on the algebra of observables $\mathfrak{A}$; now it is necessary to consider the symmetry properties of the states. Let G be a group represented by automorphisms $\{\alpha_g : g \in G\}$, then a state ω is said to be G-invariant if $\omega(\alpha_g A) = \omega(A) \forall A \in \mathfrak{A}$, $g \in G$ and to be G-ergodic if it is an extremal point of the convex set of G-invariant states. Recall that a state is extremal in a convex family if it cannot be represented as a nontrivial convex combination of two members of the family.

Let G be the invariance group of the Hamiltonian, then Professor Ruelle has argued that a state ω describing a pure phase of the system should be G-ergodic. We now want to show that this property of being G-ergodic leads to the relation between long-range order and broken symmetry. For homogeneous systems, G will contain a spatial translation subgroup T, for which we shall use the special notation

$$A(x) = \alpha_x A \quad \forall A \in \mathfrak{A},\ x \in T.$$

A T-invariant state ω is said to be strongly clustering if

$$\lim_{|x| \to \infty} (A(x)B) = \omega(A)\omega(B) \quad \forall A,\ B \in \mathfrak{A},$$

and weakly clustering if

$$\lim_{V \to \infty} \frac{1}{V} \int_V d^\nu x\ \omega(A(x)B) = \omega(A)\omega(B)\ \forall A,\ B \in \mathfrak{A},$$

where ν is the dimension of T. I will say that a state ω exhibits long-range-order if it is not weakly clustering. The intuitive idea of long-range order is that some pair of observables is correlated for large separations; that is, the strong cluster property is violated. The definition I have adopted is more stringent, since it excludes cases where weak clustering is satisfied but strong clustering is not. The strong definition agrees with the definitions customarily employed in discussions of Bose condensation, superconductivity, crystal formation, etc.; and it has the further advantage of yielding the usual relation between long-range order and broken symmetry. To see this, we need some results that are conveniently gathered together in Professor Ruelle's book[7].

Theorem 2.1. A T-invariant state ω is T-ergodic if and only if it is weakly clustering. In other words, a state exhibits long-range order if and only if it is not T-ergodic. The next step is to use the existence of integral decompositions for states on a C*-algebra. The principal result is:

Theorem 2.2. Every T-invariant, locally normal, KMS state ω is given by a unique integral decomposition into T-ergodic, locally normal, KMS states; i.e.,

$$\omega(A) = \int d\mu(\sigma)\sigma(A) \quad \forall\, A \in \mathfrak{A}$$

where μ is a probability measure on states carried by the T-ergodic, locally normal, KMS states. The published version of this theorem does not involve the KMS condition, but it can be included with a simple modification of the proof.

Now, suppose that we are given a state ω_o describing a pure phase so that it is G-ergodic. If ω_o exhibits long-range order, theorems 2.1 and 2.2 tell us that it has a nontrivial decomposition into T-ergodic states. If these states are G-invariant, we face a contradiction with the original assumption that ω_o is G-ergodic; therefore, the T-ergodic states cannot be G-invariant. This is the precise statement of the relation between long-range order and broken symmetry.

III. THE BOGOLIUBOV INEQUALITY

The last piece of machinery I need is the infinite-volume form of the famous Bogoliubov inequality; in the present context, this result is based on the following:

Theorem 3.1 Let ω be a KMS state, then the bilinear form

$$(A,B) = \frac{1}{\beta}\int_0^\beta d\tau\ \omega\left(A^\dagger(-i\tau)B\right)$$

defines a norm-continuous inner product on $\mathfrak{A}$.

To make sense of the statement of this theorem, I first have to explain what is meant by a complex time-translation $A \to A(z)$. The concept is defined in the paper of Haag, Hugenholtz, and Winnink[8] as follows: Let $\hat{A}(\epsilon)$ be the Fourier transform of $A(t)$ and assume that $\hat{A}(\epsilon)$ has compact support [as an operator-valued distribution], then $A(z)$ is defined by

$$A(z) = \int d\epsilon\ e^{-i\epsilon z}\,\hat{A}(\epsilon)$$

for any complex z. It can be shown that the subalgebra $\tilde{\mathfrak{A}} = \{A : \hat{A} \text{ has compact support}\}$ is norm-dense in $\mathfrak{A}$, and that $A \to A(z)$ is an automorphism on $\mathfrak{A}$ satisfying

$$[A(z)]^\dagger = A^\dagger(z^*).$$

For $A, B \in \tilde{\mathfrak{A}}$, (A,B) is well defined.

To show that (A,B) is an inner product, it is necessary to recall that the KMS condition automatically implies invariance under time translations[9]. For $A \in \tilde{\mathfrak{A}}$ this extends to invariance under complex time-translations; that is $\omega(A(z)) = \omega(A)$. Using the properties already established, it is easy to verify the defining properties for an inner product,

$$(A,B)^* = \frac{1}{\beta}\int_0^\beta d\tau\ \omega(B^\dagger A(i\tau)),$$

$$= \frac{1}{\beta}\int_0^\beta d\tau\ \omega(B^\dagger(-i\tau)A),$$

$$= (B,A).$$

$$(A,A) = \frac{1}{\beta}\int_0^\beta d\tau\ \omega\left(A^\dagger\left(-\frac{i\tau}{2}\right)A\left(\frac{i\tau}{2}\right)\right)$$

$$= \frac{1}{\beta}\int_0^\beta d\tau\ \omega\left(A\left(\frac{i\tau}{2}\right)^\dagger A\left(\frac{i\tau}{2}\right)\right) \geq 0.$$

So far I have only used the KMS condition to obtain time-translation invariance, which, in turn, implied that (A,B) is an inner product on $\mathfrak{A}$; I now have to use the KMS condition explicitly to get a special property of the inner product. We can estimate (A,A) as follows:

$$(A,A) = \frac{1}{\beta}\int_0^\beta d\tau\ \omega\ (A^\dagger A(i\tau))$$

$$= \frac{1}{\beta}\int_0^\beta d\tau \int d\epsilon\ \ e^{\tau\epsilon}\ \omega(A^\dagger \hat{A}(\epsilon))$$

$$= \int d\epsilon\ \frac{e^{\beta\epsilon} - 1}{\beta\epsilon}\ \ \omega(A^\dagger \hat{A}(\epsilon)).$$

The last line is justified by the fact that the first factor is a C^∞ function of ϵ and the second is a distribution with compact support. Furthermore, one can easily see that $\omega(A^\dagger A(t))$ is a positive-definite function, which

means, by Bochner's theorem, that $\omega(A^{\dagger}\hat{A}(\epsilon))$ is a positive distribution. The combination of this remark with the elementary inequality

$$\frac{e^x - 1}{x} \leq \frac{e^x + 1}{2}$$

yields

$$(A,A) \leq \int d\epsilon \, \frac{e^{\beta\epsilon} + 1}{2} \, \omega(A^{\dagger}\hat{A}(\epsilon)),$$

$$= \tfrac{1}{2} \left[\omega(A^{\dagger}A(i\beta)) + \omega(A^{\dagger}A)\right],$$

$$= \omega\left(\tfrac{1}{2}\{A,A^{\dagger}\}\right) .$$

To get the last line, I had to use the KMS condition in the form:

$$\omega(BA(t + i\beta)) = \omega(A(t)B),$$

which is valid for $A,B \in \mathfrak{A}$. Finally, the Schwartz inequality and the last estimate yield the norm-continuity of the inner product by

$$|(A,B)|^2 \leq (A,A)(B,B) \leq \omega\left(\tfrac{1}{2}\{A,A^{\dagger}\}\right) \omega\left(\tfrac{1}{2}\{B,B^{\dagger}\}\right) \leq ||A||^2 \, ||B||^2.$$

Thus (A,B) extends to all of $\mathfrak{A}$ by continuity, and the proof of theorem 3.1 is complete.

The Bogoliubov inequality itself is obtained by choosing $B = i\frac{\partial}{\partial t}C(t)\,|_{t=0}$; then

$$(A,B) = \frac{1}{\beta}\int_0^{\beta} d\tau \, \omega\left(A^{\dagger} \, i\frac{\partial}{\partial t}C(t + i\tau)\,|_{t=0}\right) .$$

Replace $i\partial/\partial t$ by $\partial/\partial\tau$ in the integrand and perform the integral to get

$$(A,B) = \frac{1}{\beta}\left[\omega(A^\dagger\, C(i\beta)) - \omega(A^\dagger C)\right]$$

$$= \frac{1}{\beta}\,\omega([C,A^\dagger]),$$

where the KMS condition was used to get the last line. Next, use the Schwartz inequality again, together with the previous extimate of (A,A) to get

$$\omega(\{A,A^\dagger\})\ \omega([C,(i\dot{C})^\dagger]) \geq \frac{2}{\beta}\left|\omega([C,A^\dagger])\right|^2,$$

where $\dot{C} \equiv \frac{\partial}{\partial t}\, C(t)\,|_{t=0}$. This is the well-known Bogoliubov inequality.

IV. ABSENCE OF BROKEN SYMMETRIES IN THIN FILMS

The construction of machinery is now finished, and I can proceed to the real topic of this talk, which is the absence of long-range order in thin films. I choose coordinates so that the bounding planes are given by z = 0 and z = L; also, I should remark that hard-wall boundary conditions are to be imposed on these two planes. The Hamiltonian for a homogeneous thin film is evidently invariant under rotations and translations in the x-y plane; therefore, I will take G to be the product of the two-dimensional Euclidean group and whatever internal symmetry group is present (e.g., gauge transformations, spin rotations, etc.).

Let ω_o be a G-ergodic state describing a pure phase of the film; I want to know if ω_o can exhibit long-range order. We have already seen that this is equivalent to asking if ω_o can be decomposed into T-ergodic states with a broken continuous symmetry. A negative answer to this question is provided by the following theorem, which is the central result of this whole discussion.

Theorem 4.1. Every T-invariant, locally normal, KMS state ω for a thin film system is necessarily invariant under any one-parameter group $F \subset G$ locally generated by a conserved density q(r).

That is, there are no states with broken continuous symmetries and consequently no states exhibiting long-range order.

The proof consists of choosing suitable operators to substitute into the Bogoliubov inequality. I will take

$$K = \int_S d^2 x\, e^{-ik\cdot x}\, \tilde{A}(x),$$

$$\tilde{A} \equiv A - \omega(A),$$

and

$$M = \int_{S\times L} d^3 r\, e^{-ik\cdot r}\, q(r),$$

where $A \in \mathfrak{A}_L$ and $q(r)$ is the density generating F. I have adopted the convention that x,x', etc. are vectors with vanishing z-component while r,r' are general vectors, also, the momentum k has no z-component. The integral defining M is taken over a cylindrical region with height L and cross-section S. The next step is to substitute K and M into the Bogoliubov inequality

$$\omega(\{K,K^\dagger\})\ \omega([M,(i\dot{M})^\dagger]) \geq \frac{2}{\beta}\left|\omega([K,M^\dagger])\right|^2,$$

divide by $V^2 = (S\cdot L)^2$, and let $S \to \infty$. I will just sketch the calculations involved.

$$\lim_{S\to\infty} \omega(\{K,K^\dagger\}) = \int d^2x\, e^{-ik\cdot x}\, \omega(\{\tilde{A}(x),\ \tilde{A}^\dagger\})$$

$$\equiv \hat{C}_A(k)$$

This result follows from the translation invariance of ω. I will refer to $\hat{C}_A(k)$ as the correlation function for A. In a similar way, we find

$$\lim_{S\to\infty} \frac{1}{S}\, \omega([K,M^\dagger]) = \lim_{V\to\infty} \int_V d^3 r\, e^{ik\cdot r}\, \omega([A,q(r)]).$$

Since $A \in \mathfrak{A}_L$; that is, $A \in \mathfrak{A}(V_0)$ for some finite V_0, the integrand in the last equation vanishes for r outside V_0. This means that the limit $k \to 0$ and the limit $V \to \infty$ are interchangable; consequently,

$$\lim_{k\to 0} \lim_{S\to\infty} \frac{1}{S} \omega([K,M^\dagger]) = \lim_{V\to\infty} \int_V d^3r\ \omega([A,q(r)])$$

$$= i\frac{\partial}{\partial\lambda}\ \omega(\alpha_\lambda A)_{\lambda=0}\ .$$

The last line follows from the fact that q(r) is the local generator for α_λ.

The last calculation is more complicated and involves the use of the continuity equation.

$$\lim_{S\to\infty} \frac{1}{S^2L^2}\ \omega([M,(iM)^\dagger]) = \lim_{V\to\infty} \frac{(-i)}{V^2} \int_V d^3r \int_V d^3r'\, e^{-ik\cdot(r-r')}$$

$$\times\ \omega([q(r),\ \frac{\partial}{\partial t}\ q(r',t)])_{t=0}\ .$$

The continuity equation and some integrations by parts yield

$$\lim_{V\to\infty} \frac{1}{V^2}\ \omega(M,(iM)^\dagger]) = k_\alpha \int d^3\ x\ e^{-ik\cdot x}\ \omega([\bar{q}(x),\ \bar{I}_\alpha(0)]),$$

where $\bar{q}(x) \equiv L^{-1}\int_0^L dz\ q(x,z)$ etc., and the summation convention applies to vector indices. In obtaining this result I had to drop various surface terms arising from the integration by parts. The contributions from the top and bottom of the cylinder $S \times L$ vanish by virtue of the hard-wall boundary conditions, and the contributions from the sides are eliminated, in the limit, by the factors S^{-1}. I am really only interested in this result for small $|k|$. If the interparticle potential is reasonable; for example, if it is short-ranged and has a well-behaved Fourier transform, then the integrand will fall off rapidly at large $|x|$ and the integral can be expanded as a series in k.

Assuming this to be the case, we have, for small k

$$\lim_{S\to\infty} \frac{1}{S}\, \omega([M,(i\dot{M})^{\dagger}]) = k_{\alpha} W_{\alpha\beta} k_{\beta} \leq W k^2,$$

where $W < \infty$ is the largest eigenvalue of $W_{\alpha\beta}$. ⊙

Putting all this together, we finally have

$$k^2 \hat{C}_A(k) \geq \frac{2}{\beta W} \left| \frac{\partial}{\partial\lambda}\, \omega(\alpha_\lambda A)_{\lambda=0} \right|^2 .$$

The Bogoliubov inequality thus gives information about the small-$|k|$ behavior of the correlation function for any local observable A. Now $\hat{C}_A$ is the Fourier transform of $C_A(x) = \omega(\{A(x), A^{\dagger}\})$, which is a continuous positive definite function satisfying

Theorem 4.2. (Bochner) The Fourier transform of a continuous, positive definite function R^{ν} is a finite, positive measure.

We now integrate the inequality over a sphere of radius k_o centered at the origin. If $\hat{C}_A$ contains a delta function concentrated at k=0, it can be dropped without changing the result and we denote the resulting measure by $d\mu$. For a ν-dimensional configuration space we have

$$k_o^2 \int_{k<k_o} d\mu(k) \geq \int_{k<k_o} k^2 d\mu(k) \geq \frac{2}{\beta W} \left| \frac{2}{\partial\lambda}\, \omega(\alpha_\lambda A)_{\lambda=0} \right|^2 g(\nu) k_o^{\nu}$$

where $g(\nu)k_o^{\nu}$ is the volume of the ν-sphere and $\int_{k<k_o} d\mu(k) \to 0$ as $k_o \to 0$; therefore, if $\nu \leq 2$ we get a contradiction unless

$$\frac{\partial}{\partial\lambda}\, \omega(\alpha_\lambda A)_{\lambda=0} = 0 \;\forall A \in \mathfrak{A}_L .$$

I can replace A by $\alpha_\mu A$, in which case this result implies

$$\frac{\partial}{\partial\mu}\, \omega(\alpha_\mu A) = 0 \;\forall A \in \mathfrak{A}_L, \alpha_\mu \in F$$

But this is the same as $\omega(\alpha_\mu A)=\omega(A)$; in other words, ω is F-invariant. This completes the proof of Theorem 4.1.

The simplest example of a forbidden broken symmetry is Bose condensation. In this case F is the gauge group generated by the particle density $\rho(r)$. One easily finds $W = \omega(\bar{\rho}(0))$. Since the whole argument is independent of statistics, I can equally well conclude that superconductivity is forbidden in thin films. The formation of a crystal lattice is more complicated since there is a two-stage decomposition involved. In the first stage the fully invariant state ω_0 is decomposed into states with a fixed orientation of the crystal axes but no fixed location of the lattice; in the second stage these states are further decomposed into states having a fixed location for the lattice. Our general result forbids the broken rotational invariance encountered in the first stage so the second stage can never be reached.

V. DISCUSSION

The situation as I have outlined it in these talks would be eminently satisfactory if it were not the case that real thin films exhibit behavior commonly ascribed to long-range order. Thus, thin films of helium exhibit superflow and thin-film superconductors are well known. Consequently, there is an apparent contradiction between theory and experiment. Since the theoretical arguments only require a few quite general assumptions, the most probable explanation is that the concept of long-range order appropriate to bulk systems is not applicable to thin films. One promising candidate for a new definition of long-range order is the idea of "weak" long-range order. Briefly, one assumes that the generalized susceptibility or response function for some observable diverges. This behavior is consistent with weak and even strong clustering. In any case, it is clear that there is at present no fundamental theoretical explanation for the behavior of thin films.

REFERENCES

1. P.C. Hohenberg, Phys. Rev. 158, 383 (1967)

2. N. D. Mermin and H. Wagner, Phys. Rev. Letters 17, 113 (1966).

3. N. D. Mermin, Phys. Rev. 176, 250 (1968).

4. D. Jasnow and M. E. Fisher, Phys. Rev. Letters 23, 286 (1969).

5. J. F. Fernandez, Phys. Rev. A 3, 1104, (1971).

6. M. D. Girardeau, J. Math. Phys. 10, 993 (1968).

7. D. Ruelle, Statistical Mechanics (Benjamin, New York, 1969) p. 161 ff.

8. R. Haag, N. M. Hugenholtz, and M. Winnink, Commun. Math. Phys. 5, 215 (1967).

9. M. Winnink, Cargese Lectures in Physics Vol. 4, (Gordon and Breach, New York, 1970).

FORMULATION OF THE MANY - BODY PROBLEM FOR COMPOSITE PARTICLES

M. D. Girardeau
Institute of Theoretical Science and Department of Physics
University of Oregon
Eugene, Oregon 97403

I. Atomic Second-Quantization Formalism

A. Motivation

There are two distinct problems in treating a system of interacting particles each of which is composite in the sense of being composed of several more elementary constituents. The first is the very difficult problem of finding reasonably accurate approximate solutions of any nontrivial quantum-mechanical many-body problem. The second problem, with which these lectures will deal, is that of even formulating the problem in such a way as to take account of the existence of the composite particles. It is, of course, well known that composite particles behave like elementary bosons or fermions when they are (in some reasonable sense) well separated or when the interparticle interactions are small compared to the internal excitation energies[1]). However, there are many problems in which these criteria are violated, yet the composite nature of the particles remains important. Examples are high-temperature gases and partially ionized plasmas, chemical reactions in general, and systems such as superconductors containing electron pairs or other complexes. Even in liquid helium at low temperatures, virtual excitation of the atoms is by no means negligible, since it is responsible for the van der Waals attraction which binds the system into a liquid. In these and other problems, a method of formulation in which the existence of the composite particles is treated kinematically, through use of appropriate composite-particle dynamical variables, is desirable.

It turns out to be possible, starting out from the usual Schrödinger representation of states and observables of the system of interacting "elementary" constituents (electrons and nuclei in the case of systems of atoms), to completely eliminate the explicit dependence on the dynamical variables of the constituents, representing the states and observables in terms of dynamical variables referring only to the atoms (for example, the translational wave vectors and internal excitation quantum numbers of the atoms). Once this has been done, one can introduce a second-quantization formalism in an elementary manner, in which the states and observables are represented in terms of composite particle annihilation and creation operators satisfying elementary Bose or Fermi commutation relations, in spite of the fact that the composite particles are not elementary. The price one pays for the complete elimination of dynamical variables of the constituents is that, in the first place, subsidiary conditions ensuring the correct symmetry under interatomic* exchange of constituents must be imposed in order to set up a one-one correspondence between states in the many-atom state space and those in the state space of the constituents; in the second place, all single-atom states, including the continuum states, must be included in order to obtain a complete many-atom state space. Although the many-atom representation thus obtained is exactly equivalent to a conventional representation in terms of the constituents, both the subsidiary conditions and the continuum atomic states introduce great difficulties in practical calculations. For this reason, my description of this representation, in Secs. I and II of these lectures, will be quite abbreviated. I think that the main value of this representation is that it serves as a useful preliminary to the formally more complicated representation which will be developed in Sec. III (which will occupy most of these lectures), in which only the bound states of the composite particles are represented in terms of atomic dynamical variables, leaving the unbound states to be described in terms of

*From now on our terminology and notation will be adapted to the special case where the composite particles are atoms, for the sake of definiteness. Nevertheless, the method is quite general.

dynamical variables of the constituents rather than the (essentially unknown) continuum atomic states. I think that such representations will be more useful and tractable in practical calculations in spite of being superficially more complicated. In the remainder of Secs. I and II of these lectures I will describe some of the salient features of the original work on representations in terms of both bound and continuum atomic states. Further details can be found in the literature[2),3)].

B. Expansions in Terms of Atomic Product States

For the sake of definiteness, consider a system of n identical atoms each composed of one nucleus and ℓ electrons. Let $\{\varphi_\alpha(Xx_1 \ldots x_\ell)\}$ be a set of single-atom wave functions, orthonormal and complete in the sense

$$\int \varphi_\alpha^*(Xx_1 \ldots x_\ell)\, \varphi_\beta(Xx_1 \ldots x_\ell)\, dX dx_1 \ldots dx_\ell = \delta_{\alpha\beta},$$

$$\sum_\alpha \varphi_\alpha(Xx_1 \ldots x_\ell)\, \varphi_\alpha^*(X'x_1' \ldots x_\ell')$$

$$= (\ell!)^{-1} \delta(X-X') \sum_{P'} \epsilon(P')P'[\delta(x_1-x_1') \ldots \delta(x_\ell-x_\ell')] \qquad (I.1)$$

where $x_j = (r_j, \delta_j)$ denotes both the position and spin z-component variable (=↑ or ↓) of electron j, X the position of the nucleus and also its spin z-component variable in case its total spin is not zero, $\int$ means integration over positions and summation over spins, $\delta_{\alpha\beta}$ is a Kronecker delta with respect to discrete and a Dirac delta function with respect to continuous quantum numbers, $\sum_\alpha$ is a sum over discrete and integral over continuous quantum numbers, and $\delta(X-X')$ and $\delta(x-x')$ are Dirac delta functions of position and Kronecker delta functions of spin. The form of the completeness relation takes into account the antisymmetry of the φ_α in the electron variables; $\sum_{P'}$ denotes a sum over all permutations P′ of the primed variables, $\epsilon(P')$ being +1 for even and -1 for odd permutations. Note that the inclusion of the continuum atomic states in the set

$\{\varphi_\alpha\}$ is quite essential for completeness; if only the bound states are included, no such completeness relation holds. It is this feature which leads to difficulties in applying this formalism to practical calculations.

A system of nuclei and electrons whose numbers are appropriate to an integral number n of such atoms has a wave function ψ which can be expanded as follows:

$$\psi(X_1 \ldots X_n x_1 \ldots x_{\ell n}) = \sum_{\alpha_1 \ldots \alpha_n} c(\alpha_1 \ldots \alpha_n)$$

$$\times \varphi_{\alpha_1}(X_1 x_1 \ldots x_\ell) \ldots \varphi_{\alpha_n}(X_n x_{\ell n-\ell+1} \cdots x_{\ell n}), \qquad \text{(I.2)}$$

with coefficients

$$c(\alpha_1 \ldots \alpha_n) = \int \varphi_{\alpha_1}^*(X_1 x_1 \ldots x_\ell) \ldots \varphi_{\alpha_n}^*(X_n x_{\ell n-\ell+1} \cdots x_{\ell n})$$

$$\times \psi(X_1 \ldots X_n x_1 \ldots x_{\ell n})\, dX_1 \ldots dX_n dx_1 \ldots dx_{\ell n}. \qquad \text{(I.3)}$$

It might be thought that such an expansion is unphysical because we have picked one particular assignment of nuclei and electrons to atoms, i.e., nucleus 1 and electrons $1 \ldots \ell$ to atom 1, etc. However, it follows from (I.3), the antisymmetry of ψ in its electron coordinates, and its symmetry or antisymmetry in nuclear coordinates, that either all of the coefficients c are unchanged or else all simultaneously change sign under a permutation of the assignment of nuclei and electrons to atoms, depending on the parity of the permutation. Thus the expansion (I.2) is in fact independent of the particular assignment, apart from physically unobservable constant ± 1 phase factors.

In order to transform to a representation in which dynamical variables of the nuclei and electrons are eliminated in favor of those of atoms, one can consider the coefficents $c(\alpha_1 \ldots \alpha_n)$ as new wave functions and the arguments $\alpha_1 \ldots \alpha_n$ as the atomic dynamical variables. Part of

the permutation symmetry of ψ under exchange of identical constituents (nuclei and electrons) appears explicitly in c, in that it follows from (I.3) that c is either a symmetric or an antisymmetric function of $\alpha_1 \ldots \alpha_n$, depending on whether $2J + \ell$ is even or odd, where J is the nuclear spin (at this point we are applying the spin-statistics theorem). On the other hand, the symmetry of ψ under interatomic exchange of constituents (not exchange of whole atoms) appears in c in a concealed form. If one compares the expansion (I.2) with one differing only by interatomic exchange of an electron between the pth and qth atom, one finds[2]) that the wave functions c must satisfy the linear relation

$$\sum_{\alpha\beta}(\alpha_p\alpha_q|I_{elec}|\alpha\beta)c(\alpha_1\ldots\alpha_{p-1}{}^{\alpha}\alpha_{p+1}\ldots\alpha_{q-1}{}^{\beta}\alpha_{q+1}\ldots\alpha_n) = -c(\alpha_1\ldots\alpha_n) \qquad (I.4)$$

where the electron exchange matrix is defined as

$$(\alpha_p\alpha_q|I_{elec}|\alpha\beta) \equiv \int \varphi_{\alpha_p}^*(Xx_1\ldots x_\ell)\,\varphi_{\alpha_q}^*(X'x_1'\ldots x_\ell')\,\varphi_\alpha(Xx_1'x_2\ldots x_\ell) \times \varphi_\beta(X'x_1x_2'\ldots x_\ell')\,dXdx_1\ldots dx_\ell dX'dx_1'\ldots dx_\ell'. \qquad (I.5)$$

The relation (I.4) for one particular value of p and q, say p = 1 and q = 2, together with the symmetry or antisymmetry of $c(\alpha_1\ldots\alpha_n)$, implies the relation for all values of p and q. Hence it is convenient to state (I.4) in the symmetrized form

$$\sum_{p<q}^{n}\sum_{\alpha\beta}(\alpha_p\alpha_q|I_{elec}|\alpha\beta)c(\alpha_1\ldots\alpha_{p-1}{}^{\alpha}\alpha_{p+1}\ldots\alpha_{q-1}{}^{\beta}\alpha_{q+1}\ldots\alpha_n) = -\tfrac{1}{2}n(n-1)c(\alpha_1\ldots\alpha_n). \qquad (I.6)$$

Similarly, the symmetry or antisymmetry of ψ under permutations of the nuclear arguments $X_1 \ldots X_n$ implies that c must satisfy the linear relation

$$\sum_{p<q}^{n} \sum_{\alpha\beta} (\alpha_p \alpha_q | I_{nuc} | \alpha\beta) c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \cdots \alpha_{q-1} \beta \alpha_{q+1} \cdots \alpha_n)$$

$$= (-1)^{2J} \tfrac{1}{2} n(n-1) c(\alpha_1 \ldots \alpha_n) \qquad (I.7)$$

where the nuclear exchange matrix is defined as

$$(\alpha_p \alpha_q | I_{nuc} | \alpha\beta)$$

$$\equiv \int \varphi_{\alpha_p}^*(X x_1 \ldots x_\ell) \varphi_{\alpha_q}^*(X' x_1' \ldots x_\ell') \varphi_\alpha(X' x_1 \ldots x_\ell) \varphi_\beta(X x_1' \ldots x_\ell')$$

$$\times dX dx_1 \ldots dx_\ell \, dX' dx_1' \ldots dx_\ell'. \qquad (I.8)$$

The subsidiary conditions (I.6) and (I.7), together with the condition of symmetry or antisymmetry and $c(\alpha_1 \ldots \alpha_n)$, are in fact necessary and sufficient conditions[2]) that the space of wave functions c be in one-one correspondence with the space of properly antisymmetric and symmetric ψ's. They can be interpreted as saying that the "physical state space" of c's is not the entire space of symmetric or antisymmetric c's, but the subspace of the simultaneous eigenstates of the two linear, hermitian operators defined by the left sides of (I.6) and (I.7), with eigenvalues $-\frac{1}{2}n(n-1)$ and $(-1)^{2J} \frac{1}{2}n(n-1)$, respectively. If we had started with an expansion differing from (I.2) by the inclusion of prefactors of explicit antisymmetrizing and symmetrizing or antisymmetrizing operators with respect to the electronic and nuclear variables, we would have found that the simple explicit expression (I.3) for the expansion coefficients c would have been replaced by an implicit equation for c with a nonunique solution. The same conditions (I.6) and (I.7) would nevertheless have appeared as conditions picking out a unique solution[2]).

The correspondence between the space of ψ's and the space of c's is such that inner products are preserved[2]), i.e.

$$(\psi,\psi') \equiv \int \psi^*(X_1 \ldots X_n x_1 \ldots x_{\ell n})\, \psi'(X_1 \ldots X_n x_1 \ldots x_{\ell n}) \times dX_1 \ldots dX_n\, dx_1 \ldots dx_{\ell n}$$

$$= (c,c') \equiv \sum_{\alpha_1 \ldots \alpha_n} c^*(\alpha_1 \ldots \alpha_n) c'(\alpha_1 \ldots \alpha_n). \qquad (I.9)$$

The subsidiary conditions (I.6),(I.7) play an essential role in the proof of (I.9).

C. Representation of Observables

Let T be any single-particle operator, e.g. the kinetic energy operator, which has the structure

$$T\psi(X_1 \ldots X_n x_1 \ldots x_{\ell n}) = \left[\sum_{j=1}^{n} T(X_j) + \sum_{j=1}^{\ell n} T(x_j)\right] \times \psi(X_1 \ldots X_n x_1 \ldots x_{\ell n}) \qquad (I.10)$$

on the space of ψ's. The form of T as an operator on the space of c's is easily found by expanding $T\psi$ in the manner (I.2); the result is [2])

$$Tc(\alpha_1 \ldots \alpha_n) = \sum_{p=1}^{n} \sum_{\alpha} (\alpha_p |T| \alpha) c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \ldots \alpha_n) \qquad (I.11)$$

where the atomic kinetic energy matrix elements have the expected form

$$(\alpha_p|T|\alpha) = \int \varphi_{\alpha_p}^*(Xx_1\ldots x_\ell)\,[T(X)+\sum_{j=1}^{\ell} T(x_j)\,]\varphi_\alpha(Xx_1\ldots x_\ell) \times dX dx_1\ldots dx_\ell. \qquad (I.12)$$

Similarly, a two-particle operator, of which interaction potentials are the most important examples, has the structure

$$V\psi(X_1\ldots X_n x_1\ldots x_{\ell n}) \equiv \Big[\sum_{j<k}^{n} V(X_jX_k)+\sum_{j<k}^{\ell n} V(x_jx_k)+\sum_{j=1}^{n}\sum_{k=1}^{\ell n} V(X_jx_k)\Big] \times \psi(X_1\ldots X_n x_1\ldots x_{\ell n}) \qquad (I.13)$$

on the space of ψ's. When transformed into the space of c's, it decomposes[2]) into an interatomic part V_o and an intra-atomic part V':

$$Vc(\alpha_1\ldots\alpha_n) = (V_o + V')c(\alpha_1\ldots\alpha_n),$$

$$V_o c(\alpha_1\ldots\alpha_n) = \sum_{p=1}^{n}\sum_{\alpha} (\alpha_p|V|\alpha)c(\alpha_1\ldots\alpha_{p-1}\alpha\alpha_{p+1}\ldots\alpha_n)$$

$$V'c(\alpha_1\ldots\alpha_n) = \sum_{p<q}^{n}\sum_{\alpha\beta} (\alpha_p\alpha_q|V|\alpha\beta) \times c(\alpha_1\ldots\alpha_{p-1}\alpha\alpha_{p+1}\ldots\alpha_{q-1}\beta\alpha_{q+1}\ldots\alpha_n) \qquad (I.14)$$

with

$$(\alpha_p|V|\alpha) = \int \varphi_{\alpha_p}^*(Xx_1\ldots x_\ell)\Big[\sum_{j<k}^{\ell} V(x_j x_k) + \sum_{j=1}^{\ell} V(Xx_j)\Big]$$

$$\times\ \varphi_\alpha(Xx_1\ldots x_\ell)\ dXdx_1\ldots dx_\ell, \qquad (I.15)$$

$$(\alpha_p\alpha_q|V|\alpha\beta) = \int \varphi_{\alpha_p}^*(Xx_1\ldots x_\ell)\,\varphi_{\alpha_q}^*(X'x_1'\ldots x_\ell')$$

$$\times\Big[V(XX')+\sum_{j=1}^{\ell}\sum_{k=1}^{\ell} V(x_j x_k')+\sum_{j=1}^{\ell} V(Xx_j')+\sum_{j=1}^{\ell} V(X'x_j)\Big]$$

$$\times\ \varphi_\alpha(Xx_1\ldots x_\ell)\,\varphi_\beta(X'x_1'\ldots x_\ell')\,dXdx_1\ldots dx_\ell\ dX'dx_1'\ldots dx_\ell'.$$

(I.16)

D. Atomic Second-Quantization Representation

A quantized field representation can now be introduced by any of the usual methods used for systems of elementary particles. We choose the Fock representation, in which state vectors $|c)$ are represented as

$$|c) = \begin{bmatrix} c_0 \\ c_1(\alpha_1) \\ \vdots \\ c_n(\alpha_1\ldots\alpha_n) \\ \vdots \end{bmatrix} \qquad (I.17)$$

with inner product

$$(c\,|\,c') = c_0^* c_0' + \sum_{n=1}^{\infty} \sum_{\alpha_1 \ldots \alpha_n} c_n^*(\alpha_1 \ldots \alpha_n) c_n'(\alpha_1 \ldots \alpha_n). \tag{I.18}$$

Here c_0 is the vacuum amplitude, c_1 the one-atom amplitude, etc. The atomic annihilation and creation operators are defined by [4])

$$a_\alpha \begin{bmatrix} c_0 \\ c_1(\alpha_1) \\ \cdot \\ \cdot \\ \cdot \\ c_n(\alpha_1 \ldots \alpha_n) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \equiv \begin{bmatrix} c_1(\alpha) \\ 2^{\frac{1}{2}} c_2(\alpha_1 \alpha) \\ \cdot \\ \cdot \\ \cdot \\ (n+1)^{\frac{1}{2}} c_{n+1}(\alpha_1 \ldots \alpha_n \alpha) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{I.19}$$

and

$$a_\alpha^\dagger \begin{bmatrix} c_0 \\ c_1(\alpha_1) \\ \cdot \\ \cdot \\ \cdot \\ c_n(\alpha_1 \ldots \alpha_n) \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \equiv \begin{bmatrix} 0 \\ \delta_{\alpha\alpha_1} c_0 \\ \cdot \\ \cdot \\ \cdot \\ (n^{\frac{1}{2}}/n!) \sum_P \epsilon(P) P\left[\delta_{\alpha\alpha_n} c_{n-1}(\alpha_1 \ldots \alpha_{n-1})\right] \\ \cdot \\ \cdot \\ \cdot \end{bmatrix} \tag{I.20}$$

where P runs over all permutations of $\alpha_1 \ldots \alpha_n$ and $\epsilon(P)$ is to be taken as +1 for all p if $2J+\ell$ is even, whereas if $2J+\ell$ is odd, then $\epsilon(P)$ is +1 or -1 depending on

whether P is even or odd. At this point it is customary for someone to object that the two definitions (I.19) and (I.20) should be interchanged. However, the definitions are correct as stated, as can be seen by referring to Fock[4]) or considering the special case of a state with only a vacuum amplitude c_0, which is then annihilated by a_α, whereas it is changed by $a_\alpha^\dagger$ into a state with only a one-atom amplitude. It follows directly from their definitions that these atomic annihilation and creation operators satisfy elementary Bose or Fermi commutation or anticommutation relations

$$a_\alpha a_\beta - (-1)^{2J+\ell} a_\beta a_\alpha = 0,$$

$$a_\alpha a_\beta^\dagger - (-1)^{2J+\ell} a_\beta^\dagger a_\alpha = \delta_{\alpha\beta} . \qquad \text{(I.21)}$$

These simple relations are to be contrasted with the more complicated relations

$$A_\alpha A_\beta^\dagger - (-1)^{2J+\ell} A_\beta A_\alpha = 0,$$

$$A_\alpha A_\beta^\dagger - (-1)^{2J+\ell} A_\beta^\dagger A_\alpha = \delta_{\alpha\beta} + \int ( )\, \underbrace{\psi^\dagger \ldots \psi}_{\ell}{}^\dagger \underbrace{\psi \ldots \psi}_{\ell}$$

factors factors

$$+ \ldots + \int ( )\, \psi^\dagger \psi \qquad \text{(I.22)}$$

satisfied by the more naive atomic annihilation and cretion operators defined by

$$A_\alpha^\dagger \equiv \int dX dx_1 \ldots dx_\ell \varphi_\alpha(X x_1 \ldots x_\ell)\, \psi^\dagger(X)\, \psi^\dagger(x_1) \ldots \psi^\dagger(x_\ell) \qquad \text{(I.23)}$$

where $\psi^\dagger(X)$ and $\psi^\dagger(x)$ are the usual quantized-field creation operators for a nucleus and an electron.

In terms of this atomic second-quantization representation, the state (I.17) can be represented as

$$|c) = \left[c_0 + \sum_{n=1}^{\infty} (n!)^{-\frac{1}{2}} \sum_{\alpha_1 \ldots \alpha_n} c(\alpha_1 \ldots \alpha_n) a_{\alpha_1}^{\dagger} \ldots a_{\alpha_n}^{\dagger}\right] |0) \qquad (I.24)$$

where $|0)$ is the normalized atomic vacuum state. The Hamiltonian (I.10)-(I.16) becomes

$$H = T + V_0 + V',$$

$$T = \sum_{\alpha\beta} (\alpha|T|\beta) a_{\alpha}^{\dagger} a_{\beta}, \quad V_0 = \sum_{\alpha\beta} (\alpha|V|\beta) a_{\alpha}^{\dagger} a_{\beta},$$

$$V' = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta|V|\gamma\delta) a_{\alpha}^{\dagger} a_{\beta}^{\dagger} a_{\delta} a_{\gamma} \qquad (I.25)$$

which is of the familiar form except that the matrix elements are between states of atoms rather than of "elementary" particles. If the φ_α are chosen to be free-atom energy eigenstates with eigenvalues ϵ_α, then it is not difficult to show that [2]) the single-atom part of H becomes diagonal:

$$H = H_0 + V',$$

$$H_0 = \sum_{\alpha} \epsilon_\alpha N_\alpha,$$

$$N_\alpha \equiv a_{\alpha}^{\dagger} a_{\alpha}. \qquad (I.26)$$

The subsidiary conditions (I.6) and (I.7) become in this representation

$$I_{elec} |c) = -\tfrac{1}{2} n(n-1) |c),$$

$$I_{nuc} |c) = (-1)^{2J} \tfrac{1}{2} n(n-1) |c) \qquad (I.27)$$

where

$$I_{elec} = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta|I_{elec}|\gamma\delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma,$$

$$I_{nuc} = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta|I_{nuc}|\gamma\delta)\ a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma. \qquad (I.28)$$

The zero-temperature n-atom problem is that of finding the simultaneous eigenstates of H, of the total atom- number operator

$$N = \sum_\alpha N_\alpha \qquad (I.29)$$

with eigenvalue n, and of the operators (I.28) with eigenvalues $-\frac{1}{2}n(n-1)$ and $(-1)^{2J}\frac{1}{2}n(n-1)$.

In practice, however, <u>exact</u> satisfaction of (I.27) is out of the question, since I_{elec} and I_{nuc} have the structure of interatomic interactions, and even strong ones. Furthermore, the continuum matrix elements are important in (I.27). It is possible to define projection operators for these subsidiary conditions which effectively replace them by additional exchange interaction terms in the Hamiltonian[3]), which could in principle then be treated in the same approximation as the true interatomic interaction V′. However, the continuum atomic matrix elements still cause difficulties in practice. Therefore, I shall not discuss this projection operator formalism. However, I understand that Professor Sakakura will have something to say about it next week, in connection with an approach to the formulation of a hybrid representation for systems of composite and elementary particles different from the approach which I will discuss in Sec. III of these lectures.

II. One Atom Plus One Extra Electron

A. Motivation

It is useful to examine a representation, for a system of one atom plus one extra electron, which is closely related to the many-atom representation just described, since, together with that representation, it can serve as a prototype for the more complex but also (hopefully) more useful hybrid representations which will be described in Sec. III.

B. Formulation

Consider first a single atom with ℓ electrons, described by the same complete orthonormal set $\{\varphi_\alpha(Xx_1 \ldots x_\ell)\}$ of atomic wave functions as used in Sec. I. If one extra electron is added to this system, the resultant system of one nucleus and $\ell + 1$ electrons is described by Schrödinger wave functions $\psi(Xx_1 \ldots x_{\ell+1})$ which are antisymmetric in all $\ell+1$ electron variables. Any such wave function can be expanded as follows:

$$\psi(Xx_1 \ldots x_{\ell+1}) = \sum_\alpha c_\alpha(x_{\ell+1})\ \varphi_\alpha(Xx_1 \ldots x_\ell) \qquad \text{(II.1)}$$

where

$$c_\alpha(x_{\ell+1}) = \int \varphi_\alpha^*(Xx_1 \ldots x_\ell)\, \psi(Xx_1 \ldots x_{\ell+1})\ dX dx_1 \ldots dx_\ell. \qquad \text{(II.2)}$$

ψ is automatically antisymmetric in $x_1 \ldots x_\ell$ since the φ_α are. However, the condition that it also be antisymmetric under exchanges of $x_{\ell+1}$ with any of the other x_j imposes a subsidiary condition on the amplitudes c_α analogous to (I.6) and (I.7). To derive it, note that

$$\psi(Xx_{\ell+1}x_2 \ldots x_\ell x_1) = -\psi(Xx_1 \ldots x_{\ell+1}). \qquad \text{(II.3)}$$

Inserting the expansion (II.1) in both sides, multiplying by $\varphi^*(Xx_1 \ldots x_\ell)$, and integrating, one finds that the c_α must satisfy

$$\sum_{\beta}\int K_{\alpha\beta}(x,x')c_{\beta}(x')dx' = -c_{\alpha}(x) \qquad \text{(II.4)}$$

where

$$K_{\alpha\beta}(x,x') \equiv \int \varphi_{\alpha}^{*}(Xx'x_{2}\ldots x_{\ell})\varphi_{\beta}(Xxx_{2}\ldots x_{\ell})dXdx_{2}\ldots dx_{\ell}. \qquad \text{(II.5)}$$

If we define $\underset{\sim}{c}(x)$ as the column vector with components $c_{\alpha}(x)$ and $\underset{\approx}{K}(x,x')$ as the matrix with elements $K_{\alpha\beta}(x,x')$, then (II.4) can be written as

$$\int \underset{\approx}{K}(x,x')\ \underset{\sim}{c}(x')\,dx' = -\underset{\sim}{c}(x), \qquad \text{(II.6)}$$

i.e. $\underset{\sim}{c}(x)$ must be an eigenvector of the hermitian "exchange kernel" $\underset{\approx}{K}$ with eigenvalue -1. Any $\underset{\sim}{c}(x)$ derived from a totally antisymmetric $\psi(Xx_{1}\ldots x_{\ell+1})$ according to (II.2) is automatically an eigenfunction with eigenvalue -1, i.e. (II.6) is a necessary condition for total antisymmetry of ψ. Conversely, if $\underset{\sim}{c}$ satisfies (II.6) then it follows from (II.1) and (II.2) that (II.3) is satisfied, i.e. ψ is totally antisymmetric. Thus the eigenvalue equation (II.6) is both necessary and sufficient for complete antisymmetry of ψ. Furthermore, the space of totally antisymmetric ψ's is in one-one correspondence with the space of $\underset{\sim}{c}$'s satisfying (II.6), and this correspondence preserves inner products:

$$(\psi,\psi') \equiv \int \psi^{*}(Xx_{1}\ldots x_{\ell+1})\psi'(Xx_{1}\ldots x_{\ell+1})dXdx_{1}\ldots dx_{\ell+1}$$

$$= (\underset{\sim}{c},\underset{\sim}{c}') \equiv \int \underset{\sim}{c}^{\dagger}(x)\ \underset{\sim}{c}'(x)dx$$

$$\equiv \sum_{\alpha}\int c_{\alpha}^{*}(x)\ c_{\alpha}'(x)dx. \qquad \text{(II.7)}$$

I leave the proof to you as an exercise. As in the case of (I.9), the subsidiary condition (II.6) plays an essential role in simplifying the expression for the inner product in the space of state vectors $\underset{\sim}{c}$. It might seem

surprising that the vector function $\underset{\sim}{c}(x)$ of a single electron variable x can contain precisely the same physical information as the Schrödinger wave function ψ which depends on the coördinates of a nucleus and ℓ+1 electrons. The point is, of course, that the extra information is built into $\underset{\sim}{c}$ via its dependence on the vector index α, the set of single-atom quantum numbers.

C. Representation of the Hamiltonian

In Schrödinger representation the Hamiltonian has the general form

$$H = T(X) + \sum_{j=1}^{\ell+1} T(x_j) + \sum_{j=1}^{\ell+1} V(Xx_j) + \sum_{j<k}^{\ell+1} V(x_j x_k). \quad \text{(II.8)}$$

To find the representation of H as an operator on state vectors $\underset{\sim}{c}$, let H operate on (II.1), multiply by $\varphi_\alpha^*(Xx_1 \ldots x_\ell)$, and integrate. In this way one finds that H acts as a matrix operator $\underset{\approx}{H}(x)$:

$$H \underset{\sim}{c}(x) = \underset{\approx}{H}(x)\underset{\sim}{c}(x) \quad \text{(II.9)}$$

where the notation $\underset{\approx}{H}(x)$ means that the matrix operator $\underset{\approx}{H}$ acts on the x dependence of $\underset{\sim}{c}(x)$, and the elements of this matrix operator are given by

$$H_{\alpha\beta}(x) = [\epsilon_\alpha + T(x)]\delta_{\alpha\beta} + V_{\alpha\beta}(x) \quad \text{(II.10)}$$

where the potential matrix operator $\underset{\approx}{V}(x)$ is defined as

$$V_{\alpha\beta}(x) \equiv \int \varphi_\alpha^*(Xx_1 \ldots x_\ell)\left[V(Xx) + \sum_{j=1}^{\ell} V(x_j x)\right]$$

$$\times \varphi_\beta(Xx_1 \ldots x_\ell)\, dX dx_1 \ldots dx_\ell \,. \quad \text{(II.11)}$$

For simplicity it has been assumed that the φ_α are chosen to be the free-atom energy eigenstates:

$$\Big[T(X) + \sum_{j=1}^{\ell} T(x_j) = \sum_{j=1}^{\ell} V(x_j x_k) \sum_{j<k}^{\ell} V(x_j x_k)\Big] \varphi_\alpha(X x_1 \ldots x_\ell)$$

$$= \epsilon_\alpha \varphi_\alpha(X x_1 \ldots x_\ell) \qquad \text{(II.12)}$$

as was assumed in (I.26). This assumption is not essential; if it is not made then the term $\epsilon_\alpha \delta_{\alpha\beta}$ in (II.10) is replaced by $(\alpha|T|\beta)+(\alpha|V|\beta)$, with matrix elements defined by (I.12) and (I.15).

D. Definition of Projected Hamiltonian

I would like now to sketch a method of satisfying the exchange subsidiary condition (II.6) through construction of an appropriate projection operator. The method is similar to the projection operator formalism[3]) alluded to in Sec. I, which was not discussed there because of lack of time. Here the system is simpler, so the projection operator is also simpler and I can at least very briefly sketch the ideas without giving any details of the proof.

We want to find the simultaneous eigenvectors $\underset{\sim}{c}$ of the Hamiltonian matrix $\underset{\approx}{H}$ [Eq.(II.10)] and the exchange matrix kernel $\underset{\approx}{K}$ [Eq.(II.6)], with eigenvalue -1 for the latter. One can reduce this to the problem of finding the eigenvectors of a suitable projected Hamiltonian, in which (II.6) is exactly replaced by an additional electron-atom exchange interaction. Let P be the projection operator onto the space of all eigenvectors $\underset{\sim}{c}(x)$ of (II.6) with the stated eigenvalue -1. This space is closed under the action of $\underset{\approx}{H}$ since the Schrödinger Hamiltonian is symmetric under permutations of electrons. Hence

$$[P,H] = 0 \qquad \text{(II.13)}$$

where H is defined by (II.9)-(II.11). Define a "projected Hamiltonian" $\mathcal{H}$ by

$$\mathcal{K} = PH. \tag{II.14}$$

Suppose that $\underset{\sim}{c}$ is an eigenvector of $\mathcal{K}$ with eigenvalue E:

$$\mathcal{K}\underset{\sim}{c} = E\underset{\sim}{c}. \tag{II.15}$$

Then if $E \neq 0$,

$$P\underset{\sim}{c} = E^{-1}P\mathcal{K}\underset{\sim}{c} = E^{-1}P^2H\underset{\sim}{c}$$

$$= E^{-1}PH\underset{\sim}{c} = E^{-1}\mathcal{K}\underset{\sim}{c} = \underset{\sim}{c} \tag{II.16}$$

and

$$E\underset{\sim}{c} = \mathcal{K}\underset{\sim}{c} = PH\underset{\sim}{c} = HP\underset{\sim}{c} = H\underset{\sim}{c} \tag{II.17}$$

Thus any eigenvector of $\mathcal{K}$ with a nonzero eigenvalue is necessarily also an eigenvector of P with eigenvalue 1, i.e. it satisfie (II.6); furthermore, it is also an eigenvector of H with eigenvalue E. On the other hand, the eigenstates of $\mathcal{K}$ belonging to the eigenvalue zero are in general linear combinations of eigenstates of H with eigenvalue zero and arbitrary states lying in the subspace orthogonal to the physical subspace, i.e. states annihilated by P. Thus we shall assume $E \neq 0$. This is no great loss of generality, since in application, states with E=0 will usually be of measure zero. E.g., in a scattering problem, E will be zero only if the incoming electron has kinetic energy precisely equal to the binding energy of the isolated atom. Nevertheless, we should bear in mind that when we work with the projected Hamiltonian $\mathcal{K}$, its eigenstates with eigenvalue zero will in general be physically meaningless.

E. Construction of the Projection Operator and Projected Hamiltonian

Let K be the integral operator with kernel $\underset{\sim}{\underset{\sim}{K}}$, i.e.

$$K\,\underset{\sim}{c}(x) \equiv \int \underset{\sim}{\underset{\sim}{K}}(x,x')\;\underset{\sim}{c}(x')\,dx'. \tag{II.18}$$

Define

$$K' \equiv 1 + K \tag{II.19}$$

where 1 is the unit operator. Then (II.6) can be written as

$$K'\underset{\sim}{c}(x) = 0. \tag{II.20}$$

Construct $(K')^2$:

$$(K')^2 = 1 + 2K + K^2. \tag{II.21}$$

Using the completeness relation (I.1) and some permutation algebra which I do not have time to go into here, one can show that the iterated kernel $\underset{\approx}{K}^2$ is <u>linear</u> in $\underset{\approx}{K}$, and in fact that

$$\sum_{\gamma}\int K_{\alpha\gamma}(x,x'')K_{\gamma\beta}(x'',x')$$

$$= \ell^{-1}\,[\delta_{\alpha\beta}\delta(x-x')-(\ell-1)K_{\alpha\beta}(x,x')] \tag{II.22}$$

or in terms of the notation (II.18)

$$K^2 = \ell^{-1}[1 - (\ell-1)K]. \tag{II.23}$$

Hence

$$(K')^2 = (1+\ell^{-1})K' \tag{II.24}$$

or

$$K'[K'-(1+\ell^{-1})] = 0, \tag{II.25}$$

from which it follows that K' has precisely two eigenvalues, zero and $1+\ell^{-1}$; only the eigenvectors with eigenvalue zero satisfy the subsidiary condition (II.6).

We can now easily write down the desired projection operator P:

$$P = 1-(1+\ell^{-1})^{-1}K' = (\ell+1)^{-1}-(1+\ell^{-1})^{-1}K. \tag{II.26}$$

Then one easily verifies that

$$P^2 = P, \quad K'P = PK' = 0 \tag{II.27}$$

so that (II.26) is indeed the projection operator for the subsidiary condition (II.6).

An explicit expression for the projected Hamiltonian (II.14) can then be constructed by multiplying (II.10) and (II.26) and again making use of the completeness relation for the φ_α. We give only the result:

$$\mathcal{H}\underset{\sim}{c}(x) = (\ell+1)^{-1}\underset{\approx}{H}(x)\underset{\sim}{c}(x) - (1+\ell^{-1})^{-1}\int \underset{\approx}{L}(x,x')\underset{\sim}{c}(x')dx' \tag{II.28}$$

with

$$L_{\alpha\beta}(x,x') \equiv K_{\alpha\beta}(x,x')\,[\epsilon_\beta + T(x')] + V_{\alpha\beta}(x,x'),$$

$$V_{\alpha\beta}(x,x') \equiv \int \varphi_\alpha^*(Xx'x_2\ldots x_\ell)\Big[V(Xx')+V(xx')+\sum_{j=2}^{\ell} V(x_j x')\Big] \times \varphi_\beta(Xxx_2\ldots x_\ell)\,dX dx_2\ldots dx_\ell. \tag{II.29}$$

This provides a formulation of such problems as electron-atom scattering which is in principle exact. However, the necessity of including continuum states of the atoms leads to difficulties in practical calculations. Hence this representation should be considered as the prototype of a more complex but more useful representation of the type to be discussed in Sec. III.

III. Systems of Elementary and Composite Particles

A. Historical Remarks

The representations to be discussed in this section are of very recent origin; in fact, most of the results I will describe have been obtained in the last month, and some of them are only a few days old. I apologize for exposing you to such undigested results. My excuse is that this area is currently being investigated more or less independently by several people, and the representations thus developed are likely to be applicable to a number of problems, not only such problems as partially ionized plasmas and superconductors, but also to such problems as atomic scattering and chemical and nuclear reactions. I tried many years ago, without success, to develop a hybrid representation in terms of bound states of atoms plus free-particle states of the unbound constituents (rather than continuum states of the atoms), but at the time I did not succeed. I am very much indebted to Professor Brittin for informing me of his recent work with Stolt on such a representation[5]), which convinced me that the problem is indeed soluble and motivated me to take it up again. The approach of Brittin and Stolt, based on correspondences between various Hilbert subspaces, is quite different from mine, and I do not know what the precise relationship is. I suspect, however, that the two approaches will eventually be found to be essentially equivalent. Perhaps Professor Brittin will shed some light on this in his forthcoming lectures. I have also recently learned that Professor Sakakura is working on the same problem from still a different point of view[6]), which he will describe in his lectures. Again, I suspect that his representation is essentially equivalent to mine, but this remains to be shown.

B. Motivation

As already mentioned in Sec. I, describing the unbound constituents in terms of continuum atomic states, though possible in principle, leads to difficulties in practical calculations since very little is known about such continuum states, and even when they are known (as for the hydrogen atom) they are still difficult to deal

with sufficiently accurately to provide a good approximation to the subsidiary conditions (I.6) and (I.7). Thus one is motivated to try to develop a hybrid representation in which only bound atomic states (and not necessarily even all of these bound states) are described in terms of atomic variables, with the remaining dependence of the wave functions described explicitly in terms of the constituents (e.g., in terms of plane-wave products).

C. Some Simple Cases

Start with the simplest case, n = 2 identical fermions. Let $\{\varphi_\alpha(x_1x_2)\}$ be an orthonormal (but not complete) set of antisymmetric bound-pair functions. We wish to expand a general, antisymmetric $\psi(x_1x_2)$ in terms of bound pairs and unbound fermions, i.e. we seek an expansion of the form

$$\psi(x_1x_2) = \sum_\alpha c(\alpha)\ \varphi_\alpha(x_1x_2) + c(x_1x_2) \qquad \text{(III.1)}$$

where $c(\alpha)$ is the amplitude for finding a bound pair of fermions in the state φ_α and $c(x_1x_2)$ is the amplitude for finding an unbound pair in the configuration (x_1x_2). It is obvious from the physical interpretation of amplitudes that one should choose

$$c(\alpha) \equiv \int \varphi_\alpha^*(x_1x_2)\ \psi(x_1x_2)dx_1dx_2. \qquad \text{(III.2)}$$

Then $c(x_1x_2)$ is uniquely determined as the residue, i.e. the result of subtracting off the "bound part" of ψ:

$$c(x_1x_2) = \psi(x_1x_2) - \sum_\alpha c(\alpha)\varphi_\alpha(x_1x_2), \qquad \text{(III.3)}$$

and (III.1) is satisfied as a trivial identity. Eq. (III.3) bears a strong resemblance to the definition of orthogonalized plane waves. This is no accident; one easily verifies that $c(x_1x_2)$ is indeed orthogonal to all the φ_α,

$$\int \varphi_\alpha^*(x_1x_2)\ c(x_1x_2)dx_1dx_2 = 0, \text{ all } \alpha, \qquad \text{(III.4)}$$

as it ought to be if $c(\alpha)$ and $c(x_1x_2)$ are to be interpreted as the amplitudes for bound pairs and unbound fermions, respectively.

In applications, however, it is desireable to turn the problem around, regarding $c(\alpha)$ and $c(x_1x_2)$ as the given wave functions; if one knew ψ already there would be little point in trying to explicitly introduce dynamical variables of the composite particles. The problem is then to show that the conditions that $c(x_1x_2)$ be antisymmetric and satisfy (III.4) (which we shall call the condition of bound state-continuum orthogonality) are both necessary and sufficient to uniquely determine the $c(\alpha)$ and $c(x_1x_2)$, thus establishing a one-one correspondence between the space of c's and the space of antisymmetric ψ's. In other words, we want to establish that the solution (III.2), (III.3) for the c's is the only one compatible with the antisymmetry of $c(x_1x_2)$ and its orthogonality to all the bound states. The demonstration is trivial: multiplication of (III.1) by φ_α^*, integration, and use of (III.4) yields the explicit (hence unique) expression (III.2) for $c(\alpha)$; then (III.1) yields the explicit and unique expression (III.3) for $c(x_1x_2)$. Finally, it follows from (III.1) that the inner product in the space of wave functions c is equal to the usual inner product in the space of ψ's:

$$(\psi,\psi') \equiv \int \psi^*(x_1x_2)\,\psi'(x_1x_2)\,dx_1\,dx_2 = (c,c')$$

$$\equiv \sum_\alpha c^*(\alpha)\; c'(\alpha) + \int c^*(x_1x_2)\,c'(x_1x_2)\,dx_1\,dx_2 . \qquad \text{(III.5)}$$

Having understood the trivial case n = 2, we can proceed to the less trivial case n = 3. We seek a unique expansion of a general antisymmetric $\psi(x_1x_2x_3)$ of the form

$$\psi(x_1x_2x_3)=A_3\sum_\alpha c(\alpha,x_1)\varphi_\alpha(x_2x_3) + c(x_1x_2x_3)$$

$$= \frac{1}{3}\sum_\alpha[c(\alpha,x_1)\varphi_\alpha(x_2x_3)-c(\alpha,x_2)\varphi_\alpha(x_1x_3)+c(\alpha,x_3)\varphi_\alpha(x_1x_2)]$$

$$+ c(x_1x_2x_3) \qquad \text{(III.6)}$$

where A_3 is the case n = 3 of the n-fermion antisymmetrizer

$$A_n = (n!)^{-1}\sum_P \epsilon(P)P, \qquad \text{(III.7)}$$

the sum runs over all n! permutations P of $x_1 \ldots x_n$, and $\epsilon(P)$ is +1 or -1 depending upon whether P is even or odd. The bound state-continuum orthogonality constraint analogous to (III.4) is

$$\int \varphi_\alpha^*(x_2x_3)c(x_1x_2x_3)dx_2dx_3 = 0, \text{ all } \alpha \text{ and } x_1. \qquad \text{(III.8)}$$

In the language of quantum chemistry we would say that $c(x_1x_2x_3)$ is required to be "strongly orthogonal" to all the bound states φ_α. Provided that we restrict ourselves to antisymmetric φ_α and $c(x_1x_2x_3)$, the similar relations obtained by permutation of the subscripts 1,2, and 3 are already implied by (III.8). This condition ensures that the $c(\alpha,x)$ represent only those configurations where two of the fermions are bound together, whereas $c(x_1x_2x_3)$ represents only those configurations where all three are unbound.

Multiplication of (III.6) by $\varphi_\alpha^*(x_2x_3)$, integration, and use of (III.8) and the antisymmetry of the φ_α and of $c(x_1x_2x_3)$ yields the following set of equations for the determination of the $c(\alpha,x)$:

$$\int \varphi_\alpha^*(x_2x_3)\psi(x_1x_2x_3)dx_2dx_3$$

$$= \frac{1}{3}\left[c(\alpha,x_1)-2\sum_\beta \int K(\alpha,x_1;\beta,x)\, c(\beta,x)dx\right]. \qquad \text{(III.9)}$$

Here the hermitian kernel K, which we shall call the "bound state-continuum exchange kernel" in view of the physical origin of its occurrence in (III.9), is defined as

$$K(\alpha,x_1;\beta,x_2) \equiv \int \varphi_\alpha^*(x_2x_3)\varphi_\beta(x_1x_3)dx_3. \qquad \text{(III.10)}$$

Defining

$$Kc(\alpha,x) \equiv \sum_\beta K(\alpha,x;\beta,y)\ c(\beta,y)dy, \qquad \text{(III.11)}$$

one can write (III.9) as

$$\tfrac{1}{3}(1-2K)\ c(\alpha,x) = \int \varphi_\alpha^*(x_2x_3)\,\psi(xx_2x_3)dx_2dx_3. \qquad \text{(III.12)}$$

This has a unique solution for $c(\alpha,x)$, denoted by

$$c(\alpha,x) = 3(1-2K)^{-1}\int \varphi_\alpha^*(x_2x_3)\,\psi(xx_2x_3)dx_2dx_3\ , \qquad \text{(III.13)}$$

provided only that K does not have the eigenvalue $\frac{1}{2}$, so that $(1-2K)^{-1}$ is non-singular. Then $c(x_1x_2x_3)$ is uniquely determined by (III.6) as the residue

$$\begin{aligned} c(x_1x_2x_3) &= \psi(x_1x_2x_3) - A_3 \sum_\alpha c(\alpha,x_1)\varphi_\alpha(x_2x_3) \\ &= \psi(x_1x_2x_3) - \tfrac{1}{3}\sum_\alpha [c(\alpha,x_1)\varphi_2(x_2x_3) \\ &\quad -c(\alpha,x_2)\varphi_\alpha(x_1x_3) + c(\alpha,x_3)\varphi_\alpha(x_1x_2)\,]. \end{aligned} \qquad \text{(III.14)}$$

It is easy to verify that (III.8) is indeed satisfied with these choices of $c(\alpha,x)$ and $c(x_1x_2x_3)$:

$$\int \varphi_\alpha^*(x_2x_3)\, c(x_1x_2x_3)\, dx_2dx_3$$

$$= \int \varphi_\alpha^*(x_2x_3)\, \psi(x_1x_2x_3)\, dx_2dx_3 - \tfrac{1}{3}(1-2K)\, c(\alpha,x_1) = 0. \qquad \text{(III.15)}$$

The inner product of two wave functions ψ and ψ' can be expressed in terms of the corresponding wave functions c and c', using (III.6), (III.8), (III.10), and (III.11), as

$$(\psi,\psi') \equiv \int \psi^*(x_1x_2x_3)\, \psi(x_1x_2x_3)\, dx_1dx_2dx_3$$

$$= (c,c')$$

$$\equiv \tfrac{1}{3} \sum_\alpha \int c^*(\alpha,x)\, (1-2k)\, c'(\alpha,x)\, dx$$

$$+ \int c^*(x_1x_2x_3)\, c'(x_1x_2x_3)\, dx_1dx_2dx_3. \qquad \text{(III.16)}$$

Given an antisymmetric ψ and antisymmetric φ_α, the $c(\alpha,x)$ and $c(x_1x_2x_3)$ are uniquely determined by (III.13) and (III.14). Conversely, given $c(\alpha,x)$ and an antisymmetric $c(x_1x_2x_3)$ satisfying the bound state-continuum orthogonality constraint (III.8), an antisymmetric ψ is uniquely determined via (III.6). There cannot be more than one such set of c's giving rise to a given ψ via (III.6) for suppose that there were two such sets, denoted by c_1 and c_2. Then

$$A_3 \left[\sum_\alpha c_2(\alpha,x_1)\, \varphi_\alpha(x_2x_3) \right] + c_1(x_1x_2x_3)$$

$$= A_3 \left[\sum_\alpha c_2(\alpha,x_1)\, \varphi_\alpha(x_2x_3) \right] + c_2(x_1x_2x_3). \qquad \text{(III.17)}$$

Then multiplication of both sides by $\varphi_\alpha^*(x_2x_3)$, integration, and use of (III.8) yields

$$\frac{1}{3}(1-2K)\, c_1(\alpha,x_1) = \frac{1}{3}(1-2K)\, c_2(\alpha,x_1) \qquad \text{(III.18)}$$

which, in view of the nonsingularity of the operator $(1-2K)^{-1}$, implies

$$c_1(\alpha,x_1) = c_2(\alpha,x_1). \qquad \text{(III.19)}$$

Then by III.14) and the fact that both c_1 and c_2 correspond to the same ψ, one has

$$c_1(x_1x_2x_3) = c_2(x_1x_2x_3). \qquad \text{(III.20)}$$

It follows that the space of wave functions $c(\alpha,x)$ and antisymmetric $c(x_1x_2x_3)$ satisfying (III.8) is in one-one correspondence with the space of physical (antisymmetric) ψ's.

We proceed next to the case n = 4, since some new features appear there. The generalization of (III.6) is

$$\psi(x_1..x_4) = A_4\Big[\sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)\varphi_{\alpha_1}(x_1x_2)$$

$$\times\ \varphi_{\alpha_2}(x_3x_4)+\sum_{\alpha} c(\alpha,x_1x_2)\varphi_\alpha(x_3x_4)\Big] + c(x_1..x_4)$$

$$= \frac{1}{3}\sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)\Big[\varphi_{\alpha_1}(x_1x_2)\varphi_{\alpha_2}(x_3x_4)-\varphi_{\alpha_1}(x_1x_3)\varphi_{\alpha_2}(x_2x_4)$$

$$+\ \varphi_{\alpha_1}(x_1x_4)\varphi_{\alpha_2}(x_2x_3)\Big]$$

$$+\frac{1}{6}\sum_{\alpha}\Big[c(\alpha,x_1x_2)\varphi_\alpha(x_3x_4) - c(\alpha,x_1x_3)\varphi_\alpha(x_2x_4)$$

$$+ c(\alpha,x_1x_4)\varphi_\alpha(x_2x_3) + c(\alpha,x_2x_3)\varphi_\alpha(x_1x_4)$$

$$- c(\alpha,x_2x_4)\varphi_\alpha(x_1x_3) + c(\alpha,x_3x_4)\varphi_\alpha(x_1x_2)\Big]$$

$$+ c(x_1..x_4). \qquad \text{(III.21)}$$

There are now two bound state-continuum orthogonality constraints, one on the $c(\alpha,x_1x_2)$ and one on $c(x_1..x_4)$:

$$\int \varphi_\alpha^*(x_1x_2)c(\beta,x_1x_2)dx_1dx_2 = 0, \text{ all } \alpha \text{ and } \beta;$$

$$\int \varphi_\alpha^*(x_1x_2)c(x_1x_2x_3x_4)dx_1dx_2 = 0, \text{ all } \alpha,x_3, \text{ and } x_4.$$

$$\text{(III.22)}$$

The $c(\alpha,x_1x_2)$ are required to be antisymmetric in x_1 and x_2, and $c(x_1...x_4)$ in $x_1..x_4$; also, since exchange of a fermion pair produces two sign changes, $c(\alpha_1\alpha_2)$ is required to be symmetric in α_1 and α_2. Multiplication of (III.21) by $\varphi_{\alpha_1}^*(x_1x_2)\varphi_{\alpha_2}^*(x_3x_4)$, integration, and use of (III.22) and the antisymmetry and symmetry properties of the c's yields

$$\int \varphi_{\alpha_1}^*(x_1x_2)\varphi_{\alpha_2}^*(x_3x_4)\psi(x_1..x_4)dx_1..dx_4$$

$$= \frac{1}{3}\Big[c(\alpha_1\alpha_2) - 2\sum_{\alpha\beta}(\alpha_1\alpha_2|I|\alpha\beta)c(\alpha\beta)\Big]$$

$$- \frac{2}{3}\sum_\alpha \int K(\alpha_1\alpha_2;\alpha,xy)\ c(\alpha,xy)dxdy \qquad \text{(III.23)}$$

where

$$(\alpha_1\alpha_2|I|\alpha\beta) \equiv \int \varphi_{\alpha_1}^*(x_1x_2)\varphi_{\alpha_2}^*(x_3x_4)\varphi_\alpha(x_3x_2)\varphi_\beta(x_1x_4)dx_1..dx_4 \qquad (III.24)$$

and

$$K(\alpha_1\alpha_2;\alpha,xy) \equiv \int \varphi_{\alpha_1}^*(xx_1)\varphi_{\alpha_2}^*(yx_2)\varphi_\alpha(x_1x_2)dx_1dx_2\,. \qquad (III.25)$$

Similarly, multiplication of (III.21) by $\varphi_\alpha^*(x_3x_4)$ and integration yields

$$\begin{aligned}\int \varphi_\alpha^*(x_3x_4)\psi(x_1..x_4)dx_3dx_4 &= \frac{1}{3}\Big[\sum_{\alpha_1}\varphi_{\alpha_1}(x_1x_2)c(\alpha_1\alpha) - 2\sum_{\alpha_1\alpha_2}K(\alpha,x_1x_2;\alpha_1\alpha_2)c(\alpha_1\alpha_2)\Big] \\ &+ \frac{1}{6}\Big[c(\alpha,x_1x_2) - 2\sum_\beta\int K(\alpha,x_2;\beta,x_3)c(\beta,x_1x_3)dx_3 \\ &+ 2\sum_\beta\int K(\alpha,x_1;\beta,x_3)c(\beta,x_2x_3)dx_3\Big] \qquad (III.26)\end{aligned}$$

where

$$K(\alpha,xy;\alpha_1\alpha_2) \equiv K^*(\alpha_1\alpha_2;\alpha,xy) \qquad (III.27)$$

and $K(\alpha,x_1;\beta,x_2)$ is defined by (III.10). The matrix (III.24) is the analog of the "exchange matrix" defined previously in Eq. (I.4), except that now it refers only to fermion exchange between bound states φ_α, whereas in (I.4) the φ_α included continuum states as well. The kernel (III.25), (III.27) corresponds to exchange of a pair of electrons between the continuum and a product of two bound states, one electron exchanging with each of the two bound states. It is analogous to a dynamical matrix element representing collision of two bound pairs, with one breaking up into two continuum fermions and the other remaining

bound (but in general changing its state), and the inverse process. Similarly, (III.10) is analogous to a dynamical matrix element representing collision of an unbound (continuum) fermion with a bound pair without breakup of the pair, whereas (III.24) is analogous to a dynamical matrix element representing collision of two bound pairs without breakup of either. In fact, we shall see later that these purely kinematical exchange effects give rise to terms in the second-quantized Hamiltonian which are quite analogous to dynamical terms.

Equations (III.23) and (III.26) are coupled linear, inhomogeneous equations for $c(\alpha_1\alpha_2)$ and $c(\alpha,x_1x_2)$ in terms of ψ, which can be denoted symbolically by

$$\underset{\approx}{L}\,\underset{\sim}{c} = \underset{\sim}{d} \qquad \text{(III.28)}$$

where $\underset{\sim}{c}$ is a many-component wave function which can be denoted by

$$\underset{\sim}{c} = \begin{bmatrix} c(\alpha_1\alpha_2) \\ c(\alpha,x_1x_2) \end{bmatrix} \qquad \text{(III.29)}$$

and $\underset{\sim}{d}$ is the inhomogeneity, denoted in the same representation by

$$\underset{\sim}{d} = \begin{bmatrix} d(\alpha_1\alpha_2) \\ d(\alpha,x_1x_2) \end{bmatrix} \qquad \text{(III.30)}$$

with

$$d(\alpha_1\alpha_2) = \int \varphi_{\alpha_1}^*(x_1x_2)\,\varphi_{\alpha_2}^*(x_3x_4)\,\psi(x_1..x_4)\,dx_1..dx_4,$$

$$d(\alpha,x_1x_2) = \int \varphi_{\alpha}^*(x_3x_4)\,\psi(x_1..x_4)\,dx_3dx_4. \qquad \text{(III.31)}$$

The linear matrix-integral operator $\underset{\approx}{L}$ is defined by the

right sides of (III.23) and (III.26). When we write down the expression for the inner product (c,c'), it will be easy to show, moreover, that $\underset{\approx}{L}$ is hermitian. Equations (III.23) and (III.26) possess a unique solution for $c(\alpha_1\alpha_2)$ and $c(\alpha,x_1x_2)$ provided only that $\underset{\approx}{L}$ does not possess the eigenvalue zero, so that $\underset{\approx}{L}^{-1}$ is nonsingular. The existence of such a zero eigenvalue would be an accident arising through an unfortunate choice of the φ_α, and we shall assume that they are chosen so that $\underset{\approx}{L}$ does not have a zero eigenvalue. Then $c(\alpha_1\alpha_2)$ and $c(\alpha,x_1x_2)$ are uniquely determined by ψ, and $c(x_1..x_4)$ subsequently follows uniquely from (III.21). By analogy with the verification of (III.22), one can show from (III.21), (III.23), (III.25), (III.26), and (III.10) that the bound state-continuum orthogonality constraints (III.22) are actually satisfied by the $c(\alpha_1\alpha_2)$ and $c(\alpha,x_1x_2)$ satisfying (III.23) and (III.26). Finally, the proof that there cannot be more than one choice of the c's satisfying (III.23) and (III.26), the constraints (III.22), and the proper antisymmetry and symmetry conditions and leading to the same ψ via (III.21) can be carried out in analogy with (III.17) -(III.20). Thus the space of wave functions c is in one-one correspondence with the space of physical (antisymmetric) ψ's. The inner product in the space of the c's is found to be

$$(\psi,\psi') \equiv \int \psi^*(x_1..x_4)\,\psi'(x_1..x_4)\,dx_1..dx_4$$

$$= (c,c') \equiv (\underset{\sim}{c},\ \underset{\approx}{L}\ \underset{\sim}{c}') + \int c^*(x_1..x_4)\,c'(x_1..x_4)\,dx_1..dx_4$$

$$\equiv \sum_{\alpha_1\alpha_2} c^*(\alpha_1\alpha_2)\Big[\frac{1}{3}c'(\alpha_1\alpha_2) - \frac{2}{3}\sum_{\alpha\beta}(\alpha_1\alpha_2\,|I|\,\alpha\beta)\,c'(\alpha\beta)$$

$$- \frac{2}{3}\sum_{\alpha}\int K(\alpha_1\alpha_2;\alpha,xy)\,c'(\alpha,xy)\,dxdy\Big]$$

$$+\sum_{\alpha}\int dx_1\,dx_2\,c^*(\alpha,x_1x_2)\Big[\frac{1}{6}\,c'(\alpha,x_1x_2)$$

$$- \frac{2}{3} \sum_{\beta} \int K(\alpha, x_2; \beta, x_3)\ c'(\beta, x_1 x_3)\, dx_3$$

$$- \frac{2}{3} \sum_{\alpha_1 \alpha_2} K(\alpha, x_1 x_2; \alpha_1 \alpha_2)\ c'(\alpha_1 \alpha_2) \Big]$$

$$+ \int c^*(x_1 .. x_4)\ c'(x_1 .. x_4)\, dx_1 .. dx_4 . \qquad \text{(III.32)}$$

It is easy to verify from this expression that the operator $\underset{\approx}{L}$ is hermitian, as previously stated.

As a simple example of the physical interpretation of the formalism, let us compute the c's for the case that ψ is a four-fermion function built from products of two-fermion bound states, with both pairs in the same state φ_0:

$$\psi(x_1 .. x_4) = A_4 [\varphi_0(x_1 x_2)\, \varphi_0(x_3 x_4)]$$

$$= \frac{1}{3} [\varphi_0(x_1 x_2)\, \varphi_0(x_3 x_4) - \varphi_0(x_1 x_3)\, \varphi_0(x_2 x_4)$$

$$+ \varphi_0(x_1 x_4)\, \varphi_0(x_2 x_3)]. \qquad \text{(III.33)}$$

Such a state is, in fact, of BCS form, with φ_0 playing the role of the Cooper pair wave function. Then

$$\int \varphi_{\alpha_1}^*(x_1 x_2)\, \varphi_{\alpha_2}^*(x_3 x_4)\, \psi(x_1 .. x_4)\, dx_1 .. dx_4$$

$$= \frac{1}{3} [\delta_{\alpha_1 0} \delta_{\alpha_2 0} - 2(\alpha_1 \alpha_2 | I | 00)],$$

$$\int \varphi_{\alpha}^*(x_3 x_4)\, \psi(x_1 .. x_4)\, dx_3 dx_4$$

$$= \frac{1}{3} [\delta_{\alpha 0} \varphi_0(x_1 x_2) - 2K(\alpha, x_1 x_2; 00)]. \qquad \text{(III.34)}$$

It is easily verified by substitution that the solutions of Eqs. (III.23) and (III.26) are then

$$c(\alpha_1\alpha_2) = \delta_{\alpha_1 o}\delta_{\alpha_2 o},$$

$$c(\alpha, x_1 x_2) = 0. \qquad \text{(III.35)}$$

Then by (III.21)

$$c(x_1..x_4) = 0. \qquad \text{(III.36)}$$

The state (III.33) is therefore one in which the probability of finding two bound pairs both in the state φ_o is unity, the probability of finding two bound pairs not both in the state φ_o is zero, and the probabilities of finding only one bound pair or no bound pairs is zero. This is exactly what one would naively expect for a state built only from the single pair state φ_o. The formalism is such that the complications arising from the antisymmetrization in (III.33) do not upset this naive expectation. This is not a trivial point, since an exchanged product such as $\varphi_o(x_1x_3)\varphi_o(x_2x_4)$ is not orthogonal to unexchanged products $\varphi_\alpha(x_1x_2)\varphi_\beta(x_3x_4)$ with $\alpha \neq 0$ and/or $\beta \neq 0$.

More generally, suppose that $\psi(x_1..x_4)$ is built purely from bound pairs with no continuum amplitudes, i.e.

$$\psi(x_1..x_4) = A_4 \sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)\varphi_{\alpha_1}(x_1x_2)\varphi_{\alpha_2}(x_3x_4)$$

$$= \frac{1}{3}\sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)[\varphi_{\alpha_1}(x_1x_2)\varphi_{\alpha_2}(x_3x_4) - \varphi_{\alpha_1}(x_1x_3)\varphi_{\alpha_2}(x_2x_4)$$

$$+ \varphi_{\alpha_1}(x_1x_4)\varphi_{\alpha_2}(x_2x_3)]. \qquad \text{(III.37)}$$

Then it is easily verified that (III.21), (III.23), and (III.26) are satisfied for arbitrary symmetric $c(\alpha_1\alpha_2)$ in (III.37) and the same $c(\alpha_1\alpha_2)$ in (III.21), (III.23), and (III.26), provided that

$$c(\alpha, x_1 x_2) = 0,$$

$$c(x_1 .. x_4) = 0. \qquad \text{(III.38)}$$

For two such ψ's, denoted by ψ and ψ', the inner product expression (III.32) reduces to

$$(\psi, \psi') = (c, c')$$

$$= \frac{1}{3} \sum_{\alpha_1 \alpha_2} c^*(\alpha_1 \alpha_2)\, c'(\alpha_1 \alpha_2)$$

$$- \frac{2}{3} \sum_{\alpha_1 \alpha_2} \sum_{\alpha\beta} c^*(\alpha_1 \alpha_2)(\alpha_1 \alpha_2 | I | \alpha\beta) c'(\alpha\beta). \qquad \text{(III.39)}$$

The norm of a single state ψ can thus be written as

$$(\psi, \psi) = (c, c) = \sum_{\alpha_1 \alpha_2} P(\alpha_1 \alpha_2),$$

$$P(\alpha_1 \alpha_2) = \frac{1}{3} \Big[|c(\alpha_1 \alpha_2)|^2 - c^*(\alpha_1 \alpha_2) \sum_{\alpha\beta} (\alpha_1 \alpha_2 | I | \alpha\beta) c(\alpha\beta)$$

$$- c(\alpha_1 \alpha_2) \sum_{\alpha\beta} (\alpha\beta | I | \alpha_1 \alpha_2) c^*(\alpha\beta) \Big] \qquad \text{(III.40)}$$

so that if the state ψ is normalized, then $P(\alpha_1 \alpha_2)$ has an obvious physical interpretation as the probability of finding the two bound pairs in the pair states φ_α and φ_β. This probability differs from $|c(\alpha_1 \alpha_2)|^2$ due to the effects of antisymmetrization in (III.37). The distribution function $n(\alpha)$, the mean occupation number of the pair state φ_α in the state ψ, is then

$$n(\alpha) = 2 \sum_\beta P(\alpha\beta)$$

$$= \frac{2}{3} \sum_\beta \Big[|c(\alpha\beta)|^2 - c^*(\alpha\beta) \sum_{\alpha_1 \alpha_2} (\alpha\beta | I | \alpha_1 \alpha_2) c(\alpha_1 \alpha_2)$$

$$-c(\alpha\beta)\sum_{\alpha_1\alpha_2}(\alpha_1\alpha_2|I|\alpha\beta)c^*(\alpha_1\alpha_2)\Big] \qquad \text{(III.41)}$$

where the factor 2 arises from the requirement that if four fermions are bound into two fermion pairs in various states of excitation, then one must have

$$\sum_{\alpha} n(\alpha) = 2. \qquad \text{(III.42)}$$

In the case where there are nonvanishing amplitudes for finding only one bound pair or no bound pairs present, so that the more general wave function (III.21) must be used, the expressions for mean occupation numbers are more complicated. The discussion of such cases is best postponed until after introduction of second quantization, in terms of which occupation numbers are more simply expressed.

D. Systems of Atoms, Electrons, and Nuclei

We are now in a position to deal with more realistic cases. In a gas, liquid, or solid, there are present atoms and/or molecules in various states of real or virtual excitation and translational and rotational motion. In a plasma, there are in general several species of composite particles plus "elementary" particles, namely neutral atoms and/or molecules in various states, singly-ionized ions, doubly-ionized ions,..., and unbound electrons and (at sufficiently high temperatures) nuclei. In order to avoid unnecessarily complicating the formalism before understanding the essential features, we shall restrict ourselves here to the case that the composite particles are of a single species, each composed of one nucleus, whose spin and position variables are denoted by X, and one electron, whose spin and position variables are denoted by x. Thus we have an orthonormal set $\{\varphi_\alpha\}$ of bound states, where $\varphi_\alpha=\varphi_\alpha(X\ x)$. Such a description would be applicable, e.g. to hydrogen at high enough temperatures that virtually all H_2 molecules are dissociated. There is no upper limit on the temperature, since we shall explicitly include the possibility of dissociation of the atoms. The notation for the φ_α is the same as

the case $\ell=1$ of Sec. I, except for the important difference that the set $\{\varphi_\alpha\}$ is now undercomplete, since it contains only bound states. Nevertheless, it may still be an infinite set. Thus, e.g. the atomic quantum numbers α include not only the internal quantum numbers n,ℓ,m, and s, but also the translational quantum number $\underset{\sim}{k}$, related to the total momentum $\underset{\sim}{p}$ of the atom by $\underset{\sim}{p}=\hbar\underset{\sim}{k}$. In realistic cases it will be necessary to consider a large number of $\underset{\sim}{k}$ values, proportional to the volume of the system and hence becoming infinite in the thermodynamic limit. Since the set $\{\varphi_\alpha\}$ is nevertheless undercomplete as are products of such bound states, the overcompleteness problems that plagued us in Secs. I and II will now be absent, being prevented here by the constraint of bound state-continuum orthogonality. The overcompleteness problems in Sec. I arose because the set of all atomic product states is overcomplete (as a result of the effects of exchange) provided that all continuum states of the atoms are included. Here, however, we shall include only bound states, treating the continuum states in terms of their "elementary" constituents (here nuclei and electrons).

Suppose that we are dealing with a system of n protons and n electrons, described by Schrödinger wave functions $\psi(X_1\ldots X_n x_1\ldots x_n)$. The obvious generalization of the expansion (III.21) is

$$\psi(X_1\ldots X_n x_1\ldots x_n)$$

$$= n!\, A_n^{(\mathrm{nuc})}\, A_n^{(\mathrm{elec})} \Big\{(n!)^{-\frac{1}{2}} \sum_{\alpha_1\ldots\alpha_n} c(\alpha_1\ldots\alpha_n) \times \varphi_{\alpha_1}(X_1x_1)\ldots\varphi_{\alpha_n}(X_nx_n) + \sum_{j=1}^{n-1} (j!)^{-1}[(n-j)!]^{-\frac{1}{2}} \times \sum_{\alpha_1\ldots\alpha_{n-j}} c(\alpha_1\ldots\alpha_{n-j}, X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n) \times \varphi_{\alpha_1}(X_1x_1)\ldots\varphi_{\alpha_{n-j}}(X_{n-j}x_{n-j})\Big\} + c(X_1\ldots X_n x_1\ldots x_n) \qquad \text{(III.43)}$$

where $A_n^{(nuc)}$ and $A_n^{(elec)}$ are the proton and electron antisymmetrizers defined as in (II.7), and the combinatorial factors are inserted so as to simplify the inner product expression which will be derived subsequently. The c's are required to be antisymmetric in the X's, antisymmetric in the x's, and symmetric in the α's. In addition, they are required to be strongly orthogonal to the bound states, i.e.

$$\int \varphi_\alpha^*(X_n x_n)\, c(\alpha_1 \ldots \alpha_{n-1}, X_n x_n)\, dX_n dx_n = 0,$$

$$\int \varphi_\alpha^*(X_n x_n)\, c(\alpha_1 \ldots \alpha_{n-2}, X_{n-1} X_n x_{n-1} x_n)\, dX_n dx_n = 0,$$

$$\vdots$$

$$\int \varphi_\alpha^*(X_n x_n)\, c(X_1 \ldots X_n x_1 \ldots x_n)\, dX_n dx_n = 0 \qquad \text{(III.44)}$$

as identities in the α's and the unintegrated X's and x's. These requirements are the obvious generalization of (III.22) and ensure that the dependence of the c's on the X's and x's refers only to continuum (unbound) nuclei and electrons; furthermore, they will serve to make the solution for the c's in terms of ψ unique.

The equations determining the amplitudes c can be derived in analogy with the derivation of (III.23) and (III.26), by multiplying (III.43) by $\varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_{n-j}}^*(X_{n-j} x_{n-j})$ and integrating over $X_1 \ldots X_{n-j} x_1 \ldots x_{n-j}$, for each value of j from 0 to n-1. In this way one obtains a set of n coupled linear, inhomogeneous equations for the n amplitudes $c(\alpha_1 \ldots \alpha_n)$, $c(\alpha_1 \ldots \alpha_{n-1}, X_n x_n), \ldots,$ $c(\alpha_1, X_2 \ldots X_n x_2 \ldots x_n)$ of the form

$$M\, c(\alpha_1 \ldots \alpha_n) = (n!)^{\frac{1}{2}} \int \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_n}^*(X_n x_n)\, \psi(X_1 \ldots X_n x_1 \ldots x_n) \times dX_1 \ldots dX_n dx_1 \ldots dx_n,$$

$$\vdots$$

$$M\, c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= n!(j!)^{-1}[(n-j)!]^{-\frac{1}{2}} \int \varphi_{\alpha_1}^*(X_1 x_1)\ldots\varphi_{\alpha_{n-j}}^*(X_{n-j}x_{n-j})$$

$$\times \psi(X_1\ldots X_n x_1\ldots x_n)\,dX_1\ldots dX_{n-j}\,dx_1\ldots dx_{n-j},$$

$$\vdots$$

$$M\,c(\alpha_1, X_2\ldots X_n x_2\ldots x_n) = n\int \varphi_{\alpha_1}^*(X_1 x_1)\,\psi(X_1\ldots X_n x_1\ldots x_n)\,dX_1\,dx_1 .$$

(III.45)

Here M is a linear, hermitian operator on the c's whose explicit form* is to be obtained by substitution of (III.43) into the right sides of the equations and evaluation of the integrals, dropping terms which vanish as a result of (III.44). At this point we do not need the explicit expressions; it is sufficient to realize that as before, the solution for the c's is unique, i.e. M^{-1} is non-singular** and uniquely defined. Hence, imposition of the requirements that the c's have the proper symmetry and antisymmetry and satisfy the bound state-continuum orthogonality constraints (III.44) serves to uniquely determine the c's and establish a one-one correspondence between the space of physical ψ's and the space of c's. Eqs. (III.45) only determine the c's with $0 \le j \le n-1$; as before, $c(X_1\ldots X_n x_1\ldots x_n)$ is then uniquely determined by (III.43).

We shall require, however, explicit expressions for the most important terms in the inner product (c,c'). In

---

* Note that the explicit expression for $M\,c(\alpha_1\ldots\alpha_{n-j}, X_{n-j+1}\ldots X_n x_{n-j+1}x_n)$ consists of terms linear in the various $c(\alpha_1\ldots n_{n-k}, X_{n-k+1}\ldots X_n x_{n-k+1}\ldots x_n)$, not merely $k = j$.

** Note that if the constraints (III.44) were not imposed, then the solution would not be unique, i.e. M would have a zero eigenvalue.

evaluating (ψ,ψ'), one need include only one factor of $A_n(\text{nuc})$ and one factor of $A_n(\text{elec})$ in the integrand, since these antisymmetrizing operators are projection operators and hence idempotent. Thus by (III.43), the inner product is

$$(c,c') \equiv (\psi,\psi')$$

$$= \int dX_1 \ldots dX_n dx_1 \ldots dx_n \Big\{ (n!)^{\frac{1}{2}} \sum_{\alpha_1 \ldots \alpha_n}$$

$$\times\ c^*(\alpha_1 \ldots \alpha_n) \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_n}^*(X_n x_n) + n! \sum_{j=1}^{n-1} (j!)^{-1} [(n-j)!]^{-\frac{1}{2}}$$

$$\times \sum_{\alpha_1 \ldots \alpha_{n-j}} c^*(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$\times\ \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_{n-j}}^*(X_{n-j} x_{n-j}) + c^*(X_1 \ldots X_n x_1 \ldots x_n) \Big\}$$

$$\times\ A_n(\text{nuc}) A_n(\text{elec}) \Big\{ (n!)^{\frac{1}{2}} \sum_{\beta_1 \ldots \beta_n} c'(\beta_1 \ldots \beta_n) \varphi_{\beta_1}(X_1 x_1) \ldots \varphi_{\beta_n}(X_n x_n)$$

$$+ n! \sum_{\ell=1}^{n-1} (\ell!)^{-1} [(n-\ell)!]^{-\frac{1}{2}} \sum_{\beta_1 \ldots \beta_{n-\ell}}$$

$$\times\ c'(\beta_1 \ldots \beta_{n-\ell}, X_{n-\ell+1} \ldots X_n x_{n-\ell+1} \ldots x_n)$$

$$\times\ \varphi_{\beta_1}(X_1 x_1) \ldots \varphi_{\beta_{n-\ell}}(X_{n-\ell} x_{n-\ell}) + c'(X_1 \ldots X_n x_1 \ldots x_n) \Big\}$$

(III.46)

This inner product is conveniently decomposed as follows:

$$(c,c')=(c,c')_0+(c,c')_1+\ldots\ . \qquad \text{(III.47)}$$

Here $(c,c')_0$ arises from all those terms in (III.46) in which the arguments (X_jx_j) of the various φ_{α_j} factors are either not permuted at all, or else the pairs (X_jx_j) are permuted bodily between φ_α's without breakup. Also, the X and x arguments of the c's may be permuted freely, but permutations exchanging arguments of the c's with those of the φ_α's are excluded. Taking proper account of the combinatorial factors arising in this way, one finds

$$\begin{aligned}(c,c')_0 &= \sum_{\alpha_1\ldots\alpha_n} c^*(\alpha_1\ldots\alpha_n)c'(\alpha_1\ldots\alpha_n)\\ &+\sum_{j=1}^{n-1}\sum_{\alpha_1\ldots\alpha_{n-j}} c^*(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)\\ &\times c'(\alpha_1\ldots\alpha_{n-j},X_{n=j+1}\ldots X_n x_{n-j+1}\ldots x_n)\,dX_{n-j+1}\ldots dX_n\\ &\times dx_{n-j+1}\ldots dx_n\\ &+\int c^*(X_1\ldots X_n x_1\ldots x_n)c'(X_1\ldots X_n x_1\ldots x_n)\,dX_1\ldots dX_n\,dx_1\ldots dx_n.\end{aligned} \qquad \text{(III.48)}$$

The combinatorial factors in (III.43) were chosen so that such prefactors do not appear in (III.48).

The term $(c,c')_1$ in (III.47) is defined to be the sum of all terms in (III.46) expressible solely in terms of exchange matrices and kernels arising from single exchange of protons or electrons between atoms, or single exchange of a proton or electron between an atom and the

continuum. Such exchanges may be depicted schematically as in Fig. 1, in which the circles represent atoms, the heavy dots protons, the light dots electrons, and the lines with arrows permutation cycles of length 2 (single interchanges). In addition, arbitrary permutations of proton-electron pairs between atoms, and arbitrary permutations of protons and electrons entirely within the continuum, are allowed. These are not indicated in the figure since they do not change the value of the matrix elements (although they do contribute to the combinatorial factors).

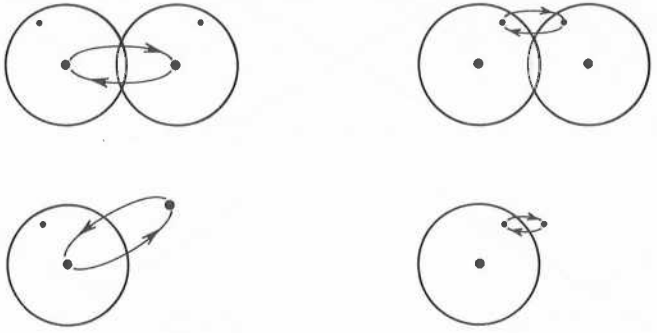


Fig. 1. Permutations contributing to $(c,c')_1$. Their contribution is only appreciable when atoms overlap each other or a continuum particle.

All such terms are expressible linearly in terms of the exchange matrix

$$(\alpha_1\alpha_2|I|\alpha\beta) \equiv \int \varphi_{\alpha_1}^*(X_1x_1)\,\varphi_{\alpha_2}^*(X_2x_2)\,\varphi_{\alpha}(X_1x_2)\,\varphi_{\beta}(X_2x_1) \times dX_1\,dX_2\,dx_1\,dx_2 \qquad \text{(III.49)}$$

and the exchange kernels

$$K(\alpha,X_1;\beta,X_2) \equiv \int \varphi_\alpha^*(X_2x)\,\varphi_\beta(X_1x)\,dx,$$

$$K(\alpha,x_1;\beta,x_2) \equiv \int \varphi_\alpha^*(Xx_2)\,\varphi_\beta(Xx_1)\,dX,$$

$$K(\alpha_1\alpha_2;\alpha,Xx) \equiv \int \varphi_{\alpha_1}^*(Xx_1)\,\varphi_{\alpha_2}^*(X_1x)\,\varphi_\alpha(X_1x_1)\,dX_1\,dx_1,$$

$$K(\alpha,Xx;\alpha_1\alpha_2) \equiv K^*(\alpha_1\alpha_2;\alpha,Xx) \qquad \text{(III.50)}$$

analogous to (III.24), (III.10), (III.25), and (III.27).

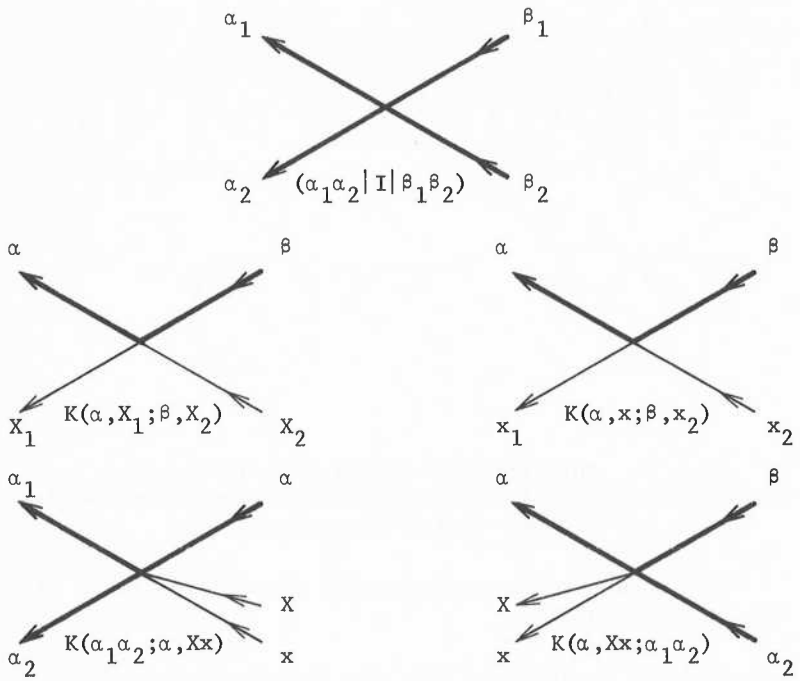


Fig. 2. Dynamical analogs of binary exchange process contributing to $(c,c')_1$. Heavy lines denote atoms; light lines, protons or electrons.

Non-negligible contributions to the exchange matrix (III.49) arise only from those regions of configuration space where two atoms collide (a binary collision). Similarly, non-negligible contributions to $K(\alpha,X_1;\beta,X_2)$ and $K(\alpha,x_1;\beta,x_2)$ arise only from regions of configuration space where one atom collides with one proton or one electron, and non-negligible contributions to $K(\alpha,Xx;\alpha_1\alpha_2)$ arise only from regions of configuration space where two atoms collide, with one breaking up into a proton and an electron (exchange cannot induce simultaneous breakup of both of the colliding atoms, although such terms will be found to occur in the Hamiltonian as a result of true dynamical interaction). The representation of these contributions to $(c,c')_1$ in terms of diagrams is shown in Fig. 2. Note that $K(\alpha_1\alpha_2;\alpha,Xx)$ corresponds to three-body collisions (one atom, one proton, and one electron). Nevertheless, we have chosen, rather arbitrarily, to include its contribution in $(c,c')_1$, since the inverse process, corresponding to $K(\alpha,Xx;\alpha_1\alpha_2)$, corresponds to only binary collisions (two atoms). The conbinatorics required to evaluate the coefficients of these exchange matrices and kernels are rather involved, so we give only the results:*

$$(c,c')_1 = -\tfrac{1}{2}n(n-1)\sum_{\alpha_i\ldots\alpha_n}\sum_{\beta_1\beta_2} c^*(\alpha_1\ldots\alpha_n)(\alpha_1\alpha_2|I|\beta_1\beta_2) \times c'(\beta_1\beta_2\alpha_3\ldots\alpha_n)$$

$$-\sum_{j=1}^{n-2}\tfrac{1}{2}(n-j)(n-j-1)\sum_{\alpha_1\ldots\alpha_{n-j}}\sum_{\beta_1\beta_2} \times \int c^*(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)$$

*Some errors in the combinatorial coefficients have been corrected since the lectures were presented, and the discussion of (III.51) has been modified accordingly.

$$\times(\alpha_1\alpha_2|I|\beta_1\beta_2)c'(\beta_1\beta_2\alpha_3\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)$$

$$\times dX_{n-j+1}\ldots dX_n dx_{n-j+1}\ldots dx_n$$

$$-\sum_{j=1}^{n-1} j(n-j)\sum_{\alpha_1\ldots\alpha_{n-j}}\sum_{\beta}$$

$$\times\int c^*(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)$$

$$\times\Big[\int K(\alpha_1,X_n;\beta,X)c'(\beta\alpha_2\ldots\alpha_{n-j},X_{n-j+1}\ldots X_{n-1}Xx_{n-j+1}\ldots x_n)dX$$

$$+\int K(\alpha_1,x_n;\beta,x)c'(\beta\alpha_2\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_{n-1}x)dx\Big]$$

$$\times dX_{n-j+1}\ldots dX_n dx_{n-j+1}\ldots dx_n$$

$$-\sum_{j=0}^{n-2}(j+1)(n-j)^{\frac{1}{2}}(n-j-1)\sum_{\alpha_1\ldots\alpha_{n-j-1}}\sum_{\beta_1\beta_2}$$

$$\times\int c^*(\alpha_1\ldots\alpha_{n-j-1},X_{n-j}\ldots X_n x_{n-j}\ldots x_n)$$

$$\times K(\alpha_{n-j-1},X_{n-j}x_{n-j};\beta_1\beta_2)$$

$$\times c'(\alpha_1\ldots\alpha_{n-j-2}\beta_1\beta_2,X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)$$

$$\times dX_{n-j}\ldots dX_n dx_{n-j}\ldots dx_n$$

$$- \sum_{j=0}^{n-2} (j+1)(n-j)^{\frac{1}{2}}(n-j-1) \sum_{\alpha_1 \ldots \alpha_{n-j}} \sum_{\beta}$$

$$\times \int c^*(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$\times K(\alpha_1 \alpha_2; \beta, X_{n-j} x_{n-j}) c'(\beta \alpha_3 \ldots \alpha_{n-j}, X_{n-j} \ldots X_n x_{n-j} \ldots x_n)$$

$$\times dX_{n-j} \ldots dX_n dx_{n-j} \ldots dx_n. \qquad \text{(III.51)}$$

The next term, $(c,c')_2$, in (III.47) is more complicated. It arises from both ternary exchange and singly-iterated binary exchange. By "ternary exchange" we mean terms in (III.46) arising from permutation cycles running through three atoms; such terms only become effective upon ternary collisions of atoms. An example of such a contribution, arising from permutation cycles of length three running through three atoms involved in the amplitudes $c(\alpha_1 \ldots \alpha_n)$, is

$$\frac{n(n-1)(n-2)}{3!} \sum_{\alpha_1 \ldots \alpha_n} \sum_{\beta_1 \beta_2 \beta_3} c^*(\alpha_1 \ldots \alpha_n)(\alpha_1 \alpha_2 \alpha_3 | I | \beta_1 \beta_2 \beta_3)$$

$$\times c'(\beta_1 \beta_2 \beta_3 \alpha_4 \ldots \alpha_n) \qquad \text{(III.52)}$$

where the ternary exchange matrix is defined as

$$(\alpha_1 \alpha_2 \alpha_3 | I | \beta_1 \beta_2 \beta_3) \equiv \int \varphi_{\alpha_1}^*(X_1 x_1) \varphi_{\alpha_2}^*(X_2 x_2) \varphi_{\alpha_3}^*(X_3 x_3)$$

$$\times \left[\varphi_{\beta_1}(X_1 x_2) \varphi_{\beta_2}(X_2 x_3) \varphi_{\beta_3}(X_3 x_1) + \right.$$

$$+ \varphi_{\beta_1}(X_1 x_3)\varphi_{\beta_2}(X_2 x_1)\varphi_{\beta_3}(X_3 x_2)\Big] \, dX_1 dX_2 dX_3 dx_1 dx_2 dx_3 . \tag{III.53}$$

The contribution (III.52) includes the combinatorial factor, i.e. it has already been summed over all electron and proton cycles of length three.

The "singly-iterated binary exchange" contributions to $(c,c')_2$ arise from terms in (III.46) involving two disjoint permutation cycles of length two. As an example, such contributions involving only the amplitudes $c(\alpha_1 \ldots \alpha_n)$ and arising from two disjoint exchanges of electrons and/or protons sum to

$$\frac{n(n-1)(n-2)(n-3)}{(2!)^2} \sum_{\alpha_1 \ldots \alpha_n} \sum_{\beta_1 \ldots \beta_4} \times c^*(\alpha_1 \ldots \alpha_n)(\alpha_1 \alpha_2 | I | \beta_1 \beta_2)(\alpha_3 \alpha_4 | I | \beta_3 \beta_4) \times c'(\beta_1 \ldots \beta_4 \ \alpha_5 \ldots \alpha_n). \tag{III.54}$$

Let us next estimate the magnitudes of these various contributions. To simplify the discussion, assume that $c(\alpha_1 \ldots \alpha_n) = c'(\alpha_1 \ldots \alpha_n) = \delta_{\alpha_1 o} \ldots \delta_{\alpha_n o}$, with all other c's and c''s vanishing*. Then $(c,c')_0$ is unity, whereas $(c,c')_1$ reduces simply to $-\frac{1}{2}n(n-1)(00|I|00)$. To estimate this exchange matrix element, note that if the system is subject to periodic boundary conditions with periodicity volume Ω, then the bound states $\varphi_\alpha(Xx)$ can be labelled by a wave vector $\underset{\sim}{k}$ and internal quantum numbers ν, i.e. $\alpha = (\underset{\sim}{k}, \nu)$, where, e.g. $\nu = (n, \ell, m, s)$. Then the φ_α will have $\underset{\sim}{\sim}$ range $\sim a_o$ as a function of $|\underset{\sim}{R}-\underset{\sim}{r}|$, where a_o is the Bohr radius, and will be of order $a_o^{-3/2}\Omega^{-\frac{1}{2}}$ within this range. Then it is easy to see from (III.49) that $(\alpha_1 \alpha_2 | I | \alpha\beta)$ will be of order (a_o^3/Ω) when nonvanishing, whereas it will vanish unless the sum of momenta on the

* This is the case of extreme Bose-Einstein condensation, for which one expects exchange effects to be largest.

left is equal to the sum of momenta on the right. Thus in the special case $\alpha_1 = \alpha_2 = \alpha = \beta = 0$, $(00|I|00) \sim (a_0^3/\Omega)$ and $(c,c')_1 \sim -\frac{1}{2} n \rho a_0^3$, where $\rho = n/\Omega$ and we are assuming that n and Ω are large but with ratio ρ independent of n and Ω (macrospcopic system). By a similar argument one estimates that the ternary exchange contribution (III.53) is of order $(3!)^{-1} n (\rho a_0^3)^2$, whereas the singly-iterated binary exchange contribution (III.54) is of order $(2!)^{-2} n^2 (\rho a_0^3)^2$. More generally, any connected contribution, i.e. one arising from a single permutation cycle of length ℓ, is expected to be of order $n(\rho a_0^3)^{\ell-1}$, whereas a disconnected contribution, arising from a permutation decomposable into more than one disjoint cycle, will be of order of the products of such factors, one for each cycle, and will hence be proportional to n^m where m is the number of disjoint cycles.

It is clear from the above estimates that the series (III.47) is seriously divergent* for a macroscopic system ($n \sim 10^{23}$), terms of higher order in exchange involving higher powers of n. The situation in this respect is similar to the behavior of the Rayleigh-Schrödinger perturbation expansion for the ground state energy of a many-body system, or the similar expansions for the equilibrium statistical mechanics (Mayer expansion in classical statistical mechanics or quantum-statistical perturbation theory.) There useful expansions are obtained by appropriate reordering, through introduction of Ursell functions or some equivalent (linked cluster perturbation theory). One expects that by the use of similar methods, one can represent the sum of all connected and disconnected contributions to (III.46) as the exponential of a sum of connected contributions only. Then the inner product (III.46) will depend exponentially on n for large n. In fact, there are very general arguments[7]) that this is the case for well-behaved many-particle wave functions.

Since such a rearrangement of the series (III.47) has not in fact yet been accomplished, we shall content ourselves here with a more pragmatic approach. It is known

* Strictly speaking, the series is not divergent for finite n, since it terminates. However, it "converges abruptly."

that after such Ursell rearrangements have been performed in the case of statistical mechanical or quantum mechanical perturbation theory, the results* obtained by use of the rearranged series agree with those obtained from the original one to lowest (first) order in the density, the rearrangement only affecting the higher-order contributions. We shall assume that the same is true of the series (III.47). More specifically, we shall assume that in calculating the contribution of exchange to many-body energies, the first two terms of the series may be used, and the results obtained will then be correct to first order in the density. The actual way in which this will be done will be through introduction of an appropriate "metric operator" which generates the various terms in (III.47) and will also be found to be amenable to second quantization. This will be evaluated only up to binary exchange terms, i.e. those terms which contribute to $(c,c')_1$. We shall call this the "binary exchange approximation". The contributions of exchange to the second-quantized Hamiltonian will be evaluated only up to binary exchange terms.

We now introduce notation which will motivate the definition of the metric operator and will also prove useful in the subsequent transition to second quantization. Define the "state vector" $|c)$ as the set of all amplitudes c; this may be conveniently thought of as a column vector:

$$|c) = \begin{bmatrix} c(\alpha_1 \ldots \alpha_n) \\ c(\alpha_1 \ldots \alpha_{n-1}, X_n x_n) \\ \vdots \\ c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n) \\ \vdots \\ c(\alpha_1, X_2 \ldots X_n x_2 \ldots x_n) \\ c(X_1 \ldots X_n x_1 \ldots x_n) \end{bmatrix} \qquad \text{(III.55)}$$

* This statement is, of course, only true for some results.

Define the inner product of two such state vectors as

$$(c\,|\,c') \equiv \sum_{\alpha_1 \cdots \alpha_n} c^*(\alpha_1 \ldots \alpha_n) c'(\alpha_1 \ldots \alpha_n)$$

$$+ \sum_{j=1}^{n-1} \sum_{\alpha_1 \ldots \alpha_{n-j}} \int c^*(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$\times c'(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$\times dX_{n-j+1} \ldots dX_n dx_{n-j+1} \ldots dx_n$$

$$+ \int c^*(X_1 \ldots X_n x_1 \ldots x_n) c'(X_1 \ldots X_n x_1 \ldots x_n) dX_1 \ldots dX_n dx_1 \ldots dx_n . \qquad \text{(III.56)}$$

This inner product is not equal to the inner product (c,c'), Eqs. (III.46)-(III.48) and (III.51), although we see from (III.48) that it is equal to the direct term, $(c,c')_0$, in (c,c'). However, there exists a linear, hermitian operator M on the space of state vectors $|c)$, such that

$$(c,c') = (c\,|M|\,c') \qquad \text{(III.57)}$$

For obvious reasons, we shall call M the "metric operator". It is defined implicitly by (III.46) and (III.56). In fact, it follows from (III.45) and (III.46) that M is the same operator as occurs in (III.45), provided that Eqs. (III.45), which define M only for $0 \leq j \leq n-1$, are supplemented by

$$Mc(X_1 \ldots X_n x_1 \ldots x_n) = c(X_1 \ldots X_n x_1 \ldots x_n) . \qquad \text{(III.58)}$$

It follows from (III.47), (III.48), and (III.51) that M is given up through binary exchange terms by

$$M = 1 - B + \ldots \tag{III.59}$$

where the "binary exchange operator" B is defined by

$$\begin{aligned} Bc(\alpha_1 \ldots \alpha_n) &= \sum_{p<q}^{n} \sum_{\alpha\beta} (\alpha_p \alpha_q | I | \alpha\beta) \\ &\quad \times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \cdots \alpha_{q-1} \beta \alpha_{q+1} \cdots \alpha_n) \\ &+ n^{-\frac{1}{2}} \sum_{p<q}^{n} \sum_{\alpha} \int \left[K(\alpha_p \alpha_q ; \alpha, Xx) + K(\alpha_q \alpha_p ; \alpha, Xx) \right] \\ &\quad \times c(\alpha_1 \ldots \alpha_{p-1} \alpha_{p+1} \cdots \alpha_{q-1} \alpha_{q+1} \cdots \alpha_n \alpha, Xx) dX dx, \end{aligned} \tag{III.60}$$

$$\begin{aligned} &Bc(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \cdots X_n x_{n-j+1} \cdots x_n) \\ &= \sum_{p<q}^{n-j} \sum_{\alpha\beta} (\alpha_p \alpha_q | I | \alpha\beta) \\ &\quad \times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \cdots \alpha_{q-1} \beta \alpha_{q+1} \cdots \alpha_{n-j}, \\ &\qquad X_{n-j+1} \cdots X_n x_{n-j+1} \cdots x_n) \\ &+ (j+1)(n-j)^{-\frac{1}{2}} \sum_{p<q}^{n-j} \sum_{\alpha} \\ &\quad \times \int [K(\alpha_p \alpha_q ; \alpha, Xx) + K(\alpha_q \alpha_p ; \alpha, Xx)] \end{aligned}$$

$$\times\, c(\alpha_1 \circ\cdots\alpha_{p-1}\alpha_{p+1}\cdots\circ\alpha_{q-1}\alpha_{q+1}\cdots\alpha_{n-j}\alpha,$$

$$X_{n-j+1}\cdots X_n X x_{n-j+1}\cdots x_n x)\,dX dx$$

$$+ \sum_{p=1}^{n-j}\sum_{q=n-j+1}^{n}\sum_{\alpha}\Big[\int K(\alpha_p, X_q;\alpha, X)$$

$$\times\, c(\alpha_1\cdots\alpha_{p-1}\alpha\alpha_{p+1}\cdots\circ\alpha_{n-j},$$

$$X_{n-j+1}\cdots X_{q-1}X X_{q+1}\cdots X_n x_{n-j+1}\cdots x_n)\,dX$$

$$+ \int K(\alpha_p, x_q;\alpha, x)\, c(\alpha_1\cdots\alpha_{p-1}\alpha\alpha_{p+1}\cdots\alpha_{n-j},$$

$$X_{n-j+1}\cdots X_n x_{n-j+1}\cdots x_{q-1}x x_{q+1}\cdots x_n)\quad dx\Big]$$

$$+ j^{-1}(n-j+1)^{\frac{1}{2}}\sum_{p=1}^{n-j}\sum_{q=n-j+1}^{n}\sum_{s=n-j+1}^{n}\sum_{\alpha\beta} K(\alpha_p, X_q x_s;\alpha\beta)$$

$$\times\quad c(\alpha_1\cdots\alpha_{p-1}\alpha_{p+1}\cdots\alpha_{n-j}\alpha\beta,$$

$$X_{n-j+1}\cdots X_{q-1}X_{q+1}\cdots X_n x_{n-j+1}\cdots x_{s-1}x_{s+1}\cdots\circ x_n),$$

$$1 \le j \le n-1, \qquad \text{(III.61)}$$

and

$$Bc(X_1\cdots X_n x_1\circ\cdots x_n) = 0, \qquad \text{(III.62)}$$

where care has been taken to ensure that B preserves the proper symmetry and antisymmetry of the c's.

E. Representation of Observables

We wish to determine how operators representing physical observables transform into the space of state vectors $|c)$. This can, in principle, be done in two steps, first finding how the operators act on the space of amplitudes c and then transforming them into the space of state vectors $|c)$ with the aid of (III.57). Let A be any operator defined on wave functions Ψ. Then A can be defined as an operator on the amplitudes c by use of (III.43), (III.44), and appropriate algebraic manipulations. Denoting the amplitudes thus determined by Ac, one has by (III.57)

$$(c,Ac') = (c|MA|c'). \qquad \text{(III.63)}$$

The metric operator M is the representation, in the space of state vectors $|c)$, of the antisymmetrizing operator $A_n^{(nuc)}A_n^{(elec)}$ in (III.46). Thus, since Schrödinger operator representing physical observables are invariant under permutations of identical particles (i.e., they commute with $A_n^{(nuc)}A_n^{(elec)}$, the corresponding operators on amplitudes c commute with M.* Thus one can also write (III.63) as

$$(c,Ac') = (c|AM|c') \qquad \text{(III.64)}$$

provided that A really corresponds to a physical observable. For such an operator, one can combine (III.63) and (III.64) into

$$(c,Ac') = (c|\mathbf{A}|c') \qquad \text{(III.65)}$$

where $\mathbf{A}$, the operator on the space of state vectors $|c)$ corresponding to the operator A on the amplitudes c, can be written in several equivalent forms:

$$\mathbf{A} = MA = AM = \tfrac{1}{2}(MA+AM) = M^{\frac{1}{2}}AM^{\frac{1}{2}} = \ldots . \qquad \text{(III.66)}$$

*This is not a trivial matter, i.e. it is not valid for arbitrary operators A invariant under permutations of α's, X's, and x's. Instead, it is a special property of those particular A's derived from a Schrödinger operator invariant under all permutations (including atom-breaking ones). For such A's, there are relations between the terms $Ac(\alpha_1 \ldots \alpha_{n-j}, X_{n-j} \ldots X_n)$ with various j's, such that $[A,M]=0$.

Thus defined, $\mathbf{A}$ is hermitian, since both M and A are.

Defining Ac to be the expansion coefficients of $A\psi$, in analogy with the procedure used for transforming observables into the space of c's in Sec. I, one can obtain formal expressions for these amplitudes by multiplication of (III.45) from the left by M^{-1}:

$$Ac(\alpha_1 \ldots \alpha_n) = (n!)^{\frac{1}{2}} M^{-1} \int \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_n}^*(X_n x_n)$$

$$\times A\psi(X_1 \ldots X_n x_1 \ldots x_n)\, dX_1 \ldots dX_n dx_1 \ldots dx_n,$$

$$Ac(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= n!(j!)^{-1}[n-j)!]^{-\frac{1}{2}} M^{-1} \int \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_{n-j}}^*(X_{n-j} x_{n-j})$$

$$\times A\psi(X_1 \ldots X_n x_1 \ldots x_n) dX_1 \ldots dX_{n-j} dx_1 \ldots dx_{n-j},$$

$$Ac(\alpha_1, X_2 \ldots X_n x_2 \ldots x_n) = nM^{-1} \int \varphi_{\alpha_1}^*(X_1 x_1) A\psi(X_1 \ldots X_n x_1 \ldots x_n)$$

$$\times dX_1 dx_1. \quad \text{(III.67)}$$

Then by (III.66) the metric operator M cancels out of the equations determining the operation of $\mathbf{A}$:

$$\mathbf{A}c(\alpha_1 \ldots \alpha_n) = (n!)^{\frac{1}{2}} \int \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_n}^*(X_n x_n)$$

$$\times A\psi(X_1 \ldots X_n x_1 \ldots x_n)\, dX_1 \ldots dX_n dx_1 \ldots dx_n,$$

$$\vdots$$

$$\mathbf{A}c(\alpha_1 \ldots \alpha_{n-j} X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= n!(j!)^{-1}\,[(n-j)!]^{-\frac{1}{2}} \int \varphi_{\alpha_1}^*(X_1 x_1) \ldots \varphi_{\alpha_{n-j}}^*(X_{n-j} x_{n-j})$$

$$\times A\psi(X_1 \ldots X_n x_1 \ldots x_n) dX_1 \ldots dX_{n-j} dx_1 \ldots dx_{n-j},$$

$$\vdots$$

$$\mathbf{A}c(\alpha_1, X_2 \ldots X_n x_2 \ldots x_n) = n \int \varphi_{\alpha_1}^*(X_1 x_1) A \psi(X_1 \ldots X_n x_1 \ldots x_n) \times dX_1 dx_1 \qquad \text{(III.68)}$$

To determine $\mathbf{A}c(X_1 \ldots X_n x_1 \ldots x_n)$ we note that in the first place, by (III.58), $\mathbf{A}$ is equal to A when acting on $c(X_1 \ldots X_n x_1 \ldots x_n)$, and in the second place $Ac(X_1 \ldots X_n x_1 \ldots x_n)$ is the residue left after subtracting from $A_{Sch}\psi$ the sum of all its bound and partially bound parts. Here A_{Sch} is the Schrödinger operator, i.e. the operator on the space of ψ's. Then by (III.43) and (III.66) one finds

$$\mathbf{A}c(X_1 \ldots X_n x_1 \ldots x_n) = A_{Sch} c(X_1 \ldots X_n x_1 \ldots x_n)$$

$$+ n! A_n^{(nuc)} A_n^{(elec)} \Big\{ (n!)^{-\frac{1}{2}} \sum_{\alpha_1 \ldots \alpha_n} \Big[c(\alpha_1 \ldots \alpha_n) \times A_{sch} \varphi_{\alpha_1}(X_1 x_1) \ldots \varphi_{\alpha_n}(X_n x_n) - \varphi_{\alpha_1}(X_1 x_1) \ldots \varphi_{\alpha_n}(X_n x_n) M^{-1} \mathbf{A}c(\alpha_1 \ldots \alpha_n) \Big]$$

$$+ \sum_{j=1}^{n-1} (j!)^{-1} [(n-j)!]^{-\frac{1}{2}} \times \sum_{\alpha_1 \ldots \alpha_{n-j}} \Big[A_{Sch} c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n) \times \varphi_{\alpha_1}(X_1 x_1) \ldots \varphi_{\alpha_{n-j}}(X_{n-j} x_{n-j}) -$$

$$-\varphi_{\alpha_1}(X_1x_1)\ldots\varphi_{\alpha_{n-j}}(X_{n-j}x_{n-j})M^{-1}\mathbf{A}c(\alpha_1\ldots\alpha_{n-j},$$

$$X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_n)\Big]\Big\}$$

(III.69)

Here the $\mathbf{A}c$ involved are to be substituted after evaluating the expressions (III.68); then multiplication by M^{-1} can be effected by taking the -1 power of the series (III.59), which, in the binary exchange approximation, gives

$$M^{-1} = 1 + B + \ldots \quad . \tag{III.70}$$

As a preliminary to obtaining more explicit general expressions for the kinetic and potential energy operators $\mathbf{T}$ and $\mathbf{V}$, it is helpful, as before, to first consider a few special cases of small values of n. For n = 1, one has

$$\psi(Xx) = \sum_\alpha c(\alpha)\ \varphi_\alpha(Xx) + c(Xx). \tag{III.71}$$

Any single-particle operator T on ψ has the structure

$$T\psi(Xx) \equiv T_{Sch}\psi(Xx) \equiv [(T(X) + T(x)]\psi(Xx). \tag{III.72}$$

Then by (III.71)

$$T\psi(Xx) = \sum_\alpha c(\alpha)[T(X)+T(x)]\ \varphi_\alpha(Xx)+[T(X)+T(x)]c(Xx). \tag{III.73}$$

Putting n=1 in (III.68), one finds

$$\mathbf{T}c(\alpha) = \sum_\beta (\alpha|T|\beta)c(\beta) \quad + \int \varphi_\alpha^*(Xx)[T(X)+T(x)]c(Xx)\,dXdx \tag{III.74}$$

where the atom-atom matrix elements are

$$(\alpha|T|\beta) \equiv \int \varphi_\alpha^*(Xx)\,[T(X) + T(x)]\,\varphi_\beta(Xx)\,dXdx, \quad \text{(III.75)}$$

as in (I.12). For the case n=1, M=1, so that (III.69) gives

$$\mathbf{T}c(Xx) = [T(X) + T(x)]c(Xx)$$

$$- \int \Delta(Xx,X'x')\,[T(X')+T(x')]c(X'x')\,dX'dx'$$

$$+\sum_\alpha (Xx|T|\alpha)c(\alpha) \quad \text{(III.76)}$$

where the "bound state kernel" Δ is defined as

$$\Delta(Xx,X'x') \equiv \sum_\alpha \varphi_\alpha(Xx)\,\varphi_\alpha^*(X'x') \quad \text{(III.77)}$$

and the atom-continuum matrix elements* as

$$(Xx|T|\alpha) \equiv [T(X)+T(x)]\varphi_\alpha(Xx)$$

$$- \int \Delta(Xx,X'x')\,[T(X')+T(x')]\varphi_\alpha(X'x')\,dX'dx'. \quad \text{(III.78)}$$

It follows from (III.77) and (III.78) that

$$\int \varphi_\alpha^*(Xx)\,(Xx|T|\beta)\,dXdx = 0, \text{ all } \alpha \text{ and } \beta. \quad \text{(III.79)}$$

Then it is easy to see that $\mathbf{T}c(Xx)$ is orthogonal to all the φ_α, as it ought to be. It is also easy to verify that the amplitudes (III.74) and (III.76) do in fact generate $T\psi$ when substituted into (III.71). Any two-particle

---

* We shall see later that such matrix elements cancel between $\mathbf{T}$ and $\mathbf{V}$ in the special case that φ_α are taken to be energy eigenstates. However, in the general case $\mathbf{T}$ can induce breakup of an atom.

operator V on ψ has the structure

$$V\psi(Xx) \equiv V_{Sch}\psi(Xx) \equiv V(Xx)\psi(Xx) \qquad \text{(III.80)}$$

where the notation means that V(Xx) operates on both the X and x dependence of ψ (it is not necessarily an ordinary potential, i.e. it may be a nontrivial operator). One finds in analogy with the derivation of (III.74) and (III.76)

$$\mathbf{V}_c(\alpha) = \sum_\beta (\alpha|V|\beta)c(\beta) + \int \varphi_\alpha^*(Xx)V(Xx)c(Xx)\,dXdx,$$

$$\mathbf{V}_c(Xx) = V(Xx)c(Xx) - \int \Delta(Xx,X'x')V(X'x')c(X'x')\,dX'dx'$$

$$+ \sum_\alpha (Xx|V|\alpha)c(\alpha) \qquad \text{(III.81)}$$

where the atom-atom matrix elements are

$$(\alpha|V|\beta) \equiv \int \varphi_\alpha^*(Xx)V(Xx)\varphi_\beta(Xx)\,dXdx \qquad \text{(III.82)}$$

in analogy with (I.15), and the atom-continuum matrix elements are

$$(Xx|V|\alpha) \equiv V(Xx)\varphi_\alpha(Xx)$$

$$- \int \Delta(Xx,X'x')V(X'x')\varphi_\alpha(X'x')\,dX'dx' \qquad \text{(III.83)}$$

in analogy with (III.78). Again, one easily verifies

$$\int \varphi_\alpha^*(Xx)(Xx|V|\beta)\,dXdx = 0, \text{ all } \alpha \text{ and } \beta, \qquad \text{(III.84)}$$

and checks that $\mathbf{V}_c$ (X x) is orthogonal to the φ_α and that the amplitudes (III.81) do in fact generate $V\psi$.

Next consider the case n = 2. Eq. (III.43) reduces to

$$\psi(X_1X_2x_1x_2) = 2A_2^{(nuc)}A_2^{(elec)}\Big[2^{-\frac{1}{2}}\sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)$$

$$\times \varphi_{\alpha_1}(X_1x_1)\varphi_{\alpha_2}(X_2x_2) + \sum_{\alpha} c(\alpha,X_2x_2)\varphi_{\alpha}(X_1x_1)\Big] + c(X_1X_2x_1x_2)$$

$$= 2^{-\frac{1}{2}} \sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)\Big[\varphi_{\alpha_1}(X_1x_1)\varphi_{\alpha_2}(X_2x_2)$$

$$-\varphi_{\alpha_1}(X_1x_2)\varphi_{\alpha_2}(X_2x_1)\Big]$$

$$+ \tfrac{1}{2}\sum_{\alpha}\Big[c(\alpha,X_2x_2)\varphi_{\alpha}(X_1x_1) - c(\alpha,X_2x_1)\varphi_{\alpha}(X_1x_2)$$

$$- c(\alpha,X_1x_2)\varphi_{\alpha}(X_2x_1) + c(\alpha,X_1x_1)\varphi_{\alpha}(X_2x_2)\Big]$$

$$+ c(X_1X_2x_1x_2) \qquad \text{(III.85)}$$

where the amplitudes satsify the strong orthogonality constraints (III.44):

$$\int \varphi_{\alpha}^*(Xx)c(\beta,Xx)\,dXdx = 0, \text{ all } \alpha \text{ and } \beta;$$

$$\int \varphi_{\alpha}^*(X_2x_2)c(X_1X_2x_1x_2)dX_2dx_2 = 0, \text{ all } \alpha, X_1, \text{ and } x_1. \qquad \text{(III.86)}$$

A single-particle operator T has the structure

$$T\psi(X_1X_2x_1x_2) \equiv T_{Sch}\psi(X_1X_2x_1x_2)$$

$$\equiv \Big[T(X_1)+T(X_2)+T(x_1)+T(x_2)\Big]\psi(X_1X_2x_1x_2) \qquad \text{(III.87)}$$

on the space of ψ's. Then by (III.85)

$$T\psi(X_1X_2x_1x_2) = 2^{\frac{1}{2}} A_2^{(\mathrm{nuc})}A_2^{(\mathrm{elec})}$$

$$\times \sum_{\alpha_1\alpha_2} c(\alpha_1\alpha_2)\{\varphi_{\alpha_2}(X_2x_2)[T(X_1)+T(x_1)]\varphi_{\alpha_1}(X_1x_1)$$

$$+ \varphi_{\alpha_1}(X_1x_1)[T(X_2)+T(x_2)]\varphi_{\alpha_2}(X_2x_2)\}$$

$$+ 2A_2^{(\mathrm{nuc})}A_2^{(\mathrm{elec})} \sum_{\alpha}\{c(\alpha,X_2x_2)[T(X_1) + T(x_1)]\varphi_{\alpha}(X_1x_1)$$

$$+ \varphi_{\alpha}(X_1x_1)[T(X_2) + T(x_2)]c(\alpha,X_2x_2)\}$$

$$+ [T(X_1) + T(X_2) + T(x_1) + T(x_2)]c(X_1X_2x_1x_2) \quad \text{(III.88)}$$

since T commutes with $A_2^{(\mathrm{nuc})}$ and $A_2^{(\mathrm{elec})}$. Then on putting n = 2 in (III.68) one finds*

$$\mathbf{T}c(\alpha_1\alpha_2) = \sum_{\alpha}[(\alpha_1|T|\alpha)c(\alpha\alpha_2)+(\alpha_2|T|\alpha)c(\alpha_1\alpha)]$$

$$- \sum_{\alpha\beta}(\alpha_1\alpha_2|IT|\alpha\beta)c(\alpha\beta)$$

$$+ 2^{-\frac{1}{2}}\int[(\alpha_1|T|(Xx)'c(\alpha_2,Xx)+(\alpha_2|T|Xx)'c(\alpha_1,Xx)]dXdx$$

$$- 2^{-\frac{1}{2}} \sum_{\alpha}\int\{(\alpha_1\alpha_2|IT|\alpha,Xx)' + (\alpha_2\alpha_1|IT|\alpha,Xx)' \quad \text{(III.89)}$$

$$+ [K(\alpha_1\alpha_2;\alpha,Xx)+K(\alpha_2\alpha_1;\alpha,Xx)][T(X)+T(x)]\}c(\alpha,Xx)\,dXdx$$

---

* Note that the operators $A_2^{(\mathrm{nuc})}$ and $A_2^{(\mathrm{elec})}$ may be shifted so as to operate on the product of φ_α* factors.

where

$$(\alpha_1\alpha_2|IT|\alpha\beta)\equiv\int\varphi_{\alpha_1}^*(X_1x_1)\,\varphi_{\alpha_2}^*(X_2x_2)\,[T(X_1)+T(X_2)+T(x_1)+T(x_2)]$$

$$\times\,\varphi_\alpha(X_1x_2)\,\varphi_\beta(X_2x_1)\,dX_1\,dX_2\,dx_1\,dx_2$$

$$=\int[K(\alpha_1X_2;\beta,X_1)T(X_1)K(\alpha_2,X_1;\alpha,X_2)$$

$$+\,K(\alpha_2,X_1;\beta,X_2)T(X_2)K(\alpha_1,X_2;\alpha,X_1)\,]dX_1\,dX_2$$

$$+\int[K(\alpha_1,x_2;\beta,x_1)T(x_1)K(\alpha_2,x_1;\alpha,x_2)$$

$$+\,K(\alpha_2,x_1;\beta,x_2)T(x_2)K(\alpha_1,x_2;\alpha,x_1)\,]dx_1\,dx_2\,,$$

$$(Xx|T|\alpha)'\equiv[T(X)+T(x)\,]\varphi_\alpha(Xx)\,,\ (\alpha|T|Xx)'\equiv[\,(Xx|T|\alpha)'\,]^*,$$

$$(\alpha_1\alpha_2|IT|\alpha,Xx)'\equiv\int\varphi_{\alpha_1}^*(Xx')\,\varphi_{\alpha_2}^*(X'x)\,[T(X')+T(x')\,]\varphi_\alpha(X'x')$$

$$\times\,dX'dx'.\qquad\text{(III.90)}$$

Similarly, one finds

$$\mathbf{T}c(\alpha,Xx)=\sum_\beta(\alpha|T|\beta)c(\beta,Xx)\;+\;[T(X)+T(x)\,]c(\alpha,Xx)$$

$$-\sum_\beta\int[(\alpha,X|IT|\beta,X')+K(\alpha,X;\beta,X')T(X')\,]c(\beta,X'x)\,dX'$$

$$-\sum_\beta\int[(\alpha,x|IT|\beta,x')+K(\alpha,x;\beta,x')T(x')\,]c(\beta,Xx')\,dx'$$

$$+\sum_\beta\varphi_\beta(Xx)\int(\alpha|T|X'x')'c(\beta,X'x')\,dX'dx'$$

$$+ 2^{\frac{1}{2}} \sum_{\beta} (Xx|T|\beta)' c(\beta\alpha)$$

$$+ 2^{\frac{1}{2}} \sum_{\alpha_1 \alpha_2} \left\{ (\alpha|T|\alpha_1) \varphi_{\alpha_2}(Xx) - (\alpha, Xx|IT|\alpha_1 \alpha_2)' \right.$$

$$\left. - [T(X)+T(x)] K(\alpha, Xx; \alpha_1 \alpha_2) \right\} c(\alpha_1 \alpha_2)$$

$$+ 2 \int (\alpha|T|X'x')' c(XX'xx') dX' dx' \qquad \text{(III.91)}$$

with

$$(\alpha, X|IT|\beta, X') \equiv \int \varphi_\alpha^*(X'x) [T(X)+T(x)] \varphi_\beta(Xx) dx,$$

$$(\alpha, x|IT|\beta, x') \equiv \int \varphi_\alpha^*(Xx') [T(X)+T(x)] \varphi_\beta(Xx) dX,$$

$$(\alpha, Xx|IT|\alpha_1 \alpha_2)' \equiv [(\alpha_1 \alpha_2|IT|\alpha, Xx)']'. \qquad \text{(III.92)}$$

Finally, for completeness we should exhibit the expression for $\mathbf{T} c(X_1 X_2 x_1 x_2)$. This is not determined by (III.68), but by the more involved expression (III.69). However, even for this simple case of n = 2, there are so many terms that it hardly seems worthwhile to write out the expression explicitly. Actually, in a macroscopic system ($n \sim 10^{23}$) the probability of finding <u>all</u> atoms dissociated is negligibly small, so one may safely restrict oneself to a subspace in which the totally unbound amplitude is zero, and neglect matrix elements of observables connecting this restricted subspace with the "totally unbound" subspace.

Next consider the form of the operator $\mathbf{V}$ for the case n = 2, where V is a two-particle operator:

$$V\psi(X_1 X_2 x_1 x_2) = [V(X_1 X_2)+V(x_1 x_2)+V(X_1 x_1)+V(X_2 x_2)$$

$$+ V(X_1 x_2)+V(X_2 x_1)] \psi(X_1 X_2 x_1 x_2). \qquad \text{(III.93)}$$

Derivations similar to those of (III.89) and (III.91)

give the form of $\mathbf{V}$ on the space of state vectors $|c)$. As in Sec. I, it is convenient to separate $\mathbf{V}$ into a single-atom part $\mathbf{V}_o$ and an interaction part $\mathbf{V}'$. One finds that

$$\mathbf{V} = \mathbf{V}_o + \mathbf{V}' \qquad \text{(III.94)}$$

where

$$\mathbf{V}_o\, c(\alpha_1\alpha_2) = \sum_\alpha [(\alpha_1|V|\alpha)c(\alpha\alpha_2) + (\alpha_2|V|\alpha)c(\alpha_1\alpha)]$$
$$+ 2^{-\frac{1}{2}} \int [(\alpha_1|V|Xx)'c(\alpha_2,Xx)$$
$$+ (\alpha_2|V|Xx)'c(\alpha_1,Xx)]dXdx \qquad \text{(III.95)}$$

and*

$$(Xx|V|\alpha)' \equiv V(Xx)\varphi_\alpha(Xx), (\alpha|V|Xx)' \equiv [(Xx|V|\alpha)']^*. \qquad \text{(III.96)}$$

Similarly, one finds

$$\mathbf{V}'c(\alpha_1\alpha_2) = \sum_{\alpha\beta} [(\alpha_1\alpha_2|V|\alpha\beta) - (\alpha_1\alpha_2|IV|\alpha\beta)]c(\alpha\beta)$$
$$+ 2^{-\frac{1}{2}} \sum_\alpha \int \{(\alpha_1\alpha_2|V|\alpha,Xx)' + (\alpha_2\alpha_1|V|\alpha,Xx)'$$
$$- (\alpha_1\alpha_2|IV|\alpha,Xx)' - (\alpha_2\alpha_1|IV|\alpha,Xx)'$$

*The prime distinguishes this matrix element from the previously-defined one (III.83), which contains an additional orthogonalization term. The same remark applies to the distinction between (III.90) and (III.78).

$$-[K(\alpha_1\alpha_2;\alpha,Xx)+K(\alpha_2\alpha_1;\alpha,Xx)]V(Xx)\Big\}\, c(\alpha,Xx)\,dXdx$$

$$+2^{\frac{1}{2}}\int(\alpha_1\alpha_2|V|X_1X_2x_1x_2)'c(X_1X_2x_1x_2)\,dX_1dX_2dx_1dx_2 \quad \text{(III.97)}$$

where

$$(\alpha_1\alpha_2|V|\alpha\beta)\equiv\int\varphi_{\alpha_1}^*(X_1x_1)\,\varphi_{\alpha_2}^*(X_2x_2)\,[V(X_1X_2)+V(x_1x_2)$$

$$+V(X_1x_2)+V(X_2x_1)]\varphi_\alpha(X_1x_1)\,\varphi_\beta(X_2x_2)\,dX_1dX_2dx_1dx_2 \quad \text{(III.98)}$$

as in (I.16), and

$$(\alpha_1\alpha_2|IV|\alpha\beta)\equiv\int\varphi_{\alpha_1}^*(X_1x_1)\,\varphi_{\alpha_2}^*(X_2x_2)\,[V(X_1X_2)+V(x_1x_2)$$

$$+V(X_1x_1)+V(X_2x_2)+V(X_1x_2)+V(X_2x_1)]\varphi_\alpha(X_1x_2)\,\varphi_\beta(X_2x_1)$$

$$\times\, dX_1dX_2dx_1dx_2,$$

$$(\alpha_1\alpha_2|V|\alpha,Xx)'\equiv\int\varphi_{\alpha_1}^*(Xx)\,\varphi_{\alpha_2}^*(X'x')\,[V(XX')+V(xx')+V(Xx')$$

$$+V(X'x)]\,\varphi_\alpha(X'x')\,dX'dx',$$

$$(\alpha_1\alpha_2|IV|\alpha,Xx)'\equiv\int\varphi_{\alpha_1}^*(Xx')\,\varphi_{\alpha_2}^*(X'x)\,[V(XX')+V(xx')+V(X'x')$$

$$+V(Xx')+V(X'x)]\varphi_\alpha(X'x')\,dX'dx',$$

$$(X_1X_2x_1x_2|V|\alpha_1\alpha_2)'\equiv[V(X_1X_2)+V(x_1x_2)+V(X_1x_2)+V(X_2x_1)]$$

$$\times\,\varphi_{\alpha_1}(X_1x_1)\,\varphi_{\alpha_2}(X_2x_2),$$

$$(\alpha_1\alpha_2|V|X_1X_2x_1x_2)'\equiv[(X_1X_2x_1x_2|V|\alpha_1\alpha_2)']^*. \quad \text{(III.99)}$$

Similarly, the following separation of $\mathbf{V}c(\alpha,Xx)$ will prove convenient:

$$\mathbf{V}_o\, c(\alpha,Xx) = \sum_\beta (\alpha|V|\beta)c(\beta,Xx)$$

$$+ \sum_\beta \varphi_\beta(Xx) \int (\alpha|V|X'x')\, c(\beta,X'x')dX'dx'$$

$$+ 2\int (\alpha|V|X'x')\, c(XX'xx')dX'dx' \qquad \text{(III.100)}$$

and

$$\mathbf{V}'c(\alpha,Xx) = V(Xx)c(\alpha,Xx)$$

$$+ \sum_\beta [(\alpha,X|V|\beta,X)+(\alpha,x|V|\beta,x)]c(\beta,Xx)$$

$$+ \sum_\beta \varphi_\beta(Xx) \int (\alpha,Xx|V|XX'xx')\, c(\beta,X'x')dX'dx'$$

$$- \sum_\beta \int (\alpha,Xx|IV|\beta,X'x')\, c(\beta,X'x)dX'$$

$$- \sum_\beta \int (\alpha,Xx|IV|\beta,Xx')\, c(\beta,Xx')dx'$$

$$+ 2\int [(\alpha,X|V|X'x'X)' + (\alpha,x|V|X'x'x)$$

$$+ (\alpha,Xx|V|XX'xx')']c(XX'xx')dX'dx'$$

$$+ 2^{\frac{1}{2}} \sum_\beta (Xx|V|\beta)'c(\beta\alpha)$$

$$+ 2^{\frac{1}{2}} \sum_{\alpha_1 \alpha_2} \Big\{ (\alpha|V|\alpha_1)\varphi_{\alpha_2}(Xx) + (\alpha, Xx|V|\alpha_1\alpha_2)' - (\alpha, Xx|IV|\alpha_1\alpha_2)' - V(Xx)K(\alpha, Xx; \alpha_1\alpha_2)\Big\} c(\alpha_1\alpha_2) \qquad \text{(III.101)}$$

with

$$(\alpha, X|V|\beta, X) \equiv \int \varphi_\alpha^*(X'x')[V(X'X)+V(Xx')]\varphi_\beta(X'x')dX'dx',$$

$$(\alpha, x|V|\beta, x) \equiv \int \varphi_\alpha^*(X'x')[V(x'x)+V(X'x)]\varphi_\beta(X'x')dX'dx',$$

$$(XX'xx'|V|\alpha, Xx)' \equiv [V(XX')+V(xx')+V(Xx')+V(X'x)]\varphi_\alpha(X'x'),$$

$$(\alpha, Xx|V|XX'xx')' \equiv [(XX'xx'|V|\alpha, Xx)']^*,$$

$$(\alpha, Xx|IV|\beta, X'x)' \equiv \int \varphi_\alpha^*(X'x')[V(XX')+V(xx')+V(Xx)+V(X'x') + V(Xx')+V(X'x)]\varphi_\beta(Xx')dx',$$

$$(\alpha, Xx|IV|\beta, Xx')' \equiv \int \varphi_\alpha^*(X'x')[V(XX')+V(xx')+V(Xx)+V(X'x') + V(Xx')+V(X'x)]\varphi_\beta(X'x)dX',$$

$$(\alpha, X|V|X'x'X)' \equiv \Big\{[V(XX')+V(Xx')]\varphi_\alpha(X'x')\Big\}^*$$

$$(\alpha, x|V|X'x'x)' \equiv \Big\{[V(xx')+V(X'x)]\varphi_\alpha(X'x')\Big\}^*. \qquad \text{(III.102)}$$

For completeness, we also define

$$\mathbf{V}_0 c(X_1X_2x_1x_2) \equiv 0, \qquad \text{(III.103)}$$

i.e. $\mathbf{V} = \mathbf{V}'$ when acting on $c(X_1X_2x_1x_2)$,

with $\mathbf{V}'$ determined implicitly by (III.69).

These expressions simplify in the special case that the φ_α are chosen to be single-atom energy eigenstates, i.e.

$$[T(X)+T(x)+V(Xx)]\varphi_\alpha(Xx) = \epsilon_\alpha \varphi_\alpha(Xx). \qquad \text{(III.104)}$$

Then decomposing $\mathbf{T}$ in a manner analogous to $\mathbf{V}$, i.e.

$$\mathbf{T} = \mathbf{T}_0 + \mathbf{T}',$$

$$\mathbf{T}_0 c(\alpha_1\alpha_2) = \sum_\alpha [(\alpha_1|T|\alpha)c(\alpha\alpha_2)+(\alpha_2|T|\alpha)c(\alpha_1\alpha)]$$

$$+ 2^{-\frac{1}{2}}\int [(\alpha_1|T|Xx)'c(\alpha_2,Xx)+(\alpha_2|T|Xx)'c(\alpha_1,Xx)]dXdx,$$

$$\mathbf{T}_0 c(\alpha,Xx) = \sum_\beta (\alpha|T|\beta)c(\beta,Xx)+[T(X)+T(x)]c(\alpha,Xx)$$

$$+ \sum_\beta \varphi_\beta(Xx)\int (\alpha|T|X'x')'c(\beta,X'x')dX'dx'$$

$$+ 2\int (\alpha|T|X'x')'c(XX'xx')dX'dx' \qquad \text{(III.105)}$$

and for completeness*

$$\mathbf{T}_0 c(X_1X_2x_1x_2) = [T(X_1)+T(X_2)+T(x_1)+T(x_2)]c(X_1X_2x_1x_2), \qquad \text{(III.106)}$$

---

* It is clear from (III.69) that $\mathbf{T}_0 c(X_1X_2x_1x_2)$ is one term in $\mathbf{T}c(X_1X_2x_1x_2)$, with $\mathbf{T}' c(X_1X_2x_1x_2)$ being the sum of all other terms.

one finds with the aid of (III.104) and (III.86) that $\mathbf{H}_o$ simplifies:

$$\mathbf{H}_o \equiv \mathbf{T}_o + \mathbf{V}_o ,$$

$$\mathbf{H}_o c(\alpha_1 \alpha_2) = (\epsilon_{\alpha_1} + \epsilon_{\alpha_2})\, c(\alpha_1 \alpha_2),$$

$$\mathbf{H}_o c(\alpha, Xx) = [\epsilon_\alpha + T(X) + T(x)]\, c(\alpha, Xx),$$

$$\mathbf{H}_o c(X_1 X_2 x_1 x_2) = [T(X_1) + T(X_2) + T(x_1) + T(x_2)]\, c(X_1 X_2 x_1 x_2).$$

(III.107)

$\mathbf{H}_o$ has three types of eigenstates:

$$c(\alpha_1 \alpha_2) = \delta_{\alpha_1 \alpha} \delta_{\alpha_2 \beta} + \delta_{\alpha_1 \beta} \delta_{\alpha_2 \alpha},$$

$$c(\alpha, Xx) = c(X_1 X_2 x_1 x_2) = 0, \qquad \text{(III.108)}$$

with energy eigenvalue $\epsilon_\alpha + \epsilon_\beta$;

$$c(\alpha, Xx) = \delta_{\alpha \alpha_o} f_\epsilon (Xx),$$

$$c(\alpha_1 \alpha_2) = c(X_1 X_2 x_1 x_2) = 0, \qquad \text{(III.109)}$$

with energy eigenvalue $\epsilon_{\alpha_o} + \epsilon$, where f_ϵ is a product of free proton and free electron orbitals (plane waves) with total energy ϵ;

$$c(X_1 X_2 x_1 x_2) = f_\epsilon (X_1 X_2 x_1 x_2),$$

$$c(\alpha_1 \alpha_2) = c(\alpha, Xx) = 0, \qquad \text{(III.110)}$$

with energy eigenvalue ϵ, where f_ϵ is an antisymmetrized product of free proton and free electron orbitals with total energy ϵ. Only the eigenstates of type (III.108) satisfy the bound state-continuum orthogonality constraints (III.86), and hence lie in the space of physically allowed states. The failure of the eigenstates of types (III.109) and (III.110) to satisfy these constraints* is an inescapable consequence of the fact that products of free-particle orbitals are not orthogonal to the bound states, and the related fact that the interactions between continuum particles cannot be turned off without also turning off the intra-atomic interactions, causing the atoms to disintegrate. Nevertheless, a decomposition of $\mathbf{H} = \mathbf{T} + \mathbf{V}$ into $\mathbf{H}_0 + \mathbf{H}'$ would be useful in a perturbation treatment in which atom-atom and atom-continuum exchange effects are treated perturbatively along with the actual atom-atom, atom-continuum, and continuum-continuum interactions. Such a treatment is expected to be useful at low densities. An interesting aspect of the decomposition (III.105) is that $\mathbf{T}$, like $\mathbf{V}$, contains an interaction part $\mathbf{T}'$. It is clear from (II.89) and (III.91) that $\mathbf{T}'$ contains atom-atom, atom-proton, and atom-electron interaction parts. These arise from coupling between exchange and kinetic energy.

If one undertakes the exercise of verifying that $\mathbf{T}\, c(\alpha, Xx)$ and $\mathbf{V}\, c(\alpha Xx)$ satisfy the orthogonality constraint (III.86), one may be shocked to discover that in general they do not, even if $c(\alpha, Xx)$ does. More generally, $\mathbf{A}\, c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$, for general n and a general $\mathbf{A}$ derived from a Schrödinger operator A_{Sch}, does not in general satisfy (III.44) even if c

* In fact, we shall see in the next paragraph that the eigenstates of the full $\mathbf{H}$ also fail to satisfy these constraints.

does. The point is that the metric operator M in (III.63) and (III.66) in general takes a state satisfying (III.44) into one that does not, except in the trivial case n = 1, when M = 1; this is clear from (III.45), since the right sides of these equations are in general not zero. Thus it is convenient to make a distinction between extended and restricted state spaces. We shall call the space of all state vectors $|c)$, restricted only by proper symmetry and antisymmetry under permutations of the arguments of the c's, the "ideal state space" $\mathcal{I}$. The subspace of $\mathcal{I}$ consisting of state vectors $|c)$ satisfying (III.44) will be called the "physical subspace" $\mathcal{P}$. Observables A act on $\mathcal{P}$ and leave $\mathcal{P}$ invariant*, but M and operators $\mathbf{A} = MA$ act on $\mathcal{I}$ and in general take a state vector in $\mathcal{P}$ into one having a physical component (the component in $\mathcal{P}$) and an unphysical component (the component in the orthogonal subspace $\mathcal{I}-\mathcal{P}$, the "totally unphysical" subspace.) Far from being a drawback, this is an advantage, since it will enable us to introduce second quantization easily without worrying about the fact that products of free atom, free proton, and free electron states are in general not in $\mathcal{P}$. We shall see in Sec. III.F. how observables can be projected onto the physical subspace $\mathcal{P}$, in analogy with the procedure used in Secs. II.D. and II.E. Finally, we note that the fact that $\mathbf{A}\,c(X_1 \ldots X_n x_1 \ldots x_n)$ in general contains many more terms than do the other $\mathbf{A}\,c$'s may be regarded as a result of the fact that M=1 on the "completely unbound" subspace, so that $\mathbf{A}\,c(X_1 \ldots X_n x_1 \ldots x_n)$ is strongly orthogonal to all the φ_α if c is; hence it automatically contains all of the many orthogonalization terms necessary to achieve this orthogonality.

It is now a fairly straightforward matter (although algebraically tedious) to generalize the previous derivations so as to obtain the representation of an observable A acting on Schrödinger wave functions ψ as an operator $\mathbf{A}$ on the ideal state space $\mathcal{I}$, for arbitrary n. If we

---

* This is a consequence of relationships between the effect of A on c's with different values of j, following from the fact that A_{Sch} commutes with $A_n^{(nuc)}$ and $A_n^{(elec)}$.

restrict ourselves to the binary exchange approximation, i.e. consider, in the evaluation of the integrals (III.68) after substitution of $A\psi$, only the same type of permutations as were included in evaluating (III.46) up to the binary collision term $(c,c')_1$, then the terms arising for general n are rather obvious generalizations of the ones occurring already for n = 2. Even if one restricts oneself to the binary exchange approximation, terms representing collisions of arbitrary numbers of atoms, protons, and electrons occur (note that even for n = 2, some of the matrix elements represent 3-body collisions). It is in the spirit of the binary exchange approximation to also drop matrix elements representing true dynamical collisions of more than two particles*, so we shall henceforth do so. We shall call this approximation, in which matrix elements representing multiple collisions, whether truly dynamical, exchange, or both, are omitted, the "binary interaction" approximation.**

We shall, as before, consider one and two-particle observables

$$T\psi(X_1\ldots X_n x_1\ldots x_n) \equiv T_{Sch}\psi(X_1\ldots X_n x_1\ldots x_n)$$
$$\equiv \Big[\sum_{j=1}^{n} T(X_j) + \sum_{j=1}^{n} T(x_j)\Big]\psi(X_1\ldots X_n x_1\ldots x_n),$$

$$V\psi(X_1\ldots X_n x_1\ldots x_n) \equiv V_{Sch}\psi(X_1\ldots X_n x_1\ldots x_n)$$
$$\equiv\Big[\sum_{j<k}^{n} V(X_jX_k)+\sum_{j<k}^{n} V(x_jx_k)+\sum_{j=1}^{n}\sum_{k=1}^{n} V(X_jx_k)\Big]\psi(X_1\ldots X_n x_1\ldots x_n). \quad \text{(III.111)}$$

* Here "particle" means atom, unbound proton, or unbound electron.

** Note that even with only single binary collision terms in the Hamiltonian, iterated collision effects occur in its eigenstates.

The explicit expressions for $\mathbf{T}$ and $\mathbf{V}$ up to the binary interaction approximation then follow upon substituting (III.43) and III.111) into (III.68), retaining only direct and binary exchange terms, and also dropping many-body interaction terms. The combinatorics are rather involved, so we state only the results. One finds*

$$\mathbf{H} = \mathbf{T} + \mathbf{V} = \mathbf{H}_0 + \mathbf{H}' + \mathbf{H}_{\text{spont}} \qquad \text{(III.112)}$$

where

$$\mathbf{H}_0 c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \cdots X_n x_{n-j+1} \ldots x_n)$$

$$= \sum_{p=1}^{n-j} \sum_{\alpha} \left[(\alpha_p | T | \alpha) + (\alpha_p | V | \alpha) \right] \times$$

$$\times\, c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \cdots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$+ \sum_{p=n-j+1}^{n} \left[T(X_p) + T(x_p) \right] c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n$$

$$\times\, x_{n-j+1} \cdots x_n)$$

---

* In the actual lectures only a schematic description of $\mathbf{H}$ was given. The expressions given here, and all of the subsequent analysis and discussion, are more complete.

$$+\Big\{(j+1)(n-j)^{-\frac{1}{2}} \sum_{p=1}^{n-j} \int \Big[(\alpha_p|T|Xx)' + (\alpha_p|V|Xx)'\Big]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1}\alpha_{p+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n X x_{n-j+1} \ldots x_n x)$$

$$\times dXdx + \text{h.c.}\Big\}. \qquad \text{(III.113)}$$

Here "h.c." denotes the hermitian conjugate; for the case of the term coupling the (n-j)-atom amplitude to the (n-j-1)-atom amplitude, its hermitian conjugate couples the (n-j)-atom amplitude to the n-j+1-atom amplitude. The inclusion of such off-diagonal terms in $\mathbf{H}_o$ rather than $\mathbf{H}'$ is purely a convention; we have chosen to do so since they vanish by (III.44) in case the φ_α are chosen to be single-atom energy eigenstates, Eq. (III.104). Then (III.113) reduces to

$$\mathbf{H}_o c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= \Big\{ \sum_{p=1}^{n-j} \epsilon_p + \sum_{p=n-j+1}^{n} \Big[T(X_p) + T(x_p)\Big] \Big\} \times$$

$$\times c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n). \qquad \text{(III.114)}$$

The interaction part $\mathbf{H}'$ contains a large number of terms even if terms beyond the binary interaction approximation are discarded. One finds

$$\mathbf{H}' c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= \sum_{1\le p<q\le n-j} \sum_{\alpha\beta} \left[(\alpha_p \alpha_q | V | \alpha\beta) - (\alpha_p \alpha_q | IT | \alpha\beta) - (\alpha_p \alpha_q | IV | \alpha\beta) \right]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \ldots \alpha_{q-1} \beta \alpha_{q+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n$$

$$\times x_{n-j+1} \ldots x_n)$$

$$+ \left\{ \sum_{n-j+1\le p<q\le n} \left[V(X_p X_q) + V(x_p x_q) \right] + \sum_{p=n-j+1}^{n} \sum_{q=n-j+1}^{n} V(X_p x_q) \right\}$$

$$\times c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$+ \sum_{p=1}^{n-j} \sum_{q=n-j+1}^{n} \sum_{\alpha} \left\{ \left[(\alpha_p, X_q | V | \alpha, X_q) + (\alpha_p, x_q | V | \alpha, x_q) \right] \right.$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$- \int \left[(\alpha_p, X_q | IT | \alpha, X) + K(\alpha_p X_q ; \alpha, X) T(X) \right]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_{q-1} X X_{q+1} \ldots X_n$$

$$\times x_{n-j+1} \ldots x_n) dX$$

$$- \int \left[(\alpha_p, x_q | IT | \alpha, x) + K(\alpha_p, x_q ; \alpha, x) T(x) \right]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha\alpha_{p+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_{q-1} x x_{q+1} \ldots x_n)$$

$$\times dx\Big\}$$

$$+\Big\{(j+1)(n-j)^{-\frac{1}{2}} \sum_{p=1}^{n-j} \sum_{q=n-j+1}^{n} \int \Big[(\alpha_p, X_q|V|XxX_q)' + (\alpha_p, x_q|V|Xxx_q)'\Big]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha_{p+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n X x_{n-j+1} \ldots x_n x)\, dXdx + h.c.\Big\}$$

$$+\Big\{(j+1)(j+2)\Big[(n-j)(n-j-1)\Big]^{-\frac{1}{2}} \sum_{1 \le p < q \le n-j} \int (\alpha_p \alpha_q|V|XX'xx')'$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha_{p+1} \ldots \alpha_{q-1} \alpha_{q+1} \ldots \alpha_{n-j},$$

$$, X_{n-j+1} \ldots X_n XX' x_{n-j+1} \ldots x_n xx')\, dXdX'dxdx' + h.c.\Big\}$$

$$+\Big\{(j+1)(n-j)^{-\frac{1}{2}} \sum_{1 \le p < q \le n-j} \sum_{\alpha} \int \{(\alpha_p \alpha_q|V|\alpha, Xx)' + (\alpha_q \alpha_p|V|\alpha, Xx)'$$

$$-(\alpha_p \alpha_q|IT|\alpha, Xx)' - (\alpha_q \alpha_p|IT|\alpha, Xx)' - (\alpha_p \alpha_q|IV|\alpha, Xx)'$$

$$-(\alpha_q \alpha_p|IV|\alpha, Xx)'$$

$$-\Big[K(\alpha_p \alpha_q; \alpha, Xx) + K(\alpha_q \alpha_p; \alpha, Xx)\Big]\Big[T(X) + T(x) + V(Xx)\Big]\}$$

$$\times c(\alpha_1 \ldots \alpha_{p-1}\alpha_{p+1} \ldots \alpha_{q-1}\alpha_{q+1} \ldots \alpha_{n-j}\alpha,$$

$$,X_{n-j+1} \ldots X_n X x_{n-j+1} \ldots x_n x)\,dX dx + \text{h.c.}\Big\}$$

$$+ \ldots \qquad \text{(III.115)}$$

where "..." indicates terms beyond the binary interaction approximation, and all the matrix elements occurring have been defined previously. In the terms beyond the binary interaction approximation there occur "disconnected" contributions to $\mathbf{H}'$, i.e. contributions in which the matrix elements factorize. The following 3-body term, representing coupling between kinetic energy of unbound particles and interatomic exchange, is one example:

$$- \sum_{1\le p<q\le n-j} \sum_{\alpha\beta} (\alpha_p \alpha_q |I| \alpha\beta) \sum_{r=n-j+1}^{n} [T(X_r)+T(x_r)]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1}\alpha\alpha_{p+1} \ldots \alpha_{q-1}\beta\alpha_{q+1} \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n$$

$$\times x_{n-j+1} \ldots x_n). \qquad \text{(III.116)}$$

We assume that all such disconnected terms can be canceled by series rearrangement (introduction of appropriate Ursell functions as matrix elements) as in the approach of Sakakura[6]). After this has been done, the m-particle interaction terms with $m \geq 3$ will all be connected, as are the two-particle ones, and hence may validly be neglected compared to the two-particle terms except at high densities. As in more familiar applications of the Ursell rearrangement method, one expects that introduction of such Ursell functions will not affect the two-particle interaction terms.

The part of $\mathbf{H}$ denoted by $\mathbf{H}_{\text{spont}}$ in (III.112) will be called the "spontaneous breakup" part. It is a

sum of disconnected terms in which one or more factors in each matrix element are bound state wave functions. One finds

$$\mathbf{H}_{\text{spont}}\, c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$$

$$= \sum_{p=1}^{n-j} \sum_{q=n-j+1}^{n} \sum_{r=n-j+1}^{n} \sum_{\alpha} \varphi_\alpha(X_q x_r)$$

$$\int \left[(\alpha_p | T | (Xx)' + (\alpha_p | V | (Xx)' \right]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1}, \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_{q-1} X X_{q+1} \ldots X_n x_{n-j+1} \ldots$$

$$x_{r-1} x x_{r+1} \ldots x_n)\, dX dx$$

$$+ j^{-1}(n-j+1)^{\frac{1}{2}} \sum_{p=1}^{n-j} \sum_{q=n-j+1}^{n} \sum_{r=n-j+1}^{n} \sum_{\alpha\beta} \varphi_\beta(X_q x_r)$$

$$\times \left[(\alpha_p | T | \alpha) + (\alpha_p | V | \alpha) \right]$$

$$\times c(\alpha_1 \ldots \alpha_{p-1} \alpha \alpha_{p+1} \ldots \alpha_{n-j} \beta, X_{n-j+1} \ldots X_{q-1} X_{q+1} \ldots X_n$$

$$\times x_{n-j+1} \ldots x_{r-1} x_{r+1} \ldots x_n)$$

$$+ j^{-1}(n-j+1)^{\frac{1}{2}} \sum_{p=n-j+1}^{n} \sum_{q=n-j+1}^{n} \sum_{\alpha} \varphi_\alpha(X_p x_q)$$

$$\times \left[H_{\text{Sch}} \text{ of unbound particles in } c \right]$$

$$\times c(\alpha_1 \ldots \alpha_{n-j}{}^{\alpha}, X_{n-j+1} \cdots X_{p-1} X_{p+1} \ldots X_n x_{n-j+1} \cdots x_{q-1} x_{q+1} \ldots x_n)$$

$$+ \ldots \qquad \text{(III.117)}$$

where "..." indicates more complicated direct terms, as well as terms differing from the direct terms in that the interaction factors involve exchange (but still contain factors of φ_α). $\mathbf{H}_{\text{spont}}$ is quite unphysical. It is not hermitian, the terms conjugate to the ones shown vanishing by the bound state-continuum orthogonality constraints (III.44). In fact, it is not difficult to see that $\mathbf{H}_{\text{spont}}$ always leads to breakup of atoms, never formation (this will be more evident when we introduce second quantization). Thus if $\mathbf{H}_{\text{spont}}$ were actually effective it would lead to an instability in which all atoms would dissociate spontaneously. However, one notes in the first place that $\mathbf{H}_{\text{spont}}$ has vanishing matrix elements between any two states in the physical subspace $\mathcal{P}$ of the ideal state space $\mathcal{I}$. Furthermore, we shall see later that when the appropriate projection operator for the constraints (III.44) (the projector onto $\mathcal{P}$) has been introduced, the projected Hamiltonian will contain additional orthogonalization terms which will exactly cancel $\mathbf{H}_{\text{spont}}$. From the physical point of view, spontaneous breakup of atoms does not occur because the unbound particles move in orbitals orthogonal to the bound states.

E. Second Quantization

It is now a simple matter to introduce ideal atomic annihilation and creation operators a_α and $a_\alpha{}^\dagger$, and unbound proton and electron annihilation and creation operators $\psi(X), \psi^\dagger(X), \psi(x)$, and $\psi^\dagger(x)$, by the Fock representation method of Sec. I. It is not practical to use an explicit matrix representation such as (I.17) for the state vectors, since the matrices would have to be represented in three dimensions (one for atomic, one for nuclear, one

for electronic variables). Hence we define the Fock vectors implicitly by their amplitudes c. As usual, the amplitudes must be generalized by allowing arbitrary and independent numbers of atoms, protons, and electrons. Thus we consider the amplitudes $c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_n)$ so far employed as special cases of the more general amplitudes $c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$ with m not necessarily equal to n, and allow j, n, and m to range over all integral values satisfying

$$0 \leq n < \infty, \quad 0 \leq m < \infty, \quad 0 \leq j \leq n. \tag{III.118}$$

An amplitude c_0 with no arguments represents the vacuum amplitude, where the vacuum state $|0)$ is defined as the Fock vector with amplitudes

$$c_0 = 1,$$

$$c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) = 0 \quad \text{unless } n = m = n-j = 0. \tag{III.119}$$

The inner product expression (III.56) is generalized to

$$(c|c') \equiv \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{j=0}^{n} \sum_{\alpha_1 \ldots \alpha_{n-j}} \int c^*(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) \times c'(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) \times dX_{n-j+1} \ldots dX_n dx_{n-j+1} \ldots dx_m \tag{III.120}$$

where the summand is to be interpreted as $c_0^* c_0'$ in case $n = m = n-j = 0$.

In analogy with (I.19), we define the annihilation operators by

$$a_\alpha|0) \equiv \psi(X)|0) \equiv \psi(x)|0) \equiv 0 \qquad \text{(III.121)}$$

and

$$\begin{aligned}
&a_\alpha c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m)\\
&\quad\equiv (n-j+1)^{\frac{1}{2}}\, c(\alpha_1\ldots\alpha_{n-j}\alpha,X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m),\\
&\psi(X)\; c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m)\\
&\quad\equiv (j+1)^{\frac{1}{2}}\, c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X X_n x_{n-j+1}\ldots x_m),\\
&\psi(x)\; c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m)\\
&\quad\equiv (m-n+j+1)^{\frac{1}{2}}\, c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m x).
\end{aligned} \qquad \text{(III.122)}$$

Similarly, the creation operators are defined* by

$$\begin{aligned}
&a_\alpha^\dagger c(\alpha_1\ldots\alpha_{n-j},X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m)\\
&\quad\equiv (n-j)^{-\frac{1}{2}} \sum_{p=1}^{n-j} \delta_{\alpha_p\alpha} c(\alpha_1\ldots\alpha_{p-1}\alpha_{p+1}\ldots\alpha_{n-j},\\
&\qquad\qquad X_{n-j+1}\ldots X_n x_{n-j+1}\ldots x_m),
\end{aligned}$$

---

* Strictly speaking, one should not use the notations $a_\alpha^\dagger$, $\psi^\dagger(X)$, and $\psi^\dagger(x)$ until it has been shown that these creation operators are indeed the hermitian conjugates of (III.122). This is not difficult to do, using the definition (III.120) of the inner product and the symmetry of the c's.

$$\psi^{\dagger}(X)\ c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$$

$$\equiv j^{-\frac{1}{2}} \sum_{p=n-j+1}^{n} (-1)^{n-p} \delta(X_p - X) c(\alpha_1 \ldots \alpha_{n-j},$$

$$X_{n-j+1} \ldots X_{p-1} X_{p+1} \ldots X_n x_{n-j+1} \ldots x_m),$$

$$\psi^{\dagger}(x)\ c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$$

$$\equiv (m-n+j)^{-\frac{1}{2}} \sum_{p=n-j+1}^{m} (-1)^{m-p} \delta(x_p - x) c(\alpha_1 \ldots \alpha_{n-j},$$

$$X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_{p-1} x_{p+1} \ldots x_m). \qquad \text{(III.123)}$$

The commutation relations

$$[a_\alpha, a_\beta]_- = 0,\ [a_\alpha a_\beta^{\dagger}]_- = \delta_{\alpha\beta};$$

$$[\psi(X), \psi(X')]_+ = 0,\ [\psi(X), \psi^{\dagger}(X')]_+ = \delta(X - X');$$

$$[\psi(x), \psi(x')]_+ = 0,\ [\psi(x), \psi^{\dagger}(x')]_+ = \delta(x - x');$$

$$[a_\alpha, \psi(X)]_- = [a_\alpha, \psi^{\dagger}(X)]_- = [a_\alpha, \psi(x)]_- = [a_\alpha, \psi^{\dagger}(x)]_- = 0;$$

$$[\psi(X), \psi(x)]_- = [\psi(X), \psi^{\dagger}(x)]_- = 0 \qquad \text{(III.124)}$$

follow* directly from the definitions (III.122) and (III.123); here $[\ ]_+$ denotes the anticommutator and $[\ ]_-$ the commutator. The atom number operators N_α and unbound proton and electron number density operators $\rho(X)$ and $\rho(x)$ are defined by

$$N_\alpha \equiv a_\alpha^\dagger a_\alpha, \quad \rho(X) \equiv \psi^\dagger(X)\psi(X), \quad \rho(x) \equiv \psi^\dagger(x)\psi(x) \qquad \text{(III.125)}$$

and satisfy

$$N_\alpha\,|0) = \rho(X)\,|0) = \rho(x)\,|0) = 0,$$

$$N_\alpha c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$$

$$= \sum_{p=1}^{n-j} \delta_{\alpha_p \alpha} c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m),$$

$$\rho(X) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$$

$$= \sum_{p=n-j+1}^{n} \delta(X_p - X) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m),$$

$$\rho(x) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m)$$

---

* Note that the proton field operators commute (not anti-commute) with the electron field operators. This is not assumed; it follows from (III.122) and (III.123). The proton operators could have been made to anticommute with the electron ones by using more complicated phase factors in the definitions. However, there seems to be little point in doing so. In any event, physical results are independent of whether the proton operators commute or anticommute with the electron ones.

$$= \sum_{p=n-j+1}^{m} \delta(x_p - x) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) . \qquad \text{(III.126)}$$

The total atom number operator N_{at}, total unbound proton number operator N_{nuc}, and total unbound electron number operator N_{elec} are defined as

$$N_{at} \equiv \sum_{\alpha} N_{\alpha}, \quad N_{nuc} \equiv \int dX \rho(X), \quad N_{elec} \equiv \int dx \rho(x) \qquad \text{(III.127)}$$

and satisfy

$$N_{at} c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) = (n-j) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m),$$

$$N_{nuc} c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) = j c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m),$$

$$N_{elec} c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) = (m-n+j) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) . \qquad \text{(III.128)}$$

An arbitrary state vector $|c)$ in the ideal state space* $\mathcal{I}$ can be represented in terms of the Fock vacuum $|0)$ and the creation operators defined above as

---

* The meaning of $\mathcal{I}$ has now been extended to allow general values of n, m and j satisfying (III.118), not merely those with m = n.

$$|c) = \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} \sum_{j=0}^{n} [n-j)!j!(m-n+j)!]^{-\frac{1}{2}}$$

$$\sum_{\alpha_1 \ldots \alpha_{n-j}} \int dX_{n-j+1} \ldots dX_n dx_{n-j+1} \ldots dx_m$$

$$\times c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) a_{\alpha_1}^{\dagger} \ldots a_{\alpha_{n-j}}^{\dagger}$$

$$\times \psi^{\dagger}(X_{n-j+1}) \ldots \psi^{\dagger}(X_n) \psi^{\dagger}(x_{n-j+1}) \ldots \psi^{\dagger}(x_m) |0) \quad \text{(III.129)}$$

where the term with n=m=n-j=0 is to be interpreted as $c_0|0)$. The state vectors in the physical subspace $\mathcal{P}$ all have m=n (with n the same for all states in $\mathcal{P}$) and hence satisfy

$$(N_{at} + N_{nuc})|c) = (N_{at} + N_{elec})|c) = n|c), \text{ all } |c) \in \mathcal{P}. \quad \text{(III.130)}$$

This is the source of a superselection rule which will be discussed presently.

The state vectors in $\mathcal{P}$ must also satisfy the bound state-continuum othogonality constraints (III.44). In $\mathcal{J}$ these take the simple form

$$A_{\alpha}|c) = 0, \text{ all } \alpha \text{ and all } |c) \in \mathcal{P}, \quad \text{(III.131)}$$

where*

$$A_{\alpha} \equiv \int dXdx\ \varphi_{\alpha}^{*}(Xx)\psi(x)\psi(X). \quad \text{(III.132)}$$

* $A_{\alpha}^{\dagger}$ can be interpreted as the creation operator for an "unbound proton-electron pair in the bound state φ_{α}," and conditions (III.131) ensure that such unphysical pairs do not exist.

In fact, upon substituting (III.129) into (III.131), applying the normal-ordering theorem or (III.124), and equating the various linearly independent components separately to zero one concludes that

$$\int \varphi_\alpha^*(X_n x_m) c(\alpha_1 \ldots \alpha_{n-j}, X_{n-j+1} \ldots X_n x_{n-j+1} \ldots x_m) dX_n dx_m = 0. \tag{III.133}$$

In the special case n = m, Eqs. (III.133) reduce to (III.44). Conversely, upon multiplying (III.133) by

$$[j!(m-n+j)!/(n-j)!]^{\frac{1}{2}} a_{\alpha_1}^\dagger \ldots a_{\alpha_{n-j}}^\dagger \times \psi^\dagger(X_{n-j+1}) \ldots \psi^\dagger(X_{n-1}) \psi^\dagger(x_{n-j+1}) \ldots \psi^\dagger(x_{m-1}) |0),$$

summing over $\alpha_1 \ldots \alpha_{n-j}$, and integrating over $X_{n-j+1} \ldots X_{n-1}$ and $x_{n-j+1} \ldots x_{m-1}$, one deduces (III.131). The conditions (III.131) are in turn equivalent to

$$A_\alpha^\dagger A_\alpha |c) = 0, \text{ all } \alpha \text{ and all } |c) \in \mathcal{P}. \tag{III.134}$$

It is obvious that Eqs. (III.134) follow from (III.133); conversely, (III.134) implies

$$(c|A_\alpha^\dagger A_\alpha|c) = \|A_\alpha|c)\|^2 = 0, \text{ all } \alpha \text{ and all } |c) \in \mathcal{P}, \tag{III.135}$$

which is true if and only if (III.131) is satisfied. Finally, the set of conditions (III.134) is equivalent to the single condition

$$\sum_\alpha A_\alpha^\dagger A_\alpha |c) = 0, \text{ all } |c) \in \mathcal{P}, \tag{III.136}$$

since the operators $A_\alpha^\dagger A_\alpha$ are all positive semidefinite. Using (III.77), this can be written as

$$\Delta|c) = 0, \text{ all } |c) \in \mathcal{P} \qquad \text{(III.137)}$$

where

$$\Delta \equiv \int dXdX'dxdx' \psi^{\dagger}(X)\psi^{\dagger}(x)\Delta(Xx,X'x')\psi(x')\psi(X'). \qquad \text{(III.138)}$$

The physical subspace $\mathcal{P}$ is the subspace of $\mathcal{J}$ satisfying both (III.130) and (III.137). A method of satisfying (III.137) by the introduction of an appropriate projection operator will be discussed in Sec. III. G.

We now exhibit the representation of the Hamiltonian* as an operator in $\mathcal{J}$, in terms of annihilation and creation operators. It follows from (III.113), (III.122), and (III.123) that

$$\mathbf{H}_o = \sum_{\alpha\beta} [(\alpha|T|\beta) + (\alpha|V|\beta)] a_{\alpha}^{\dagger} a_{\beta}$$

$$+ \int dX \psi^{\dagger}(X) T(X) \psi(X) + \int dx \psi^{\dagger}(x) T(x) \psi(x)$$

$$+ \left\{ \sum_{\alpha} \int dXdx [(\alpha|T|Xx)' + (\alpha|V|Xx)'] a_{\alpha}^{\dagger} \psi(x)\psi(X) + \text{h.c.} \right\}$$

(III.139)

In case the φ_α are chosen to be energy eigenstates so that (III.114) is valid, this simplifies to

$$\mathbf{H}_o = \sum_{\alpha} \epsilon_{\alpha} N_{\alpha} + \int dX \psi^{\dagger}(X) T(X) \psi(X) + \int dx \psi^{\dagger}(x) T(x) \psi(x). \qquad \text{(III.140)}$$

---

* In fact, since we have not been explicit regarding the forms of $T(X), T(x), V(XX')$, $V(xx')$ and $V(Xx)$, this will implicitly define any one or two-particle observable as an operator on $\mathcal{J}$.

The physical interpretation is obvious: $\mathbf{H}_0$ is the sum of energies of noninteracting atoms, protons, and electrons. Similarly, one finds for the interaction part (III.115)

$$\mathbf{H}' = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} [(\alpha\beta|V|\gamma\delta) - (\alpha\beta|IT|\gamma\delta) - (\alpha\beta|IV|\gamma\delta)] \times a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma$$

$$+ \tfrac{1}{2} \int dX dX' \psi^\dagger(X) \psi^\dagger(X') V(XX') \psi(X') \psi(X)$$

$$+ \tfrac{1}{2} \int dx dx' \psi^\dagger(x) \psi^\dagger(x') V(xx') \psi(x') \psi(x)$$

$$+ \int dX dx \, \psi^\dagger(X) \psi^\dagger(x) V(Xx) \psi(x) \psi(X)$$

$$+ \sum_{\alpha\beta} \int dX a_\alpha^\dagger \psi^\dagger(X) (\alpha, X|V|\beta, X) \psi(X) a_\beta$$

$$+ \sum_{\alpha\beta} \int dx a_\alpha^\dagger \psi^\dagger(x) (\alpha, x|V|\beta, x) \psi(x) a_\beta$$

$$- \sum_{\alpha\beta} \int dX dX' a_\alpha^\dagger \psi^\dagger(X) [(\alpha, X|IT|\beta, X') + K(\alpha, X; \beta X') T(X')] \times \psi(X') a_\beta$$

$$- \sum_{\alpha\beta} \int dx dx' a_\alpha^\dagger \psi^\dagger(x) [(\alpha, x|IT|\beta, x') + K(\alpha, x; \beta, x') T(x')] \times \psi(x') a_\beta$$

$$+ \Big[\sum_\alpha \int dX dx dX' a_\alpha^\dagger \psi^\dagger(X') (\alpha, X'|V|XxX') \times \psi(X') \psi(x) \psi(X) + \text{h.c.} \Big]$$

$$+\Big[\sum_{\alpha}\int dX dx dx' a_{\alpha}^{\dagger}\psi^{\dagger}(x')(\alpha,x'|V|Xxx')'\psi(x')\psi(x)\psi(X) + \text{h.c.}\Big]$$

$$+\Big[\tfrac{1}{2}\sum_{\alpha\beta}\int dX dX' dx dx' a_{\alpha}^{\dagger}a_{\beta}^{\dagger}(\alpha\beta|V|XX'xx')'\psi(x')\psi(x)\psi(X')\psi(X) + \text{h.c.}\Big]$$

$$+\Big\{\sum_{\alpha\beta\gamma}\int dX dx\, a_{\alpha}^{\dagger}a_{\beta}^{\dagger}\Big\{(\alpha\beta|V|\gamma,Xx)' - (\alpha\beta|IT|\gamma,Xx)' - (\alpha\beta|IV|\gamma,Xx)' - K(\alpha\beta;\gamma,Xx)[T(X)+T(x)+V(Xx)]\Big\}\psi(x)\times\psi(X)a_{\gamma} + \text{h.c.}\Big\}$$

$$+\ldots \qquad \text{(III.141)}$$

where, as usual, "..." denotes terms beyond the binary interaction approximation. Note that although the terms $a^{\dagger}a^{\dagger}\psi\psi a$ and $a^{\dagger}a^{\dagger}\psi\psi\psi\psi$ represent 3-particle and 4-particle collisions, their hermitian conjugates, $a^{\dagger}\psi^{\dagger}\psi^{\dagger}aa$ and $\psi^{\dagger}\psi^{\dagger}\psi^{\dagger}\psi^{\dagger}aa$, represent <u>binary</u> atomic collisions, with the two atoms breaking up into (one atom + one proton + one electron) or (no atoms + two protons + two electrons). Hence such terms are included in the binary interaction approximation.

Finally, the "spontaneous breakup" Hamiltonian is found to be

$$\mathbf{H}_{\text{spont}} = \sum_{\alpha\beta}\int dX dx a_{\alpha}^{\dagger}A_{\beta}^{\dagger}[(\alpha|T|Xx)' + (\alpha|V|(Xx)')]\times\psi(x)\psi(X)a_{\beta}$$

$$+ \sum_{\alpha\beta\gamma} a_\alpha^\dagger [(\alpha|T|\beta)+(\alpha|V|\beta)] A_\gamma^\dagger a_\gamma a_\beta$$

$$+ \sum_\alpha \int dX\ \psi^\dagger(X) A_\alpha^\dagger T(X) \psi(X) a_\alpha$$

$$+ \sum_\alpha \int dx\ \psi^\dagger(x) A_\alpha^\dagger T(x) \psi(x) a_\alpha + \ldots \qquad \text{(III.142)}$$

where, as stated previously, we have only written down enough terms to indicate the general structure of $\mathbf{H}_{\text{spont}}$. Because all terms in $\mathbf{H}_{\text{spont}}$ contain at least one factor $A_\alpha^\dagger$ on the left, all matrix elements of $\mathbf{H}_{\text{spont}}$ between states in $\mathcal{P}$ vanish by (III.131). For essentially the same reason, we shall see later that $\mathbf{H}_{\text{spont}}$ is annihilated by the projector onto $\mathcal{P}$.

The diagramatic representations of the direct interaction terms in $\mathbf{H}'$ are shown in Fig. 3; the notation is the same as in Fig. 2. The terms in $\mathbf{H}'$ involving exchange (more precisely, coupling of exchange to kinetic or potential energy) are of the same general structure, as is clear from (III.141). The representation of the first term in the "spontaneous breakup" Hamiltonian $\mathbf{H}_{\text{spont}}$ is shown in Fig. 4. Since the spontaneous breakup vertex does not involve any true interaction, the nonexistence of any potential causing the breakup is symbolized by a dashed line connecting the vertex to the letter S (for "spontaneous"). Note that the diagram is disconnected, as is the case for all terms in $\mathbf{H}_{\text{spont}}$.

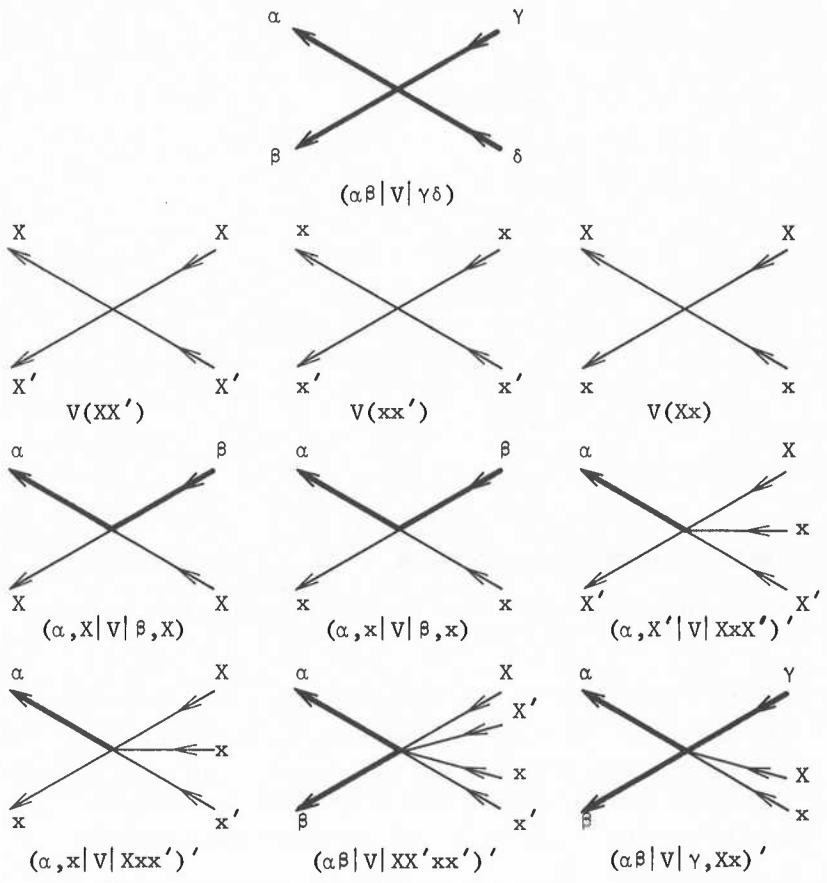


Figure 3.

Direct interaction processes included, together with their hermitian conjugates and similar exchange processes, in the binary interaction approximation to $\mathbf{H}'$.

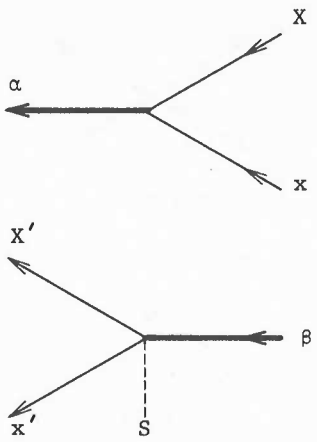


Figure 4.

Representation of one term in $\mathbf{H}_{\text{spont}}$. The top half of the diagram arises from the true dynamical processes $(\alpha|T|Xx)'$ and $(\alpha|V|Xx)'$.

The terms in $\mathbf{H}$ are the same* as those occurring in the Hamiltonian $\tilde{H}$ of Brittin and Stolt[5]), obtained by a different method, except that $\tilde{H}$ does not contain an analog of our $\mathbf{H}_{\text{spont}}$, and that all the exchange terms in our $\mathbf{H}$ have the opposite sign from the corresponding exchange terms in $\tilde{H}$. This sign difference is

---

* Actually, the terms involving $(\alpha\beta|V|XX'xx')$ and its conjugate do not appear in the expression for H given by Brittin and Stolt, but they would undoubtedly appear in a more accurate expression.

apparently connected with the fact that $\tilde{H}$ is constructed to have the same eigenvalues as the Schrödinger Hamiltonian (to the given approximation), whereas $\mathbf{H}$ is constructed to give the correct matrix elements. We shall see in Sec. III,F. that the Hamiltonian H, related to $\mathbf{H}$ by (III.66), has the same eigenvalues as the Schrödinger Hamiltonian, and is therefore more closely related to Brittin and Stolt's Hamiltonian $\tilde{H}$. Our $\mathbf{H}$ also resembles another Hamiltonian obtained, by a still different method, by Sakakura[e]). The precise relationship between $\mathbf{H}$, H, $\tilde{H}$, and Sakakura's Hamiltonian has not yet been elucidated.

We next discuss the superselection rules related to (III.130). It is obvious from the physics, and can also be verified directly from (III.139)-(III.142), that the processes in which atoms, protons, or electrons appear or disappear are such that whenever an atom disappears, an unbound proton-electron pair appears, and vice versa. Hence

$$\left[\ \mathbf{H}\ , (N_{at} + N_{nuc})\right] = \left[\ \mathbf{H}\ , (N_{at} + N_{elec})\right] = 0. \tag{III.143}$$

More generally, if $\mathbf{A}$ is any operator on $\mathcal{J}$ derived from a Schrödinger operator A_{Sch} representing a physical observable, then

$$\left[\ \mathbf{A}\ , (N_{at} + N_{nuc})\right] = \left[\ \mathbf{A}\ , (N_{at} + N_{elec})\right] = 0. \tag{III.144}$$

It follows that every observable $\mathbf{A}$ has vanishing matrix elements between any two states belonging to

different eigenvalues of $N_{at} + N_{nuc}$, or different eigenvalues of $N_{at} + N_{elec}$. In fact, states in $\mathcal{P}$ corresponding to n-proton, n-electron Schrödinger wave functions are simultaneous eigenstates of $N_{at} + N_{nuc}$ and $N_{at} + N_{elec}$ with eigenvalue n for both; this is the meaning of (III.130). However, (III.144) is valid on the entire ideal state space $\mathcal{I}$.

The metric operator M, which is important in the probability interpretation because of (III.57), and will also be employed in evaluating the second-quantized H, can also be represented in terms of annihilation and creation operators. It follows from (III.59)-(III.62), (III.122), and (III.123), or by comparison of (III.59)-(III.62) with (III.113), (III.115), (III.139), and (III.141), that

$$M = 1 - B + \dots ,$$

$$B = \tfrac{1}{2} \sum_{\alpha\beta\gamma\delta} (\alpha\beta|I|\gamma\delta) a_\alpha^\dagger a_\beta^\dagger a_\delta a_\gamma$$

$$+ \sum_{\alpha\beta} \int dX dX' K(\alpha,X;\beta,X') a_\alpha^\dagger \psi^\dagger(X) \psi(X') a_\beta$$

$$+ \sum_{\alpha\beta} \int dx dx' K(\alpha,x;\beta,x') a_\alpha^\dagger \psi^\dagger(x) \psi(x') a_\beta$$

$$+ \sum_{\alpha\beta\gamma} \int dX dx \left[K(\alpha\beta;\gamma,Xx) a_\alpha^\dagger a_\beta^\dagger a_\gamma \psi(x) \psi(X) \right.$$

$$\left. + K(\gamma,Xx;\alpha\beta) \psi^\dagger(X) \psi^\dagger(x) a_\gamma^\dagger a_\beta a_\alpha \right]. \qquad \text{(III.145)}$$

The first line of this expression is the same as the exchange operator* I of Sec. I, Eq. (I.28), except that now only bound states are included in the summations. The remaining lines of (III.145) represent explicit bound state-continuum exchange effects which were implicit in (I.28) because of the inclusion of continuum atomic states. Eq. (I.28) also includes continuum-continuum exchange effects which are absent in (III.145), being now accounted for automatically by the antisymmetry of the c's or the anti-commutation relations satisfied by the $\psi(X)$ and $\psi(x)$ operators and their hermitian conjugates.

F. Energy Spectrum and Statistical Mechanics

It is convenient at this point to introduce Dirac notation for the space** $\mathcal{S}$ of n-proton, n-electron Schrödinger wave functions $\psi(X_1 \ldots X_n x_1 \ldots x_n)$. We have shown that the c's in (III.43) are uniquely determined by ψ provided that the symmetry and antisymmetry requirements on the c's are complemented by the strong orthogonality constraints (III.44); conversely, every such set of c's uniquely determines a properly antisymmetric ψ by (III.43). Thus a state in $\mathcal{S}$ can be denoted by $|c\rangle$, the Dirac ket notation standing in this case for the set of all such c's, and the inner product $\langle c|c'\rangle$ between two such states [previously denoted by (c,c')] is given by (III.46). The vectors $|c) \in \mathcal{P}$ are in one-one correspondence $|c) \leftrightarrow |c\rangle$ with the vectors $|c\rangle \in \mathcal{S}$. Let $\{|c_i)\}$ be any complete orthonormal basis in $\mathcal{P}$:

$$(c_i|c_j) = \delta_{ij}, \qquad \sum_i |c_i)(c_i| = 1_{\mathcal{P}} \tag{III.146}$$

where $1_{\mathcal{P}}$ is the unit operator on $\mathcal{P}$ (not on the ideal state space $\mathcal{J}$, of which $\mathcal{P}$ is a proper subspace). This

---

* For the case of hydrogen atoms, $I_{nuc} = I_{elec} \equiv I$.

** We assume that $\mathcal{S}$ is a separable Hilbert space. This will be the case if the particles are confined to a finite volume Ω, with either box enclosure or periodic boundary conditions on the ψ's.

orthonormal basis in $\mathcal{P}$ determines a set $\{|c_i\rangle\}$ of vectors in $\mathcal{S}$ in one-one correspondence $|c_i\rangle \leftrightarrow |c_i)$ according to (III.43) and (III.45). The set $\{|c_i\rangle\}$ is not orthonormal, due to the occurrence of the metric operator M in (III.57). On the other hand, it is complete and linearly independent*.

The operators $\mathbf{A}$ defined by (III.66) and evaluated according to (III.68) are constructed in such a way that, by (III.65) and (III.46),

$$(c_i| \mathbf{A} |c_j) = \langle c_i|A_{Sch}|c_j\rangle. \qquad \text{(III.147)}$$

In particular, denoting the matrix elements of the Hamiltonian $\mathbf{H}$ in the orthonormal basis $\{|c_i)\}$ by $\mathbf{H}_{ij}$, one has

$$\mathbf{H}_{ij} \equiv (c_i| \mathbf{H} |c_j) = \langle c_i|H_{Sch}|c_j\rangle. \qquad \text{(III.148)}$$

However, the eigenvalues of the n-proton, n-electron Hamiltonian H_{Sch} are not in general** equal to those of $\mathbf{H}$, since the matrix elements of H_{Sch} are evaluated in a non-orthonormal basis, in contrast with those $\mathbf{H}_{ij}$

---

* If the strong orthogonality constraints had not been imposed, one could not have made this statement. In fact, the correspondence between $\mathcal{S}$ and $\mathcal{P}$ would not have been one-one.

**

**They do, however reduce to those of $\mathbf{H}_o + \mathbf{H}'$ in the limit of zero density.

of $\mathbf{H}$. Furthermore, $\mathbf{H}$ acts on the entire ideal state space $\mathfrak{H}$ and does not leave its physical subspace $\mathfrak{P}$ invariant, so that $\mathbf{H}$ does not necessarily have any eigenstates lying in $\mathfrak{P}$.

Consider, on the other hand, the operator H given, according to (III.66), by

$$H = M^{-1}\,\mathbf{H} \tag{III.149}$$

or evaluated directly from (III.67) with $A = H_{Sch}$ on the right side and $A = H$ on the left side. It can be shown, in analogy with Dyson's theory of spin-wave interactions[9] that the eigenvalues of H are identical with those of H_{Sch}. Note first that since the set $\{|c_i\rangle\}$ is complete, one can expand

$$H_{Sch}|c_i\rangle = \sum_j H_{ji}|c_j\rangle. \tag{III.150}$$

On the other hand, since the set $\{|c_i\rangle\}$ is not orthonormal, one has

$$H_{ji} \neq \langle c_j|H_{Sch}|c_i\rangle, \tag{III.151}$$

i.e. the H_{ji} are not the matrix-elements of H_{Sch} occurring in (III.148). It follows from the definition of H that*

---

* In order to interpret the c's in (III.67) directly as constituting the basis $\{|c_i\rangle\}$, this basis must be reinterpreted as a basis on $\mathfrak{H}$, at the expense of lengthening the discussion. The conclusion is the same as that of the abbreviated discussion above. The essential point is that H leave $\mathfrak{P}$ invariant, whereas $\mathbf{H}$ does not.

$$H|c_i) = \sum_j H_{ji}|c_j) \qquad \text{(III.152)}$$

with the same coefficients H_{ji} as occur in (III.150). Furthermore, one has by (III.146)

$$H_{ji} = (c_j|H|c_i). \qquad \text{(III.153)}$$

Let $|E\rangle$ be an eigenstate of H_{Sch} and expand it in terms of the nonorthogonal set $\{|c_i\rangle\}$:

$$|E\rangle = \sum_i (c_i|E)\ |c_i\rangle; \qquad \text{(III.154)}$$

again, $(c_i|E) \neq \langle c_i|E\rangle$. Then the eigenvalue equation

$$H_{Sch}|E\rangle = E|E\rangle \qquad \text{(III.155)}$$

is equivalent to

$$\sum_{ij} (H_{ji} - E\delta_{ji})\ (c_i|E)\ |c_j\rangle = 0 \qquad \text{(III.156)}$$

or, since the $|c_j\rangle$ are linearly independent,

$$\sum_i (H_{ji} - E\delta_{ji})\ (c_i|E) = 0. \qquad \text{(III.157)}$$

The eigenvalues of H_{Sch} are the values of E for which this set of equations has a nontrivial solution for the amplitudes $(c_i|E)$, and the corresponding eigenstates are given by (III.154). Now define the state $|E) \in \mathcal{P}$ by

$$|E) \equiv \sum_i (c_i|E)\,|c_i) \qquad \text{(III.158)}$$

with the same amplitudes $(c_i|E)$. Since the set $\{|c_i)\}$ is orthonormal, $(c_i|E)$ is the inner product of $|c_i)$ with $|E)$. Furthermore, by (III.152) one has

$$H|E) = \sum_{ij} H_{ji}\,(c_i|E)\,|c_j). \qquad \text{(III.159)}$$

It follows that the solutions E of (III.157) are also eigenvalues of H, with the corresponding eigenfuntions being given by (III.158). We conclude that the eigenvalue spectrum of H determined in the ideal state space $\mathcal{J}$ includes all eigenvalues of H_{Sch}. More specifically, those eigenvectors $|E)$ of H which lie in the physical subspace $\mathcal{P}$ have eigenvalues equal to eigenvalues of H_{Sch}, and conversely, every eigenvalue of H_{Sch} is equal to an eigenvalue of H with an eigenvector $|E) \in \mathcal{P}$. The relation between the eigenvectors $|E\rangle$ of H_{Sch} and those $|E)$ of H is given by (III.154) and (III.158) with the same coefficients $(c_i|E)$. The analog of (III.154) with $(c_i|E)$ replaced by $\langle c_i|E\rangle$ is false. If H has any eigenvectors $|E) \notin \mathcal{P}$, they do not correspond to eigenvalues and eigenvectors of H_{Sch}.

These results can now be used to relate the partition function Z evaluated over the n-proton, n-electron Schrödinger space $\mathcal{S}$ to ones evaluated over the ideal state space $\mathcal{J}$ and its physical subspace $\mathcal{P}$. Denote the traces over $\mathcal{S}$, $\mathcal{P}$, and $\mathcal{J}$ by $\mathrm{Tr}_{\mathcal{S}}$, $\mathrm{Tr}_{\mathcal{P}}$, and $\mathrm{Tr}_{\mathcal{J}}$, and let P be the projector onto $\mathcal{P}$. Then it follows from the above results that

$$Z \equiv \mathrm{Tr}_{\mathcal{S}}\; e^{-\beta H_{Sch}} = \mathrm{Tr}_{\mathcal{P}}\, e^{-\beta H} = \mathrm{Tr}_{\mathcal{J}}\; (Pe^{-\beta H}). \qquad \text{(III.160)}$$

Since H leaves $\mathcal{P}$ invariant, it commutes with P. Thus if we define a projected Hamiltonian $\mathcal{H}$ by

$$\mathcal{H} \equiv PH \qquad \text{(III.161)}$$

as in Sec. II.D., one can use the same argument (II.14)-(II.17) to conclude that the eigenvalue spectra of $\mathcal{K}$ and H_{Sch} coincide except for the spurious eigenvalue zero of $\mathcal{K}$. The eigenstates of $\mathcal{K}$ with $E \neq 0$ are eigenstates of H with the same eigenvalue and lying in $\mathcal{P}$, whereas eigenstates of $\mathcal{K}$ with eigenvalue zero are in general linear combinations of eigenstates of H with eigenvalue zero (if there are any such) and <u>arbitrary</u> states in the "completely unphysical" subspace $\mathcal{J} - \mathcal{P}$. It follows that if $\mathrm{Tr}' e^{-\beta\mathcal{K}}$ is defined as the trace of $e^{-\beta\mathcal{K}}$ over the ideal state space $\mathcal{J}$ but with the contribution from the eigenvalue zero excluded, then one has

$$Z \approx \mathrm{Tr}' e^{-\beta\mathcal{K}} \tag{III.162}$$

where $\approx$ denotes exact equality in case H does not have the eigenvalue zero, and asymptotic equality in the thermodynamic limit in case it does*. The same conclusion (III.162) also follows from (III.160) and the identity

$$Pe^{-\beta H} = e^{-\beta\mathcal{K}} - 1 + P \tag{III.163}$$

which is easily proved upon noting that P commutes with $\mathcal{K}$ and that on $\mathcal{J}-\mathcal{P}$ one has $\mathcal{K} = 0$, $e^{-\beta\mathcal{K}} = 1$, and $P = 0$.

According to (III.149) and (III.161),

$$\mathcal{K} = PM^{-1}\,\mathbf{H}\,. \tag{III.164}$$

The explicit form of P in the binary interaction approximation can be found by examining the algebraic properties of the bound state kernel (III.77) and the corresponding second-quantized operator Δ, Eq. (III.138). The identity

---

* We assume that if H has the eigenvalue zero, its degeneracy does not grow exponentially with the volume Ω.

$$\int \Delta(Xx,X''x'')\,\Delta(X''x'',X'x')\,dX''dx'' = \Delta(Xx,X'x') \quad \text{(III.165)}$$

is a trivial consequence of the definition (III.77) and orthonormality of the φ_α. Then using the normal ordering theorem one finds that the operator (III.138) satisfies

$$\Delta^2=\Delta + \int (\ )\,\psi^\dagger\psi^\dagger\psi^\dagger\psi\psi\psi + \int (\ )\,\psi^\dagger\psi^\dagger\psi^\dagger\psi^\dagger\psi\psi\psi\psi \quad \text{(III.166)}$$

where $(\ )\,\psi^\dagger\psi^\dagger\psi^\dagger\psi\psi\psi$ symbolizes a sum of normally ordered terms each containing three creation operators on the left and three annihilation operators on the right; such terms behave like three-body interactions, i.e. they only contribute when three unbound particles (protons and/or electrons) collide. Similarly, the terms $(\ )\,\psi^\dagger\psi^\dagger\psi^\dagger\psi^\dagger\psi\psi\psi\psi$ behave like four-body interactions. It is consistent with the binary interaction approximation to drop such three-body and four-body interaction terms, in which case

$$\Delta^2 = \Delta + \ldots \quad \text{(III.167)}$$

where, as usual, "..." stands for terms beyond the binary interaction approximation. One concludes that within the binary interaction approximation, Δ is a projection operator, with eigenvalues zero and unity. Any state vector $|c) \in \mathcal{J}$ can be resolved into eigenstates of Δ with these eigenvalues:

$$|c) = |c_0) + |c_1),$$

$$\Delta|c_0) = 0, \quad \Delta|c_1) = |c_1). \quad \text{(III.168)}$$

By (III.137), $|c_0) \in \mathcal{P}$; then, since $|c_1)$ is orthogonal to $|c_0)$ because it belongs to a different eigenvalue, one concludes that $|c_1) \in \mathcal{J}\text{-}\mathcal{P}$. Thus Δ is the projector onto

$\mathcal{J} - \mathcal{P}$, and hence* $1 - \Delta$ is the projector P onto the physical subspace $\mathcal{P}$:

$$P = 1 - \Delta + \ldots, \qquad \text{(III.169)}$$

the dots denoting terms beyond the binary interaction approximation.

The leading terms in H, can be found by substitution of (III.70) and (III.145) into (III.149) and algebraic manipulations arising in the application of the normal-ordering theorem. Similarly, the leading terms in $\mathcal{K}$ can be found** by substitution of (III.169) into (III.161) or (III.164) and appropriate contractions. The resultant expression are lengthy and will not be written out here. Both H and $\mathcal{K}$ consist of terms with the same structure as those in $\mathbf{H}_o + \mathbf{H}'$, but the detailed matrix elements

---

*P is the projector onto all of $\mathcal{P}$, not a proper subspace, since, by definition, $\mathcal{P}$ is the space of all eigenstates of Δ with eigenvalue zero. We assume that $\mathcal{J}$ has already been restricted by imposition of the constraints (III.130), i.e. these constraints are not included in P. Note that Δ commutes with N_{at}, N_{nuc}, and N_{elec}, and hence with both $N_{at} + N_{nuc}$ and $N_{at} + N_{elec}$. In an application to statistical mechanics one could impose (III.130) only as thermal averages by using a grand ensemble, replacing $\mathcal{K}$ in (III.162) by $\mathcal{K} - \mu_{nuc}(N_{at} + N_{nuc}) - \mu_{elec}(N_{at} + N_{elec})$.

**A more complete expression than (III.142) for $\mathbf{H}_{spont}$ is $\mathbf{H}_{spont} = \sum_\alpha A_\alpha^\dagger(\mathbf{H}_o + \mathbf{H}')a_\alpha + \ldots$. It follows that $P\,\mathbf{H}_{spont} = (1 - \Delta + \ldots)\,\mathbf{H}_{spont} = 0 + \ldots$.

are different*.

G. Discussion

We have shown how a second-quantization formalism for many-atom systems in which unbound nuclei, unbound electrons, and bound atoms are represented by "elementary particle" field operators can be constructed by a sequence of changes of representation, starting with the usual Schrödinger representation in terms of the dynamical variables of all nuclei and electrons. This representation is similar to those arrived at in different ways by Brittin and Stolt[5]) and by Sakakura[6]); however, more analysis is needed in order to clarify the relationship. Still a different approach to construction of such a representation can be based on the "redundant mode" method, in which fictitious "elementary atoms" are introduced and then given physical meaning by an appropriate unitary transformation. The most useful transformation for this purpose appears to be a generalization of one employed by Tani[9]) in his analysis of scattering from a one-particle bound state. The Hamiltonian obtained in this way has both the same eigenvalues and matrix elements as the original Schrödinger Hamiltonian H_{Sch}, to given order in the density**. Such an approach will be discussed elsewhere.[10]

---

*In addition to the exchange terms in $\mathbf{H}_0 + \mathbf{H}'$, H contains additional exchange terms arising from the prefactor $M^{-1} = (1 + B + \ldots)$, which "renormalize" the exchange matrix elements in $\mathbf{H}'$. Additional "orthogonality interaction" terms occur in $\mathcal{K}$, which can again be incorporated by appropriate redefinition of the matrix elements.

**The additional freedom necessary to achieve this is provided by an appropriate small modification of the constraint (III.137).

It is clear that similar representations adapted to the treatment of scattering of electrons by atoms, or nucleons by nuclei, can be derived by similar methods, as can representations useful in a "first-principles" approach to theories of chemical and nuclear reactions.

REFERENCES

1. P. Ehrenfest and J. R. Oppenheimer, Phys. Rev. 37, 333 (1931).

2. M. D. Girardeau, J. Math. Phys. 4, 1096 (1963); 11, 68 (1970).

3. M. D. Girardeau, J. Math. Phys. 12, 1799 (1971).

4. V. Fock, Z.f. Physik 75, 622 (1932).

5. R. H. Stolt and W. E. Brittin, Phys. Rev. Letters 27, 616 (1971) and this volume, 55.

6. A. Y. Sakakura, Phys. Rev. Letters 27, 822 (1971) and this volume, 483.

7. See, e.g., N. M. Hugenholtz, Physica 23, 481 (1957), 510.

8. F. J. Dyson, Phys. Rev. 102, 1217 (1956), in particular Secs. 3-5.

9. S. Tani, Phys. Rev. 117, 252 (1960).

10. M. D. Girardeau, Phys. Rev. Letters 27, 1416 (1971).

INFINITE QUANTUM SYSTEMS

Rudolf Haag
II. Institut für Theoretische Physik
Hamburg University
Hamburg, Germany

I. Introduction.

There are two areas in physics in which a quantum theoretical treatment of infinitely extended systems (or systems with an infinite number of degrees of freedom) becomes necessary (or at least desirable). These areas are

1) Quantum Statistical Mechanics
2) Quantum Field Theory or, to put it into a slightly more general context: The theory of elementary particles as approached within the frame of local, relativistic Quantum Physics.

The reasons why need a brief explanation. Statistical Mechanics wants to describe the gross properties of some 10^{20} particles enclosed in a container of the size of some cubic centimeters. Thus, in reality, the number of degrees of freedom as well as the volume are finite. But in a precise quantum theoretical treatment of such a complicated system we would not find the basic qualitative predictions of statistical mechanics. There is no irreversibility, there are no thermodynamical equilibrium states etc. These features arise only from an approximation in which, at the appropriate places, one omits terms which vanish in the limit when the particle number N and the volume V approach infinity. Thus, if one wants to have a mathematical frame in which the laws of statistical mechanics are strictly contained one has to start from a situation in which the "thermodynamic limit"

$$N \to \infty; \qquad V \to \infty; \qquad \frac{N}{V} \text{ fixed} \tag{1.1}$$

has been taken at the outset, i.e. one has to consider a system with infinitely many particles occupying infinite space with a finite mean density.

In elementary particle theory one could also claim that we do not really need to consider infinitely extended space-time and unlimited energies. But again, if we "enclose the system in a box" some of the simple concepts disappear (i.e. they can then only be approximately defined). Among these are the notion of particles, the S-matrix, the invariance properties.

In both areas therefore the consideration of infinite systems is an idealization which simplifies the conceptual structure. The price one pays for this is an increase in the sophistication of the mathematical apparatus. As it is the case with every bargain: one has to be prepared to pay the price.

This price involves on the one hand a distinction between various kinds of convergence. On the other hand, to obtain a satisfactory and natural setting, we have to start from the algebraic version of the mathematical formalism of Quantum Physics instead of the more widely known Hilbert space formulation. I shall describe these two versions and their relation in section II. A glossary of mathematical terms and some relevant theorems is given in section III in the hope that this will provide a quick access to the language used in many recent papers in the two areas mentioned. The lectures of Prof. Hugenholtz will give examples of this approach to problems in Statistical Mechanics. Therefore I shall not discuss this area. In section IV we shall then sketch one topic in elementary particle physics, the structure of the set of charge quantum numbers and its relation to the statistics of particles (Bose-Fermi-parastatistics).

Before going into details it is perhaps worthwhile to indicate the salient feature which is typically (though not exclusively) associated with infinite systems. It may be called the appearance of "superselection rules".* The set of physical states decomposes into families which are distinguished by parameters which have no quantum

* The term "superselection rule" was introduced and illustrated by examples in [1].

fluctuations. Such parameters are the "macroscopic observables" (for instance the temperature) in the case of Statistical Mechanics, they are the charge quantum numbers in elementary particle physics. These parameters arise in a natural way if one uses the algebraic formulation of Quantum Physics; they correspond to the classification parameters of inequivalent representations of one and the same abstract algebra.

II. Mathematical Formalism of General Quantum Physics

A. Hilbert space formulation

I suppose that the majority of the audience is familiar with J. von Neumann's famous book, Mathematical Foundations of Quantum Mechanics [2]. It contains among other things a description of the concepts used in general quantum physics (i.e. concepts which are supposed to apply not only to quantum mechanics but equally well to quantum field theory) and describes the mathematical objects corresponding to these concepts. Let us briefly recall this conceptual and mathematical structure.

The basic physical concepts are "states" and "observables." A "state" (of the physical system under consideration) is--at least for practical purposes--a statistical ensemble of identical systems, produced (or "prepared") by some piece of experimental equipment called the source of the state. An "observable" is an apparatus which subjects each system of the ensemble to a "measurement" thereby registering a number, "the measured value," and--in the ideal case--releasing the system after the measurement for further subsequent observations. The observable thus decomposes the original ensemble into a collection of subensembles according to the different measured values.

One may distinguish "pure states" and "mixtures." If we have any two states s_1 and s_2, then we can always obtain another state s by mixing s_1 and s_2 with arbitrary (positive) weights λ_1, λ_2 ($\lambda_1 + \lambda_2 = 1$). This means that we prepare an ensemble containing, say, N systems by using the source of s_1 for $\lambda_1 N$ of

the cases and the source of s_2 in $\lambda_2 N$ of the cases. Clearly in any subsequent experiment on this ensemble the probability for any result will be $\lambda_1 p_1 + \lambda_2 p_2$ where p_1, p_2 are the respective probabilities for the same result in the states s_1, s_2. Let us express this relation by

$$s = \lambda_1 s_1 + \lambda_2 s_2 \tag{2.1}$$

Conversely, given a state s one may ask whether there exists a pair of other states, s_1 and s_2, and weights λ_1, λ_2, such that s, s_1 and s_2 are in the relation (2.1). If it is impossible to find such a pair, different from s itself, then s is called "pure." It is one of the essential characteristics of Quantum Physics that the decomposition of an ensemble into subensembles by an observable (as described above) does not correspond to a relation of the type (2.1). If we throw together all the subensembles released by the observable the resulting mixture will be a different state from the one which existed before the measurement. The observable has not only recorded numbers but also changed the state.

Mathematically, pure states are represented by vectors in a Hilbert space $\mathcal{H}$ or, more precisely, by rays in this space (the vectors Ψ and $c\Psi$ corresponding to one and the same state). Observables are represented by self-adjoint operators acting on $\mathcal{H}$. The possible measured values which can occur in the measurement of the observable T are the spectral values of the operator T. To each distinct spectral value t_n there is a spectral projector p_n. If Ψ describes the state before the measurement then the sub-ensemble consisting of those systems after the observation for which the value t_n has been recorded is given by $p_n\Psi$. Thus the probability for finding the value t_n in the observation T on the ensemble is:

$$p_n = \frac{\|p_n \Psi\|^2}{\|\Psi\|^2} \tag{2.2}$$

If F is a real monotonic function with an inverse then the operator F(T) has the same spectral projections as T and the spectral values are $F(t_n)$. Therefore the observables T and F(T) are measured by the same apparatus, only the "scale" has been changed to indicate $F(t_n)$ instead of t_n. This remark may be used to achieve a formal simplification of general arguments in two ways. First, one defines the "expectation value" of the observable T in the state Ψ by:

$$\langle T\rangle_{\Psi} = \Sigma p_n t_n = \frac{(\psi, T\psi)}{(\Psi, \Psi)} \qquad (2.3)$$

The spectral values t_n of the observable T and the probabilities p_n for their occurrence in a state Ψ can be reconstructed if we have the expectation values $(\Psi, F(T)\Psi)$ for a sufficiently large set of functions F of the observable T. Secondly, we may be satisfied by considering instead of T (which possibly may be an unbounded operator, an operator whose spectrum extends to infinite values) bounded functions of T, i.e., bounded operators. Instead of one observable T we then deal with a commutative algebra of bounded operators and the physical statements about measured values and probabilities are incorporated in the knowledge of the expectation values of all these bounded observables.

Two more comments on this formalism:

1. Impure states are described by "density matrices." A density matrix ρ is a positive self adjoint operator with finite trace. The expectation value of the observable T in such a state is given by

$$\langle T\rangle_{\rho} = \frac{\text{Trace } T\rho}{\text{Trace } \rho} \qquad (2.4)$$

The special case of a pure state and the expression (2.3) result when ρ is the projection operator on the direction of the single vector Ψ.

2. We shall always use the "Heisenberg picture" to describe the development in time. This means that a state prepared by some source is described by a fixed vector or density matrix in Hilbert space irrespective of the time at which we intend to make an observation on it. The physical specification of an observable must then include not only the description of the apparatus and its placement in the laboratory but also the time at which it is used to make the measurement.

B. Algebraic Approach

While Hilbert space is the most natural mathematical setting for Schrödinger's wave mechanics the initial work of Heisenberg-Born-Jordan and in particular Dirac's notion of "q-numbers" is closer in spirit to the algebraic approach. The mathematical description of an observable is not primarily given by an operator on a Hilbert space but rather by a "q-number" (i.e. an element of an abstract algebra) whose relevant properties are determined by algebraic relations with other q-numbers.

A mathematically precise and complete formulation of this point of view was given by I.E. Segal [3]. The physical meaning of the terms "observable" and "state" will be the same as described before, but the primary mathematical object is the algebra $\mathfrak{A}$ in which the set of observables is embedded. The general description involves algebraic and topological aspects. We start with the former and defer the discussion of the topology to the end. Let A, B,... denote elements of $\mathfrak{A}$ and α, β,... complex numbers with $\overline{\alpha}$, $\overline{\beta}$ their complex conjugates. The algebraic operations which can be performed within $\mathfrak{A}$ are

1) addition of elements: $A + B$
2) multiplication of an element by a complex number: αA
3) multiplication of elements (in general non commutative): $A B$
4) involution (corresponding to the adjoint for an operator on a Hilbert space): $A \to A^*$.

The laws prevailing among these operations are familiar. With respect to 1) and 2) $\mathfrak{A}$ is a linear space. The multiplication 3) is associative, distributive with respect to 1) and commutes with 2). The involution satisfies

$$(\alpha A + \beta B)^* = \bar{\alpha}\, A^* + \bar{\beta}\, B^*$$

$$(AB)^* = B^*\, A^*$$

$$A^{**} = A$$

The significance of the involution is that it allows us to define real (or self adjoint) elements and positive elements in $\mathfrak{A}$. $A \in \mathfrak{A}$ is called real, if $A^* = A$, it is called positive if it can be written in the form $A = B^* B$ with some $B \in \mathfrak{A}$. Observables correspond to real elements of $\mathfrak{A}$.* States correspond to "expectation functionals" over $\mathfrak{A}$. An "expectation functional" ω (also called a "positive, linear form") is a linear** function from the algebra to the complex numbers taking positive values (including possibly zero) on positive elements of $\mathfrak{A}$. In other words ω assigns a complex number $\omega(A)$ to each $A \in \mathfrak{A}$ such that

$$\omega(\alpha\, A + \beta\, B) = \alpha\, \omega(A) + \beta\, \omega\, (B) \qquad (2.5)$$

and

$$\omega(A^* A) \geq 0. \qquad (2.6)$$

Usually we shall deal with an algebra which contain a unit element 1. In that case we may add as a normalization convention on the state ω the condition

$$\omega(1) = 1 \qquad (2.7)$$

---

* For simplicity we shall pretend that every real element corresponds to an observable.

** Since the addition of two noncommuting observables does not have a simple operational meaning the reason why an expectation functional should be linear over $\mathfrak{A}$ does not lie on the surface.

The structure described so far is still too loose and allows many pathologies or unwanted features. For instance, it could happen that there is a real element which cannot be expressed as a difference of two positive elements. In order to exclude such features we have to add topological considerations. The simplest kind of a topology in a linear space is provided by assigning to each element A a norm $\|A\|$ i.e. a positive number satisfying

$$\|\alpha\, A\| = |\alpha|\, \|A\|, \tag{2.8}$$

$$\|A + B\| \leq \|A\| + \|B\|. \tag{2.9}$$

In the case of an algebra with involution the norm should also satisfy

$$\|A\, B\| \leq \|A\| \cdot \|B\| \tag{2.10}$$

$$\|A^*\| = \|A\|. \tag{2.11}$$

Suppose we have found a norm on $\mathfrak{A}$ satisfying these requirements (2.8) to (2.11). Then we can complete $\mathfrak{A}$ by adding to it the limit points of Cauchy sequences. The completion of $\mathfrak{A}$, denoted by $\bar{\mathfrak{A}}$ will still be a *-algebra (algebra with an involution) and, considering its linear structure it is a Banach space. It is therefore appropriately called a Banach *-algebra. Given $\mathfrak{A}$ as a *-algebra there will be in general many different possible choices of a norm. Under good circumstances there is, however, one preferred choice called the "minimal regular norm", and this norm is determined by the algebraic structure of $\mathfrak{A}$. The line of argument, due to I. M. Gelfand and M. A. Naimark is the following (see for instance [4],[5]). The set of positive linear forms over $\mathfrak{A}$ does not depend on the choice of a norm on $\mathfrak{A}$. If we assign a norm we can ask whether all positive linear forms over $\mathfrak{A}$ are continuous in that norm topology. If so, we call the norm <u>regular</u>. One finds that if a regular norm exists at all then there is also a minimal regular norm $\|\ \|_c$, uniquely determined by

$$\|A\|_c \leq \|A\|_r \qquad \text{all } A \in \mathfrak{A}$$

where $\| \ \|_r$ stands for an arbitrary regular norm. Moreover, the minimal regular norm has the property

$$\|A^* A\|_c = \|A\|^2. \tag{2.12}$$

$\mathfrak{A}$, equipped with its minimal regular norm (assuming its existence) and completed in that topology will be called a C*-algebra. Mathematically, a C*-algebra (like Hilbert space) is a very natural object in the sense that all structural assumptions just fit together to entail a beautiful mathematical theory. For instance, there are sufficiently many positive elements in a C*-algebra so that every real element is a difference between two positive ones and hence any positive linear form is real on the real part of $\mathfrak{A}$. Furthermore the spectrum of a real element A is well defined and $\|A\|_c$ is just the supremum of the absolute values in the spectrum.

Therefore Segal's postulate that $\mathfrak{A}$ should be a C*-algebra appears mathematically very natural. Let us see how it is to be understood in the example of quantum mechanics of 1 degree of freedom. There we are dealing with the algebra generated in some sense by two observables p and q satisfying the commutation relations

$$[p,q] = -i \tag{2.13}$$

One could consider the algebra $\mathfrak{A}_1$ consisting of finite linear combinations of $p^n q^m$. The product of such polynomials can be rearranged, using the commutation relations, to be again of the form $\Sigma\, C_{nm} p^n q^m$. The adjoint will be defined by

$$(p^n q^m)^* = q^m p^n$$

which can again be reordered, shifting the powers of p to the left using (2.13). One has therefore a *-algebra. But one finds that this algebra possesses no regular norm. This reflects, of course, the fact that the spectra of p and q extend to infinity. Noting, however, as remarked in section II A that a bounded function of an observable A is essentially the same physical measurement as A we may construct

the algebra, as Weyl suggested, from the elements

$$U(a, b) = e^{i(ap + bq)} \quad (a, b \text{ real parameters}). \tag{2.14}$$

The equivalent of (2.13) are then the commutation relations in the Weyl form

$$U(a,b)\ U(a',b') = e^{\frac{i}{2}(ab'-a'b)}\ U(a+a',\ b+b'). \tag{2.15}$$

The involution is defined by

$$U(a,b)^* = U(-a,-b). \tag{2.16}$$

Thus the set of finite linear combinations $\Sigma\ C_{a_i b_i}\ U(a_i,b_i)$ forms a *-algebra, say $\mathfrak{A}_2$, and so does the set $\mathfrak{A}_3$ consisting of elements

$$U(f) = \int f(a,b)\ U(a,b)\ d\,a\,d\,b$$

when f runs through the complex valued, absolutely integrable functions of the 2 real variables a and b. Both $\mathfrak{A}_2$ and $\mathfrak{A}_3$ possess regular norms and hence can be completed to become C^*-algebras. The C^*-norm of an element of $\mathfrak{A}_2$ or $\mathfrak{A}_3$ coincides with the operator norm of the corresponding Hilbert space operator in the Schrödinger representation.

One sees from this example that the relation between the canonical quantities p,q and the possible choices for a C^*-algebra of observables is very analoguous to that between the generators of a Lie group and the group elements (or the group algebra).

C. Infinite Systems and Locality

In the case of the infinite systems with which we are concerned some further general structure of the algebra must be added to the one discussed under II B in order to achieve a reasonable physical interpretation. The main point is that we admit as elements of $\mathfrak{A}$ only observables of essentially local character. We assume that there are observables

which can be measured within finitely extended space-time regions. If $\mathfrak{B}$ is a space-time region, $\mathfrak{A}(\mathfrak{B})$ shall denote the algebra generated by all observables measurable within $\mathfrak{B}$. Let $\mathfrak{R}$ denote some basic set of simple, finitely extended regions. To fix the ideas: In the nonrelativistic case (application to statistical mechanics), where measurements at a sharp time are deemed possible, we may take $\mathfrak{R}$ to consist of 3-dimensional balls or cubes at a time t. In the relativistic case it is convenient to choose $\mathfrak{R}$ to consist of (finitely extended) double cones. Such a double cone is determined by two points x_1, x_2 in Minkowski space with x_2-x_1 a positive timelike vector. Then K_{x_1,x_2} is the set of points which lie both inside the backward cone of x_2 and the forward cone of x_1.

We regard then the net of local algebras $\mathfrak{A}(K)$ with K ranging through $\mathfrak{R}$ as the mathematical object which fixes the theory. This point of view was developed within the Hilbert space frame in [6],[7],[8],[9] and in the purely algebraic form in [10]. The algebra $\mathfrak{A}$ of section II B is then the C^*-algebra generated by all the $\mathfrak{A}(K)$, i.e. it is the norm completion of $\cup\mathfrak{A}(K)$. Similarly, for any infinite region $\mathfrak{B}$ we shall define

$$\mathfrak{A}(\mathfrak{B}) = \bigcup_{K\subset\mathfrak{B}} \mathfrak{A}(K) \qquad (2.17)$$

We call $\mathfrak{A}$ the algebra of quasilocal observables, since every element of it can be uniformly approximated by local quantities.

It is perhaps remarkable that the purely geometric resolution of $\mathfrak{A}$ into the net $\{\mathfrak{A}(K)\}$ (i.e. the identification of classes of observables according to the space-time-regions in which they may be measured) is sufficient to fix the physical interpretation of the theory. In other words: given a net $\{\mathfrak{A}(K)\}$ which satisfies the principles of locality, causality, covariance the physical phenomena predicted by this "theory" are determined unambiguously. (See [10] and the last section of [11])

III Mathematical Glossary

This section is intended for the benefit of the physicist who is not familiar with the branch of mathematics used here and who nevertheless insists on reading these notes. It is, of course, not clear that such individuals exist but I hope so. I try to give a brief exposition of the most important concepts, definitions and some of the most relevant mathematical results (listed as R_k, k running from 1 to 12). For further (and better) information the imaginary reader is referred to the book by Naimark [5].

1. Convergence of sequences of vectors in Hilbert space

Notation: $\mathcal{H}$ a Hilbert space; Ψ a vector in $\mathcal{H}$,

$$\|\Psi\| = (\Psi,\Psi)^{\frac{1}{2}} \qquad \text{its norm (length).}$$

There are two distinct notions of convergence, both relevant to problems in physics (e.g. scattering theory): The sequence Ψ_n converges strongly towards Ψ if $\|\Psi-\Psi_n\| \to 0$, it converges weakly if $|(\Phi,\Psi-\Psi_n)|\to 0$ for every fixed $\Phi\in\mathcal{H}$. (Φ chosen arbitrarily but independent of n). To these two notions of convergence correspond two topologies of $\mathcal{H}$, the strong and the weak topology.

R_1. The unit ball of $\mathcal{H}$ is weakly (but not strongly) compact.

Or: every infinite set of vectors $\{\Psi\}$ with $\|\Psi\| \leq 1$ has at least one weak limit point. For most practical purposes: every sequence Ψ_n with $\|\Psi_n\|$ uniformly bounded has a weakly convergent subsequence.

2. Bounded linear operators

Q a linear operator acting on $\mathcal{H}$. Its norm is defined by $\|Q\| = \sup_{\Psi\in\mathcal{H}} \frac{\|Q\Psi\|}{\|\Psi\|}$. If $\|Q\|$ exists (i.e. if it is finite) Q is called bounded. The set of all bounded, linear operators on $\mathcal{H}$ is denoted by $B(\mathcal{H})$. The topology determined by the above norm is called the uniform topology in $B(\mathcal{H})$.

R_2. The operator norm $\|Q\|$ is a C^*-norm, i.e. it satisfies relations (2.8) through (2.12).

Since $B(\mathfrak{H})$ is closed under the algebraic operations II B 1) through 4) and complete with respect to the norm one has

R_3. $B(\mathfrak{H})$ is a C^*-algebra.

With Q_n a sequence in $B(\mathfrak{H})$ we consider three notions of convergence to a limit $Q \in B(\mathfrak{H})$

a) Uniform: $\|Q_n - Q\| \to 0$,

b) Strong: For every $\Psi \in \mathfrak{H}$ the sequence of vectors $Q_n\Psi$ converges strongly (as defined in III 1) to $Q\Psi$,

c) Weak: For every $\Psi \in \mathfrak{H}$ the sequence of vectors $Q_n\Psi$ converges weakly to $Q\Psi$.

To each of these notions of convergence there corresponds a topology on $B(\mathfrak{H})$, the uniform (or norm) topology, the strong and the weak topology. They are decreasing in strength. Thus a weakly closed set in $B(\mathfrak{H})$ is a forteriori strongly closed and a strongly closed set is always uniformly closed. (The weaker the topology, the more limit points are added in performing the closure). Corresponding to R_1 one has

R_4. The unit ball of $B(\mathfrak{H})$ (i.e. the set with $\|Q\| \leq 1$) is weakly compact.

3. Operator algebras

R_5. The weak and the strong closures of a *-algebra in $B(\mathfrak{H})$ coincide.

A strongly (or weakly) closed *-algebra in $B(\mathfrak{H})$ containing the unit operator 1 is called a von Neumann ring, a uniformly closed *-algebra in $B(\mathfrak{H})$ we call a concrete C^*-algebra. Obviously (by the ordering of the strengths of the topologies) every von Neumann ring is also a concrete C^*-algebra but the converse is not true in general.

The commutant of any set $S \subset B(\mathfrak{H})$ is denoted by S'. It consists of all those elements of $B(\mathfrak{H})$ which commute with every member of S. For the commutant of the commutant, i.e. for $(S')'$ we write S''. The adjoint set (consisting of the adjoint operators of all elements of S) is denoted by S^*.

R_6. If S is a self adjoint set in $B(\mathcal{H})$ (i.e. $S=S^*$) then
a) S' is a von Neumann ring
b) S'' is the von Neumann ring generated from S and 1 (by algebraic operations and weak closure).
c) $S''' = S'$

4. Factors

Let R be a von Neumann ring. The subalgebra

$$Z = R \cap R'$$

is called the center of R. If Z is trivial, i.e. if it consists only of multiples of the unit operator, then R is called a factor.

This terminology comes from the following

R_7. If $\mathcal{H}$ is finite demensional and R a factor from $B(\mathcal{H})$ then $\mathcal{H}$ can be written as a direct product of two Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ such that

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2; \quad R = B(\mathcal{H}_1)\otimes 1; \quad R' = 1 \otimes B(\mathcal{H}_2). \tag{3.1}$$

For infinite dimensional Hilbert spaces the situation is more complicated. The analysis of F. J. Murray and J. von Neumann [12] leads to a classification of factors into three basic types, only one of which (type I) is the analogon of the finite dimensional situation described under R_7.

The analysis starts from an ordering of the projection operators which are contained in the factor R. Two such projectors P_1 and P_2 are called equivalent (in symbols $P_1 \underset{R}{\sim} P_2$) if there is an operator $U \in R$ which maps the subspace $P_1\mathcal{H}$ isometrically on $P_2\mathcal{H}$. Similarly one defines $P_1 \underset{R}{\geq} P_2$ (existence of $U \in R$ which maps $P_1\mathcal{H}$ isometrically onto a proper subspace of $P_2\mathcal{H}$). One finds

R_8. With R a factor, P_1, P_2 projectors from R one has always precisely one of the three possibilities

i) $P_1 \underset{R}{>} P_2$

ii) $P_1 \underset{R}{<} P_2$

iii) $P_1 \underset{R}{\sim} P_2$

Corresponding to this ordering of projectors $P \in R$ one can introduce a relative dimension of the subspaces $P\mathcal{H}$. Dim P is a non negative number (possibly ∞) which satisfies Dim P_1 = Dim P_2 iff $P_1 \underset{R}{\sim} P_2$, Dim P_1 < Dim P_2 iff $P_1 \underset{R}{<} P_2$ and Dim P = 0 iff P = 0. Further, if P_1 and P_2 are mutually orthogonal ($P_1P_2 = 0$) then Dim $(P_1 + P_2)$ = Dim P_1 + Dim P_2.

Apart from an arbitrary (positive) normalization factor the dimension function is uniquely determined by these requirements.

R_9. A factor R is either of

Type I: R contains minimal projectors (P_1 is minimal if $P_1 \neq 0$ and $P_2 \underset{R}{<} P_1$ implies $P_2 = 0$). Normalizing the dimension function so that it takes the value 1 on the minimal projectors, Dim P ranges through all integer values up to a maximal value n or up to and including ∞. If n is a finite integer R is called of type I_n, otherwise of type I_∞.

Type II: Dim P ranges through a continuum of values which may be either normalized to be the closed interval 0,1 (type II_1) or it may be the whole real line including 0 and ∞ (type II_∞).

Type III: Dim P takes only the values 0 and ∞. In that case all nonvanishing projectors from R are equivalent (in the sense iii).

Remark. In the finite dimensional example discussed under R_7 we have a factor of type I_n where n is the ordinary dimension of the space $\mathcal{H}_1$ and Dim P is the ordinary dimension of $P\mathcal{H}_1$. In other words, the multiplicity provided by the dimension of $\mathcal{H}_2$ plays no role.

The relative trace over R. If R is a factor of finite type (type I_n or II_1) then there exists one unique positive linear form over R, denoted by Tr which is invariant:

$$\mathrm{Tr}(UAU^{-1}) = \mathrm{Tr}A \qquad \text{for any } A\in R \text{ and unitary } U\in R \tag{3.2}$$

and normalized by

$$\mathrm{Tr}\ 1 = 1 \tag{3.3}$$

If R is a factor of seminfinite type (type I_∞ or II_∞) then the relative trace may be defined not on all operators of R but on a weakly dense subalgebra which does not contain 1. It is then an unbounded positive linear form on this subset, satisfying (3.2) but we cannot normalize it by (3.3). It is unique, apart from a normalization convention. In a factor of type III no trace may be defined at all. But recently in the theory of modular operators (Tomita-Takesaki theory) [13],[14],[15] mappings between factors of type III and type II have been discovered which may possibly be useful to introduce "pseudo-traces" also in the type III case.

Remarks. It appears that in the physical applications which concern us here (see introduction) we meet most frequently with factors of type III. The mathematical theory of a finer classification of factors seems to be at this moment in a period of major progress and in close contact with examples and principles in physics.

5. Representations

Let now $\mathfrak{A}$ be a C^*-algebra. By a representation π of $\mathfrak{A}$ we mean a mapping from $\mathfrak{A}$ to the bounded operators

of some Hilbert space $\mathcal{H}$ which conserves the algebraic structure (II B, operations 1 to 4). If for every $A \neq 0$ the representing operator $\pi(A) \neq 0$ the representation is called faithful.

R10. If π is faithful then the operator norm of $\pi(A)$ equals the C^*-norm of A i.e. the mapping π establishes an isomorphism between the abstract C^*-algebra $\mathfrak{A}$ and the concrete C^*-algebra $\pi(\mathfrak{A})$.

If π is not faithful $\|\pi(A)\| \leq \|A\|$ and π is a homomorphism from $\mathfrak{A}$ to $\pi(\mathfrak{A})$. Again $\pi(\mathfrak{A})$ is uniformly closed, i.e. a concrete C^*-algebra.

Cyclic representation: There exist a vector $\Psi \in \mathcal{H}$ (a cyclic vector) such that $\pi(\mathfrak{A})\Psi$ is dense in $\mathcal{H}$.

Irreducible representation: $\mathcal{H}$ contains no proper subspace which is transformed into itself by $\pi(\mathfrak{A})$. This is equivalent to (Schur's lemma)

$$\pi(\mathfrak{A})'' = B(\mathcal{H}) \text{ or} \tag{3.4}$$

$$\pi(\mathfrak{A})' = \{\lambda\ 1\} \tag{3.5}$$

Primary representation: $\pi(\mathfrak{A})''$ is a factor, i.e.

$$\pi(\mathfrak{A})'' \cap \pi(\mathfrak{A})' = \{\lambda\ 1\} \tag{3.6}$$

Unitary equivalence between $\pi_1(\mathfrak{A}) \subset B(\mathcal{H}_1)$ and $\pi_2(\mathfrak{A}) \subset B(\mathcal{H}_2)$: Existence of a unitary mapping V from $\mathcal{H}_1$ onto $\mathcal{H}_2$ such that

$$V\,\pi_1(A) = \pi_2(A)V \text{ for all } A \in \mathfrak{A} \tag{3.7}$$

Quasiequivalence between π_1 and π_2: The von Neumann rings $\pi_1(\mathfrak{A})''$ and $\pi_2(\mathfrak{A})''$ are isomorphic (regarding the mapping $\pi_1(A) \to \pi_2(A)$ and its extension to the weak closure).

6. Families of States

Let $\mathfrak{A}$ be a C^*-algebra, $\pi(\mathfrak{A}) \subset B(\mathcal{H})$ a representation and W an arbitrary positive operator of trace class

from $B(\mathcal{H})$ (not necessarily from $\pi(\mathfrak{A}'')$ normalized by

$$\mathrm{Tr}\ W = 1 \tag{3.8}$$

Then QW is again of trace class for every $Q \in B(\mathcal{H})$ and W defines a state ω_W over the C^*-algebra $\mathfrak{A}$ by

$$\omega_W(A) = \mathrm{Tr}\ (\pi(A)W) \tag{3.9}$$

since ω_W obviously satisfies the requirements (2.5), (2.6), (2.7). When W ranges through all positive, normalized trace class operators in $B(\mathcal{H})$ we obtain in this way a family of states over $\mathfrak{A}$. This family is called the set of normal states with respect to the representation π. We shall use the symbol $\mathcal{S}_\pi$ for it. A subset of this family are the vector states of π, which result if W runs through all 1-dimensional projections in $B(\mathcal{H})$. We may write these as

$$\omega_\Psi(A) = (\Psi, \pi(A)\Psi) \tag{3.10}$$

with Ψ a unit vector from $\mathcal{H}$. (Compare (2.3) and (2.4)). The normal states of a primary representations we call primary states.

R_{11}. If π_1 and π_2 are two primary representations then either $\mathcal{S}_{\pi_1} = \mathcal{S}_{\pi_2}$ in which case π_1 and π_2 are quasiequivalent, or $\mathcal{S}_{\pi_1} \cap \mathcal{S}_{\pi_2}$ is empty, i.e. the two families of states are completely disjoint.

Therefore, it amounts to the same whether we consider quasiequivalence classes of primary representations or families of primary states. One such family is generated from a single member ω by the operations

i) translation with elements in the algebra

$$\omega \to \omega_B;\ \omega_B(A) = \frac{\omega(B^*AB)}{\omega(B^*B)};\ B \in \mathfrak{A},$$

ii) Finite convex combinations (mixture),

iii) closure in the norm topology of state space.

The GNS-construction (Gelfand, Naimark, Segal)

Given an arbitrary state ω over $\mathfrak{A}$. This construction yields a cyclic representation π_ω so that ω occurs as a vector state of this representation (related to a cyclic vector Ω in the representation space by $\omega(A) = (\Omega, \pi(A)\Omega)$. This starts from the remark that $\mathfrak{A}$ is itself a linear space in which, given ω, a semidefinite scalar product may be defined by

$$\langle A,B\rangle = \omega\ (A^*\ B) \tag{3.11}$$

On the other hand we may regard each element A also as a linear operator acting on the space $\mathfrak{A}$ (the image of B under the operation A being A B). Finally one has to come from the space $\mathfrak{A}$ with its semidefinite metric (3.11) to a space with positive definite metric which may then be completed to become a Hilbert space. One defines the set $J \subset \mathfrak{A}$ by

$$Z \in J \quad \text{if } \omega(Z^*Z) = 0 \tag{3.12}$$

and checks that J is a linear space and a left ideal of $\mathfrak{A}$, i.e. if $Z \in J$ and $A \in \mathfrak{A}$ then $A\ Z \in J$. We call J the GNS - ideal of the state ω. Taking $\mathfrak{K} = \mathfrak{A}/J$ (set of classes of $\mathfrak{A}$ modulo J) one obtains a representation space with positive definite metric. In it the class of elements $1 + J$ is a cyclic vector Ω satisfying

$$(\Omega,\pi(A)\Omega) = \omega(1^*\ A\ 1) = \omega(A)$$

R_{12}. The vector states of an irreducible representation are pure states. The GNS - construction, starting from a pure state leads to an irreducible representation.

IV. The Superselection Rules of Particle Physics. Statistics.

If we compare the framework sketched in section II C with conventional quantum field theory we see as the most striking difference that the latter uses (in realistic models) field quantities which can not be associated with the algebra of observables $\mathfrak{A}$. An example is the Dirac field. There are two simple reasons showing immediately that this

field Ψ cannot be associated with $\mathfrak{A}$. The first argument (due to Wigner): A rotation by 360° is equivalent to doing nothing as far as any physical situation is concerned but it changes the sign of Ψ. The second argument: The commutator between $\Psi(x)$ and $\Psi(y)$ does not decrease as $x - y \to \infty$ in a spacelike direction and the fact that the anticommutator vanishes is of no help for an observable; a measurement of the Hermitean part of Ψ in one region, if possible, would have a disturbing effect on a similar measurement in another region and the disturbance would not decrease with the distance.

The use of unobservable fields is tied to the existence of superselection rules within the manifold of states one wants to consider. Thus, the first argument above as used in [1] demands a superselection rule between states of integer spin and those of half integer spin; the second argument demands a superselection rule between states of Bose type and those of Fermi type.

In conventional field theory (Hilbert space formalism with unobservable fields) the superselection structure and the relation between fields and observables is governed by a group $\mathfrak{G}$. An element g of this group corresponds to a transformation which (like the rotation by 360° mentioned above) does not produce any change of the physical situation but is nevertheless represented by a non trivial operator U(g) acting on the Hilbert space ($U(g) \neq \lambda\, 1$). An observable must be invariant under this group, i.e. it must commute with all U(g). The Hilbert space may be decomposed into subspaces ("superselection sectors") $\mathfrak{K}_i$ according to the "spectrum of the group $\mathfrak{G}$". In the simplest and most important case when $\mathfrak{G}$ is Abelian this corresponds to a simultaneous diagonalization of all the U(g) so that in each subspace $\mathfrak{K}_i$ the U(g) act like multiples of the identity, but like different multiples in the different $\mathfrak{K}_i$. Then a coherent superposition of state vectors is possible (physically meaningful) only if these vectors belong to the same sector $\mathfrak{K}_i$.

The two examples of superselection rules given (integer-half integer spin; Bose-Fermi type) are the most unrefutable cases and actually, because of the spin-statistics theorem, they are tied together. The group which is relevant here has only one element besides the identity. This

transforms $\Psi(x)$ into $-\Psi(x)$. It is, however, rather generally believed that the fundamental conservation laws of electric charge, baryon number, lepton number are also linked to transformation groups ("gauge transformations") which leave all observables invariant. This implies then superselection rules between states differing in any of these charge quantum numbers. Empirically the Bose-Fermi alternative is tied to the charges. With the usual assignment of quantum numbers a state is of Fermi type if the sum of baryon and lepton number is odd, it is of Bose type if this sum is even. We shall therefore use the term (generalized) charge for the set of all parameters which are needed to distinguish the different superselection sectors and the term "gauge group" for the group $\mathcal{G}$.

Let us now look at these features starting from the frame of general local quantum physics. The theory is then characterized by the net of algebras of observables $\{\mathfrak{A}(K)\}$. Does this net determine superselection rules, unobservable fields, a gauge group and statistics? What is the structure allowed by the principles of relativistic locality and causality? Are Bose - and Fermi statistics the only possibilities? Various aspects of these questions have been discussed in [16], [17], [18], [19], [20] (see also [9] and [10]).

It has already been mentioned in the introduction that in an algebraic formulation of the theory superselection rules arise naturally because the algebra will in general allow inequivalent representations. Our problem here is, however, an embarrassment of riches: The algebra has far too many inequivalent irreducibe representation*. The interesting superselection rules for particle physics correspond to a tiny subset of these and we have first to motivate and formulate the criteria which single out among all possible representations the relevant ones. The origin of this restriction is the idealization always adopted in elementary particle physics: We consider only states which look (asymptotically) like the vacuum for observations in far away regions of space. Therefore, it is convenient to start from one distinguished representation π_0, the "vacuum representation" which is obtained by the GNS-construction

* This is a consequence of the fact that $\mathfrak{A}$ is an inductive limit of the net of local algebras.

from the vacuum state ω_o (see section III.6). All other representations of interest to us will be closely related to it. The precise criteria one should adopt in singling them out have been formulated in a satisfactory, usable form only for the case when there are no long range forces (no zero mass particles) in the theory*. In that case we are led to consider the subset of representations for which the corresponding families of states (see III.6) contain strictly localized states (for a discussion of this see [19 I]). Here we use Knight's definition of strict localization [21] and call the state ω localized in the double cone K if it coincides completely with the vacuum ω_o for all observations in the space-like complement K' of K, in other words if $\omega|_{\mathfrak{A}(K')} = \omega_o|_{\mathfrak{A}(K')}$**. Moreover one finds that a localized state ω may be related to the vacuum by means of a localized morphism ρ of the algebra:

$$\omega(A) = \omega_o(\rho(A)) \tag{4.1}$$

By a morphism we mean a mapping $A \to \rho(A)$ from $\mathfrak{A}$ into $\mathfrak{A}$ which preserves the algebraic structure and the norm. The image $\rho(\mathfrak{A})$ may be the whole algebra (then ρ is an automorphism) but the case where $\rho(\mathfrak{A})$ is a proper subalgebra of $\mathfrak{A}$ is also possible and of interest. We call ρ localized in K if it acts trivially on the algebra of the space-like complement K' i.e. if $\rho(A) = A$ for $A \in \mathfrak{A}(K')$.

The set $\mathfrak{R}$ of interesting representations is then related to the vacuum representation π_o in the following way: $\pi \in \mathfrak{R}$ if (up to equivalence) we can write

$$\pi(A) = \pi_o(\rho(A)) \tag{4.2}$$

with ρ a localized morphism. The state (4.1) is a vector state of the representation π. We may devide the localized morphisms in equivalence classes calling ρ_1 and ρ_2 equivalent representations. We shall call ρ pure if it generates

* One may hope that the superselection structure derived below for the case of short range forces retains validity also in electrodynamics (although it will not be complete there) since we may consider this as the limiting case, starting from a finite photon mass.

** $\omega|_{\mathfrak{A}_1}$ denotes the restriction of ω to the subalgebra $\mathfrak{A}_1$

by (4.2) an irreducible representation (i.e. if the state (4.1) is pure). The interesting superselection sectors (generalized charges) correspond to equivalence classes of pure, localized morphisms.

To discuss the structure of this set (of charges) we need some simple properties of localized morphisms:

i) morphisms with mutually spacelike localization region commute.
ii) the product of morphisms respects the class division, i.e. the class of the product $\rho_1\rho_2$ does not change if ρ_1 and ρ_2 vary within their classes.
iii) if ρ is an automorphism then all positive and negative powers ρ^n are defined and they are all pure. If ρ is pure but not an automorphism then ρ^2 is never pure. Finally,
iv) it is reasonable to assume that the Poincaré transform of a localized morphism ρ is in the same class as ρ^*.

By ii) we have a product composition of classes (composition of charges). It is commutative because of i) and iv). The "simple charges", which correspond to automorphism classes, form an Abelian group. For the others we have a more complicated composition law. The product of two pure but not simple charges is a mixture (reducible representation) which can be decomposed again into a finite direct sum of irreducibles. In this sense the composition law of two such charges ξ_1, ξ_2 can be written in the form

$$\xi_1\xi_2 = \Sigma\, \xi_i \qquad (4.3)$$

An example of such a situation would be the composition law of isospin. Indeed, in a theory with strict isospin invariance in which there exists no observable distingishing different members of a multiplet the magnitude of the isospin is a charge quantum number. A sector with nonvanishing

* This means that the Poincaré transformations should be implementable by unitary operators in the representation $\pi_\rho(\rho(\cdot))$. If there should exist localized morphisms for which this is not true we exclude such representations from $\mathfrak{R}$.

isospin is then generated from the vacuum by a ρ which is not an automorphism.

Besides charge composition we have charge conjugation. In the case of simple charges this is trivial: If ξ belongs to the class of ρ, then the class of ρ^{-1} is the conjugate $\bar{\xi}$. For non simple charges the construction of the conjugate is more interesting. If ρ_x is the morphism arising from translation of ρ by the space-time vector x then ρ_x and ρ are equivalent and, in fact, there is a unitary $U_x(\rho) \in \mathfrak{A}$ so that

$$\rho_x(A) = U_x(\rho)\rho(A)\ U_x^{-1}(\rho)$$

We may interpret $U_x(\rho)$ as a charge transfer: If ρ is localized in K and ξ denotes the class of ρ then $U_x(\rho)$ shifts a charge ξ from the region K to K + x. Thus the state ω_x defined by

$$\omega_x(A) = \omega_o(U_x(\rho)\ A\ U_x^{-1}(\rho))$$

has charge zero but locally, within K + x it appears to have charge ξ which is compensated by a "hole" in the region K. As we let x tend spacelike to infinity only the hole remains and ω_x converges weakly to a state with the conjugate charge $\bar{\xi}$ localized in K.

Except for the case of simple charges (which is, of course the most important case) the charges do not form a group. But the structure of the set of charges discussed above suggests that this set may be the dual of a compact group (to be called the gauge group). By the dual of $\mathfrak{G}$ we mean the collection of equivalence classes of irreducible unitary representations of $\mathfrak{G}$ equipped with the composition law given by the Clebsch-Gordon series for the direct product and with the conjugation (complex conjugation of representations). It is probable but has not been shown so far that the information about the structure of the set of charges is sufficient to determine $\mathfrak{G}$ [22].

Let us look finally at the problem of statistics. Suppose $\omega_1, \ldots \omega_n$ are all pure states, all having the same charge ξ but localized in regions $K_1, \ldots K_n$ which lie spacelike to each other. So they are generated in the sense of

(4.1) by equivalent morphisms $\rho_1,\ldots\rho_n$ with mutually spacelike support. The state

$$\omega(\cdot) = \omega_o(\rho_1\rho_2\ldots\rho_n(\cdot))$$

has charge ξ^n and corresponds to the physical picture of a charge ξ sitting in every one of the regions $K_1\ldots K_n$ (and vacuum in the spacelike complement of $\cup K_i$). We call this state ω the "product state" of the ω_i and write

$$\omega = \omega_1 \times \omega_2 \ldots \times \omega_n \tag{4.4}$$

Since morphisms with mutually spacelike supports commute the state (4.4) does not change if we permute the ω_i. Thus there is only one localized product state, irrespective of the order of the factors. Let us now look at the state vectors instead of the states! To get a reference point we pick some morphism ρ in the class ξ. There are unitary elements $U_i \in \mathfrak{A}$ which relate ρ_i to ρ (equivalence $\rho_i \sim \rho$)

$$\rho_i(A) = U_i\rho(A)\ U_i^{-1}. \tag{4.5}$$

Quite generally (irrespective of the position of the supports of the ρ_i) (4.5) implies the equivalence of $\prod \rho_i$ with ρ^n:

$$\rho_1\ldots\rho_n(A) = U\rho^n(A)\ U^{-1} \tag{4.6}$$

where

$$U = U_1\ \rho(U_2)\ldots\rho^{n-2}(U_{n-1})\rho^{n-1}(U_n). \tag{4.7}$$

Consider now the three representations

$$A \to \pi_o(A);\ A \to \pi_1(A) = \pi_o(\rho(A));\quad A \to \pi_n(A) = \pi_o(\rho^n(A))$$

corresponding to charges o, ξ, ξ^n respectively. They act all in the same Hilbert space; the physical interpretation of a vector of this space depends on the representation to be considered. Thus the vector Ω_o which corresponds to the vacuum state in the representation π_o corresponds to the state $\omega_o(\rho(\cdot))$ in π_1 and to the state $\omega_o(\rho^n(\cdot))$ in π_n. We see that ω_i may be represented in π_1 by the vector

$$\Psi_i = \pi_o(U_i^{-1})\Omega_o \tag{4.8}$$

and ω of eq. (4.4) in π_n by

$$\Psi = \pi_o(U^{-1})\ \Omega_o \tag{4.9}$$

with U given by (4.7). Symbolically we can write

$$\Psi = \Psi_1 \times \Psi_2 \times \ldots \times \Psi_n \tag{4.10}$$

as a short hand for the relations (4.7),(4.8), (4.9). It It turns out now that the product (4.10) or, alternatively, the expression (4.7) is not independent of the order of factors (order in which the ρ_i appear) even when the supports of the ρ_i are mutually spacelike If P is a permutation of n objects and U_P denotes the expression (4.7) with ρ_i replaced by $\rho_{P(i)}$ then

$$\varepsilon_\rho^{(n)}(P) = U^{-1}(P)\ U \tag{4.11}$$

will in general differ from 1. One finds that it is independent of the choice of the ρ_i as long as they are all in the equivalence class of ρ and have mutually spacelike supports. Furthermore the $\varepsilon_\rho^n(P)$ with fixed ρ and n form a unitary representation of the permutation group and they commute with $\rho^n(A)$ for all $A\in\mathfrak{A}$. Thus, a change in the order of factors in (4.10) changes the state vector by some $\pi_o(\varepsilon_\rho^{(n)}(P)$ but it does not change the state ω; in other words, all these vectors arising from permutations define the same expectation functional over the algebra. If ξ is a simple charge then ρ^n is pure and the representation $\pi_o(\rho^n(\cdot))$ is irreducible. By Schur's lemma then $\varepsilon_\rho^{(n)}(P)$ can only be a multiple of the identity. In that case $\varepsilon_\rho^{(n)}(P)$ must be either the totally symmetric or the totally antisymmetric representation of the permutation group. If ξ is not simple then ρ^n is reducible and other representations of the permutation group can occur.

The "statistics" of states with charge ξ means the characterization up to quasiequivalence of the $\varepsilon_\rho^{(n)}$ for fixed ρ in the class ξ and all positive integers n. (Variation of ρ within its class does not change the quasiequivalence class of $\varepsilon_\rho^{(n)}$). Practically this is done by giving the Young tableaux which occur in the decomposition of the representation $\varepsilon_\rho^{(n)}$ of the permutation group into irreducibles. The analysis in [19,I] shows that this is governed by a single parameter λ, the "statistics parameter" of the

charge ξ. Apart from the somewhat pathological case $\lambda = 0$ one obtains precisely the possibilities suggested by H. S. Green [23], namely para-Bose statistics of some definite order d or para-Fermi statistics of some order d. The sign of λ gives the distinction between para-Bose and para-Fermi, $|\lambda|^{-1}$ gives the order of the parastatistics. Ordinary Bose or Fermi statistics (d = 1) occur if and only if ξ is a simple charge. Parastatistics of order $d \neq 1$ appears whenever we have a non-Abelian gauge group (when $\mathfrak{A}$ allows localized morphisms which map onto a proper subset of $\mathfrak{A}$).

Usually one regards parastatistics as an exotic possibility. But it is really very harmless since we have to bear in mind that it depends on the knowledge of what is observable. If we have a theory with a net $\{\mathfrak{A}_1(K)\}$ which has a non Abelian gauge group (for instance the isospin group) then it may be that this theory is just an approximation to a situation with a richer supply of observables, described by a net $\{\mathfrak{A}(K)\}$ for which the gauge group is Abelian. On the level of the approximate theory with SU_2 as gauge group a sector with isospin I gives rise to parastatistics of order (2 I + 1), but the more precise theory splits this sector into (2 I + 1) different ones, each carrying ordinary Bose - or Fermi statistics. In the example we may say that in the less precise theory there is no way to distinguish a single proton from a single neutron. So the particle is a nucleon and it is a parafermion of order 2; similarly we have a π-meson as a paraboson of order 3 etc. If on the other hand we have observables which distinguish the different electric charge states in the isospin multiplets then instead of one parafermion of order 2 we have two ordinary Fermions, etc. [20].

We have seen how a theory of local observables leads naturally to the concept of statistics and how the type of statistics is related to the structure of the set of charge quantum numbers (superselection structure). Combining this analysis with the work of H. Epstein [24] one finds that the connection between spin and statistics holds in the sense that the sign of λ is positive or negative depending on whether the spins in the sector are integer or half integer [19, II]. This gives a somewhat deeper understanding of the relation between spin and statistics because the statistics were not introduced at the outset by means of commutation relations of unobservable fields but result as a structural

property of the observable algebra. One also finds symmetry between charge conjugate sectors: They have the same energy - momentum spectrum and the same statistics parameter. Finally one can show that the statistics parameters determine the metric of the scattering states and thereby enter into expressions for cross sections.

REFERENCES

1. Wick, J. C., A. S. Wightman and E. P. Wigner: "Intrinsic Parity of Elementary Particles", Phys. Rev. 88, 101 (1952).

2. von Neumann, J.: "Mathematical Foundations of Quantum Mechanics", Princeton, English Translation (1955).

3. Segal, I. E.: "Postulates for General Quantum Mechanics", Ann. Math. 48, 930 (1947).

4. Gelfand, I. M. and M. A. Naimark: Mat. Sb. 12, 197 (1943).

5. Naimark, M. A.: Normed Rings, Groningen, R. Nordhoff (1964).

6. Haag, R: Discussion des axiomes et des propriétés asymptotiques d'une théorie des champs locaux avec particules composées", Colloque Int. Lille 1957, published by C.N.R.S., Paris (1959).

7. Haag, R. and B. Schroer: "Postulates of Quantum Field Theory", J. Math. Phys. 3, 248 (1962).

8. Araki, H.: Lecture Notes "Einführung in die axiomatische Feldtheorie", Zürich (1961/62).

9. Haag, R.: "Bemerkungen zum Nahwirkungsprinzip in der Quantenphysik", Annalen d. Phys. 11, 29 (1963).

10. Haag, R. and D. Kastler: "An Algebraic Approach to Quantum Field Theory", J. Math. Phys. 5, 848 (1964).

11. Araki, H. and R. Haag: "Collision cross sections in terms of local observables", Commun. Math Phys. 4, 77 (1967).

12. Murray, F. J. and J. von Neumann: "On Rings of Operators" I to III Ann. Math. 37, 116 (1936), Transact. Amer. Math. Soc. 41, 208 (1937).
Ann. Math. 41, 94 (1940).
Ann. Math. 44, M16 (1943).

13. Tomita, M.: "Quasistandard von Neumann Algebras" mimeographed notes (1967).

14. Takesaki, M.: "Tomita's theory of modular Hilbert algebras and its applications", Lecture Notes in Mathematics Nr. 128, Springer Verlag, Heidelberg (1970).

15. Takesaki, M.: Report in the Conference on Math. Problems of Quantum Field Theory and Quantum Statistics, Moscow, Dec. (1972).

16. Borchers, H. J.: Local Rings and the Connection of Spin with Statistics, Commun. Math. Phys. 1, 281 (1965).

17. Federbush, P.: "Local Operators in the Presence of Super-selection Rules", J. Math. Phys. 9, 1718 (1968).

18. Doplicher, S., R. Haag and J. E. Roberts: "Fields, Observables and Gauge Transformations",
Part I, Commun. Math. Phys. 13, 1 (1969)
Part II, Commun. Math. Phys. 15, 173 (1969).

19. Doplicher, S., R. Haag and J. E. Roberts: "Local Observables and Particle Statistics",
Part I, Commun. Math. Phys. 23, 199 (1971).
Part II, in preparation.

20. Drühl, R. Haag and J. E. Roberts: "On Parastatistics" Commun. Math. Phys. 18, 204 (1970).

21. Knight, J. M.: "Strict Localization in Quantum Field Theory", J. Math. Phys. 2, 459 (1961).

22. Roberts, J. E.: Report in the Conference on Math. Problems of Quantum Field Theory and Quantum Statistics, Moscow, Dec. (1972).

23. Green, H. S.: "A Generalized Model of Field Quantization", Phys. Rev. 90, 270 (1953).

24. Epstein, H.: "CTP Invariance of the S-Matrix in a Theory of Local Observables", J. Math. Phys. 8, 750 (1967).

EQUILIBRIUM STATES IN STATISTICAL MECHANICS
C*-ALGEBRAIC APPROACH

N. M. Hugenholtz
University of Groningen
Groningen, The Netherlands

1. Introduction.

Statistical mechanics is the theory of systems composed of a large number of identical subsystems. The aim is to derive thermodynamical laws from the known interaction between the individual subsystems. To give a precise mathematical description of a large (macroscopic) system as opposed to a small system (atom, nucleus etc.) one considers the model of an infinitely extended system composed of infinitely many identical subsystems. In the traditional approach to statistical mechanics, the prescription for calculating properties of equilibrium states is only applicable to finite systems. This implies that one can get results for infinite systems only at the end of the calculation, by taking the thermodynamical limit.

In the algebraic approach, discussed in these lectures, one deals with infinite systems, in particular, infinite quantum systems. It is therefore necessary to reformulate quantum mechanics in such a way that it may be applied both to finite and to infinite systems. This program necessitates the use of some mathematical techniques, with which most physicists are not very familiar. For a detailed study of operator theory we refer to books by Naimark [1] and Dixmier [2,3]. A compilation of the most important definitions and theorems can be found in an appendix to Ruelle's book on statistical mechanics [4].

Algebraic methods can be applied not only to statistical mechanics but also to quantum field theory. This is due to the great formal similarity between quantum field theory and, e.g., the many body theory. In both cases all observables are expressible in terms of quantized fields

like the canonical pair $\pi(x,t)$ and $\Phi(x,t)$ with the commutation relations

$$[\Phi(x,t), \Phi(y,t)] = [\pi(x,t), \pi(y,t)] = 0$$

and

$$[\pi(x,t), \Phi(y,t)] = -i\delta(x-y) \tag{1.1}$$

or the annihilation and creation operators $\psi(x,t)$ and $\psi^*(x,t)$ for a particle in the space-point x at time t, with the well-known commutation relations:

a. for bosons:

$$[\psi(x,t), \psi(y,t)] = [\psi^*(x,t), \psi^*(y,t)] = 0$$

and

$$[\psi(x,t), \psi^*(y,t)] = \delta(x-y) \tag{1.2}$$

b. for fermions:

$$\{\psi(x,t), \psi(y,t)\} = \{\psi^*(x,t), \psi^*(y,t)\} = 0$$

and

$$\{\psi(x,t), \psi^*(y,t)\} = \delta(x-y) \tag{1.3}$$

Mathematically the essential difference between finite and infinite systems is the occurence of many inequivalent irreducible representations of the observables as operators in a Hilbert-space in the case of an infinite system. On the other hand, von Neumann [5] has proven that all irreducible representations of the canonical commutation relations

$$[p_i,p_j] = [q_i,q_j] = 0, \; [p_i,q_j] = -i\delta_{ij}$$

for i and j = 1, 2, ... , $s < \infty$

as selfadjoint operators in a Hilbert-space $\mathfrak{H}$ are unitarily equivalent. In other words: whereas for a finite system, all irreducible representations are unitarily, and hence physically equivalent, this is not the case for infinite systems. That this fact is physically relevant is

illustrated by the fact that, as will be shown, the ground states of a gas at two different densities lead to inequivalent irreducible representations. Similarly, the representations corresponding to equilibrium states of a gas at two different temperatures are inequivalent, even disjoint [6].

In the next section we shall give a simple but typical example of inequivalent irreducible representations.

2. Inequivalent representations of the C.A.R.

In this section we consider a simple example of inequivalent representations of (1.3) the canonical anti-commutation relations (C.A.R.). Before proceeding we shall rewrite (1.3) in terms of the smeared out operators $\psi(f)$ and $\psi(f)^*$, where f is a complex square-integrable function of x:

$$\psi(f) = \int d^3x \; \psi(x) \; f(x)^*.$$

We then get the commutation relations

$$\{\psi(f), \psi(g)\} = \{\psi(f)^*, \psi(g)^*\} = 0$$

and

$$\{\psi(f), \psi(g)^*\} = (f,g), \tag{2.1}$$

where

$$(f,g) = \int d^3x \; f(x)^* \; g(x).$$

The first representation of (2.1) we shall consider is the Fock-representation. The representation space $\mathcal{H}_F(V)$ for a given volume V is given by the direct sum

$$\mathcal{H}_F(V) = \sum_{n=0}^{\infty} \oplus \mathcal{H}_n$$

where $\mathcal{H}_n$ is the Hilbert space of anti-symmetrical n-particle wave-functions in V. An arbitrary vector $\Psi \in \mathcal{H}_F(V)$ is determined by its components $\Psi^{(0)}$, $\Psi^{(1)}$, $\Psi^{(2)}$ etc. Since $||\Psi||^2 = \sum_n ||\Psi^{(n)}||^2$, Ψ must always satisfy the condition

$$\Sigma \; ||\Psi^{(n)}||^2 < \infty .$$

In this space we define the operators $\psi(f)$ and $\psi(f)^*$ as follows

$$(\psi(f)^*\Psi)^{(n)}(x_1 x_2 .. x_n) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (-1)^{i-1} f(x_i) \Psi^{(n-1)}(x_1 .. \not{x}_i .. x_n), \tag{2.2}$$

where the wave-function $\Psi^{(n-1)}$ is a function of all $x_1, x_2, \ldots, x_n$ with the exception of x_i.

$$(\psi(f)\Psi)^{(n)}(x_1 \ldots x_n) = \sqrt{n+1} \int_V d^3x \; f(x)^* \Psi^{(n+1)}(x \; x_1 \ldots x_n). \tag{2.3}$$

One proves, without difficulty, that the operators $\psi(f)$ and $\psi(f)^*$ defined in (2.2) and (2.3) are each-others adjoint, that they are bounded operators with norm

$$||\psi(f)|| = ||\psi(f)^*|| = ||f||,$$

satisfying the commutation relations (2.1). Hence we have a representation of the C.A.R., called the Fock-representation. Let Φ_0 be the vacuum-state, i.e. the vector in $\mathfrak{H}_F$ with all components = 0 except $\Phi_0^{(0)} = 1$. Clearly Φ_0 satisfies the relation

$$\psi(f)\Phi_0 = 0, \text{ for all } f, \tag{2.4}$$

and Φ_0 is the only vector in Fock-space with that property.

From the Fock-vacuum Φ_0 one can build up the whole Fock-space in the following manner: let $\{f_\alpha, \alpha = 1,2,\ldots\}$ be a basis in $L^2(V)$. Then the vectors $\psi(f_\alpha)^*\Phi_0$, $\psi(f_\alpha)^*\psi(f_\beta)^*\Phi_0$, $\psi(f_\alpha)^*\psi(f_\beta)^*\psi(f_\gamma)^*\Phi_0$, etc. form a basis in $\mathfrak{H}_F(V)$. Φ_0 is therefore a cyclic vector.

The particle number operator N(V) is defined on any vector with only a finite number of non-zero components, as follows:

$$(N(V)\Psi)^{(n)} = n\Psi^{(n)}.$$

N(V) is an unbounded operator, and one can prove that for

any basis $\{f_\alpha,\ \alpha = 1,2,\ldots\}$ in $L^2(V)$

$$N(V) = \sum_\alpha \psi(f_\alpha)^* \psi(f_\alpha). \tag{2.5}$$

Using the completeness relation for $\{f_\alpha\}$ one finds the formal expression

$$N(V) = \int_V d^3x\ \psi^*(x)\psi(x).$$

Without proof we mention that the Fock-representation is irreducible. Moreover, a representation of the C.A.R. with the properties

1. There is a vector Φ_0 such that

$$\Psi(f)\Phi_0 = 0, \text{ for all } f.$$

2. The representation is irreducible.

is unitarily equivalent to the Fock-representation. In other words, the properties 1 and 2 determine the representation uniquely, up to equivalence.

In the case of a cubic box of volume V, with periodic boundary conditions one can choose for the basis wavefunctions $f_\alpha(x)$ the plane waves

$$f_{\vec{k}}(x) = V^{-\frac{1}{2}} e^{i\vec{k}\vec{x}},$$

with

$$k_i = \frac{2\pi}{V^{\frac{1}{3}}} n_i,\ n_i = 0, \pm 1, \pm 2, \ldots .$$

We then write

$$a(\vec{k}) = \psi(f_{\vec{k}}) = \frac{1}{\sqrt{V}} \int d^3x\ \psi(x)\ e^{-i\vec{k}\vec{x}}.$$

$a(\vec{k})$ and $a(\vec{k})^*$ are the annihilation and creation operators for a particle of momentum k. The commutation relations are

$$\{a(\vec{k}),\ a(\vec{k}')\} = \{a(\vec{k})^*,\ a(\vec{k}')^*\} = 0$$

and

$$\{a(\vec{k}),\ a(\vec{k}')^*\} = \delta_{\vec{k},\vec{k}'}.$$

Until now the volume V could be chosen either finite or infinite. We shall now consider the special case that $V = R^3$. We can then introduce translations in space. Let $\Psi \in \mathfrak{H}_F(R^3)$. The equations

$$(U(a)\Psi)^{(n)}(x_1 \ldots x_n) = \Psi^{(n)}(x_1+a, x_2+a \ldots x_n+a)$$

determine a unitary operator, the translation operator. It follows from the definition, that Φ_0 is invariant for translations, and is the only vector in $\mathfrak{H}_F$ with that property.

Another advantage of taking $V = R^3$ is the possibility of defining densities. Let V_0 be a finite subvolume of R^3, and compare the operators

$$N(V_0) = \int_{V_0} d^3x \psi^*(x)\psi(x) \text{ and } N,$$

the total particle number operator. Clearly, one has

$$N(V) \leq N;$$

as an operator inequality. We now define a density operator n, as follows: let Ψ be in the domain of N, then

$$n\Psi = \lim_{V\to\infty} \frac{1}{V} N(V)\Psi.$$

The limit of the r.h.s. vanishes, since

$$\left|\left|\frac{1}{V} N(V)\Psi\right|\right| \leq \left|\left|\frac{1}{V} N\Psi\right|\right| \underset{V\to\infty}{\to} 0.$$

Hence the operator n, when applied to a dense set of vectors in $\mathfrak{H}_F(R^3)$, gives zero. The density operator is therefore identically zero, which means that the vectors in Fock-space describe states of vanishing particle-density. For this reason, it is obvious, that the Fock-represntation is not suited for describing the states of systems with non-vanishing particle-density.

We shall now construct another representation of the C.A.R. by means of a Bogoliubov-transformation. We define,

for finite V, new creation and annihilation operators by means of the transformation

$$a(k)' = \cos(\alpha(k))a(k) - \sin(\alpha(k))a(-k)^*. \qquad (2.6)$$

Formally, i.e. disregarding convergence questions, this transformation can be obtained by means of the unitary transformation

$$a(k)' = e^{iQ}a(k)e^{iQ},$$

with

$$Q = \frac{i}{2}\sum_k \alpha(k)[a(-k)a(k) - a(k)^*a(-k)^*]. \qquad (2.7)$$

As a consequence of the fact that $a(k)$ and $a(-k)$ anti-commute, it is natural to assume that $\alpha(k) = -\alpha(-k)$.

If we deal with spin $\frac{1}{2}$ particles, k stands for momentum and spin $(\vec{k},s)$. In that case this relation is satisfied if $\alpha(\vec{k},s)$ depends on the length $|k|$ only, and

$$\alpha(|k|,+) = -\alpha(|k|,-).$$

Let us now consider Q. If Q is a self-adjoint operator in Fock space, then the transformation (2.6) is unitary and both representations are unitarily equivalent. To see whether Q is a well-defined operator we apply Q to an arbitrary vector Ψ in $\mathfrak{H}_F$ with a finite number of non-vanishing components. We have

$$(a(k)^*\Psi)^{(n)}(x_1..x_n) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}(-1)^{i-1}\frac{e^{ikx_i}}{\sqrt{V}}\Psi^{(n-1)}(x_1..\not{x}_i..x_n)$$

and thus

$$\left(\sum_k \alpha(k)a(k)^*a(-k)^*\Psi\right)^{(n)}(x_1...x_n) =$$

$$= \frac{1}{\sqrt{n(n-1)}}\sum_{\substack{i=1\\ i\neq j}}^{n}\sum_{j=1}^{n}(-1)^{i+j}\hat{\alpha}(x_i-x_j)\Psi^{(n-2)}(x_1..\not{x}_i..\not{x}_j..x_n),$$

where

$$\hat{\alpha}(x_i-x_j) = \frac{1}{V}\sum_k \alpha(k)\, e^{ik(x_i-x_j)}.$$

Clearly a necessary and sufficient condition for this operator to be defined on a dense subset of $\mathfrak{H}_F(V)$ is, that

$$\hat{\alpha}(x_i - x_j) \in L^2(V^2),$$

which condition is equivalent with

$$\sum_k |\alpha(k)|^2 < \infty. \tag{2.8}$$

In the finite case it depends on the sequence $\alpha(k)$ whether or not this condition is satisfied. In the thermodynamical limit ($V \to \infty$) however, the sum in (2.8) becomes proportional to V, so that (2.8) can never be satisfied in the infinite volume case. This means that in the infinite case, two representations of the C.A.R. connected by a non-trivial Bogoliubov transformation are inequivalent.

We could have reached this conclusion also without calculations. Suppose that both representations were equivalent. Then there would be in $\mathfrak{H}_F(V)$ a Fock-vacuum Φ_0' with respect to $a(k)'$. Let us consider translations in space:

$$\psi(x) \to \psi(x+a)$$

or

$$a(k) \to a(k)e^{ika}.$$

It follows then from (2.6) that also

$$a(k)' \to a(k)'e^{ika}.$$

In other words the Bogoliubov-transformation commutes with space translations. Hence the Fock-vacuum Φ_0' is also invariant for translations. Hence Φ_0 and Φ_0' must be identical, which implies immediately that the Bogoliubov transformation must be the identity ($\alpha(k) \equiv 0$).

We shall finish this discussion by considering the following special case:

$$\alpha(\vec{k},s) = 0 \text{ for } |k| > k_F,$$

$$\alpha(\vec{k},\pm) = \pm \frac{\pi}{2} \text{ for } |k| < k_F,$$

We then get the transformation

$$a'(\vec{k},s) = a(\vec{k},s) \qquad \text{for } |k| > k_F$$

$$a'(\vec{k},+) = -a(-\vec{k},-)^* \qquad \text{for } |k| < k_F.$$
$$a'(\vec{k},-) = a(-k,+)^*$$

Clearly, for V finite, we get an equivalent representation. The Fock-vacuum of the primed operators is the state in which all particle-states of energy less than $\frac{k_F^2}{2m}$ are occupied, whereas all other particle-states are unoccupied. It is the groundstate of a system of N free fermions, where N equals the number of single-particle states within a sphere of radius k_F in momentum space.

If we now take the limit $V \to \infty$ we conclude on the basis of our previous discussion, that the ground state of a gas of non-interacting fermions in an infinite volume and of finite particle density gives rise to a representation, which is inequivalent with the Fock-representation.

3. States and observables.

In order to allow for infinite quantum systems our quantum mechanical notion of state has been generalized: A state of a system is a prescription which assigns a number to each observable of the system. This functional has to be linear and positive (i.e. to each positive observable there corresponds a positive number).

We shall now give some physical arguments to justify this definition. Let us consider first an ordinary (finite) quantum system. The observables are then (self-adjoint) operators acting on a Hilbert space $\mathfrak{H}$. A pure state is a vector Ψ in $\mathfrak{H}$. This vector defines the positive linear form $\langle A \rangle = (\Psi, A\Psi)$. The unpure or mixed states (ensembles) are defined by means of a density operator ρ, i.e. a positive operator ρ with $\mathrm{Tr}\ \rho = 1$. The average value $\langle A \rangle$ is now defined by $\langle A \rangle = \mathrm{Tr}\ \rho A$. We see again that we have obtained a positive linear form. This shows that our new definition of a state is a generalization of the well-known notion in quantum mechanics.

We consider next the situation which arises in quantum statistical mechanics. One takes a large but finite system and characterizes the state of the system by means of a density operator ρ_V, where V is the volume. For an equilibrium state one can take the grand-canonical ensemble, for which ρ_V has the form

$$\rho_V = \frac{e^{-\beta(H_V - \mu N_V)}}{\mathrm{Tr}\ e^{-\beta(H_V - \mu N_V)}}, \tag{3.1}$$

with H_V the Hamilton operator of the system and N_V the particle number operator. We are now able to calculate ensemble averages of our observables:

$$\langle A \rangle_V = \mathrm{Tr}\ \rho_V A. \tag{3.2}$$

Since we are interested in infinite systems we must take the so-called thermodynamical limit, i.e. must let the volume V tend to infinity. By taking this limit in (3.2) keeping A fixed we get the expectation value or average $\langle A \rangle$ of A for the infinite system. It is implied by this procedure that A must be a local observable, i.e. an observable which refers to a bounded region in space. The thermodynamical limit of (3.2) does not always exist for all local observables A. On physical grounds we expect this limit to exist provided there are no phase-transitions for the given values of β and μ. The existence of the limit does however not imply that the density operator ρ_V has a limit ρ for $V \to \infty$ and that $\langle A \rangle = \lim_{V\to\infty} \langle A \rangle_V = \mathrm{Tr}\ \rho A$. On the contrary, we expect that $\lim_{V\to\infty} \rho_V = 0$. Indeed, consider the denominator in (3.1). This is the partition function Z_V. The pressure of the system is derived from Z_V by the formula

$$p = kT \lim_{V\to\infty} \frac{1}{V} \log Z_V.$$

Consequently we can write, for large V,

$$Z_V = e^{\frac{pV}{kT}},$$

so that Z_V diverges exponentially for $V \to \infty$.

From this example we learn the following: By taking the thermodynamical limit one is able to find the average or expectation value of any local observable. This means that we have a prescription assigning a number to every local observable. Again this functional is positive and linear, since these properties are conserved in the thermodynamical limit.

We have already remarked that observables must have some locality property, because real observations always are carried out in some finite region of space. Before specifying the set of observables for a few specific infinite systems, we shall mention some general features.

Since observables may be added and multiplied the set of observables is an algebra. If we limit ourselves to bounded observables, we can assign a norm $||A||$ to each observable A. We therefore have a normed algebra, the completion of which is the algebra $\mathfrak{A}$ of quasi-local observables. This is a C^*-algebra *). In the most simple cases (C.A.R. algebra and lattice systems) the algebra $\mathfrak{A}$ has the following quasi-local structure. To each bounded volume V there is assigned an algebra $\mathfrak{A}(V)$. This is a C^*-algebra or even a W^*-algebra (= Von Neumann algebra). These local algebras satisfy the conditions

1. If $V_1 \subset V_2$, then $\mathfrak{A}(V_1) \subset \mathfrak{A}(V_2)$. This is called isotony.
2. If $V_1 \cap V_2 = \emptyset$, $[\mathfrak{A}(V_1), \mathfrak{A}(V_2)] = 0$. This is called causality.

As a consequence of 1. the set $\mathfrak{A}_L = \cup\, \mathfrak{A}(V)$ is an algebra, the algebra of all strictly local observables. The norm-closure $\mathfrak{A}$ is then the C^*-algebra of quasi-local observables.

Transformations are automorphisms of the C^*-algebra $\mathfrak{A}$. As we shall see in the examples, space-translations and time-evolutions are commutative groups of automorphisms of $\mathfrak{A}$. Obviously, in the case of a lattice system, the group of space translations is a discrete group, in a continuous system it is a continuous group. Let $\underline{a}$ be a vector, then

---

*) For the definition of C^*-algebra and other mathematical objects used in these notes see the lecture notes by R. Haag in this volume.

$\alpha_a A$ is the observable obtained from A by a translation over $\underline{a}$. Similarly $\alpha_t A$ is obtained from A by evolution over the time t. For each $\underline{a}$ and each t α_a and α_t are automorphisms of $\mathfrak{A}$. The group property is expressed by the relations

$$\alpha_{a_1}\alpha_{a_2} = \alpha_{a_1+a_2} \quad \text{and} \quad \alpha_{t_1}\alpha_{t_2} = \alpha_{t_1+t_2}.$$

A state ω of a C^*-algebra $\mathfrak{A}$ is a normed positive linear form, i.e. a form satisfying the conditions:

a. $\omega(\lambda_1 A_1 + \lambda_2 A_2) = \lambda_1\omega(A_1) + \lambda_2\omega(A_2)$, for all complex λ_1 and λ_2 and for all A_1 and $A_2 \in \mathfrak{A}$,

b. If $A \in \mathfrak{A}$ is positive, then $\omega(A) \geq 0$.

c. $\omega(1) = 1$, where 1 is the unit-element of $\mathfrak{A}$.

The state ω is said to be invariant under a certain automorphism α ($A \to \alpha A$) if $\omega(\alpha A) = \omega(A)$ for all A. A state, which is invariant for all translations will be called homogeneous. If a state is invariant for time-evolution we shall say that it is a stationary state.

4. Observable algebras for certain simple cases.

I. Fermi-gas. Let $\mathfrak{H}_F(R^3)$ be the Fock-space for the canonical anti-commutation relations (C.A.R.). We define

$$\mathfrak{A}(V) = \{\psi(f), \psi^*(g), \text{ with supp. } f,g \text{ in } V\},$$

the *algebra generated by $\psi(f)$ and $\psi^*(g)$ for all f and g with support in V. $\mathfrak{A}(V)$ is then the local algebra corresponding to V. As follows from the definition the isotony condition is satisfied, but due to the fact that the ψ's anticommute the causality condition is not.

The C*-algebra defined by $\overline{\mathfrak{A}=\bigcup_V \mathfrak{A}(V)}$, where the bar means completion in the norm, is called the C.A.R.-algebra. It is easy to prove that $\mathfrak{A}$ contains the operators $\psi(f)$, with $f \in L_2(R^3)$. Consequently $\mathfrak{A}$ is generated by all $\psi(f), \psi^*(g)$ with f and $g \in L_2$. We may consider two subalgebras:

$\mathfrak{A}_e$ is the C*-algebra generated by polynomials containing only even powers of ψ's and ψ^*'s. All elements of this algebra are clearly invariant for the replacement $\psi \to -\psi$.

$\mathfrak{A}_g$ is the C*-algebra generated by polynomials consisting of terms containing an equal number of factors ψ and ψ^*. The elements of $\mathfrak{A}_g$ are invariant for gauge-transformations (see further in this section). Since all physically measurable quantities are gauge-invariant $\mathfrak{A}_g$ is the actual observable algebra. We have the inequalities

$$\mathfrak{A}_g \subset \mathfrak{A}_e \subset \mathfrak{A}.$$

The algebras $\mathfrak{A}_g(V)$ and $\mathfrak{A}_e(V)$, which are defined in obvious analogy with $\mathfrak{A}(V)$ both satisfy the causality condition.

The C.A.R.-algebras $\mathfrak{A}$ and its subalgebra $\mathfrak{A}_g$ are not the only possible choice for an observable algebra for a system of fermions. The local algebra $\mathfrak{A}(V)$ is weakly dense in $B\{\mathfrak{H}_F(V)\}$, the set of all bounded operators on the Fock-space $\mathfrak{H}_F(V)$ of the volume V. By locally including the weak limit points one may replace $\mathfrak{A}(V)$ by $B\{\mathfrak{H}_F(V)\}$. This leads to a larger algebra

$$\mathfrak{A}' = \bigcup_V \{B\ \mathfrak{H}_F(V)\}.$$

In connection with the time-evolution we shall come back to this point.

We shall now define some transformations.

1. Space translations. Let f_a be the function obtained from f by shifting over the distance a: $f_a(x) = f(x-a)$. We then define the mapping α_a by $\alpha_a\psi(f) = \psi(f_a)$. It is immediately seen that the commutation relations of the operators $\psi(f)$ and $\psi(g)^*$ are not affected by α_a. This means that the mapping $A \to \alpha_a A$, which is defined in an obvious manner for all polynomials in $\psi(f)$ and $\psi(g)^*$ and hence, by continuity, for all $A \in \mathfrak{A}$ is an automorphism of $\mathfrak{A}$. That α_a is a strongly continuous automorphism follows from

$$\lim_{a\to 0} ||\alpha_a \psi(f) - \psi(f)|| = \lim_{a\to 0} ||f_a - f|| = 0.$$

2. Gauge transformations. The mapping α_ν is defined by $\alpha_\nu \psi(f) = \psi(f)e^{i\nu}$. This leads in an obvious manner to an automorphism α_ν of $\mathfrak{A}$. In contrast to space-translation and time-evolution (see hereafter) gauge-transformations are local, in the sense that $\alpha_\nu \mathfrak{A}(V) = \mathfrak{A}(V)$.

3. Time-evolution. If the notion of an abstract C^*-algebra of observables is physically meaningful, time-evolution ought to be an automorphism of that algebra. It follows that the proper choice of $\mathfrak{A}$ may depend on the dynamics of the system. We first study the case of free fermions. We then define the mapping α_t by the equation $\alpha_t \psi(f) = \psi(f_t)$, where f_t is the solution of the one-particle-wave equation which at $t = 0$ coincides with f. Again we easily prove that α_t defines a strongly continuous one parameter group of automorphisms of $\mathfrak{A}$.

In the case of interaction the situation is less clear. It is likely that for a realistic two-particle interaction the algebra is not invariant for time-evolution. If that is the case time-evolution is not an automorphism of the C.A.R. algebra and the problem is to define another observable algebra for which time-evolution is an automorphism. It is not clear how this may be achieved.

II. Bose-gas. The operators $\psi(f)$ and $\psi(g)^*$ in the Fock-space $\mathfrak{H}_F$ of Bose particles are unbounded in contrast to the Fermi-case. We must therefore proceed somewhat differently. One may define the s.a. fields Φ and π by the equations

$$\Phi(f) = \frac{1}{\sqrt{2}} (\psi(f) + \psi(f)^*),$$

$$\pi(f) = \frac{-i}{\sqrt{2}} (\psi(f) - \psi(f)^*).$$

One then defines unitary operators $U(f) = e^{i\Phi(f)}$ and $V(f) = e^{i\pi(f)}$, which satisfy the multiplication rules $U(f_1)U(f_2) = U(f_1+f_2)$; $V(f_1)V(f_2) = V(f_1+f_2)$ and

$U(f)V(g) = V(g)U(f)\, e^{-i(f,g)}$. The local algebra $\mathfrak{A}(V)$ is generated by all $U(f)$ and $V(g)$ with support of f and g in V. Various inequivalent definitions of the observable algebra are now possible:

$$\mathfrak{A}_1 = \overline{\cup\, \mathfrak{A}(V)};$$

$$\mathfrak{A}_2 = \overline{\cup\, B\{\mathfrak{H}_F(V)\}};$$

$$\mathfrak{A}_3 = \overline{\{U(f), V(g), \text{ for all } f,g \in S(R^3)\}};$$

$$\mathfrak{A}_4 = \overline{\{U(f), V(g), \text{ for all } f,g \in L_2(R^3)\}}.$$

There are the inequalities

$$\mathfrak{A}_1 \not\subset \mathfrak{A}_3 \not\subset \mathfrak{A}_4 \not\subset \mathfrak{A}_2.$$

Space-translations and gauge-transformations are defined as in the case of fermions. They are automorphisms of all four algebras defined above. Time-evolution of a free boson system is defined by $\alpha_t \Psi(f) = \Psi(f_t)$. It can be shown that $\mathfrak{A}_1$ and $\mathfrak{A}_2$ are not invariant for this mapping. On the other hand $\mathfrak{A}_3$ and $\mathfrak{A}_4$ are invariant. On this ground $\mathfrak{A}_1$ and $\mathfrak{A}_2$ cannot be considered proper choices for free boson systems. The case of interacting bosons is even less clear, and will not be discussed in these lecture notes.

III. Quantum lattice systems. In many respects lattice systems are much simpler than continuous systems. A lattice system is perhaps the simplest non-trivial example of a system with infinitely many degrees of freedom. They provide, therefore, excellent testing ground for the algebraic treatment of infinite systems. The following well-known models are special cases of quantum lattice systems: the quantum lattice gas, the Ising model and the Heisenberg model.

Let Z^ν be a ν-dimensional cubic lattice, the points of which we shall call x,y etc.. To each lattice point x there is assigned a p-dimensional Hilbert space $\mathfrak{H}_x$. In most applications the dimension $\nu = 1$, 2 or 3, $p = 2$, which corresponds to a spin $\frac{1}{2}$ at each lattice point. If $\Lambda \subset Z^\nu$ is a finite sublattice of Z^ν with $N(\Lambda)$ lattice

points, we assign to Λ the Hilbert space $\mathfrak{H}_\Lambda$ which is defined as the tensor $\prod_{x\in\Lambda} \otimes \mathfrak{H}_x$. Clearly $\mathfrak{H}_\Lambda$ is $p^{N(\Lambda)}$ dimensional. The set of operators ($p^{N(\Lambda)}$ x $p^{N(\Lambda)}$ matrices) on $\mathfrak{H}_\Lambda$ will be denoted by $B(\mathfrak{H}_\Lambda)$. In a similar fashion we define $\mathfrak{H}_{Z^\nu}$ as the infinite tensor product $\prod_{x\in Z^\nu} \otimes \mathfrak{H}_x$. Each operator $A \in B(\mathfrak{H}_\Lambda)$ may be considered to be an operator on $\mathfrak{H}_{Z^\nu}$ in the following obvious way. If $\Lambda_c = Z^\nu - \Lambda$, we can write

$$\mathfrak{H}_{Z^\nu} = \mathfrak{H}_\Lambda \otimes \mathfrak{H}_{\Lambda_c}.$$

If we now identify the operator $A \in B(\mathfrak{H}_\Lambda)$ with the tensor product $A \otimes 1$ it becomes an operator on $\mathfrak{H}_{Z^\nu}$.

We now define, for each finite Λ,

$$\mathfrak{A}(\Lambda) = B(\mathfrak{H}_\Lambda).$$

These local algebras are considered as subalgebras of the set $B(\mathfrak{H}_{Z^\nu})$. They then satisfy the isotony condition and the causality condition. Indeed, if $\Lambda_1 \subset \Lambda_2$ we have $\mathfrak{A}(\Lambda_1) \subset \mathfrak{A}(\Lambda_2)$. If on the other hand $\Lambda_1 \cap \Lambda_2 = \emptyset$, then $\mathfrak{A}(\Lambda_1)$ and $\mathfrak{A}(\Lambda_2)$ commute, as can be checked easily. In accordance with section 3 we obtain the algebra $\mathfrak{A}$ of quasi-local observables by taking the closure of the union of all $\mathfrak{A}(\Lambda)$:

$$\mathfrak{A} = \overline{\bigcup_\Lambda \mathfrak{A}(\Lambda)}.$$

We next define some important groups of transformations in lattice systems.

1. Space-translations. For each $a \in Z^\nu$ $\mathfrak{H}_x$ is mapped isometrically onto $\mathfrak{H}_{x+a}$. This defines in a natural manner an isometric mapping of $\mathfrak{H}_\Lambda \to \mathfrak{H}_{\Lambda+a}$ with a corresponding mapping $A \in B(\mathfrak{H}_\Lambda) \to \alpha_a A \in B(\mathfrak{H}_{\Lambda+a})$. This mapping extends to an automorphism α_a of $\mathfrak{A}$, which maps $\mathfrak{A}(\Lambda)$ onto $\mathfrak{A}(\Lambda+a)$.

2. Internal transformations. In each point x we select a unitary operator on $\mathfrak{H}_x$ in such a way, that $\alpha_a U_x = U_{x+a}$. We then define the unitary operator U_Λ on $\mathfrak{H}_\Lambda$ by

$$U_\Lambda = \prod_{x\in\Lambda} \otimes U_x.$$

The mapping $A \in \mathfrak{A}(\Lambda) \to A' = U_\Lambda A U_\Lambda \in \mathfrak{A}(\Lambda)$ defines an outer automorphism of $\mathfrak{A}$ which is an inner automorphism of each $\mathfrak{A}(\Lambda)$. In the case of $p = 2$ these transformations correspond to rotations of the direction of the spin in each lattice point.

3. Time-evolution. We assume that to every bounded $\Lambda \subset Z^\nu$ is assigned a Hamilton-operator $H(\Lambda) \in \mathfrak{A}(\Lambda)$. This hamiltonian is translationally invariant:

$$\alpha_a H(\Lambda) = H(\Lambda+a).$$

Potentials may be defined by the equations (for each Λ)

$$H(\Lambda) = \sum_{X\subset\Lambda} \Phi(X).$$

We then have

$$\alpha_a \Phi(X) = \Phi(X+a).$$

The potential Φ is said to have finite range if the number of sets X such that $0 \in X$ and $\Phi(X) \neq 0$ is finite. The union Δ of all such sets X is called the range of Φ.

Let Λ_1 and Λ_2 be so far separated that $(\Lambda_1 - \Lambda_2) \cap \Delta = \emptyset$, then

$$H(\Lambda_1 \cup A_2) = H(\Lambda_1) + H(\Lambda_2).$$

There is then no interaction between the sublattices Λ_1 and Λ_2.

Before studying the time-evolution of the infinite lattice we consider two well-known examples of lattice systems, the Ising-model and the Heisenberg-model. The hamiltonians have the following form:

Ising: $$H(\Lambda) = \sum_{x,y} J(x,y)\sigma_3(x)\sigma_3(y);$$

Heisenberg: $$H(\Lambda) = \sum_{x,y} J(x,y)\, \vec{\sigma}(x)\vec{\sigma}(y).$$

The potentials are

Ising: $\Phi(X) = 0$ for $N(X) \neq 2$

$$\Phi(x,y) = J(x,y)\sigma_3(x)\sigma_3(y);$$

Heisenberg: $\Phi(X) = 0$ for $N(X) \neq 2$

$$\Phi(x,y) = J(x,y)\vec{\sigma}(x)\vec{\sigma}(y).$$

The hamiltonian $H(\Lambda)$ defines an inner automorphism $A \to A_t^\Lambda = e^{iH(\Lambda)t} A\, e^{-iH(\Lambda)t}$ of $\mathfrak{A}(\Lambda)$. Let us now take $A \in \mathfrak{A}(\Lambda_0)$, $\Lambda_0 \subset \Lambda$, and consider the limit of A_t^Λ for $\Lambda \to \infty$. If the potential has finite range, one would expect on physical grounds this to exist. Indeed, it has been shown by Robinson [7] that this limit exists in the norm topology for every t:

$$\text{norm} \lim_{\Lambda\to\infty} A_t^\Lambda = \alpha_t A,$$

and for every $A \in \mathfrak{A}(\Lambda_0)$, with Λ_0 arbitrary finite. The mapping $A \to \alpha_t A$ can be extended by continuity to a strongly continuous group of automorphisms of $\mathfrak{A}$. Ruelle [4] has shown that the restriction to potentials of finite range can be relaxed somewhat to potentials that vanish rapidly at large distances.

5. The ground state. *)

In this section we study the thermodynamical limit of the ground state and some of its properties. To be specific we shall restrict our considerations to quantum lattice systems. If we have some prescription to assign a state ω_Λ of $\mathfrak{A}(\Lambda)$ to every bounded sublattice Λ, and if there is a state ω of $\mathfrak{A}$, such that for each $A \in \mathfrak{A}(\Lambda_0)$ and each bounded Λ_0

$$\lim_{\Lambda\to\infty} \omega_\Lambda(A) = \omega(A)$$

*) In order to avoid too much overlap between these lecture notes and the notes of lectures given at the conference on mathematical methods of comtemporary physics in London, 1971, most proofs have been omitted in this section and the next.

we call ω the thermodynamic limit of ω_Λ. In general that limit will not exist, but it can be shown that the sequence will have one or more accumulation points and that one can find subsequences converging to these limit points.

We now apply this to the ground state. Consider a finite sublattice Λ, and let $\Psi_0(\Lambda)\in\mathfrak{H}_\Lambda$ be the ground state of $H(\Lambda)$. For simplicity we shall assume that for all Λ the smallest eigenvalue of $H(\Lambda)$ is non-degenerate. We now define the state ω_Λ of $\mathfrak{A}(\Lambda)$ by

$$\omega_\Lambda(A) = (\Psi_0(\Lambda), A\Psi_0(\Lambda)).$$

Before taking the thermodynamical limit of ω_Λ we mention an important property of ω_Λ.

Lemma 5.1: Let A and $B\in\mathfrak{A}(\Lambda)$, then the expression $\omega_\Lambda(A_t^\Lambda B)$ can be extended to an entire function of a complex variable t which is uniformly bounded for $\mathrm{Im}\ t \leq 0$.

For the proof one writes ω_Λ as the expectation value of the vector $\Psi_0(\Lambda)$, and uses the fact that $\Psi_0(\Lambda)$ is the ground state of $H(\Lambda)$.

We shall now define what we mean by a ground state of the infinite system. Any state which is the thermodynamical limit of the sequence of ground states ω_Λ or of some subsequence of these is a ground state of the infinite system. Some properties of such states will now be given.

Theorem 5.2: Let ω be a ground state of the infinite lattice and $A,B\in\mathfrak{A}$. Then

i. ω is invariant for α_t.
ii. the expression $\omega(\alpha_t AB)$ can be extended to a function of a complex variable t which is analytic for $\mathrm{Im}\ t < 0$ and uniformly bounded for $\mathrm{Im}\ t \leq 0$.

The proof is based on the fact that these properties hold for each Λ and that, as discussed in section 4, norm $\lim_{\Lambda\to\infty} A_t^\Lambda = \alpha_t A$.

The analyticity condition expressed in theorem 5.2 has some farreaching consequences as we shall show presently. One might even use that condition as the defining property of the ground state of an infinite system, although it is not yet known whether this new definition is equivalent with the earlier one. To discuss some implications of this condition let us introduce the spectrum condition. A state ω which is invariant for α_t satisfies the spectrum condition if in the representation determined by ω the hamiltonian (i.e., the generator of the unitary group U_t which implements α_t) has a non-negative spectrum. We then have the

Theorem 5.3: A state which satisfies the analyticity condition of theorem 5.2 is invariant and satisfies the spectrum condition, and vice versa.

The next theorem concerns the unitary operator U_t.

Theorem 5.4: Let the state ω satisfy the analyticity condition of theorem 5.2, and let U_t be the unitary group representing α_t in the representation π determined by ω, then $U_t \in \pi(\mathfrak{A})''$, where $\pi(\mathfrak{A})''$ is the bicommutant of $\pi(\mathfrak{A})$.

An important consequence of this theorem is expressed by the

Corollary 5.5: If ω is the only state satisfying the analyticity condition of theorem 5.2, then ω is pure.

6. The equilibrium state at $T \neq 0$.

The discussion in this section will be in many respects similar to that in section 5. We start by defining an equilibrium state for a finite lattice Λ. Let $\beta=(kT)^{-1}$, where k is Boltzmann's constant and T the temperature, then the state ω_Λ defined by

$$\omega_\Lambda(A) = \mathrm{Tr}\ \rho_\Lambda A,$$

with

$$\rho_\Lambda = e^{-\beta H(\Lambda)}/\mathrm{Tr}\ e^{-\beta H(\Lambda)},$$

is the state of our lattice Λ corresponding to the canonical ensemble at temperature T (Gibbs-state). Using the definition of ω_Λ and the fact that the trace is invariant for cyclic permutation of factors, one proves

Lemma 6.1: Let $A,B\in\mathfrak{A}(\Lambda)$, then $\omega_\Lambda(A_t^\Lambda B)$ can be extended to an entire function of a complex variable t which is uniformly bounded in the strip $-\beta \le \mathrm{Im}\ t \le 0$; similarly $\omega_\Lambda(BA_t^\Lambda)$ can be extended to an entire function of t which is uniformly bounded in $0 \le \mathrm{Im}\ t \le \beta$; in addition

$$\omega_\Lambda(A_t^\Lambda B) = \omega_\Lambda(BA_{t+i\beta}^\Lambda) \quad \text{for } -\beta \le \mathrm{Im}\ t \le 0.$$

We now take the thermodynamical limit of ω_Λ. This limit, or the limit of any converging subsequence of ω_Λ is, by definition, an equilibrium state of the infinite lattice at $T = (k\beta)^{-1}$. The following theorem expresses how many of the properties of lemma 6.1 are conserved by taking the thermodynamical limit.

Theorem 6.2: Let ω be an equilibrium state of the infinite lattice, and let $A,B\in\mathfrak{A}$. Then

i. ω is invariant for α_t.
ii. $\omega(\alpha_t AB)$ can be extended to a function of a complex variable t, which is analytic for $-\beta < \mathrm{Im}\ t < 0$ and uniformly bounded for $-\beta \le \mathrm{Im}\ t \le 0$; similarly $\omega(B\alpha_t A)$ can be extended to a function, which is analytic for $0 < \mathrm{Im}\ t < \beta$ and uniformly bounded for $0 \le \mathrm{Im}\ t \le \beta$; in addition

$$\omega(\alpha_t AB) = \omega(B\alpha_{t+i\beta}A) \quad \text{for } -\beta \le \mathrm{Im}\ t \le 0.$$

The proof of this theorem is very analogous to that of theorem 5.2.

Definition: If a state has the property expressed in theorem 6.2ii, we say that the state satisfies the K.M.S.-condition.

According to theorem 6.2 every equilibrium state satisfies the K.M.S.-condition. Like in the case of the ground state we would be tempted to use the K.M.S.-condition as the defining property of an equilibrium state of the infinite lattice. It has the

advantage that it does not refer to the thermodynamical limit. Again it is not known whether the two definitions of an equilibrium state are identical. For finite systems, however, one can prove without difficulty that any state which satisfies the K.M.S.-condition is a Gibbs-state.

The K.M.S.- condition is in many respects the analogue for $T \neq 0$ of the analyticity condition of theorem 5.2. Also the following theorem is quite similar to one in section 5.

Theorem 6.3: A state which satisfies the K.M.S.-condition is invariant for α_t. There is also a theorem that corresponds to theorem 5.4. According to theorem 5.4 an element of the commutant $\pi(\mathfrak{A})'$ commutes with U_t, since $U_t \in \pi(\mathfrak{A})''$. This means that the commutant is pointwise invariant. In the case $T \neq 0$ we have

Theorem 6.4: If ω satisfies the K.M.S.-condition, then in the representation π determined by ω, all elements of the center $\pi(\mathfrak{A})'' \cap \pi(\mathfrak{A})'$ are invariant.

An important consequence of this theorem is the

Corollary 6.5: If ω is the only state satisfying the K.M.S.-condition, then ω is primary.

In addition to these results which have their counterpart in the ground state case, there are other results that have no such counterpart. In particular, the K.M.S.- condition has farreaching consequences with regard to the representation space determined by ω as expressed by the following theorem [8].

Theorem 6.6: If ω satisfies the K.M.S.-condition, then the cyclic representation $(\pi, \mathfrak{H}, \Omega)$ determined by ω has the following structure

i. There exists an involution operator J such that $J\pi(\mathfrak{A})'J = \pi(\mathfrak{A})''$ and $J\Omega = \Omega$.
ii. J commutes with U_t
iii. The positive operator $T = U_{\frac{1}{2}i\beta}$ satisfies the equation $T\pi(A)\Omega = J\pi(A)^*\Omega$.

An immediate and trivial consequence of this theorem is the fact that a state which satisfies the

K.M.S.-condition is not pure. Indeed, since J maps the Von Neumann-algebra onto its commutant, the commutant is non-trivial, so that π is reducible and ω not pure.

7. Heisenberg-model with long range interaction.

In this section we shall show by means of a simple model how the K.M.S.-condition may be used to calculate all correlation functions of the equilibrium state.

Our model is a molecular field approximation to the Heisenberg model. We obtain the hamiltonian of our model from that of the Heisenberg model

$$H(\Lambda) = \tfrac{1}{2} \sum_{x,y\in\Lambda} J(x,y)\vec{\sigma}(x)\vec{\sigma}(y),$$

by replacing $J(x,y)$ by $J/N(\Lambda)$, where $N(\Lambda)$ is the number of lattice points in Λ. We see that in our model the interaction between two lattice points is independent of their position. The potential Φ depends, however, on Λ so that the results of section 4 concerning the time-evolution cannot be applied.

Our model hamiltonian for the sublattice Λ is then

$$H(\Lambda) = \tfrac{1}{2} \frac{J}{N(\Lambda)} \Big(\sum_{x\in\Lambda} \vec{\sigma}(x) \Big)^2.$$

The operator $\sum_{x\in\Lambda} \vec{\sigma}(x)$ has a simple meaning. It is (disregarding a numerical factor) the operator for the total magnetization. The hamiltonian $H(\Lambda)$ gives rise to an inner automorphism of $\mathfrak{A}(\Lambda)$, from which we get the equations of motion

$$\frac{d}{dt}\sigma_i(x) = i[H(\Lambda),\sigma_i(x)] = i \frac{J}{N(\Lambda)} \sum_{x'\in\Lambda} \vec{\sigma}(x')[\vec{\sigma}(x),\sigma_i(x)] \tag{7.1}$$

We notice that the operator $\frac{1}{N(\Lambda)} \sum_{x\in\Lambda} \vec{\sigma}(x)$ is the operator for the average magnetization per lattice site.

We shall now investigate what happens if we try to take the thermodynamical limit. In this limit the equations of motion (7.1) would have a well-defined meaning

provided the space average occurring at the right hand side has a limit in the C*-algebra $\mathfrak{A}$. This is, however, not the case. This means, among other things, that the hamiltonian in this model does not define an automorphism group of the algebra of observables. On the other hand, it is known that in certain representations space averages of the infinite system exist, as, for example, expressed by the following theorem, which is due to Doplicher, Kastler and Robinson [9].

Theorem 7.1: If $\mathfrak{A}$ is asymptotically abelian with respect to translations (in the sense that for all A and $B \in \mathfrak{A}$ $[\alpha_a A, B] \to 0$ for $|a| \to \infty$) and ω is invariant for space translations, then in the representation determined by ω the space average of an observable exists in the strong operator topology and is a c-number (multiple of the identity) if and only if ω is extremal invariant.

We shall apply this theorem and assume that our system is in a state ω which is extremal invariant. We then consider the representation determined by ω. For simplicity we shall use the same notation for the observables and their representants in $B(\mathfrak{H})$. We can now replace the space average

$$\lim_{\Lambda\to\infty} \frac{1}{N(\Lambda)} \sum_{x\in\Lambda} \vec{\sigma}(x)$$

by the c-number $\vec{m}$. The equations of motion are now

$$\frac{d}{dt}\sigma_i(x) = iJ\vec{m}[\vec{\sigma}(x),\sigma_i(x)]. \qquad (7.2)$$

In this equation $\vec{m}$ is the average magnetization per lattice point in the representation determined by ω. Clearly, the equations (7.2) are representation-dependent. They give rise to the time-evolution

$$\sigma_i(x,t) = e^{iJ\vec{m}\vec{\sigma}(x)t}\sigma_i(x)e^{-iJ\vec{m}\vec{\sigma}(x)t}. \qquad (7.3)$$

This is an inner automorphism of $\mathfrak{A}$ which is, however, dependent on the choice of ω. As we see, this automorphism is even a local automorphism, in the sense that it is an automorphism of each $\mathfrak{A}_x = B(\mathfrak{H}_x)$.

Using properties of the Pauli-spinmatrices we can rewrite the equations (7.3) in the form

$$\sigma_i(x,t)=\cos^2(Jmt)\sigma_i(x)+i\,\sin(Jmt)\cos(Jmt)\,\frac{\vec{m}}{m}\,[\vec{\sigma}(x),\sigma_i(x)]\,+$$

$$+\,\sin^2(Jmt)\,(-\sigma_i(x)\,+\,2\,\frac{m_i\vec{m}}{m^2}\,\vec{\sigma}(x)) \qquad (7.4)$$

As we see immediately from (7.3) we have

$$\vec{m}\,\vec{\sigma}(x,t)\,=\,\vec{m}\,\vec{\sigma}(x). \qquad (7.5)$$

Since we want to calculate correlation functions of the system in thermal equilibrium at a given temperature T, we shall assume that ω satisfies the K.M.S.-condition. As we shall see the state is now completely determined.

Using the equation

$$m_i\,=\,\lim_{\Lambda\to\infty}\frac{1}{N(\Lambda)}\sum_{x\in\Lambda}\sigma_i(x) \qquad (7.6)$$

and the fact that ω is assumed to be translationally invariant, we find that

$$m_i\,=\,<\sigma_i(x)>. \qquad (7.7)$$

We shall now make use of the K.M.S.-condition, applied to the operators $A=\sigma_1(x)$ and $B=\sigma_1(x)$. This gives

$$<\sigma_1(x)\sigma_1(x,i\beta)>\,=\,<\sigma_1(x)\sigma_1(x)>\,=\,1. \qquad (7.8)$$

According to the equations of motion (7.4) we have

$$\sigma_1(x,i\beta)=\cosh^2(Jm\beta)\sigma_1(x)-\sinh(Jm\beta)\cosh(Jm\beta)\frac{\vec{m}}{m}[\vec{\sigma}(x),\sigma_1(x)]\,-$$

$$-\,\sinh^2(Jm\beta)\,(-\sigma_1(x)+\,2\frac{m_1\vec{m}}{m^2}\,\vec{\sigma}(x)).$$

If we substitute this in (7.8) we get the equation

$$1 = \cosh^2(Jm\beta) -$$
$$- \sinh(Jm\beta)\cosh(Jm\beta)\,\frac{\vec{m}}{m}\langle\sigma_1(x)[\vec{\sigma}(x),\sigma_1(x)]\rangle -$$
$$- \sinh^2(Jm\beta)\left(-1+2\,\frac{m_1\vec{m}}{m^2}\langle\sigma_1(x)\vec{\sigma}(x)\rangle\right),$$

which after some simple calculations and using (7.7) may be reduced to the form

$$\tanh(Jm\beta) = -m. \tag{7.9}$$

Equation (7.9) determines the modulus m of the average magnetization of the infinite lattice. There is always a solution $m = 0$, and if J is positive, that is the only solution. However, for negative values of J another solution exists provided $-\beta J > 1$ or $kT < -J$. If we define the critical temperature T_c by means of the equation $kT_c = -J$, then equation (7.9) admits a solution $m \neq 0$ if $T < T_c$. Hence, below the critical temperature there are two solutions, $m = 0$ and another solution $m \neq 0$. If one calculates the free energy corresponding to both solutions one will find that the solution with average magnetization $m \neq 0$ has the smallest free energy. The other solution is therefore unphysical. One should notice that equation (7.9) determines the magnitude but not the direction of the spontaneous magnetization. The direction remains arbitrary.

We still have not calculated all correlation functions, which are of the form

$$\langle\sigma_{i_1}(x_1)\sigma_{i_2}(x_2)\ \ldots.\rangle.$$

In order to do so, we consider the expression

$$\frac{\langle A0\rangle}{\langle 0\rangle} \tag{7.10}$$

where $A\in\mathfrak{A}_x$ and $0\in\mathfrak{A}_\Lambda$, with $x\notin\Lambda$. Let 0 be positive; then (7.10) is a state of $\mathfrak{A}_x$ which, as one can check easily, satisfies the K.M.S.-condition. Now the lattice site x alone is certainly a finite system.

According to section 6 the K.M.S.-condition applied to a finite system determines the state uniquely. Therefore this state of $\mathfrak{A}_x$ is equal to ω, so that

$$\frac{\langle A0 \rangle}{\langle 0 \rangle} = \langle A \rangle,$$

or

$$\langle A0 \rangle = \langle A \rangle \langle 0 \rangle.$$

From this we conclude immediately to the result, that for different points x_i

$$\langle \sigma_{i_1}(x_1) \sigma_{i_2}(x_2) \ldots \rangle = m_{i_1} m_{i_2} \ldots .$$

We have now determined our state ω completely.

REFERENCES

1. Naimark, M. A., (1964), "Normed Rings", Noordhoff, Groningen.

2. Dixmier, J., Les algèbres d'operateurs dans l'espace hilbertien (Algèbres de Von Neumann)(Gauthier-Villars, Paris 1957).

3. Dixmier, J., Les C*-algèbres et leurs représentations (Gauthier-Villars, Paris 1964).

4. Ruelle, D., (1969), "Statistical Mechanics", Benjamin, New York.

5. von Neumann, J., (1931), Math. Ann. 104, 507.

6. Takesaki, M., (1970 a), Commun. Math. Phys. 17, 33-41.

7. Robinson, D. W., (1967), Commun. Math. Phys. 6, 151-160

8. Haag, R., N. M. Hugenholtz and M. Winnink (1967), Commun. Math. Phys. 5, 215-236.

9. Doplicher, S., D. Kastler and D. W. Robinson (1966), Commun. Math. Phys., 3, 1.

WAVE-MECHANICAL FORMULATIONS OF EQUATIONS OF MOTION FOR GRAVITATIONAL SYSTEMS OF MANY PARTICLES

Herbert Jehle
George Washington University
Washington, D. C.

Many attempts have been made to give a hydrodynamical interpretation of wave-mechanics in view of explaining the latter in terms of hidden variables. This is not the objective of the present paper. Rather, we ask whether the hydrodynamical equations of motion for a large number of particles in purely gravitational interaction (i.e. a gravitational system) may be usefully reformulated in terms of a Schroedinger-type wave equation. The role of the fundamental constant h^3, which characterizes the uncertainty principle of ordinary wave mechanics is now taken by a macroscopic phase space volume σ^3 which characterizes the coarse-grainedness of the density and streaming fields of the gravitational system. That phase space volume, rather than being a fundamental constant, is a quantity characterizing the systems in question.

This wave-mechanical tool may usefully be applied in considering the recent development of the solar system, assuming that the present solar system was preceeded by a large set of smaller masses under gravitational interaction. It is suggested that the interacting set of particles undergo changes toward preferential orbital elements, changes which may be discussed in wave-mechanical terms. A study of the statistical aspects of such a purely gravitational system should be the first step towards a discussion of the regularities of orbital elements, a study preceeding that of more complex models based on hydromagnetic considerations and gas laws in addition to gravitational effects.

Diverse efforts have already been made in this direction. The most promising recent efforts have been done on the basis of computer calculations of a large number of mass points in interaction. Different assumptions about those systems have led to interesting but quite different results. Stan Ulam has given an analysis with the assumption of mergings of any two mass points when they make too close an approach; the formation of double stars etc. is studied. [1]

Another approach uses conventional orbital perturbation calculations to look into questions of stability of distributions and tendencies in the distribution of orbital elements. Such calculations involve extremely complex sets of assumptions when it comes to a problem such as the development of the distribution of orbital elements in the solar system. So complex are these assumptions that only in the problem of distribution around the gaps in the system of minor planets has progress been made, in particular by Dirk Brouwer, Wm. H. Jefferys, V. Szebehely, J.H. Bartlett, K. Franklin, C. Sagan, M. Lecar and many other colleagues. [2]

In this note we concern ourselves with density - and streaming - field approaches (sometimes called hydrodynamical) to gravitational systems, [3] in particular to the developing solar system. As an earlier example we remind ourselves of the von Weizsaecker model of whirls. With the present note we want to propose a more specific model of the hydrodynamical category. We start from the recognition that in a statistical dynamical system, the density and streaming fields are only defined in a coarse-grained way.

To this effect, we consider a system of masses of various sizes, single or multiple masses or subclusters, and we consider each of these statistically independent elements to move in the same smoothed-out potential $\bar{U}(x,y,z,t)$, with superimposed potential fluctuations. For gravitational systems where the interaction is long-range, such a picture is particularly appropriate; close encounter terms, so important in the Boltzmann approach to the dynamical theory of gases, play here only a very subordinate role. As, indeed, the large fluctuations of potential (and corresponding force fluctuations δF_ν per unit

mass, from the smooth field of force $-\nabla_y \bar{U}$) have the predominant effect, the deflections of individual particles caused by these force-fluctuations are almost independent of the sizes of the masses of these deflected particles. For this reason, the force or potential fluctuation concept is particularly adequate for gravitational systems. The fluctuations cause a diffusion superimposed upon the smooth streaming field in a manner similar to that calculated by Einstein, Smoluchowski, Kramers, Uhlenbeck and others.[4]

Instead of a large number of assumptions needed to carry through calculations of the orbital type, the basic assumption about coarse-grainedness in the present approach is as follows. For simplicity of illustration, let us consider the statistically independent elements to be all of equal mass and consider the phase-space volume per independent element to be

$$\delta x\, \delta y\, \delta z\, \delta v_x\, \delta v_y\, \delta v_z = \sigma^3 = (2\pi\bar{\sigma})^3 . \qquad (1)$$

By Liouville's theorem, applied to the smooth streaming, this volume stays constant, following a particular trajectory. As however, gravitational systems may usually be thought of as implying a fairly rapid mixing of phase-space motion, (on the coarse grained scale), that particular volume becomes one constant characterizing the entire system.

The situation therefore resembles pretty closely that of a Thomas-Fermi atom. In the Thomas-Fermi atom, however, all volumes corresponding to the lower stationary states are filled up to a certain level. But the situation is quite different in gravitational systems for which thermal types of considerations are not applicable. We may expect, e.g., in a developing solar system, to have the lowest states represented as well as groups of states corresponding to the predecessors of the various planets and satellites. For each group it is a large number of states which represent the large number of independent particles which are expected to have then formed the respective planets and satellites.

The coarse-grainedness also implies a limit to the strongest flattening of the system: consider a δv_z of the order of the root mean square deviation of the system, Δv_z, times $(4\pi)^{\frac{1}{2}}$. Then the system's thickness, $\Delta z(4\pi)^{\frac{1}{2}}$ has to be greater than or equal to the average distance δz between independent elements. Otherwise we would get a physically unrealistic picture of a perfectly flat system with random distribution of elements in the plane of the system. Accordingly,

$$\Delta z \, \Delta v_z \geq \tfrac{1}{2} \bar{\sigma}, \tag{2}$$

where it turns out that this system constant is of the order of (M=mass of the system)

$$\bar{\sigma} \approx MG/c \tag{3}$$

There is a certain analogy of that assignment to the quantum mechanical one.[5] Consider

$$\begin{aligned} \delta x \, \delta v_x/c &\approx h/mc = hc/mc^2 = 137 \times 2\pi e^2/mc^2 \\ \delta x \, \delta v_x/c &\approx \sigma/c = 2\pi MG/c^2 = 2\pi M^2 G/Mc^2 \end{aligned} \tag{4}$$

As mentioned in the introduction, the circumstance expressed in (3),(2) invites the description of a gravitational system in terms of a Schroedinger-type wave equation. The hydrodynamical analogy of wave mechanics was first considered by Madelung and by Van Vleck who showed the equivalence of a Schroedinger equation to a continuity equation and an Euler-Bernouilli equation[6], cf. (5),(6),(7) and who also showed that there is a perfect analogy to Liouville's phase-space conservation in quantum mechanics: if a wave packet is interpreted in terms of a distribution of appropriately chosen functions of position and momentum, the time development of such a wave packet follows the classical analog of invariance of the corresponding phase-space volume due to Liouville's theorem. Further investigations on these relationships have been made by Bohm, DeBroglie, Wilhelm, Takabayasi, Fenyes, Weizel, Nelson, de la Pena-Auerbach, and many others.[7]

We are not here concerned with the hydrodynamical interpretation of wave-mechanics, but rather with the wave

mechanical formulation of statistical dynamics, especially of coarse-grained hydrodynamical systems. They imply the laws of classical dynamics, plus statistical hypotheses. The latter ones are crucial. They might be formulated in terms of assumptions about orbital elements' distributions, about moments of velocity distributions, or, alternatively, they may be formulated in a more appropriate way in terms of coarse-grainedness, thus leading to the uncertainty of definition of the hydrodynamical fields. This indeterminacy makes significant the continuity and the Euler-Bernouilli equations alone, because the relations involving higher moments of velocity are expected to be overshadowed by the indeterminacies.

Whereas in ordinary quantum mechanics, stationary states play a predominant role, here we have to deal with a set of many stationary states corresponding to a set of many particles. Discreteness effects which are so characteristic for quantum mechanics, are therefore not discernable except for the coarse-graining phenomena like the above mentioned phenomenon of strongest flattening of a gravitational system.

In this approach, superposition of probability amplitudes ψ, rather than that of probabilities ρ, is assumed. One may ask what the role of such amplitude superposition is in the present theory. Interference effects arise in a trivial fashion in a gravitational setting which produces tidal effects. Interferences between initial ψ states and those resulting from time-dependent perturbation theory due to a quadrupole gravitational field reproduce the common tidal effects. In a more subtle setting of perturbations, discussed in figures 1,2,3 and Eqs. (9),(14),(15), interferences cause time-dependent perturbation potentials which lead to transition relations of the type:

$$(C_3 - C_4)\ \text{pickup} = \pm\ (C_3 - C_4)\ \text{release}$$

Those relations are of the type, so familiar in atomic physics between energy levels; they have their counterpart in the orbital pick-up and release arguments done with more clumsy classical statistical tools.

In the present paper, the point is made that the hydrodynamical equations for gravitational systems may appropriately be handled by wavemechanical techniques. Whether or not the developing solar system, in its recent development stages, satisfies the conditions of being representable simply as a gravitational system is an entirely different question. The numerical results shown in the table give, however, much credence to the feasibility of considering the orbital regularities to be explicable in such terms. Thus, if the assumptions about a system of masses in gravitational interaction are applicable, the wave-mechanical formulation provides for an immensely useful and reliable tool in formulating statistical dynamics. The basic inderterminacy assumption which replaces the former unwieldly set of assumptions is reminiscent of the situation is quantum mechanics in 1925, where the former complex set of hypotheses in terms of a Bohr orbital theory was replaced by quantum mechanics.

The basic formalism relating wavemechanics to hydrodynamics[8] is given by splitting the wave equation

$$(+\tfrac{1}{2}\,\bar{\sigma}^2\nabla^2+i\bar{\sigma}\,\partial/\partial t-c^2-U)\,|\psi|\exp\,(iS/\bar{\sigma}) = 0 \qquad (5)$$

into

$$+\sum_{1}^{3}\nabla_\nu(|\psi|^2\nabla_\nu S) + (\partial/\partial t)\;|\psi|^2 = 0 \qquad (6)$$

and a generalized Euler-Bernouilli equation including the rest energy c^2.

$$\partial S/\partial t+\sum_{1}^{3}\tfrac{1}{2}(\nabla_\nu S)^2 - (\bar{\sigma}^2/2|\psi|)\nabla^2|\psi|+U+c^2 = 0 \qquad (7)$$

The third term represents pressure potential due to mean square deviation of velocities.

A developing solar system in our terminology is understood to mean the following: from a certain time in the past onward, hydromagnetic effects and gas laws did no longer play a role important for the development of the present-day distribution of orbital elements. Such systems have been discussed in Poincaré's "Lecons sur les

Hypothèses Cosmogoniques":[9] Particles moving with high eccentricity or inclination experienced more encounters than other particles and thus fell towards the central region which left only direct orbits of moderate inclinations and excentricities. Gravitational interaction and rare encounters were considered then to have been responsible for the development towards preferential orbits.

We may then assume that in this already flattened early solar system, a subcluster was formed at the location of Saturn and particles with Jupiter's orbital elements had also been present in great abundance. Saturn and Sun being considered as second and first bodies in a restricted three-body system (orbiting circularly about each other), we may ask what influence they may have on the Jupiter particles. It may easily be seen that Jupiter particles, under the influence of the Saturn cluster, make preferential transitions toward orbital elements which are approximately those of Uranus. In other words, Jupiter particles not only form Jupiter, but also may be thrown into Uranus orbits.

In order to see this we have to take recourse to the discussion of the restricted three body problem,[2] given in the appendix. Using a rotating coordinate system, rotating with angular velocity Ω about the center of mass of Saturn and Sun, the gravitational potential is time independent. The Jacobi constant is an integral of the equation of motion, the value of the Hamiltonian for that system's small third bodies. Diagrammatically, it is shown in Fig. 1 that with Saturn considered as the second body, Uranus and Jupiter have approximately the same Jacobi constant $C=E-\Omega L_z$. It is however by no means obvious that this may permit a transition of Jupiter particles to Uranus because the Hill limit surfaces prevent such third body transitions. Indeed, the regions below the curved lines $-EL_z^2=\text{const}$ (which represents circular orbits in the energy-angular momentum diagram) are unphysical regions. They correspond to regions forbidden by the Hill condition $\dot{r}^2 \geq 0$ (eq. (11)).

In the orbital picture to which we shall refer in greater detail in the appendix, this transition from Jupiter to Uranus may be considered as follows: A highly

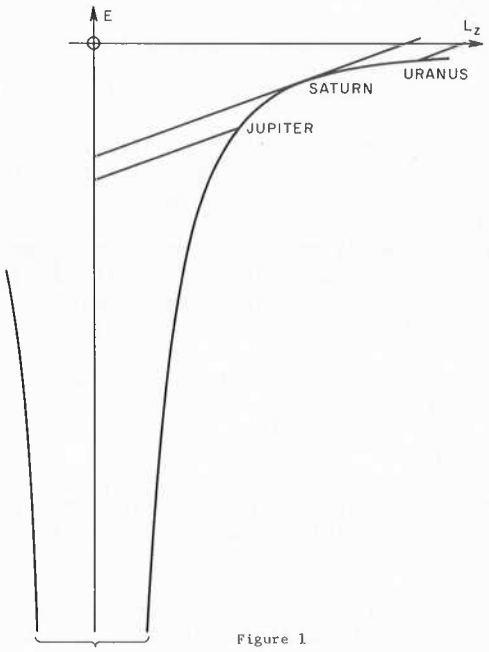


Figure 1

Fig. 1 expresses a relationship between Jacobi-constant levels; in this case the Jacobi-constants, with respect to Saturn chosen as second body in the restricted three body problem, are drawn as the parallel lines. It is seen that Jupiter and Uranus lie approximately on the same level. The relationship is formalized in Eq. (15). Direct and retrograde circular orbits are on the lines $-EL_z^2$ = constant. The diagrams are drawn with the first body considered as predominantly massive; energy E and angular momentum L_z are given in equation (12).

eccentric and quite massive particle of the Saturn swarm dives deeply into the Jupiter region, picks up a small Jupiter particle, which then moves in a hyperbolic orbit relative to that Saturn particle, and may thereafter be released from the latter upon reaching its aphelion region. Under appropriate assumptions, the released particle is then expected to have the same Jacobi-constant with respect to Saturn as it had before the pick-up. That released particle in general does not have a circular orbit with respect to the Sun. The transition to orbital elements closely resembling those of present day Uranus is assumed to be a subsequent process in which these released Uranus particles move along the line of constant value of Jacobi-constant until they reach the somewhat circular orbit, i.e. the intersection with the curved line, Fig. 1. The statistical assumptions, needed to calculate this process, as assumptions needed for any orbital type transition, are very complex; it is well known to those who have worked on such statistical problems that it is very difficult to make trustworthy assumptions because of the enormous complexity and arbitrariness of the situation. We have therefore found it more convincing to discuss this issue of equality of Jacobi-constants of Jupiter and Uranus (both with respect to Saturn) on the basis of wave-mechanical-hydrodynamical description.

In the wave-mechanical description there is evidently no tunnel effect which leads Jupiter particles to pass through the unphysical region to Uranus particles. The aforementioned eccentricity of some Saturn particles (and presumably also of some Jupiter particles) implies that the wave functions of Saturn particles and Jupiter particles overlap, leading to a gravitation potential whose time-dependence is chacterized by the beat frequencies of those overlapping wave functions. That gravitational potential, in turn, will cause transitions from the Saturn wave functions to Uranus wave functions, i.e. transitions characterized by the same (Jacobi-constant) level distance

$$C_{Saturn}-C_{Uranus} = C_{Saturn}-C_{Jupiter} \qquad (8)$$

Indeed, in the rotating coordinate system, the Schrödinger type equation corresponding to equation (5)

has the form, cf. (12),

$$[-\tfrac{1}{2}\,\bar{\sigma}^2(\nabla_x^2+\nabla_y^2+\nabla_z^2)+U-i\,\bar{\sigma}\Omega(y\nabla_x-x\nabla_y)+c^2]\psi = +i\,\bar{\sigma}\partial\psi/\partial t \quad (9)$$

where, $\bar{\sigma}\omega=C$ the Jacobi-constant for a ψ function of frequency ω, and Ω is the angular velocity of the Saturn-Sun system. The wave function of such a system is a superposition of a very wide range of energy eigenstates of what corresponds to quantum numbers ranging in the order of 10^5 in the cases of our planetary and satellite systems; we do not expect coherence effects to arise.

The aforementioned level relationships are the outcome of ordinary time-dependent perturbation theory. Besides the above-mentioned relationships,

$$C_{\text{second body}}-C_{\text{pickup particle}}=C_{\text{second body}}-C_{\text{release particle}}, \quad (10)$$

which is pictured in figure 1, we have also to expect similar relationships with opposite signature to occur, such as shown on the level schemes in figures 2 and 3.

Appendix

The Jacobi integral (for a small third body moving in the field of the first and second bodies which go in circular orbits about each other) is obtained by using a reference frame, x,y,z, which rotates with the constant angular velocity Ω about the center of mass of the first and second bodies. In that frame, the potential energy $U(x,y,z)$ is time independent, $\partial U/\partial t=0$. Multiplying the equations of motion $\ddot{x}-2\Omega\dot{y}-\Omega^2 x=-\partial U/\partial x$ etc., by $\dot{x}$, etc, one obtains immediately the first form of the Jacobi integral

$$\tfrac{1}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)\ -\tfrac{1}{2}\,\Omega^2(x^2+y^2)\ +\ U(x,y,z)\ =\ C \quad (11)$$

along the path of the third body. As the term $\tfrac{1}{2}(\dot{x}^2+\dot{y}^2+\dot{z}^2)$ cannot be negative, a third body of given Jacobi constant C is limited in its motion to regions in x,y,z space bounded by the "Hill limit surfaces" $\dot{x}^2+\dot{y}^2+\dot{z}^2=0$ in Eq.(11). The corresponding Langrangian defines the momenta $P_x = \dot{x}-\Omega y$, $P_y = \dot{y}+\Omega x$, $P_z = \dot{z}$; these are the (third body's)

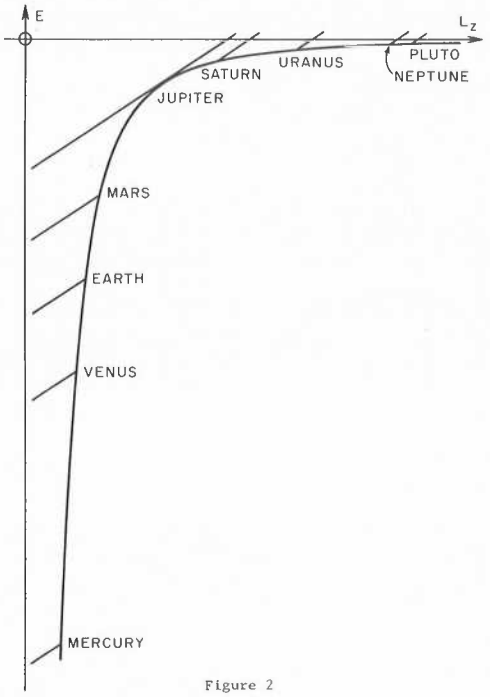


Figure 2

Fig. 2 shows the Jacobi-constants of the major planets all referred to Jupiter as the second body.

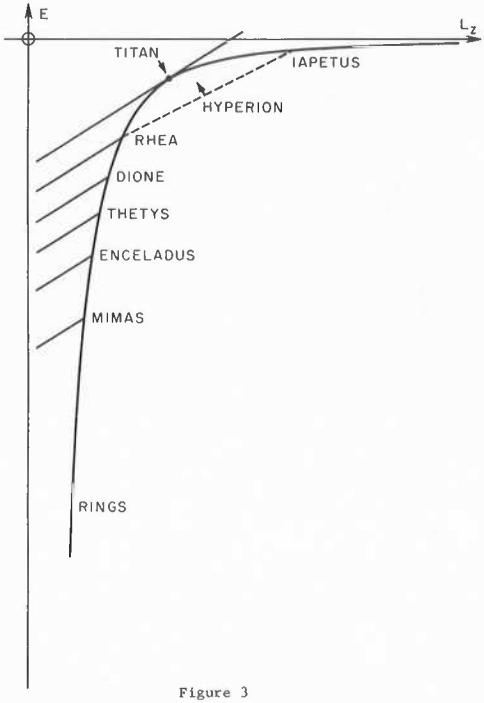


Figure 3

Fig. 3 shows the Jacobi-constants of Saturn's satellites with respect to Titan as second body.

x,y,z components of velocity in the inertial system. Expressing (11) in terms of P_x, P_y, P_z one obtains

$$\tfrac{1}{2}(P_x^2+P_y^2+P_z^2) + U-\Omega(xP_y-yP_x) \equiv E - \Omega L_z = C \qquad (12)$$

(12)denotes the Hamiltonian H which, since $\partial U/\partial t=0$, is constant. For circular orbit approximation about a predominant first body of mass M, $E = -\tfrac{1}{2}GM/a$, $L_z = (GMa)^{\frac{1}{2}}$ which are applicable to a third or to the second body. We may plot E and L_z in a diagram often employed by B. Lindblad[10] (Fig. 1). Parabolic orbits are at E = 0, circular direct and circular retrograde orbits are

$$-EL_z^2 = \text{constant } (=\tfrac{1}{2}(GM)^2). \qquad (13)$$

Elliptic orbits lie between these two curves (13); the region left and right of that pair of curves is unphysical. The slope $\Omega=(GMa_2^{-3})^{\frac{1}{2}}$ of a line $E-\Omega L_z=C$ in this diagram is equal to that of a line tangent to (13) at the point which would correspond to E, L_z of the second body, because (13) implies $dE/dL_z = -2E/L_z = \Omega$. The constant C labels the different parallel lines. Figure 1 is crudely simplified in that it does not show a tip at the location of the second body.

As an illustration of the application of the Jacobi integral we may refer to a note on the evolution of comets.[11] In that case Jupiter was considered as the second body in the restricted three-body problem for obvious reasons.

We are now in the position to formulate in detail the consequences of the above-mentioned transitions, illustrated in Fig. 1. This relationship between Jupiter and Uranus, due to Saturn acting as a second body implies that the tangent line of Saturn, which has the slope Ω, and the line connecting Jupiter and Uranus (in line with both the Jupiter and Uranus Jacobi line), have the same slope. Accordingly, in terms of the circular orbital elements of the three planets, we have[5]

$$(GM)^{\frac{1}{2}}a_{sat}^{-3/2} \approx (dE/dL_z)_{sat} = \Omega = \Delta E/\Delta L_z \approx (GM)^{\frac{1}{2}}\,\tfrac{1}{2}(a_{Ju}^{-1}-a_{Ur}^{-1})/(a_{Ur}^{\frac{1}{2}}-a_{Ju}^{\frac{1}{2}}) \qquad (14)$$

$$a_{sat}\sqrt{a_{sat}} \approx 2\, a_{Ju}a_{Ur}/(\sqrt{a_{Ju}} + \sqrt{a_{Ur}}) \qquad (15)$$

with Saturn considered as the second body.

We may now ask what occurs if the swarm of Jupiter particles is considered as the second body and the interaction with Saturn particles is considered. Transitions will occur to wave functions corresponding to particles located in the minor planet region; if they are assumed to adjust to approximately circular orbital condition, they would be expected to be found at the location 446×10^{11} cm, as indicated in table one. Similarly, if the particles forming a Uranus swarm are considered to represent the second body, they lead to transitions from Saturn to Pluto regions. In other words, the mechanism described in Fig. 1 may be considered to lead to sequence of orbital elements[5] given in the second column of Table 1. There are two more such planetary sequences, both corresponding to the transitions illustrated in Fig. 1. We note that these theoretical sequences nicely correspond to the actual sequences. In the first of these sequences we have started with the accurate semi-major axes of Jupiter and Saturn; the second sequence with accurate values for Saturn and Neptune; the third represents a crude interpolation sequence. With the present masses of minor planets, the second sequence would not be meaningful in regard to connecting the inner planets with the outer planets, but there may have been a larger amount of mass present in the minor planet region at the time to which we refer when the transition occurred.

Very beautiful is the coincidence of some of the theoretical sequences for Saturn satellites with the observed data. Also of interest are the sequences for Jupiter's and Uranus's satellites. The figures given in parentheses correspond to non-observed satellites.

Figures 2 and 3 correspond to situations in which only the heaviest planet (Jupiter) or heaviest satellite (Titan) are considered as second bodies. The approximate matching level spacings may again be an expression for interactions of the second type: for example, the overlap between Jupiter and Mars particles causing a time-dependent potential, which in turn causes transtion from Mars particles to Earth particles.

TABLE 1

We list semi-major axes of major planets and satellites; the transitions shown in Fig. 1 lead to theoretical sequences for major axes, sequences in which any three consecutive entries have the relationship shown in equation (15), (i.e. fig. 1), the middle entry as the second body.

System of Planets, semi major axes to 10^{11} cm.

	(Observed)	(Three theoretical sequences)		
Mercury	57.85		55.5	
Venus	108.11		108.5	108
Earth	149.45			155
Mars	227.7		230	227
Minor Pl.	(414)	446	538	341
Jupiter	777.6	777.6		
Saturn	1428	1428	1428	
Uranus	2873	2896		
Neptune	4501		4501	
Pluto	5900	5920		

Saturn's Satellites, semi major axes in 10^{8} cm.

	(Observed)	(Three theoretical sequences)		
Rings	120	120	120	
Janus	160			150
Mimas	188.5		189	188
Enceladus	252.5	239		240
Thetys	299		311	307
Dione	383			399
Rhea	534	520	535	525
Titan	1240	1268		
Hyperion	1503			
Lapetus	3580	3574		
Phoebe	13031 retrogr.			

(Continued next page)

Table 1 (cont'd)

Jupiter's Satellites, semi major axes in 10^8 cm.

	(Observed)	(Three theoretical sequences)		
V	181	184	181	185
				(275)
I Io	420	419		420
II Europe	667			663
III Ganymede	1068	1068	1068	1080
IV Gallisto	1877			1825
		(3166)		(3065)
VI	11430			
VII	11890	11690	11730	
X	12010			
VIII etc	24000 retrogr.			

Uranus Satellites, semi major axes in 10^8 cm.

	(Observed)	(Two theoretical sequences)	
Miranda	131	132	
Ariel	193		195
			(253)
Umbriel	268	268	
			(332)
Titania	440		440
Oberon	588.5	588	588

ACKNOWLEDGMENTS

I would like to express my deep gratitude for the hospitality at the Université Libre de Bruxelles, and in particular, for the encouragement and advice I received from Professor J. F. Cox. I would like to thank my colleague Professor Wm. C. Parke for his many suggestions. This work was supported by The George Washington University Committee on Research through an NSF institutional grant. For this generous assistance I am very grateful.

REFERENCES

1. S. Ulam, Motion Picture about Dynamics of a System of Masses in Gravitational Interaction.

2. J. H. Bartlett and C. A. Wagner, K. Danske Vid. Selsk. 2, #7 (1964); 3, #1 (1965);
D. Brouwer, Ann. Rev. Astr. and Aph. 1, 219 (1963);
Astr. J. 68, 152 (1963);
D. Brouwer and Wm. H. Jefferys III, Astr. J. 71, 543 (1966);
F. A. Franklin and G. Colombo, Icarus 12, 338 (1970);
M. Lecar, Astr. J. 71, 706 (1966; Mem. Soc. R. Sci. Liège 11, 227 (1967);
C. Sagan and W. W. Kellog, Ann. Rev. Ast. and Aph. 6, 2 (1968);
P. S. Shelus and S. S. Kumar, Astr.J. 75, 315 (1970);
V. Szebehely, Theory of Orbits (Acad. Press 1967).

3. I. Prigogine and M. J. Haggerty, Superclusters, Cosmology and Irr. Processes, priv. comm.; I. Prigogine and G. Severne, Physica 32, 1376 (1966);
Wm. C. Saslaw, M. N. 150, 299 (1970); 141, 77 (1968); 143, 437 (1969); 147, 253 (1970); C. F. von Weizsaecker Ap. J. 114, 165 (1951).

4. H. A. Kramers, Physica 7, 284 (1940); G. E. Uhlenbeck and G. W. Ford, Lect. in Stat. Mech., Am. Math. Soc. (1963).

5. H. Jehle, Z. F. Astroph. 15, 182 (1938); 18, 116 (1939); 19, 132 (1939); 19, 225 (1940);
Wm. C. Saslaw, priv. comm.

6. K. M. Case, Inst. f. Def. Analys. Res. Paper, 402 (1968)
K. M. Case, in "Statistical Models and Turbulence" ed. by M. Rosenblatt and C. Van Atta, Lecture Notes in Physics, Vol. 12, Springer (1972);
H. Froehlich, Proc. Phys. Soc. 87, 330 (1966); E. Madelung, Z. f. Phys. 40, 322 (1927);
J.H. Van Vleck, Phil. of Science 8, 275 (1941).

7. D. Bohm and J. P. Vigier, Phys. Rev. 96, 208 (1954); L. de Broglie, Etude critique, (Gautiers-Villars, Paris 1963); I. Fenyes, Z. f. Phys. 132, 81 (1952); E. Nelson, Phys. Rev. 150, 1079 (1966); Dyn. Th. of Brown. Motion (Princeton U. Press 1967); T. Takabayasi, Progr. Theor. Phys. 14, 283 (1955); W. Weizel, Z. f. Phys. 136, 582 (1954);
H. E. Wilhelm, Phys. Rev. D 1, 2278 (1970); Prog. Theor. Phys. 43, 861 (1970).
L. de la Pena-Auerbach, J. Math. Phys. 12, 453 (1971); 10, 1620 (1968); Revisita Mexicana Fig. 19, #2, 193 (1971).

8. H. Jehle, Phys. Rev. 70, 538 (1946).

9. G. Contopoulos, I. A. U. Symp. 25, (1965); R.H. Dicke, Nature 192, 440 (1961);
S. S. Kumar, Icarus 6, 136 (1967); Ap. J. 137, 1126 (1963);
Z. f. Aph. 58, 248 (1964); D. Layzer, Ann. Rev. Astr. and Aph. 2, 341 (1964); Astroph. and Gen. Rel., Brandeis Univ. Summer Inst. Theor. Phys. 1968, 2, 155 (G. & B. 1971); Int. Conf. on Thermodyn., Cardiff 1970 Ap. J. 137, 351 (1963).
T. Nakano, Progr. Theor. Phys., (July 1970); P.J.E. Peebles, J. Roy, Astr. Soc. Canada 63, 4 (1969);
H. Poincare, Lecons sur les Hyp. Cosmogoniques (Hermann 1913); D. W. Sciama and M. J. Rees, Nature 217, 511 (1968);
D. ter Haar, Ann. Rev. of Astr. and Aph. 5, 267 (1967).

10. B. Lindblad, Handb. d. Aph. 5, pt. 2, 937 (Springer 1933); Stockh. Obs. Ann. 20, #4 (1958).

11. P. Bougeois et J. F. Cox, Bull. Astr. 9, fasc. 7,8,9, 349 (1936);
J. F. Cox, and H. Jehle, Bull. Astr. de l'Obs. Roy. de Belgique 3, 248 (1944);
H. Jehle, Ap. J. 91, 471 (1940); L. Schulhof, Astr. Nachr. 124, 193 (1890);
P. van Halteren, Bull. Ac. Roy. de Belg., 312 (1947).

FUNCTIONAL TECHNIQUES AND THEIR APPLICATION IN QUANTUM FIELD THEORY

John R. Klauder
Bell Telephone Laboratories, Incorporated
Murray Hill, New Jersey

1. INTRODUCTION

In the broadest sense, functional methods permeate all of quantum theory, and the reason for their appeal is self-evident. Whether it be in the Schrödinger representation, the utilization of a specific group or algebra representation, the path integral approach of Feynman, the generating functional for time-ordered operator fields, or otherwise, functional techniques serve to bring quantum theory into the realm of analysis and subject therefore to the powerful tools of that discipline.

Yet for all the powerful techniques we still seem to be in the dark as to the proper formulation (let alone a meaningful solution!) of many problems in quantum field theory. One may adopt the approach of elaborating an elegant -- but largely formal -- structure purporting to solve a given field theory along more-or-less standard lines. Or one may proceed otherwise. Quantum field theory is, after all, the quantum theory of an infinite number of degrees of freedom, and covariant field theories are an important, but nevertheless small, class of such theories. Models with other invariance groups, particularly noncompact invariance groups, frequently have instructive lessons to offer if existence questions with regard to their construction can be overcome.

The analysis of certain classes of models is a primary goal of these lectures. The key to their analysis lies in a close interplay between functional techniques, probability theory and Hilbert space methods some parts of which may be unfamiliar to the average reader.

Consequently, these methods are reviewed and developed to the extent needed in the next two chapters, Chapters 2 and 3. The basic model field theories to be discussed -- the ultralocal scalar field theories -- are developed in Chapter 4. (Some material, especially that on scale-invariant models and scale dimension of the fields, is presented here for the first time.)*

In Chapter 5 a new class of model field theories is discussed, a class which contains covariant fields as examples, and about which a surprising amount of structual information can be ascertained. Insofar as these models are obtained from a "base theory" by extension to an additional dimension -- and also in virtue of the impressive geological surroundings in the Boulder area -- we adopt a suitable nomenclature from the geological sciences and refer to these as diastophic models.**

In the remainder of this chapter we revisit some of the classic functional formulations of quantum field theory, and examine some elementary features of an infinite number of degrees of freedom. To demonstrate that new wealth still lies within the province of a functional formalism it is incumbent upon us to show the relationship of our results to those predicted, or at least suggested, by one or more of the classic functional techniques surviving -- due to their elemental and fundamental truths -- from earlier eras. This brief comparison is discussed in Chapter 6. Hopefully, insight into the Great Problems may be won from analyses such as those presented in these notes.

---

* General references for the material presented in these notes appear in the Reference section at the end.

** diastrophism, n. The process or processes by which the earth's crust is deformed, producing continents and ocean basins, plateaus and mountains, folds of strata; also the results of these processes. diastropic, adj. : Webster's Third New International Dictionary (G. and C. Merriam Company, Springfield, Mass., 1959).

1.1 CLASSIC FUNCTIONAL APPROACHES TO QUANTUM FIELD THEORY

By way of illustration suppose we deal with a scalar field, $\varphi(\underset{\sim}{x},t)$ defined for $\underset{\sim}{x}\in R^s$, an s-dimensional configuration space. The purpose of quantum field theory, it may be stated, is to provide a realization for the local field operator $\varphi(\underset{\sim}{x},t)$. Supposedly, one has the Hamiltonian $\mathcal{H}$ and the space-translation generator $\underset{\sim}{\mathcal{P}}$ obtained by "correspondence" with some particular c-number theory. For instance, in a self-interacting covariant example we suspect that

$$\mathcal{H} = \int\{\tfrac{1}{2}[\pi^2+(\underset{\sim}{\nabla}\varphi)^2 + m_0^2\varphi^2] + V[\varphi]\}d\underset{\sim}{x}\ , \qquad (1\text{-}1)$$

$\pi(\underset{\sim}{x}) = \dot{\varphi}(\underset{\sim}{x})$, and

$$[\varphi(\underset{\sim}{x}),\ \dot{\varphi}(\underset{\sim}{y})] = i\delta(\underset{\sim}{x}-\underset{\sim}{y}).$$

Stated otherwise, an equation of motion, such as

$$(\Box + m_0^2)\varphi(x) = -V'[\varphi(x)], \qquad (1\text{-}2)$$

should determine the space-time dependence of the field operator.

An analogue of Schrödinger's formulation of quantum mechanics arises when we formally diagonalize the field operator $\varphi(\underset{\sim}{x})$ (at $t = 0$) and represent it by multiplication, say by $\Lambda(x)$. The conjugate operator $\pi(\underset{\sim}{x})$ becomes $-i\delta/\delta\Lambda(\underset{\sim}{x})$; and Schrödinger's equation, $\mathcal{H}\Psi = E\Psi$, is a functional differential equation in which

$$\mathcal{H} = \int\{\tfrac{1}{2}[-\delta^2/\delta\Lambda^2 + (\underset{\sim}{\nabla}\Lambda)^2 + m_0^2\Lambda^2] + V[\Lambda]\}d\underset{\sim}{x}. \qquad (1\text{-}3)$$

Elegantly simple, and heuristically almost preordained, such a prescription nevertheless requires "renormalization rules", which are usually deduced by the insertion of one or another cutoff.

The situation is little different in the other

classic functional formulations. Take, for example, the system of coupled (time-ordered) Green's function equations, or for conceptual simplicity that equation satisfied by the (Schwinger) T-product generating functional

$$\Omega\{j\} = \langle 0|Te^{i\int\varphi(x)j(x)dx}|0\rangle,$$

where $j(x)$ is a c-number (test) function. In a standard fashion the equation of motion (1 - 2) may be cast into an equation for Ω, namely

$$\left\{(\Box + m_0^2)\frac{\delta}{i\delta j(x)} + V'\left[\frac{\delta}{i\delta j(x)}\right] - j(x)\right\}\Omega\{j\} = 0.$$

A formal solution may be obtained using a functional Fourier transform,

$$\Omega\{j\} = \int e^{i\int j(x)\xi(x)dx}\tilde{\Omega}(\xi)\,\mathcal{D}\xi,$$

which leads to the solution

$$\Omega\{j\} = N\int e^{i\int j(x)\xi(x)dx + iI(\xi)}\,\mathcal{D}\xi,$$

where

$$\begin{aligned} I(\xi) &= \int\{\tfrac{1}{2}[(\partial_\mu\xi)^2 - m_0^2\xi^2] - V[\xi]\}dx \\ &\equiv I_0(\xi) - \int V[\xi]dx \end{aligned}$$

denotes the classical action and N is chosen so that $\Omega\{0\} \equiv 1$. This equally elegant, but quite formal path integral solution may be re-expressed according to

$$\begin{aligned} \Omega\{j\} &= e^{-i\int V[\delta/i\delta j]dx}\, N\int e^{i\int j(x)\xi(x)dx + iI_0(\xi)}\,\mathcal{D}\xi \\ &= N' e^{-i\int V[\delta/i\delta j]dx}\, e^{-\frac{1}{2}\int j(x)\Delta_F(x-y)j(y)dxdy}, \end{aligned} \qquad (1\text{-}4)$$

which is a useful heuristic expression, where as usual (in

four-dimensional space time)

$$\Delta_F(x) = \frac{i}{(2\pi)^4} \int \frac{e^{-ip\cdot x} d^4p}{p^2 - m_0^2 + i\epsilon}$$

Cutoffs Require Caution

How simple field theory would be if these formal notions really worked! The traditional remedy for the malaise contained in the preceding formulas is an application of cutoffs, both covariant and noncovariant. One device, for example, replaces Δ_F in (1-4) by

$$\Delta_{F,\Lambda}(x) = \frac{i}{(2\pi)^4} \int \frac{e^{-ip\cdot x}}{(p^2 - m_0^2 + i\epsilon)} \left[\frac{-\Lambda^2}{(p^2 - \Lambda^2 + i\epsilon)}\right]^2 d^4p$$

a bounded, continuous function so that the interaction term is locally defined. Such modifications in the functional formalism need not respect the underlying quantum theory, and correspond, in this case, to indefinite metrics. Eventually one links m_0, the coupling constants, etc., to Λ and attempts to pass to the limit $\Lambda \to \infty$.

Although conceptually reasonable -- and of unbelievable accuracy in the case of quantum electrodynamics -- there is just no a priori assurance that such procedures are physically reasonable. Consider the hypothetical class of models with classical Hamiltonians

$$H_N = \tfrac{1}{2}\Sigma_1^N (p_n^2 + \omega^2 q_n^2) + \lambda(\Sigma_1^N q_n^2)^2. \qquad (1\text{-}5)$$

When $N < \infty$, the quantum mechanics is straightforward; however, the limit of those straighforward quantum constructions as $N \to \infty$ exists only so long as $\lambda \to 0$ leading at the same time to a free thoery. Nevertheless, the quantum theory for $N = \infty$ can be formulated directly (via symmetry arguments) without passing through a sequence of cutoff theories, and is, as it should be, <u>not</u> a free theory ($\lambda \neq 0$). The qualitative differences between the true theory and each member of the sequence of cutoff theories are extensive and of unexpected varieties (e.g., reducible

representation of the canonical field and momentum operators; Hamiltonian inexpressible solely as a function of the canonical operators; etc.) Such gross differences between the true no-cutoff theory and the nonexistent limit of the conventional cutoff theories suggest an attitude of caution with regard to cutoffs: they work when they work, and they don't work when they don't! In the absence of definitive information a conservative viewpoint is advisable -- and this is just the attitude we adopt in the analysis of the model field theories which we discuss in later chapters.

To criticize cutoffs is by no means to criticize the concepts of renormalization, that is, the recognition that the naive construction of the Hamiltonian (say) is often incorrect and that counterterms of various types may be required. For instance, in some simple examples the counterterms have the principal effect of enforcing a specific representation of the field operators to be used in constructing the Hamiltonian. Examples of this type demonstrate the existence and relevance of inequivalent representations, and illustrate their relationship to renormalization effects.

1.2. ELEMENTARY EXAMPLES OF INEQUIVALENT REPRESENTATIONS OF FIELD OPERATORS

Consider the example of a collection of independent, identical harmonic oscillators with the Hamiltonian

$$\mathcal{H} = \tfrac{1}{2} \sum_{n=1}^{\infty} (P_n^2 + \omega^2 Q_n^2 - \omega)$$

together with the commutation relation

$$[Q_n, P_m] = i\,\delta_{nm}; \quad n,m = 1,2,\ldots \quad .$$

It follows that the operator

$$F_N = \frac{2}{N} \sum_{n=1}^{N} P_n^2$$

commutes with all the P_m and fulfills $[F_N, Q_m] = -i4P_m/N$ provided $N \geq m$. Thus as $N \to \infty$, $\lim F_N = F_\infty$ commutes with all

the P_m and Q_m, and it ought to be a multiple of unity for an irreducible representation. To evaluate F_∞ we may employ the ground state $|0>$ of the system as follows:

$$F_\infty = \lim \langle 0| F_N |0\rangle = \lim \frac{2}{N} \sum_{n=1}^{N} \langle 0|P_n^2|0\rangle = \omega.$$

Here we see a dynamical quantity, the energy level, derived from kinematical variables. For two different dynamical systems of this type characterized by ω and $\omega' \neq \omega$, the associated canonical operators cannot be unitarily equivalent. For, in the contrary case, where $P_n' = VP_nV^{-1}$ holds for all n, it would follow that

$$\omega' = F_\infty' = \lim F_N' = V \lim F_N V^{-1} = V\omega V^{-1} = \omega$$

which is manifestly incorrect. Thus these distinct dynamical systems require inequivalent representations of the canonical operators. There is just no escaping this fact, even in so simple an example as an infinite number of independent harmonic oscillators!

Once it is recognized that inequivalent representations arise, examples can be envisaged in which many such representations appear simultaneously. Consider the hypothetical model with Hamiltonian

$$\mathcal{H} = \tfrac{1}{2}\sum_{n=1}^{\infty} [P_n^2 + (P_0^2 + Q_0^2)^2 Q_n^2 - (P_0^2 + Q_0^2)] + \tfrac{1}{2}(P_0^2 + Q_0^2 - 1)$$

in which P_0 and Q_0 are an additional canonical pair. Here it is clear that $P_0^2 + Q_0^2$ is a constant of the motion and can be diagonalized to have the value $2p+1$, $p=0,1,2,\ldots$. In the subspace characterized by p, the remaining oscillator variables are characterized by a representation of the preceding type with $\omega = 2p+1$, and hence, in each subspace, by an inequivalent representation.

Another elementary example is given by

$$\mathcal{H} = \tfrac{1}{2}\sum_{n=1}^{\infty}[P_n^2 + (Q_n + c)^2 - 1].$$

For each real value c we again deal with an inequivalent representation of P_n and Q_m. Consider the operator

$$G_N = \frac{1}{N} \sum_{n=1}^{N} Q_n$$

which commutes with all the Q_m and fulfills $[G_N, P_m] = i/N$, $N \geq m$. Hence $G_\infty = \lim G_N$ commutes with all canonical operators and is a multiple of unity. This representation "tag" can be evaluated, much as before, by

$$G_\infty = \lim < 0 \; |G_N| \; 0 > = \lim \frac{1}{N} \sum_{n=1}^{N} < 0 \; |Q_n| \; 0 > = -c.$$

Again, for two different models, characterized by c and $c' \neq c$, we must have inequivalence of the canonical operators. For, in the contrary case, where $Q_n' = VQ_nV^{-1}$ holds for all n, it follows that

$$-c' = G_\infty' = \lim G_N' = V \lim G_N V^{-1} = V(-c)V^{-1} = -c$$

violating our assumption. We may extend this model in an evident fashion by "promoting" the constant c to a dynamical variable. Imagine that c is replaced by $\psi^\dagger\psi$, the number operator corresponding to a single fermion degree of freedom, and that we choose for the Hamiltonian

$$\mathcal{H} = \tfrac{1}{2} \sum_{n=1}^{\infty} [P_n^2 + (Q_n + \psi^\dagger\psi)^2 - 1] + \psi^\dagger\psi.$$

Here, $\psi^\dagger\psi$ is a constant of the motion having eigenvalues 0 and 1; in each of the corresponding subspaces an inequivalent representation of the canonical boson operators appears. This example is evidently similar to a static nucleon coupled by a Yukawa coupling to a meson field. Note the need for an infinite nucleon mass renormalization (when the bracket is expanded out), which compensates the perturbation theoretic diagrams

corresponding to the "infinite nucleon self mass."

A broader class of primer models is provided by

$$\mathcal{H} = \sum_{n=-\infty}^{\infty} \mathcal{H}_1(P_n, Q_n)$$

where, e.g.,

$$\mathcal{H}_1(P_n, Q_n) = \tfrac{1}{2}(P_n^2 + Q_n^2) + \lambda Q_n^4 - E$$

and E is chosen so that the ground state φ of $\mathcal{H}_1$ has zero eigenvalue. An essential feature of this example (like the earlier ones) is a noncompact invariance group ("translation"),

$$TP_nT^{-1} = P_{n+1}, \quad TQ_nT^{-1} = Q_{n+1},$$

which leaves the Hamiltonian invariant. Formally it is clear that

$$T^r e^{ipQ_o} T^{-r} = e^{ipQ_r}$$

converges as $r \to \infty$ to a multiple of unity,

$$A(p) = \lim e^{ipQ_r},$$

which can be evaluated by

$$A(p) = \lim \langle 0 | e^{ipQ_r} | 0 \rangle = \int e^{ipx} \varphi^2(x)\, dx.$$

Here $\varphi(x)$ (assumed real) is the Schrödinger representative of the ground state $\mathcal{H}_1$. Equivalence of two representations

involves equality of their A values for all real p, and hence, essentially, equality of the ground state $\varphi(x)$.

Model Building: Embellishments and Diastrophisms

The simple example given above illustrates a rudimentary type of "embellishment" of a simpler underlying problem, namely the single degree of freedom problem. Classically one could assert that

$$H_1(p,q) = \tfrac{1}{2}(p^2 + q^2) + \lambda q^4$$

and its straightforward discrete embellishment

$$H = \sum_{n=-\infty}^{\infty} H_1(p_n,q_n) = \sum_{n=-\infty}^{\infty} \{\tfrac{1}{2}(p_n^2 + q_n^2) + \lambda q_n^4\}$$

made sense. However quantum mechanically as we know this is not quite right since the minimum energy would be $\Sigma\, E=\infty$. An energy adjustment is the only "renormalization" required in discussing the quantum theory for discrete embellishments, which are therefore quite straightforward.

However, continous embellishments are a different matter altogether! By continuous embellishment we have in mind that p and q are "promoted" to functions of a real variable $w\in R$, $p\to p(w)$, $q\to q(w)$, and the embellished Hamiltonian is taken as

$$H = \int H_1\ (p(w),\ q(w))\ dw. \qquad (1\text{-}6)$$

Such a straightforward classical embellishment has, for its counterpart, a complicated quantum embellishment which will be the subject of considerable study on our part in these notes (Chapters 4 and 5). It is this process of model building to which we associate the geological term "diastrophism," corresponding to extension into a new dimension. The restriction $a<w<b$ defines a certain "stratum," and it is clear that different solutions may be running their course simultaneously, but quite independently, in different strata. Finally, while we have illustrated a diastrophic model (1-6)

based on a single degree of freedom problem, the base theory H_1 may also be a field theory, say a covariant one with Hamiltonian given in (1-1). The diastrophism of such field theories is the subject of Chapter 5.

2. HILBERT SPACE AND PROBABILITY THEORY

2.1 VECTORS AND OPERATORS: SELECTED PROPERTIES

Although the main ideas of Hilbert space used in quantum theory are generally well known, it is useful to emphasize some of the less common aspects that we will find especially useful. In dealing with a specific problem it is often of great utility to focus on a specific realization of Hilbert space matched to the problem in one way or another. For example, the practical utility of $L^2(R^3)$ for the Schrödinger equation of a particle in three dimensions is clear, although, in principle, $L^2(R)$ is "equivalent." Aspects of a problem unseen in one realization may become self-evident in a more appropriate realization. Thus it becomes not only useful to be aware of the abstract features of Hilbert space analysis, but also, if possible, to appreciate their concrete aspects in specific realizations.

Dense Sets and Total Sets

Abstractly a dense set of vectors $\mathcal{D}$ contained in a separable Hilbert space $\mathcal{H}$ is characterized by the property that for each $\psi\in\mathcal{H}$ and $\varepsilon>0$, there exists a vector $\varphi\in\mathcal{D}$ such that $||\varphi-\psi||<\varepsilon$. This may be accomplished by a countable set of vectors in $\mathcal{D}$, but this need not be the case. A total set of vectors $\mathcal{T}$ is characterized by the property that finite linear combinations of the vectors in a total set are dense, or, by the property that $(\varphi,\psi)=0$ for all $\varphi\in\mathcal{T}$ implies that $\psi=0$. Again, a total set may be countable (e.g. a complete orthonormal basis) but this need not be the case.

A Useful Convergence Criterion

Convergence of operators under various conditions frequently arises. Often one deals with a sequence of uniformly bounded operators B_N and questions whether or not

this sequence converges weakly to an operator B, that is whether

$$(\varphi, B_N \psi) \to (\varphi, B\psi)$$

for all φ, $\psi \in \mathfrak{H}$. An especially useful Lemma -- which we do not prove but frequently use -- is the

Silver Lemma*: For a sequence $\{B_N\}$ of uniformly bounded operators to converge weakly to an operator B it is necessary and sufficient that the sequence of complex numbers $(\varphi_i, B_N \psi_j)$ converge for arbitrary members φ_i, ψ_j in total sets.

Numerous examples of the Silver Lemma will appear subsequently, but it is useful to give one example here. Suppose that we employ the Hilbert space $\mathcal{l}^2$ composed of square summable sequences $\{z_n\}$, and that $(B_N)_{mn} = \delta_{mN}\delta_{nN}$, which is uniformly bounded. In the total set of orthonormal vectors of that basis it is clear that $(B_N)_{mn} \to 0$ as $N \to \infty$, and thus we are assured that (w=weak)

$$\text{w-lim } B_N = 0.$$

Operators and Forms

However what happens if B_N is not uniformly bounded? Suppose as an example we choose $(B_N)_{mn} = N^2 \delta_{mN}\delta_{nN}$ in a given basis in $\mathcal{l}^2$. Clearly, as $N \to \infty$, $(B_N)_{mn} \to 0$, but the conditions of the Silver Lemma do not apply. In particular, consider the dense set of vectors $\{z_n\}$ for which

$$z_n = d_n/n$$

where $\lim d_n = d_\infty$ exists. Between vectors of this type

$$\sum_{m,n=1}^{\infty} z_m^* (B_N)_{mn} z_n' = d_N^* d_N' \to d_\infty^* d_\infty';$$

---

* So named because of its great utility and fundamental role, and in analogy with the "Golden Rules" of quantum mechanics.

namely the matrix elements converge but observe that they are sensitive to the details of the "final entry" of the vectors $\{z_n\}$ and $\{z'_n\}$. If $\lim B_N$ were an operator, then the resultant expression would have to be continuous in the vector $\{z_n\}$. In particular, if $\lim B_N$ were an operator, it should follow that

$$\lim_{M\to\infty} \lim_{N\to\infty} \sum_{\substack{m=M \\ n=1}}^{\infty} z^*_m (B_N)_{mn} z'_n = 0,$$

but this limit equals $d^*_\infty d'_\infty$ which is in general different from zero. Such behavior is characteristic of so-called forms, which are defined for specific (dense to be useful) sets of kets and bras. We will occasionally encounter forms as well as operators, and it is frequently a useful generalization. Indeed, it is often convenient to characterize operators as well as forms by total sets of matrix elements. The Riesz Representation Theorem [which states, for example, that every continuous antilinear functional $A(\{z_n\})$ on ℓ^2 can be put in the form $\Sigma\, z^*_n a_n$, where the square-summable sequence $\{a_n\}$ is determined by A] can be used to help establish whether a given expression represents the matrix elements of an operator, or instead corresponds to a form.

2.2 RANDOM VARIABLES : SELECTED PROPERTIES

Probability Distributions

The theory of probability is so closely related to various aspects of quantum theory that much profit follows from cross fertilization of the two fields. A real random variable is characterized by a probabiltiy distribution $\mu(x)$ satisfying three basic properties:

(i) Nondecreasing: $\mu(x+h) \geq \mu(x)$, $h>0$;
(ii) Right continuous: $\mu(x+0) = \mu(x) \geq \mu(x-0)$;
(iii) Normalization: $\mu(-\infty) = 0$, $\mu(\infty) = 1$.

Such functions admit a canonical decomposition into discrete (d), singular (s) and absolutely continuous (ac) components: $\mu = a\mu_d + b\mu_s + c\mu_{ac}$, where $a\geq 0$, $b\geq 0$, $c\geq 0$ and $a+b+c = 1$. The discrete portion contains the discontinuiies of μ; the reminder, $\mu - a\mu_d$, is continuous. All discrete distributions are of the form

$$\mu_d(x) = \Sigma\, p_j\, \theta(x-x_j),$$

where $p_j>0$, and $\theta(x)$ is the degenerate distribution having unit discontinuity at $x=0$. A singular distribution $\mu_s(x)$ is continuous, but has a derivative for which $\mu_s'(x) = 0$ almost everwhere (a.e.). In addition, there exists a set N of Lebesque measure zero for which

$$\int_N d\mu_s(x) = 1.$$

An absolutely continuous distribution $\mu_{ac}(x)$ is the integral of its derivative, $\mu_{ac}(x) = \int_{-\infty}^{x} p(x)dx$, where the density function $p(x) \equiv \mu_{ac}'(x)\geq 0$, and is evidently $L^1(R)$, i.e. $\int |p(x)|dx<\infty$. A distribution μ which has only one component (e.g., a=1, b=c=o) is called pure (e.g., purely discrete).

Characteristic Functions

To each distribution $\mu(x)$ we associate a characteristic function $C(s)$ defined by

$$C(s) \equiv \int e^{isx} d\mu(x).$$

Every characteristic function is a continuous function, and respects the conditions $C(0) = 1$, $C(-s)=C^*(s)$ and $|C(s)|\leq 1$. The decomposition of probability distributions leads to a corresponding decomposition of characteristic functions:

$$C(s) = aC_d(s) + bC_s(s) + cC_{ac}(s).$$

Every discrete characteristic functions $C_d(s)$ is an almost periodic function and satisfies $\limsup C_d(s) = 1$ as $s \to \infty$. Every absolutely continuous characteristic function has the form

$$C(s) = \int e^{isx} p(x)dx,$$

and since $p(x)\in L^1(R)$ it follows from the Riemann-Lebesque Lemma that $\limsup C_{ac}(s) = 0$ as $s \to \infty$. [In general, this is only a necessary criterion for a distribution to be absolutely continuous since there are examples of singular distributions for which $\limsup C_s(s) = \zeta$ for any value of ζ in the range $0\leq\zeta\leq 1$.] A sufficient condition for a distribution to be absolutely continuous is that $\int |C(s)|ds<\infty$. Indeed, in that case, it even follows that

$$p(x) = (2\pi)^{-1} \int e^{-isx} C(s)ds$$

is continuous.

An important condition for the convergence of a sequence of characteristic functions is given by the Continuity Theorem: A sequence $C_n(s) = \int e^{isx} d\mu_n(x)$ of characteristic functions converges to a characteristic function $C(s) = \int e^{isx} d\mu(x)$ if and only if the functions $C_n(s)$ converge point-wise to a function C(s) which is continuous in the neighborhood of s=0. In such a case one asserts that the sequence of probability distributions $\mu_n(x)$ converges (weakly) to the probability distribution $\mu(x)$.

An important criterion for a function C(s) to be a characteristic function is contained in Bochner's Theorem: A function C(s) is the characteristic function of a probability distribution if and only if (i) C(s) is continuous, (ii) C(0) = 1, and (iii) C(s) is a "positive-definite function," i.e.

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j^* C(s_i - s_j) \geq 0$$

for all real s_i, complex α_i and $N < \infty$. The importance of the last condition follows from the desired relation

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j^* \int e^{i(s_i - s_j)x} d\mu(x) = \int \left| \sum_{i=1}^{N} \alpha_i e^{is_i x} \right|^2 d\mu(x) \geq 0.$$

2.3 RELATION OF HILBERT SPACE AND PROBABILITY THEORY

Self-Adjoint Operators

Unbounded operators abound in quantum theory, and domain questions are an unavoidable by-product. The vector $\varphi \in \mathfrak{H}$, an abstract Hilbert space, is in the domain of A, $\mathfrak{D}_A$, provided that $A\varphi \in \mathfrak{H}$. If $\mathfrak{D}_A$ is dense and the relation $(\psi, A\varphi) = (\psi^{\sim}, \varphi)$ holds for some ψ and all $\varphi \in \mathfrak{D}_A$, then we may identify $\psi^{\sim} = A^{\dagger}\psi$ where $A^{\dagger}$ is the adjoint operator. The set ψ for which such a condition holds defines the domain of $A^{\dagger}$, $\mathfrak{D}_{A^{\dagger}}$. If $\mathfrak{D}_A = \mathfrak{D}_{A^{\dagger}}$ and $A = A^{\dagger}$ on $\mathfrak{D}_A$, the operator is called self-adjoint.

In most physical applications one is interested in self-adjoint operators, for it is such operators which are the generators of unitary transformations. However, the domain conditions specified above are often complicated to verify. It is therefore of interest that an alternative characterization exists as now outlined.

Unitary One-Parameter Groups

A unitary one-parameter group is a family of operators $U(s)$ which fulfills the conditions:

(i) $U(o) = I$, the identity operator;
(ii) $U(s)U(s') = U(s+s')$;
(iii) $U(-s) = U(s)^{-1} = U(s)^{\dagger}$,

and weak continuity, i.e., continuity of the function $(\varphi, U(s)\psi)$ for all $\varphi, \psi \in \mathfrak{H}$. A fundamental theorem asserts that every unitary one-parameter group $U(s)$ is generated by some self-adjoint operator A, and that $U(s) \equiv e^{isA}$. Moreover, $\varphi \in \mathfrak{D}_A$ if and only if $(is)^{-1}[U(s)-1]\varphi$ converges in norm as $s \to 0$; and the limit is the vector $A\varphi$.

Spectral Resolution

Another fundamental theorem associated with self-adjoint operators is the spectral theorem. To each self-adjoint operator A is associated a family of projection operators $E(x)$ [i.e., $E^2(x) = E(x)$, $E(x)^{\dagger} = E(x)$] with the properties

(i) Nondecreasing: $E(x+h) \geq E(x)$, $h>0$;
(ii) Right continuous: $E(x+0) = E(x) \geq E(x-0)$;
(iii) Normalization: $E(-\infty) = 0$, $E(\infty) = I$.

In terms of this spectral family

$$A = \int x\,dE(x),$$

and similarly for various functions of A such as

$$e^{isA} = \int e^{isx}\,dE(x).$$

Relation to Characteristic Functions

Evidently the mean of the last expression in the state ψ leads to

$$(\psi,\ e^{isA}\psi) = \int e^{isx}\, d(\psi, E(x)\psi)$$
$$= \int e^{isx}\, d\mu(x) \equiv C(s),$$

which associates the characteristic function of a probability distribution to each self-adjoint operator A and state $\psi\in\mathcal{H}$. To give examples of various distributions it suffices to give the operator A and the state ψ. A discrete distribution is given by

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \equiv \sigma_z;\ \psi = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

which leads to the characteristic function $C_d(s) = \cos s$. A singular distribution is given by

$$A = \sum_{n=1}^{\infty} 3^{-n}(\sigma_z)_n\ ;\ \psi = \prod_{n=1}^{\infty} \otimes \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}_n$$

which leads to the characteristic function $C_s(s) = \prod_{n=1}^{\infty} \cos(s/3^n)$. An absolutely continuous distribution is provided by A=Q, the Schrödinger position operator, and ψ any wave function $\psi(x)$ so that $C_{ac}(s) = \int e^{isx}|\psi(x)|^2 dx$.

2.4 SPECIAL CLASSES OF PROBABILITY DISTRIBUTIONS

Statistical independence of the variables x_1 and x_2 follows from the factorization condition

$$\langle e^{is(x_1+x_2)}\rangle = \langle e^{isx_1}\rangle\ \langle e^{isx_2}\rangle,$$

or the equivalent statement that the characteristic function C(s) of the variable $x = x_1+x_2$ is given by $C_1(s)C_2(s)$. It is clear that C(s) is a characteristic function since

$$C(s) = C_1(s)C_2(s) = \iint e^{is(x_1+x_2)} d\mu_1(x_1) d\mu_2(x_2)$$
$$= \int e^{isx}\, d\mu_{1*2}(x)$$

where

$$\mu_{1*2}(x) \equiv \int \mu_1(x-y)\ d\mu_2(y)$$

which fulfills all the criteria to be a distribution function.

If the distributions μ_1 and μ_2 are equal, then it follows from the foregoing that $C(s) = [C_1(s)]^2$ defines a new characteristic function, and this property extends to an arbitrary integral power, $C(s) \equiv [C_1(s)]^m$, $m \geq 1$. Conversely, in this case, it would follow that $[C(s)]^{1/m} = C_1(s)$ would be a characteristic function, but in general such fractional powers do not lead to characteristic functions.

Infinitely Divisible Distribution Functions

There is an important class of distributions (and therefore of characteristic functions) called infinitely divisible characterized by the property that

$$[C(s)]^{1/m} \equiv C_m(s)$$

is a characteristic function for all positive integers m. An immediate consequence is that $[C(s)]^{n/m}$ is a characteristic function for all positive integers n and m, and thus by the Continuity Theorem [as the ratio $(n/m) \to \tau$] it follows that $[C(s)]^{\tau}$ is a characteristic function for all real $\tau > 0$. Examples of infinitely divisible characteristic functions are: the degenerate, $C(s) = e^{ias}$; the Gaussian, $C(s) = e^{-bs^2}$, $b>0$; and the Poisson, $C(s) = e^{\lambda(e^{is}-1)}$, $\lambda>0$.

If $\tilde{C}(s)$ denotes a characteristic function, then the expression

$$C(s) = e^{p[\tilde{C}(s)-1]}, \qquad p>0$$

defines an infinitely divisible characteristic function. An important theorem asserts that these special examples are dense in the set of all such characteristic functions; namely, we have

De Finetti's Theorem: Every infinitely divisible characteristic function is obtained as the limit

$$C(s) = \lim e^{p_m[\tilde{C}_m(s)-1]} \tag{2-1}$$

as $m \to \infty$, where $p_m > 0$ and $\tilde{C}_m(s)$ are characteristic functions

for all m*. Specifically, if C(s) is infinitely divisible then $[C(s)]^{1/m} \equiv C_m(s)$ is a characteristic function so that

$$C(s) \equiv \lim e^{m\{[C(s)]^{1/m}-1\}} = \lim e^{m[C_m(s)-1]}$$

which is a construction in the manner of the theorem.

A canonical representation, due to Lévy, asserts that C(s) is an infinitely divisible characteristic function if and only if

$$\ln C(s) = ias - bs^2 + \int_{|x|>0} (e^{isx} - 1 - \frac{isx}{1+x^2}) d\sigma(x) \quad (2\text{-}2)$$

where a, b, and σ are real, $b \geq 0$, and σ is a positive measure fulfilling

$$\int_{|x|>0} (\frac{x^2}{1+x^2}) \, d\sigma(x) < \infty.$$

It has been shown that either $b>0$ or $\int_{|x|>0} d\sigma_{ac}(x) = \infty$ is sufficient for C(s) to be absolutely continuous.

The class of symmetric $[C(-s) = C(s)]$ infinitely divisible distributions is characterized by the fact that

$$C(s) = e^{-bs^2 - \int_{|x|>0} [1-\cos(sx)] d\sigma(x)}. \quad (2\text{-}3)$$

We shall encounter this formula often in our further studies. We term the initial contribution the Gaussian component, and the latter the non-Gaussian (or Poisson) component.

Stable Distributions

A further specialization within the class of probability distributions are the so-called stable distributions. The characteristic functions of such distributions have the

---

* The converse is straightforward: Every characteristic function of the form (2-1) is infinitely divisible.

property that for every $\beta_1>0$ and $\beta_2>0$ there exist a $\beta>0$ and real γ such that

$$C(\beta_1 s)C(\beta_2 s) = e^{i\gamma s}C(\beta s). \qquad (2\text{-}4)$$

It follows from this relation that for all m $C(s) = [C(\beta_m s)]^m e^{i\gamma_m s}$ for some $\beta_m>0$ and γ_m. Thus $[C(s)]^{1/m}$ is a characteristic function for all m, and hence stable distributions are infinitely divisible.

We confine our attention only to symmetric stable distributions, $C(-s) = C(s)$. The functional relation (2-4) can be used to show for symmetric stable distributions that

$$C(s) = e^{-k|s|^\alpha},$$

where α, the exponent of the distribution, satisfies $0<\alpha\leq 2$, and $k>0$ (we exclude the uninteresting case $k=0$). On reference to (2-3) we see that $\alpha=2$ corresponds to a Gaussian distribution, while the other cases for which $0<\alpha<2$ correspond to non-Gaussian distributions (b=0) where

$$d\sigma(x) = k'|x|^{-(1+\alpha)}dx; \qquad |x|>0.$$

All stable distributions are absolutely continuous as follows from the fact that $C(s) \in L^1(R)$. (Stable distributions will arise in our analysis of scale invariant field theory models.)

2.5 FUNDAMENTAL PROPERTIES OF ANNIHILATION AND CREATION OPERATORS

Single Degree of Freedom

Many of the properties and techniques of use in a field theory have their analogue in a single degree of freedom framework. Imagine that a and $a^\dagger$ are the annihilation and creation operators of a single degree of freedom which fulfill the commutation relation $[a,a^\dagger] = 1$ (i.e., the identity operator). The state $|0>$ satisfies $a|0> = 0$ and the vectors $|n> \equiv (n!)^{-\frac{1}{2}}a^{\dagger n}|0>$ for all n form a total set in $\mathfrak{H}$. The vectors

$$|z\rangle \equiv e^{-\frac{1}{2}|z|^2} \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}} |n\rangle$$

$$= e^{-\frac{1}{2}|z|^2} e^{za^\dagger} |0\rangle,$$

for all complex z, are the coherent states and they enjoy a number of interesting and useful properties. For example, they are total in $\mathcal{H}$, and they are eigenstates of a,

$$a|z\rangle = z|z\rangle.$$

The inner product of two such states reads

$$\langle z|z'\rangle = e^{-\frac{1}{2}|z|^2 - \frac{1}{2}|z'|^2 + z^* z'},$$

and although normalized they are never orthogonal. The unitary operator

$$U(z) = e^{(za^\dagger - z^* a)}$$

defines the coherent states, $|z\rangle = U(z)|0\rangle$, and leads to the "translation formula"

$$U(z)^{-1} \, a \, U(z) = a + z.$$

Coherent state matrix elements of normally ordered operators are especially easy to compute. For if $:B(a^\dagger, a):$ expresses the operator, then

$$\langle z|:B(a^\dagger, a):|z'\rangle = B(z^*, z') \langle z|z'\rangle.$$

Moreover it follows readily that the diagonal elements $\langle z|:B(a^\dagger, a):|z\rangle = B(z^*, z)$ actually determine the operator since $B(z^*, z')$ can be constructed from a knowledge of $B(z^*, z)$.

The basic relation

$$e^{i\tau a^\dagger a} \, a^\dagger \, e^{-i\tau a^\dagger a} = e^{i\tau} a^\dagger$$

implies that

$$e^{i\tau a^\dagger a} |z\rangle = | \, e^{i\tau} z\rangle.$$

From the expression

$$\langle z|\, e^{i\tau a^\dagger a}\,|z\rangle = e^{(e^{i\tau}-1)z^*z}$$

it follows that

$$e^{i\tau a^\dagger a} = :e^{(e^{i\tau}-1)a^\dagger a}: \,. \tag{2-5}$$

A number of other properties are enjoyed by coherent states, such as a simple resolution of the identity, but they will not be of use to us in these notes.

Infinite Number of Degrees of Freedom

Let us introduce a field of formal annihilation and creation operators $A(y)$ and $A^\dagger(y)$ where $y\in R^n$ for some n, which fulfill the relation

$$[A(y),\, A^\dagger(y')] = \delta(y-y').$$

Again we assert that $A(y)|0\rangle = 0$ for all y, and that $\mathfrak{H}$ is spanned by repeated action of the formal creation operators acting on the vacuum $|0\rangle$. These requirements dictate that we deal with the Fock representation of A and $A^\dagger$. The vectors

$$|\psi\rangle = e^{-\frac{1}{2}\int|\psi(y)|^2 dy}\; e^{\int\psi(y)A^\dagger(y)dy}|0\rangle$$

for all $\psi(y)\in L^2$ are coherent states. Such states are total in $\mathfrak{H}$, but so too are any subset of coherent states in which $\psi(y)$ lies in a set dense in L^2 (in the L^2 topology). For example, the set of functions C_0^∞ (which are infinitely differentiable and of compact support) leads to a total set of coherent states. The eigenstate property reads

$$A(y)\,|\psi\rangle = \psi(y)\,|\psi\rangle.$$

The inner product of two such states is given by

$$\langle\psi|\psi'\rangle = \exp\{-\int[\tfrac{1}{2}|\psi(y)|^2 + \tfrac{1}{2}|\psi'(y)|^2 - \psi^*(y)\psi'(y)]dy\}.$$

The unitary operator

$$U(\psi) = \exp\{\int[\psi(y)A^\dagger(y) - \psi^*(y)A(y)]dy\} \tag{2-6}$$

defines the coherent state, $|\psi> = U(\psi)|0>$, and leads to the "translation formula"

$$U(\psi)^{-1} A(y) U(\psi) = A(y) + \psi(y).$$

As in the simple case above, coherent state matrix elements of normally ordered operators are simple to evaluate, and in particular

$$<\psi| : B(A^\dagger, A) : |\psi'> = B(\psi^*, \psi')<\psi|\psi'>.$$

If we set

$$\hat{W} \equiv \int A^\dagger(y)\, w\, A(y)dy, \tag{2-7}$$

where w is a self-adjoint operator acting on the variable y, then

$$e^{is\hat{W}} A^\dagger(y) e^{-is\hat{W}} = e^{isw} A^\dagger(y)$$

where e^{isw} is a unitary transformation on the variable y. It follows, in particular, that

$$e^{is\hat{W}}|\psi> = |\psi_s>,$$

where $\psi_s(y) \equiv e^{isw}\psi(y)$. The analogue of (2-5) reads

$$e^{is\int A^\dagger(y)wA(y)dy} = :e^{\int A^\dagger(y)(e^{isw}-1)A(y)dy}:$$

as may be seen directly from their action on coherent states.

Bilinear Operators

We shall be especially interested in studying operators of the form

$$W \equiv U(\xi)^{-1} \hat{W} U(\xi)$$

$$= \int [A^\dagger(y) + \xi^*(y)]\, w[A(y) + \xi(y)]dy.$$

where $U(\xi)$ is given by (2-6), assuming that $w\xi(y) \in L^2$. If we introduce

$$B(y) \equiv A(y) + \xi(y),$$

then we have the relations

$$W = \int B^{\dagger}(y) w B(y) dy$$

and

$$e^{isW} = e^{is\int B^{\dagger}(y) w B(y) dy} = :\, e^{\int B^{\dagger}(y)(e^{isw}-1)B(y)dy}\,:\,.$$

Thus it follows that

$$\langle \psi | e^{isW} | \psi' \rangle$$

$$= \exp\{\int [\psi^{*}(y)+\xi^{*}(y)](e^{isw}-1)[\psi'(y)+\xi(y)]dy\}\ \langle \psi | \psi' \rangle,$$

and in particular that

$$\langle 0 | e^{isW} | 0 \rangle = \exp\{\int \xi^{*}(y)(e^{isw}-1)\xi(y)dy\}.$$

We are especially interested in studying these latter two equations for sequences of operators W_n defined by sequences of translations $\xi_n(y)$ which are not restricted to have a limit in L^2 (not restricted to a Cauchy sequence). Let us introduce the notation

$$J_n(s) \equiv \int \xi_n^{*}(y)(e^{isw}-1)\ \xi_n(y)dy,$$

and

$$\varphi_n(y,s) \equiv (e^{isw}-1)\ \xi_n(y) \in L^2,$$

where we assume that $w\xi_n(y) \in L^2$ for all n. Now several important criteria can be distinguished.

Convergence to Characteristic Function: The sequence of characteristic functions

$$C_n(s) \equiv \langle 0 | e^{isW_n} | 0 \rangle$$

converges, in virtue of the Continuity Theorem, to a characteristic function C(s) provided that the sequence $J_n(s)$ converges pointwise to a continuous function $J(s)=\ln C(s)$. By DeFinetti's theorem C(s) will be an infinitely divisible characteristic function.

Weak Operator Convergence: The sequence of unitary operators e^{isW_n} converges weakly to an operator B, in virtue of the Silver Lemma, provided that the sequence $J_n(s)$ converges and that the vector sequence $\varphi_n(y,s)$ converges weakly to a vector $\varphi(y,s)$.

True as this last result is, we may wish to ensure that e^{isW_n} converges to a unitary one-parameter group characterized by e^{isW}. This entails further conditions on the sequences $J_n(s)$ and $\varphi_n(y,s)$ which are summarized as

Strong Operator Convergence: The sequence of unitary one-parameter groups e^{isW_n} converges weakly, hence strongly, to a unitary one-parameter group e^{isW} provided that the sequence $J_n(s)$ converges pointwise to a continuous function J(s) and that the sequence $\varphi_n(y,s)$ converges strongly to a vector $\varphi(y,s)$ continuous in the parameter s. If such is the case, we say that

$$W = \lim \int [A^{\dagger}(y)+\xi_n^{*}(y)]\ w[A(y)+\xi_n(y)]dy \qquad (2\text{-}8)$$

in the sense described above.

Note again, in these various contexts, that neither $w\xi_n(y)$ nor $\xi_n(y)$ need to be Cauchy sequences in L^2, so long as the other conditions are fulfilled. This construction will be often used explicitly and implicitly in subsequent chapers.

We conclude this chapter with an outline of the proof of the last convergence criterion. This result is readily proved if we note that (s = strong) s-lim $e^{isW_n} = U(s)$ is ensured provided

$$||(e^{isW_n} - e^{isW_m})|\psi>||^2 = 2[1-\mathrm{Re}\langle\psi|e^{-isW_n}e^{isW_m}|\psi>]$$

is a Cauchy sequence for a total set of unit vectors, say for all coherent states. From the bilinear form of W_m it follows, for each coherent state $|\psi>$, that

$$e^{isW_m}|\psi> = e^{i\theta(m)}|\psi_{(m)}>;$$

that is, apart from a phase, we obtain a new coherent state, where

$$\theta_{(m)} \equiv - \text{Im} \int \psi^*(y)(e^{-isw}-1)\xi_m(y)dy,$$

$$\psi_{(m)}(y) \equiv e^{isw}\psi(y) + (e^{isw}-1)\xi_m(y).$$

Weak convergence, implied by strong convergence, already demands that w-lim $\varphi_m(y,-s) = \varphi(y,-s)$, and thus

$$\lim \theta_{(m)} = \theta \equiv - \text{Im} \int \psi^*(y)\varphi(y,-s)dy.$$

Hence strong convergence of $e^{isW_m}|\psi>$ is fulfilled provided that

$$2[1-\text{Re}<\psi_{(n)}|\psi_{(m)}>] = || \, |\psi_{(n)}> - |\psi_{(m)}>||^2 \tag{2-9}$$

is a Cauchy sequence for a dense set of $\psi(y)\in L^2$. However, it may readily be shown, for two coherent states $|\psi>$ and $|\psi'>$, that

$$2[1-e^{-\frac{1}{2}||\psi-\psi'||^2}] \leq || \, |\psi> - |\psi'>||^2$$

$$\leq 4(||\psi||+||\psi'||)||\psi-\psi'||,$$

where $||\psi||^2=\int|\psi(y)|^2dy$. Consequently, convergence of $||\psi_{(n)} - \psi_{(m)}||$ is equivalent to convergence of (2-9). Thus we need strong convergence of $(e^{isw}-1)\xi_n(y)=\varphi_n(y,s)$ for all s. The combination law $U(s)U(s') = U(s+s')$ follows from the strong continuity, while the properties $U(o) = I$ and $U(s)^{-1}= U(-s) = U^\dagger(s)$ are straightforward. Finally, continuity of $J(s)$ and $\varphi(y,s)$ are needed to secure the weak continutiy, completing the requirements in order that $U(s)=e^{isW}$ with W self-adjoint.

3. FIELD OPERATORS AND CURRENT ALGEBRAS: GENERAL FEATURES AND SELECTED REPRESENTATIONS

3.1 EQUAL TIME COMMUTATION RELATIONS

Heuristic Formulation

A common starting point for the formulation of a quantum field theory is the introduction and analysis of various kinematical operators. In canonical theories one usually deals with a collection of field operators $\varphi_r(\underline{x})$ and conjugate momentum operators $\pi_r(\underline{x})$, r=1,...,R, that fulfill canonical commutation relations (CCR) of the form

$$[\varphi_r(\underline{x}), \pi_s(\underline{y})] = i\,\delta_{rs}\,\delta(\underline{x}-\underline{y}).$$

Frequently, one identifies the conjugate momentum $\pi_r(\underline{x})$ with $\dot{\varphi}_r(\underline{x}) = -i\,[\varphi_r(\underline{x}),\mathcal{H}]$, where $\mathcal{H}$ is the Hamiltonian operator, but there is no requirement that this be so. A further common assumption is that the kinematical operators -- φ_r and π_r in this case -- form an irreducible set of operators, and therefore that all operators including $\mathcal{H}$ can be expressed as functions of φ_r and π_r. However, this assumption is not always valid; for a generalized free field, for example, the representation of the CCR is reducible and $\mathcal{H}$ is not a function solely of the field and its conjugate.

Besides the usual canonical operators, other field algebras have recently become of importance. Current algebras usually consist of a family of local fields $j_r(\underline{x})$, r=1,...,R, which satisfy the equal-time commutation relation

$$[j_r(\underline{x}), j_s(\underline{y})] = i\,\delta(\underline{x}-\underline{y})\,c_{rst}\,j_t(\underline{x})$$

where c_{rst} are the structure constants of a Lie algebra. In such approaches one often imagines, as before, that the $j_r(\underline{x})$ form a complete set of operators, i.e., constitute an irreducible representation, and that, as a consequence, operators such as the Hamiltonian are expressible in terms of the basic fields.

Smearing Functions

The formal relations given above need to be augmented before a careful study can be initiated. Fields at a point are generally not well defined operators (but are rather forms) as is suggested by the presence of the δ function in the preceding relations. Smearing with test functions is generally required leading to the relation

$$[\varphi_r(f), \pi_s(g)] = i\,\delta_{rs}(f,g),$$

where

$$\varphi_r(f) = \int f(\underline{x})\varphi_r(\underline{x})d\underline{x}$$

etc., and

$$(f,g) \equiv \int f(\underline{x})g(\underline{x})d\underline{x}.$$

In the case of the current algebra we obtain

$$[j_r(f), j_s(g)] = i\,c_{rst}\,j_t(fg)$$

where $(fg)(\underline{x}) \equiv f(\underline{x})g(\underline{x})$. Observe that in either of the two cases if the supports of f and g lie in distinct regions of space the operators commute. For the most part we leave the class of test functions open; however they may be assumed, for convenience, to be infinitely differentiable and to have compact support.

Bounded Operator Formulation

One further refinement is traditional and useful, and that consists of replacing the unbounded field operators by bounded unitary operators (assuming that the smeared field operators are self adjoint). That is, attention is initially focussed on the operator families

$$e^{i\varphi(f)},\ e^{i\pi(g)},\ e^{ij(f)}$$

(suppressing indices) for test functions belonging to some, suitable real linear topological vector space. The commutation relations are replaced by such laws as

$$e^{i\varphi(f)}\,e^{i\pi(g)} = e^{i\pi(g)}\,e^{i\varphi(f)}\,e^{-i(f,g)},$$

which is known as the Weyl form of the CCR, and

$$e^{ij(f)}\, e^{ij(g)} = e^{ij(f\cdot g)}$$

where $f\cdot g$ represents multiplication within the group.

An essential fact, evidently, is the study of a family of self-adjoint operators we may generically call $W(f)$ and associated unitary operators $e^{iW(f)}$ for suitable test functions f. In the remainder of this chapter we primarily concentrate on Abelian and non-Abelian field algebras with only an occasional reference to the CCR's.

3.2 FUNCTIONAL CHARACTERIZATION OF CYCLIC REPRESENTATIONS

There are a number of ways to study representations of a family of unitary operators $e^{iW(f)}$. In a separable Hilbert space, such as we assume, every representation is a direct sum of cyclic representations. A cyclic representation of the operators $e^{iW(f)}$ is one for which there exists a vector -- call it $|0\rangle$ -- for which the vectors $|f\rangle \equiv e^{iW(f)}|0\rangle$ form a total set. Such a representation is uniquely determined up to unitary equivalence by the expectation functional

$$E(f) \equiv \langle 0|e^{iW(f)}|0\rangle$$

defined for all $f\in\upsilon$, some suitable space of test functions. At the very least $E(f)$ is ray continuous, i.e. continuous, for each f, in the variable s,

$$E(sf) \equiv \langle 0|e^{isW(f)}|0\rangle,$$

since $e^{isW(f)} \equiv U(s)$ is by assumption a unitary one-parameter group. However, for simplicity we shall generally assume that $E(f)$ is continuous in some suitable topology for the space υ. In most cases this is a very mild assumption.

The characterization of the representation by $E(f)$ is plausible since it follows (from the group property) that

$$\langle f''|e^{iW(f)}|f'\rangle = E(f''^{-1}\cdot f\cdot f')$$

determines a total set of matrix elements of the operators $e^{iW(f)}$. Indeed, one of the classic realizations of the representation space makes a rather direct use of the functional E(f). In this realization the Hilbert space elements are functions

$$\psi(f) \equiv \langle f|\psi\rangle \equiv \Sigma\, c_i \langle f|f_i\rangle$$

$$= \Sigma\, c_i\, E(f^{-1}\cdot f_i)$$

defined over υ. The norm of such an element is defined by

$$||\psi(\cdot)||^2 \equiv \Sigma\, c_i c_j^* E(f_j^{-1}\cdot f_i) \geq 0, \qquad (3\text{-}1)$$

which is non-negative in virture of E being a positive - definite functional. Infinite sums are admitted provided the partial sums form a Cauchy sequence in the norm as usual. The group composition law for $U[g] \equiv e^{iW(g)}$ is realized according to

$$(U[g]\psi)(f) \equiv \langle f|U[g]|\psi\rangle \equiv \langle g^{-1}\cdot f|\psi\rangle = \psi(g^{-1}\cdot f)$$

All other realizations of $e^{iW(f)}$ which have the same expectation functional are unitarily equivalent to the one just outlined.

Abelian Algebras

In spite of the equivalence of other representations it is often very convenient to have other realizations for practical applications. In general these are not easy to come by. If one deals only with an Abelian family of fields W(f) -- say just $\varphi(f)$ -- then it is possible to diagonalize that field and work in a representation analogous to the Schrödinger representation. In this case we have a relation like

$$E(f) = \langle 0|e^{i\varphi(f)}|0\rangle = \int e^{i(\Lambda,f)} d\mu(\Lambda).$$

Where Λ represents a generalized function and μ represents a type of measure on such functions. We quote without proof two theorems pertaining to such a situation. Representation Theorem 1: In order for E(f), where $f\in\upsilon$, a linear topological space, to be the Fourier transform of a

positive normalized cyclinder set measure* μ on υ' (the dual space to υ) it is necessary and sufficient that E(f) be a positive-definite functional, (sequentially) continuous and fulfill E(0)=1.

Representation Theorem 2: Any continuous, positive-definite functional E(f) on a nuclear space** υ such that E(0)=1 is the Fourier transform of a (countably additive) normalized measure on υ'.

Given a representation of E(f) as a Fourier transform in accord with either of the two preceding theorems, a natural functional Hilbert space $\mathfrak{L}^2$ emerges. The space $\mathfrak{L}^2$ is composed of (measurable) functionals $\psi(\Lambda)$ such that

$$||\psi||^2_{\mathfrak{L}^2} \equiv \int |\psi(\Lambda)|^2 d\mu(\Lambda) < \infty.$$

The cyclic vector $|0\rangle$ is represented by the functional "one", and the action of the operator $e^{i\varphi(f)}$ is given by multiplication by the functional $e^{i(\Lambda,f)}$. Other operators--for example, the conjugate momentum if it exists--would involve a functional differential operation on the functionals $\psi(\Lambda)\in\mathfrak{L}^2(\upsilon',\mu)$. It is in such a framework that

---

*A cyclinder set measure μ may be regarded as a sequence of compatible measures μ_N on R^N such that

$$\int e^{i(\Lambda,f)} d\mu(\Lambda) \equiv \lim_{N\to\infty} \int e^{i\Sigma_1^N \lambda_n (h_n,f)} d\mu_N(\{\lambda_n\})$$

where Λ restricted to R^N equals $\Sigma_1^N \lambda_n h_n(x)$.

**Let $\mathfrak{H}_r$, r=1,2,..., be a family of Hilbert spaces with elements $f\in\mathfrak{H}_r$ that may be identified such that (i) $\mathfrak{H}_{r+1}\subset\mathfrak{H}_r$, (ii) $||f||_r \le ||f||_{r+1}$ where $||f||_r$ is the norm appropriate to $\mathfrak{H}_r$ and (iii) $\mathfrak{H}_s = T_s^r \mathfrak{H}_r$, where for any r there is an s>r such that T_s^r is nuclear, i.e., $\mathrm{Tr}(T_s^{r\dagger} T_s^r)^{\frac{1}{2}} < \infty$. A nuclear space is then basically characterized as the closure of the countable Hilbert space $\bigcap_{r=1}^{\infty} \mathfrak{H}_r$ in the metric $d(f)=\Sigma_1^\infty 2^{-r} ||f||_r/(1+||f||_r)$. For example, let $\mathfrak{H}_1=\ell^2$, $f=\{z_n\}$, and define $||f||_r^2 \equiv \Sigma n^r |z_n|^2$. Then $f\in\upsilon$ if and only if $n^r z_n \to 0$ as $n\to\infty$ for all r. The sequence $f^{(k)}\to 0$ (in d), i.e., $d(f^{(k)})\to 0$, provided $||f^{(k)}||_r \to 0$ for all r, namely, provided $\sup_{(n)} n^r z_n^{(k)} \equiv A_r^{(k)} \to 0$ for all r.

one would seek to make precise the realization of the conjugate operator and the Hamiltonian heuristically presented in Chapter 1 in Eq. (1-3), but this is not our concern at this point.

3.3 TRANSLATION, CLUSTER AND SCALE INVARIANCE CONDITIONS

Translation Invariance

In an effort to limit the vast number of possible representations, let us impose several desirable physical properties. Let us assume that we deal with a translation invariant dynamical problem, and, correspondingly, a translationally invariant state $|0\rangle$. If there was a unitary family of translation operators $U(\underline{a})$ with the properties

$$U(\underline{a})^{-1}\, \varphi(\underline{x})\, U(\underline{a}) \equiv \varphi(\underline{x}-\underline{a})$$

$$U(\underline{a})\,|0\rangle = |0\rangle,$$

then it would follow, for all $f\in\upsilon$ and all $\underline{a}\in R^s$, that

$$E(f) = E(f_a),$$

where $f_a(\underline{x}) \equiv f(\underline{x}+\underline{a})$. Conversely, if $E(f) = E(f_a)$ it follows that there exists a unitary set of operators $U(\underline{a})$ with the stated properties. This already follows from the functional realization introduced earlier in which one can define

$$(U(\underline{a})\Psi)(f) = \Psi(f_a).$$

The unitarity of this set of transformations, and the existence of an invariant cyclic vector all follow from the invariance and continuity of the expectation functional $E(f)$.

Cluster Decomposition

Not only should $|0\rangle$ be translationally invariant but in many problems it should be the unique invariant state. Physically, this uniqueness reflects the expected statistical independence of experiments carried out in two remote regions of space. Mathematically, this feature takes the form of a factorization,

$$\langle f''|U(\underline{a})|f'\rangle \to \langle f''|0\rangle\langle 0|f'\rangle,$$

as $|\underline{a}|\to\infty$. Consequently, we impose on invariant expectation functionals the requirement that

$$\lim_{|\underline{a}|\to\infty} E(f''^{-1}\cdot f'_a) = E(f''^{-1})\,E(f').$$

Since this is, by hypothesis, a total set of matrix elements of a uniformly bounded sequence of operators the Silver Lemma ensures that

$$\text{w-}\lim_{|\underline{a}|\to\infty} U(\underline{a}) = |0\rangle\langle 0|.$$

This behavior holds independent of the direction in which spatial infinity is approached.

In general, invariance and clustering properties severely limit the class of interesting representations. It is useful to note an alternate characterization of these functional restrictions. Suppose we introduce

$$E(f) \equiv e^{-L(f)},$$

where $L(0) = 0$ and $L(sf)$ is finite for any $f\in\upsilon$ at least whenever s is in some neighborhood of the origin. More generally, if $f^{(k)}\to 0$ in υ and $E(f)$ is continuous in υ as we assume, then it follows that there exists a K such that $E(f^{(k)})\neq 0$, for all $k\geq K$, and consequently $L(f^{(k)})$ is well defined for all $k\geq K$. Actually, in many cases of interest $L(f)$ is well defined for all $f\in\upsilon$.

The invariance condition $E(f) = E(f_a)$ evidently requires that $L(f) = L(f_a)$. The cluster property translates into the requirement

$$\lim_{|\underline{a}|\to\infty} L(f''^{-1}\cdot f'_a) = L(f''^{-1}) + L(f').$$

It is perhaps worth noting the relationship of $L(f)$ to the truncated vacuum expectation values of the field operator. These may be introduced through a generating functional which is defined by

$$<0|e^{i\varphi(f)}|0> \equiv e^{<0|\{e^{i\varphi(f)}-1\}|0>^T}.$$

Thus, it follows that

$$-L(f) = <0|e^{i\varphi(f)}|0>^T - 1 = \sum_{n=1}^{\infty} (n!)^{-1} i^n <0|\varphi^n(f)|0>^T$$

assuming a power series expansion for purposes of illustration.

Scale Invariance

Another transformation which is occasionally introduced is the scale (or dilatation) transformation. Such a transformation makes sense for a free massless field, and it is often assumed to hold for an interacting theory in which only dimensionless parameters appear. In such cases we assume that there exists a family of unitary operators $V(S)$, $S>0$, such that $V(S)V(S')=V(SS')$, with the properties

$$V^{-1}(S)\varphi(\underline{x},t)\ V(S) = S^d\ \varphi(S\underline{x},St),$$

$$V(S)|0> = |0>,$$

where d is called the scale dimension of the field. For a canonical theory, by which we mean that

$$[\varphi(\underline{x},t),\ \dot{\varphi}(\underline{y},t)] = i\ \delta(\underline{x}-\underline{y}),$$

the scale dimension has its canonical value

$$d_c = \tfrac{1}{2}(s-1)$$

where s is the number of space dimensions. However, other values of d arise as well.

If we set $t=0$, it follows that

$$V^{-1}(S)\ \varphi(f)\ V(S) = \varphi(f_{(S)})$$

where

$$f_{(S)}(\underline{x}) = S^{d-s}\ f(S^{-1}\underline{x}).$$

Consequently invariance of the expectation functional takes the form

$$E(f) = E(f_{(S)}). \tag{3-2}$$

for all $S>0$.

Unitary Equivalence of Representations

A criterion for equivalence or inequivalence of two representations can often be established with the aid of the expectational functional. Let us assume that invariance and cluster properties hold under the translation group $U(\underline{a})$. When such is the case it follows that

$$e^{iW(f_a)} = U(\underline{a})^{-1}\, e^{iW(f)}\, U(\underline{a})$$

weakly converges as $|\underline{a}|\to\infty$ to a multiple of unity as may be seen from the total set of matrix elements

$$<f''|e^{iW(f_a)}|f'> = E(f''^{-1}\cdot f_a\cdot f')$$

and an application of the Silver Lemma. Assume, for example, that f,f', and f'' all have compact support so that for large enough $|\underline{a}|$ the factors commute and

$$E(f''^{-1}\cdot f_a\cdot f') = E(f''^{-1}\cdot f'\cdot f_a)\to E(f''^{-1}\cdot f')\ E(f).$$

From this relation we even observe that as $|\underline{a}|\to\infty$

$$\text{w-lim}\ e^{iW(f_a)} = <0|e^{iW(f)}|0> = E(f).$$

This property holds for each $f\in\upsilon$, which means that for each f, $E(f)$ is a "tag" for the representation.

If two such representations are unitarily equivalent such that

$$V^{-1}\, e^{iW(f)}\, V = e^{iW'(f)}$$

holds for all $f\in\upsilon$, then it follows that as $|\underline{a}|\to\infty$

$$E'(f) = \text{w-lim}\, e^{iW'(f_a)} = \text{w-lim}\, V^{-1} e^{iW(f_a)} V$$

$$= V^{-1} E(f)\, V = E(f).$$

But if $E'(f) = E(f)$ for all f we are already assured that the representations are equivalent. On the other hand, if $E'(f) \neq E(f)$ for any f, then no V exists and the representations are inequivalent.

In summary, therefore, distinct expectation functionals label inequivalent representations. Note this argument makes no statement regarding the reducibility or irreducibility of the representation; the only ingredients are invariance and cluster decomposition.

3.4 INFINITELY DIVISIBLE REPRESENTATIONS OF FIELD ALGEBRAS

In this section we wish to construct expectation functionals which incorporate the various features we discussed above. We do not attempt to determine all such representations but rather those which are analogues of infinitely divisible distributions. This is most conveniently carried out in the framework of exponential Hilbert spaces which is quite directly based on the Fock space methods introduced in Chapter 2.

Exponential Hilbert Space

Consider an abstract Hilbert space $\mathfrak{h}$ and the Fock space $\mathfrak{H}$ based on it,

$$\mathfrak{H} = \bigoplus_{n=0}^{\infty} (\mathfrak{h}^{\otimes n})_s,$$

where the symmetric subspace is implied. For each $\psi \in \mathfrak{h}$ we associate the unit vector $|\psi\rangle \in \mathfrak{H}$ based on the definition

$$|\psi\rangle \equiv N \bigoplus_{n=0}^{\infty} (n!)^{-\frac{1}{2}} (\psi^{\otimes n})_s$$

where $N = \exp[-\frac{1}{2}(\psi,\psi)]$ is a normalization factor such that $\| |\psi\rangle \| = 1$. To each $\varphi \in \mathfrak{h}$ we associate an annihilation operator $A(\varphi)$, antilinear in φ, with the property that

$$A(\varphi)\,|\psi\rangle = (\varphi,\psi)\,|\psi\rangle \tag{3-3}$$

Observe that the states $|\psi\rangle$ are just the coherent states again and (3-3) expresses that they are eigenstates for $A(\varphi)$. We have the relations

$$[A(\varphi),\ A(\psi)^\dagger] = (\varphi,\psi),$$

and

$$|\psi\rangle = N\, e^{A(\psi)^\dagger}|0\rangle = U(\psi)\,|0\rangle,$$

where

$$U(\psi) = \exp\,[A(\psi)^\dagger - A(\psi)].$$

A bilinear operator [compare Eq. (2-7)]

$$\hat{W} \equiv (A, wA)$$

is characterized by its coherent state matrix elements,

$$\langle\psi|\hat{W}|\psi'\rangle = (\psi, w\,\psi')\langle\psi|\psi'\rangle,$$

and it follows that

$$e^{i\hat{W}} = \bigoplus_{n=0}^{\infty} (e^{iw})^{\otimes n}. \tag{3-4}$$

If w is self adjoint on $\mathfrak{h}$, $\hat{W}$ is self adjoint on $\mathfrak{H}$.

Now consider that $w=w(f)$, $\hat{W}=\hat{W}(f)$ are smeared field operators and that the $w(f)$ correspond to some current algebra. That is, we imagine that

$$[w(f),\ w(g)] = i\, w([f,g])$$

in some unspecified but fairly self-evident notation. Then it follows that

$$[\hat{W}(f),\ \hat{W}(g)] = (A,\ [w(f),\ w(g)]A)$$
$$= i\,\hat{W}\,([f,g]),$$

namely, that $\hat{W}(f)$ also satisfies the same algebra. As a consequence $e^{iw(f)}$ and $e^{i\hat{W}(f)}$ constitute group

representations, and the latter is highly reducible even if the former is irreducible. The group representation property is preserved under a fixed unitary transformation such that

$$e^{iW(f)} \equiv U(\xi)^{-1}\, e^{i\hat{W}(f)} U(\xi),$$

which, if $w(f)\xi \in \mathfrak{h}$, means that

$$W(f) = (B, w(f)B) = ([A+\xi],\ w(f)[A+\xi]). \qquad (3\text{-}5)$$

We note that

$$\langle\psi|e^{iW(f)}|\psi'\rangle = e^{(\psi+\xi,[e^{iw(f)}-1](\psi'+\xi))}\langle\psi|\psi'\rangle$$

and in particular that

$$E(f) \equiv \langle 0|e^{iW(f)}|0\rangle = e^{(\xi,[e^{iw(f)}-1]\xi)}. \qquad (3\text{-}6)$$

We have already assumed that $e^{iw(f)}$ constitutes a group representation and have observed that $e^{iW(f)}$ constitutes such a representation for any $U(\xi)$, i.e., for any $\xi \in \mathfrak{h}$. From this it follows that $E(f)$ is infinitely divisible, that is for each positive integer m,

$$[E(f)]^{1/m} \equiv E_m(f) = e^{1/m(\xi,[e^{iw(f)}-1]\xi)}$$

is a positive definite functional [in the sense of (3-1)]. This is clear since E_m is obtained from E simply by the change of ξ to $\xi/\sqrt{m}$. On the other hand, the analog of DeFinetti's Theorem implies that all infinitely divisible representations are limits of those for which Eq. (3-6) holds. For, if we assume that $E(f)$ is infinitely divisible, then

$$m\{[E(f)]^{1/m}-1\} = m\{E_m(f)-1\}$$

$$\equiv m(\theta_m,[e^{iw(f)}-1]\theta_m),$$

for some representation $w(f)$, where $||\theta_m||=1$, from which it follows that

$$E(f) = \lim e^{m\{[E(f)]^{1/m}-1\}} = \lim e^{m(\theta_m,[e^{iw(f)}-1]\theta_m)}. \qquad (3\text{-}7)$$

With this motivation let us study the limits of expressions of the form (3-6). We divide our study into two parts.

Abelian Field Algebras

Suppose first that we have an Abelian family of operators -- say just the single field operator $\varphi(f)$. In that case we may diagonalize the field in accord with the theorems mentioned earlier, but this time we diagonalize in the exponent. That is, we express (3-6) in the form

$$E(f) = \langle 0|e^{i\varphi(f)}|0\rangle = e^{\int[e^{i(\Lambda,f)}-1]d\sigma(\Lambda)}.$$

Sequences of such expectation functionals converge to an expectation functional provided that

$$\lim E_n(f) = \lim e^{\int[e^{i(\Lambda,f)}-1]d\sigma_n(\Lambda)}$$

converges to a continuous functional which is the analogue of the Continuity Theorem. Tortrat has given an analogue of the Lévy canonical formula (2-2) which reads

$$\ln E(f) = i(a,f) - (f,bf) + \int[e^{i(\Lambda,f)}-1-\frac{i(\Lambda,f)}{1+||\Lambda||^2}]\, d\sigma(\Lambda) \qquad (3\text{-}8)$$

where a and b are real, $b\geq 0$ and $||\Lambda||^2$ can be appropriately defined. For symmetric functionals, where $E(-f)=E(f)$, the general form reads

$$E(f) = e^{-(f,bf) - \int[1-\cos(\Lambda,f)]d\sigma(\Lambda)}. \qquad (3\text{-}9)$$

In these expressions, $\int d\sigma(\Lambda)$ need not be finite, but rather

$$\int ||\Lambda||^2/(1+||\Lambda||^2)\, d\sigma(\Lambda) < \infty$$

excluding, as always, the point $\Lambda\equiv 0$.

Expressions of the type (3-9) can be readily constrained to have invariance and clustering, the latter being rather naturally and easily imposed on the measure σ.

For example,

$$E(f) = e^{-(f,bf)} - \int d\underline{z} \int [1-\cos(\Lambda, f_z)]\, d\sigma_o(\Lambda)$$

fulfills the desired properties, where $b=b(\underline{x}-\underline{y})$, $f_z=f(\underline{x}+\underline{z})$ and σ_o is concentrated on generalized functions with support "near $\underline{x}=o$."

It should be observed in the preceding construction that the sequence of field operators $\varphi_n(f)$ do not necessarily converge to the operator $\varphi(f)$ in the "same" Hilbert space $\mathfrak{H}$. For this to be true we could invoke the analogue of the Strong Operator Convergence theorem of Chapter 2. That is, besides the convergence of $E_n(f)$ to a continuous functional, we would require the strong convergence of

$$\{e^{iw(f)}-1\}\, \xi_n \to \varphi[f] \in \mathfrak{H}, \tag{3-10}$$

which excludes, for example, the appearance of a Gaussian part in $E(f)$. On the other hand, if a Gaussian term is present [or condition (3-10) otherwise fails] we would say that $\varphi_n(f) \to \varphi(f)$ in the sense of expectation functionals.

Non-Abelian Field Algebras

For a non-Abelian group representation many of the same arguments apply, except that we cannot invoke the Fourier transform representation theorem. Consider the sequence of functionals

$$E_n(f) = \langle 0 | e^{iW_n(f)} | 0 \rangle = e^{(\xi_n, [e^{iw(f)}-1]\xi_n)}$$

which are continuous and positive definite in the sense of (3-1). We have argued in Eq. (3-7) that every infinitely divisible group representation can be expressed as the limit of such functionals. Conversely, for every sequence ξ_n such that

$$E(f) = \lim e^{(\xi_n, [e^{iw(f)}-1]\xi_n)}$$

converges to a continuous functional, we generate a valid group representation, which incidently happens to be infinitely divisible. (This latter follows because $[E_n(f)]^{1/m}$

is positive definite for each n and converges to the continuous function $[E(f)]^{1/m}$.)

These representations may be made to exhibit manifest invariance and cluster properties as well. For this purpose, let

$$\mathfrak{h} = \int^{\oplus} \mathfrak{h}_z \, d\underset{\sim}{z},$$

$$w(f) = \int^{\oplus} w_o(f_z) \, d\underset{\sim}{z}$$

and suppose that

$$\xi_n = \int^{\oplus} \xi_{nz} \, d\underset{\sim}{z},$$

where, for example,

$$\xi_{nz} \equiv u_n(\underset{\sim}{z}) \, \tilde{\xi}$$

and the function

$$u_n(\underset{\sim}{z}) \equiv 1 \; ; \quad |\underset{\sim}{z}| \leq n,$$

$$\equiv 0 \; ; \quad |z| \geq n+1$$

and falls smoothly in between. Here $\tilde{\xi} \in \mathfrak{h}_z$, which we assume to be the "same" vector for all $\underset{\sim}{z}$. With f constrained to have compact support it is evident that

$$\begin{aligned} E(f) &= \lim e^{(\xi_n, [e^{iw(f)} - 1]\xi_n)} \\ &= \lim e^{\int d\underset{\sim}{z} (\xi_{nz}, [e^{iw_o(f_z)} - 1]\xi_{nz})_z} \\ &= e^{\int d\underset{\sim}{z} \, (\tilde{\xi}, [e^{iw_o(f_z)} - 1]\tilde{\xi})_z} \end{aligned} \tag{3-11}$$

exists since for large enough n the sequence becomes constant. The functional E(f) evidently is invariant, and if $w_o(\underset{\sim}{x})$ has support near the origin, E(f) will also have the cluster property. Limits of such functionals (e.g., a

non-Cauchy sequence $\tilde{\xi}_n \in \mathfrak{h}_Z$) lead to new, infinitely divisible representations which enjoy both invariance and cluster decomposition.

Generally, the limit $E_n(f)$ will lead to convergence only of expectation functionals, and not to a convergence of the operators in the original Hilbert space $\mathfrak{H}$. This may lead to the representation "growing" unwanted subrepresentations. An example of this behavior might be the appearance of a Gaussian part in $E(f)$. One way to suppress such terms is to insist on the Strong Operator Convergence condition,

$$\{e^{iw(f)}-1\}\ \xi_n \to \varphi[f] \in \mathfrak{h},$$

where $\varphi[f]$ is continuous in f. This condition will at the same time establish the convergence of the operators $W_n(f) \to W(f)$ within one and the same Hilbert space $\mathfrak{H}$. It is worth emphasizing that although the algebra of operators $W_n(f)$ constructed in the above fashion is reducible for each n, it can happen that the limit operators $W(f)$ are irreducible [and consequently would not possess an expansion equivalent to (3-4)].

An example will serve to illustrate some of the feature discussed above.

Example: Consider the field algebra characterized by

$$[\pi(\underline{x}), \pi(\underline{y})] = i\,\delta(\underline{x}-\underline{y})\,\pi(\underline{y}), \tag{3-12}$$

which is based on the two-parameter affine group with an elementary Lie algebra

$$[B,P] = iP.$$

An irreducible representation of the affine group is given on $L^2(0,\infty)$ by the prescription

$$(e^{-isP}e^{irB}\varphi)(k) = e^{-\frac{1}{2}r}e^{-isk}\varphi(e^{-r}k)$$

for all $\varphi(k) \in L^2(0,\infty)$. We identify this representation as that denoted by w_o in Eq. (3-11), leading to a field representation given by

$$E(f,g) = \langle 0|e^{-i\pi(g)}\, e^{i\kappa(f)}|0\rangle$$

$$= \exp\{-\int d\underline{z}\int_0^\infty \xi^*(k)[\xi(k)-e^{-\frac{1}{2}f(\underline{z})-ig(\underline{z})k}\xi(e^{-f(\underline{z})}k)]dk\}\ . \tag{3-13}$$

For each $\xi(k) \in L^2(0,\infty)$ this functional characterizes a representation of (3-12), and for each unequal ray the representations are inequivalent and reducible.

Consider a non-Cauchy sequence $\xi_n(k)$ defining a sequence of functionals $E_n(f,g)$ which converges to a new functional $E(f,g)$. For example, let

$$\xi_n(k) \equiv k^{-\frac{1}{2}}\exp\left[-\left(\frac{1}{k^{1/n}}-1\right)^2\right]\hat{e}(k)$$

where $\hat{e}(k)$ is C^∞ and square integrable. Then it follows that

$$\varphi_n(r,s,k) \equiv e^{-\frac{1}{2}r-isk}\xi_n(e^{-r}k)-\xi_n(k)$$

converges strongly to

$$\hat{\varphi}(r,s,k) = k^{-\frac{1}{2}}[e^{-isk}\,\hat{e}(e^{-r}k) - \hat{e}(k)],$$

while

$$J_n(r,s) = \int_0^\infty \xi_n^*(k)[e^{-\frac{1}{2}r-isk}\xi_n(e^{-r}k)-\xi_n(k)]dk$$

converges to

$$\hat{J}(r,s) = \tfrac{1}{2}\int_0^\infty [\hat{e}^*(e^r k)[e^{-isk}\hat{e}(k)-\hat{e}(e^r k)]$$

$$+\hat{e}^*(k)[e^{-isk}\hat{e}(e^{-r}k) - \hat{e}(k)]]\,\frac{dk}{k}\ .$$

This limiting expression leads to an expectation functional

$$\hat{E}(f,g) = \exp\{\int d\underline{z}\ \hat{J}[f(\underline{z}),g(\underline{z})]\}$$

characterizing a representation of (3-12) inequivalent to

those previously discussed and one which is irreducible in contrast to the preceding examples.

As an example of the growth of unwanted subrepresentations, imagine that

$$\xi_n(k) = \xi(k) + \theta_n(k)$$

where

$$\theta_n(k) = n^{\frac{1}{2}}\theta(nk).$$

It follows that w-lim $\theta_n(k) = 0$, but $\theta_n(k)$ does not converge strongly. As a consequence we can anticipate only convergence of the expectation functional, and it is a simple matter to see that

$$E_n(f,g) \to E(f,g) = e^{\int d\underline{z}\{J[f(\underline{z}),g(\underline{z})] + Y[f(\underline{z}),g(\underline{z})]\}}$$

where J is implicitly given in (3-13), while

$$Y[r,s] = \int_0^\infty \theta^*(k)[e^{-\frac{1}{2}r}\theta(e^{-r}k) - \theta(k)]dk,$$

which is independent of s [i.e., of $g(\underline{z})$]. Thus the limiting representation of $\varkappa(\underline{x})$ and $\pi(\underline{y})$ is such that, in an obvious notation,

$$\varkappa(\underline{x}) = \varkappa_J(\underline{x}) \oplus \varkappa_Y(\underline{x})$$

$$\pi(\underline{x}) = \pi_J(\underline{x}) \oplus 0,$$

namely in the "second representation space" $\pi(\underline{x})$ is represented by zero [a perfectly acceptable solution to (3-12)!].

4. ULTRALOCAL FIELD THEORIES

4.1 HEURISTIC, CLASSICAL INTRODUCTION

Covariant Motivation

Ultralocal scalar field theories are formally obtained from covariant scalar theories by suppressing the spacial gradient term in the Hamiltonian. They are characterized, therefore, by a classical Hamiltonian of the form

$$H = \int \{\tfrac{1}{2}[\pi_{c1}^2(\underset{\sim}{x}) + m_0^2\,\varphi_{c1}^2(\underset{\sim}{x})] + V[\varphi_{c1}(\underset{\sim}{x})]\}d\underset{\sim}{x} \qquad (4\text{-}1)$$

Such models are distinguished by the property that distinct spacial points characterize statistically independent fields for all times. Stated otherwise, the light cone of the covariant theories has been collapsed to a "vertical line" passing through the space point in question. Nevertheless, the topology of the original space is retained, and indeed we still insist on a space translation generator classically given by

$$\underset{\sim}{P} = \int \pi_{c1}(\underset{\sim}{x})\ \underset{\sim}{\nabla}\ \varphi_{c1}(\underset{\sim}{x})d\underset{\sim}{x}.$$

The ultralocal theories are expected to have a trivial scattering matrix (S=1), and are not, by themselves, expected to provide physically significant predictions. These models are examples of systems with an infinite number of degrees of freedom from which we hope to learn more about such systems. Viewed in the conventional fashion these models are nonrenormalizable, and their study by standard perturbation techniques is frought with ambiguities. Techniques other than the conventional ones are required to solve these models, and such techniques will be provided.

Even before determining the quantum solution, we may anticipate that it would have at least two interesting relations to the corresponding solution of a covariant theory. On the one hand, the ultralocal models and their corresponding solutions should be the limit of covariant models as the coefficient of the term $\frac{1}{2}[\underset{\sim}{\nabla}\varphi_{c1}(\underset{\sim}{x})]^2$ vanished from the Hamiltonian. On the other hand, it might be hoped that the spacial gradient term could be restored by a perturbation analysis. Some comments on the latter idea are presented in Chapter 6.

Alternative Interpretation

Although the ultralocal models are motivated by covariant models they may be interpreted in another way as well. There is an evident, classical one degree of freedom problem underlying (4-1) described by the Hamiltonian

$$H_1(p,q) = \tfrac{1}{2}(p^2+m_0^2q^2) + V[q];$$

and the Hamiltonian (4-1) itself may be considered as a continuous embellishment of H_1 in the sense of Chapter 1. It is anticipated, therefore, that the ultralocal quantum theory should be characterized, in some sense, by an underlying one degree of freedom problem, and we shall find this to be the case, although the underlying problem is necessarily different from the "obvious" one.

Our subsequent discussion of these models focuses on a general understanding of the solution with a special emphasis on certain specific aspects. Several other discussions of these models have been given elsewhere and should be consulted by the interested reader in order to round out an appreciation of the present understanding of ultralocal scalar fields.

4.2 OPERATOR SOLUTION FOR ULTRALOCAL MODELS

Our assumptions regarding the quantum solution are minimal and quite plausible. We expect, for each potential V of a large class, that their exists a Hamiltonian operator $\mathcal{H} \geq 0$, and a unique ground state $|0\rangle$ for which $\mathcal{H}|0\rangle = 0$. The state $|0\rangle$ is also the unique translationally invariant state and satisfies $\underset{\sim}{\mathcal{P}}|0\rangle = 0$, and $[\underset{\sim}{\mathcal{P}},\mathcal{H}] = 0$, where $\underset{\sim}{\mathcal{P}}$ is the space-translation generator. We suppose initially that there exists a self-adjoint field operator

$$\varphi(f) = \int f(\underset{\sim}{x})\varphi(\underset{\sim}{x})d\underset{\sim}{x}$$

defined for all $f \in C_0^\infty$. (Subsequently, we shall generalize this assumption to a space-time smearing.) For simplicity we assume the potential V is symmetric, i.e.,

$$V[-\varphi_{c1}(\underset{\sim}{x})] = V[\varphi_{c1}(\underset{\sim}{x})],$$

and bounded below as in usual models.

Determination of Expectation Functional

We focus initially on the expectation functional

$$E(f) = \langle 0|e^{i\varphi(f)}|0\rangle$$

defined, let us suppose, for all $f(\underset{\sim}{x}) \in C_0^\infty$, together with some continuity properties, the least of which is ray

continuity. The symmetry of the potential implies that $E(-f) = E(f)$; while the principal symmetry fact of the ultralocal form of the dynamics -- the independence for-all-times of the field at distinct spacial points -- implies that

$$E(f) = e^{-\int d\underset{\sim}{x}\, L[f(\underset{\sim}{x})]}.$$

From this form alone we immediately deduce that $E(f)$ is infinitely divisible. Since $E(f)$ is a positive-definite functional for all $f(\underset{\sim}{x}) \in C_0^\infty$, it follows that

$$[E(f)]^{1/m} = e^{-\frac{1}{m}\int d\underset{\sim}{x}\, L[f(\underset{\sim}{x})]}$$

$$= e^{-\int d\underset{\sim}{y}\, L[f(m^{1/s}\underset{\sim}{y})]} = E(f_{(m)}),$$

where $f_{(m)}(\underset{\sim}{x}) \equiv f(m^{1/s}\underset{\sim}{x}) \in C_0^\infty$, and thus $E(f)^{1/m}$ is positive definite for all m. Consequently, $E(f)$ necessarily has the form given in Eq. (3-9), which implies that $E(f)$ never vanishes. It follows in our case that $L[f(\underset{\sim}{x})]$ is defined for all arguments.

To prove that $L[s]$ is necessarily continuous on the basis of our minimal assumptions we can proceed as follows. Let $u(x) \in C_0^\infty$, $x \in R$, satisfy $u(x) \equiv 1$, $|x| \leq A$; $u(x) \equiv 0$, $|x| > A+1$. Define

$$u_a(x) \equiv u(x), \quad x \leq 0$$

$$\equiv u(x+a), \quad x \geq 0$$

for all a such that $0 < a < A$; clearly $u_a(x) \in C_0^\infty$ for such a. If $s=1$ let $f_1(x)=u(x)$, $f_2(x)=u_a(x)$. If $s \geq 2$ let

$$f_1(\underset{\sim}{x}) = u(|\underset{\sim}{x}|^s),$$

$$f_2(\underset{\sim}{x}) = u_a(|\underset{\sim}{x}|^s).$$

The ray continuity (weak continuity of each unitary one parameter subgroup) ensures that

$$C_{f_i}(s) = E(sf_i) = \langle 0|e^{\text{is } \varphi(f_i)}|0\rangle,$$

$i=1,2$, are both continuous functions of s, while the special form of f_1 and f_2 imply that

$$\frac{C_{f_1}(s)}{C_{f_2}(s)} = e^{-\int d\underset{\sim}{x}\{L[sf_1(\underset{\sim}{x})] - L[sf_2(\underset{\sim}{x})]\}} = e^{-\Delta L[s]},$$

where $0 < \Delta < \infty$. It follows, therefore, that $L[s]$ is continuous in s, as was to be shown.

The positivity condition

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j^* \, E(f_i - f_j) \geq 0$$

carries over by continuity to multiples of characteristic functions,

$$f_i(\underset{\sim}{x}) = s_i \, \chi_\Delta(\underset{\sim}{x})$$

which thus implies, for all $\Delta > 0$, that

$$\sum_{i,j=1}^{N} \alpha_i \alpha_j^* \, e^{-\Delta L[s_i - s_j]} \geq 0.$$

Consequently, L has the form implicit in (2-3), namely

$$L[s] = bs^2 + \int_{|\lambda|>0} [1-\cos(\lambda s)] d\sigma(\lambda).$$

We interpret the first (Gaussian) term as describing the free, ultralocal Fock theory, but this solution is not of interest to us at this point. For the second (non-Gaussian) term we adopt an absolutely continuous measure, $d\sigma(\lambda) \equiv c^2(\lambda) d\lambda$, leading to the relation

$$E(f) = e^{-\int d\underset{\sim}{x} \int [1-\cos[\lambda f(\underset{\sim}{x})]] c^2(\lambda) d\lambda}. \qquad (4\text{-}2)$$

The plausibility of this choice for $L[s]$ will become clear in our subsequent discussions. The integrability condition on $c(\lambda)$ for this expression to exist is clearly

$$\int \lambda^2/(1+\lambda^2) \; c^2(\lambda) d\lambda < \infty,$$

but we may (and in fact will) have $\int c^2(\lambda)d\lambda=\infty$. We refer to $c(\lambda)$ as the model function.

Let us anticipate the remaining properties of the model function so the reader may be allerted to the various conditions as they arise. Clearly we can choose $c(\lambda)$ as real and symmetric, $c(-\lambda)=c(\lambda)$. We shall also find that $c(\lambda)$ is nonvanishing and twice differentiable except at the origin. Finally we will require that

$$\int c^2(\lambda)d\lambda = \infty$$

suggesting the form

$$c(\lambda) \equiv \frac{e^{-y(\lambda)}}{|\lambda|^\gamma}$$

where γ is called the singularity parameter and fulfills $\frac{1}{2} \leq \gamma < \frac{3}{2}$. (When we generalize to space-time smearing we only require that $\frac{1}{2} \leq \gamma$.)

Field Operator

An operator realization for $\varphi(f)$ can be given in terms of the relations established in Chapters 2 and 3. In the notation of Section 2.5, with $y=(\underline{x},\lambda)$, we assert that

$$\varphi(\underline{x}) = \int A^\dagger(\underline{x},\lambda)\ \lambda\ A(\underline{x},\lambda)d\lambda$$

$$+ \int c(\lambda)\ \lambda\ A(\underline{x},\lambda)d\lambda$$

$$+ \int A^\dagger(\underline{x},\lambda)\ \lambda\ c(\lambda)d\lambda .$$

If we set

$$B(\underline{x},\lambda) = A(\underline{x},\lambda) + c(\lambda)$$

and observe that for symmetric $c(\lambda)$ the principal value integral $\int\lambda c^2(\lambda)d\lambda=0$, we may also put

$$\varphi(\underline{x}) = \int B^\dagger(\underline{x},\lambda)\ \lambda\ B(\underline{x},\lambda)d\lambda . \qquad (4\text{-}3)$$

[It should be remembered that any expression of this type

should be interpreted in terms of a limit in the manner of Eq. (2-8).]

It is clear from our earlier discussion that with the above choice for $\varphi(\underline{x})$ the expectation functional E(f) coincides with (4-2). Now $\varphi(\underline{x})$, being a construct of A and $A^\dagger$, operates within the usual Hilbert space $\mathfrak{H}$ build by repeated action of the creation operators $A^\dagger$ on the (unique) vacuum $|0>$. To demonstrate cyclicity of the representation for $\varphi(\underline{x})$ -- and thereby establish the uniqueness of that representation up to unitary equivalence -- we need only establish that the closed linear span of $e^{i\varphi(f)}|0>$ coincides with the Hilbert space $\mathfrak{H}$.

Functions of the Field

To facilitate the discussion of the cyclicity of the representation for $\varphi(\underline{x})$ let us investigate some local operators which can be constructed from the field operator. From (4-3) we easily see that

$$\varphi(\underline{x})\varphi(\underline{y}) = \delta(\underline{x}-\underline{y}) \int B^\dagger(\underline{x},\lambda)\ \lambda^2 B(\underline{x},\lambda)\,d\lambda + :\varphi(\underline{x})\varphi(\underline{y}):,$$

where : : denotes normal order with respect to $B^\dagger$ and B (which are inequivalent to a Fock representation). The singular coefficient of the first term (at $\underline{x}=\underline{y}$) compared to the second term permits a test function sequence to pick out the operator

$$\varphi_r^2(\underline{x}) \equiv \int B^\dagger(\underline{x},\lambda)\ \lambda^2 B(\underline{x},\lambda)\,d\lambda \equiv Z\varphi^2(\underline{x})$$

where, formally, $Z^{-1}=\delta(0)$. This is a renormalized operator product (hence the subscript r), where multiplication occurs in "λ-space", under the integral. This procedure may obviously be extended further to yield

$$\varphi_r^p(\underline{x}) \equiv \int B^\dagger(\underline{x},\lambda)\ \lambda^p\ B(\underline{x},\lambda)\,d\lambda \equiv Z^{p-1}\varphi^p(\underline{x})$$

or generally, for a broad class of functions $V[\lambda]$,

$$(V[\varphi(\underline{x})])_r \equiv \int B^\dagger(\underline{x},\lambda)V[\lambda]B(\underline{x},\lambda)\,d\lambda \equiv Z^{-1}V[Z\varphi(\underline{x})].$$

All such local operators are functions in the usual sense of $\varphi(\underline{x})$.

Cyclicity of Representation

Without difficulty we can imagine the construction of an operator

$$G = \int\int B^{\dagger}(\underline{x},\lambda)\ h(\underline{x},\lambda)\ B(\underline{x},\lambda)\ d\underline{x}d\lambda,$$

again a construct of the field, which is well defined for all $h(\underline{x},\lambda)\in C_0^{\infty}$ that also vanish in a neighborhood of $\lambda=0$. The operator e^G, still only a construct of the field, essentially takes the vacuum into a coherent state for

$$\begin{aligned} e^G|0> &= :\ e^{\int\int B^{\dagger}(\underline{x},\lambda)\ \{e^{h(\underline{x},\lambda)}-1\}B(\underline{x},\lambda)d\underline{x}d\lambda}\ :|0> \\ &= e^{g}e^{\int\int A^{\dagger}(\underline{x},\lambda)\{e^{h(\underline{x},\lambda)}-1\}c(\lambda)d\underline{x}d\lambda}|0>, \end{aligned}$$

where

$$g = \int\int c(\lambda)\{e^{h(\underline{x},\lambda)}-1\}c(\lambda)d\underline{x}d\lambda.$$

If $c(\lambda)$ never vanishes, then $\{e^{h(\underline{x},\lambda)}-1\}c(\lambda)$ covers a dense set of $L^2(R^s\times R)$ as h varies in its allowed region. It follows that the span of $e^{i\varphi(f)}|0>$ coincides with $\mathcal{H}$ as was to be proved.

We note in passing, and without proof, that the space translation generator $\underset{\sim}{P}$ is given by

$$\underset{\sim}{P} = \int d\underline{x}\int A^{\dagger}(\underline{x},\lambda)(-i\underset{\sim}{\nabla})\ A(\underline{x},\lambda)d\lambda.$$

Structure of Hamiltonian

The form of the Hamiltonian is decisively determined by the ultralocal form of the dynamics. Heuristically, we expect that

$$\mathcal{H} = \int \mathcal{H}(\underline{x})d\underline{x}$$

where the local operator $\mathcal{H}(\underline{x})$ has the form

$$\mathcal{H}(\underline{x}) = F\{A^{\dagger}(\underline{x},\cdot),\ A(\underline{x},\cdot)\}.$$

The notation in the latter expression means, formally, that

$$\mathcal{H}(\underset{\sim}{x}) = \sum_{n,m=0}^{\infty} \int h_{n,m}(\lambda_1,\ldots\lambda_n;\ \lambda_1',\ldots\lambda_m')$$

$$\times\ A^{\dagger}(\underset{\sim}{x},\lambda_1)\ldots A^{\dagger}(\underset{\sim}{x},\lambda_n)\ A(\underset{\sim}{x},\lambda_1')\ldots A(\underset{\sim}{x},\lambda_m')d^n\lambda d^m\lambda'.$$

However only the first few terms in such a series can actually enter, specifically those of the form

$$\begin{aligned}\mathcal{H}(\underset{\sim}{x}) &= \mathcal{E}_{00}+\int h_{1,0}(\lambda)\ A^{\dagger}(\underset{\sim}{x},\lambda)d\lambda \\ &\quad + \int A(\underset{\sim}{x},\lambda)\ h_{0,1}(\lambda)d\lambda \\ &\quad + \int A^{\dagger}(\underset{\sim}{x},\lambda)\ h_{1,1}(\lambda;\lambda')\ A(\underset{\sim}{x},\lambda')d\lambda d\lambda'.\end{aligned} \tag{4-4}$$

Most higher order terms are <u>forms</u> and not local operators as evident in the case

$$D(\underset{\sim}{x}) \equiv \int A^{\dagger}(\underset{\sim}{x},\lambda_1)\ A^{\dagger}(\underset{\sim}{x},\lambda_2)\ A^{\dagger}(\underset{\sim}{x},\lambda_3)\ h_{3,0}(\lambda_1,\lambda_2,\lambda_3)d^3\lambda.$$

Smearing with a space-dependent test function, $f(\underset{\sim}{x})$, and taking coherent state matrix elements yields

$$\begin{aligned}\langle\psi|D(f)|\psi'\rangle &= \langle\psi|\psi'\rangle \\ &\quad \times \int f(\underset{\sim}{x})d\underset{\sim}{x} \int \psi^*(\underset{\sim}{x},\lambda_1)\psi^*(\underset{\sim}{x},\lambda_2)\psi^*(\underset{\sim}{x},\lambda_3) \\ &\quad \times h_{3,0}(\lambda_1,\lambda_2,\lambda_3)d^3\lambda.\end{aligned}$$

However, this expression is not even continuous in the coherent state bra, for one may choose a sequence $\psi_n(\underset{\sim}{x},\lambda)$ converging strongly to $\psi(\underset{\sim}{x},\lambda)$ for which the right side fails to converge. For example, let

$$\psi_n(\underset{\sim}{x},\lambda) = [(\underset{\sim}{x}-\underset{\sim}{a})^2+n^{-2}]^{s/5}\ f(\underset{\sim}{x})h(\lambda)$$

where $\underset{\sim}{a}$ is a point of support of $f(\underset{\sim}{x})$. Thus $h_{3,0}=0$ and by <u>hermitian symmetry</u> $h_{0,3}=0$ <u>as well</u>. A similar argument applies to the remaining terms.

Only the quadratic term in (4-4) survives the integration over R^s, so that we must have

$$\mathcal{H} = \int d\underline{x} \int A^{\dagger}(\underline{x},\lambda)\, h_{1,1}(\lambda;\lambda')\, A(\underline{x},\lambda')\, d\lambda\, d\lambda'$$

$$\equiv \int d\underline{x} \int A^{\dagger}(\underline{x},\lambda)\ \mathcal{h} A(\underline{x},\lambda)\, d\lambda$$

where $\mathcal{h}$ is a self-adjoint differential operator in the λ variables alone. Observe that $\mathcal{H}|0\rangle=0$, and that $\mathcal{h}\geq 0$ implies that $\mathcal{H}\geq 0$. In order to win uniqueness of the ground state we actually need to have $\mathcal{h}>0$.

Delineation of the Hamiltonian

To further determine $\mathcal{h}$ (and thereby $\mathcal{H}$) we must inject additional physics. We propose that

$$\mathcal{h} \equiv -\tfrac{1}{2}\frac{\partial^2}{\partial\lambda^2} + v(\lambda) \equiv -\tfrac{1}{2}\frac{\partial^2}{\partial\lambda^2} + \tfrac{1}{2}\frac{c''(\lambda)}{c(\lambda)}$$

which requires that $c(\lambda)$ not vanish and that $c(\lambda)$ be twice differentiable almost everywhere. Observe that $\mathcal{h}c(\lambda)=0$ and that it is necessary that $c(\lambda)\notin L^2$ in order that $\mathcal{h}>0$, i.e., for there to be a nondegenerate ground-state of $\mathcal{H}$. Two model functions which differ simply by a scale factor, e.g., $c_1(\lambda) = N\, c(\lambda)$, lead to the same differential operator $\mathcal{h}$ and thus to the same Hamiltonian $\mathcal{H}$.

An alternate form for $\mathcal{H}$ is also useful, which is based on the fact that we can write

$$\mathcal{h} = b^{\dagger} b,$$

where

$$b \equiv \sqrt{\tfrac{1}{2}}\, c(\lambda)\, \frac{\partial}{\partial\lambda}\, c^{-1}(\lambda).$$

With this expression for $\mathcal{h}$ we can recast

$$\mathcal{H} = \int d\underline{x} \int [bA(\underline{x},\lambda)]^{\dagger}\ [bA(\underline{x},\lambda)]\, d\lambda,$$

and since the first two operations of b are "divide by c and differentiate" we can freely add $c(\lambda)$ to the operators A above. Hence, we also can write

$$\mathcal{H} = \int d\underline{x} \int [bB(\underline{x},\lambda)]^{\dagger} \, [bB(\underline{x},\lambda)] d\lambda$$
$$= \int d\underline{x} \int B^{\dagger}(\underline{x},\lambda) \, \mathcal{h} \, B(\underline{x},\lambda) d\lambda .$$

Focus attention for the present on the term $v(\lambda)$ in $\mathcal{h}$, and indeed on the monomial λ^{2n} contained therein. Such a term contributes to $\mathcal{H}(\underline{x})$ the local operator

$$\int B^{\dagger}(\underline{x},\lambda) \, \lambda^{2n} \, B(\underline{x},\lambda) d\lambda = \varphi_r^{2n}(\underline{x})$$

illustrating the central idea in constructing the potential term in $\mathcal{H}(\underline{x})$ from renormalized powers of the field $\varphi(\underline{x})$. The motivation for the second derivative term in $\mathcal{h}$ will be given subsequently.

The Hamiltonian $\mathcal{h}$ corresponds to the underlying one degree of freedom problem alluded to earlier. If we adopt the relation $c(\lambda) = |\lambda|^{-\gamma} \exp[-y(\lambda)]$, it follows that

$$v(\lambda) = \frac{\gamma(\gamma+1)}{2\lambda^2} + \frac{\gamma y'(\lambda)}{\lambda} + \tfrac{1}{2}y'^2(\lambda) - \tfrac{1}{2}y''(\lambda),$$

where, for example, if $y(\lambda)$ is an even polynomial, the latter three terms yield a polynomial contribution to $v(\lambda)$. Note there is always a singular term λ^{-2} with a strength determined by the singularity parameter γ.

The interpretation of $v(\lambda)$ is greatly helped by considering the special cases where

$$\hat{c}(\lambda) \equiv \frac{1}{|\lambda|^{\gamma}} e^{-\frac{1}{2}\mu\lambda^2}.$$

In this case, it follows that

$$v(\lambda) = \frac{\gamma(\gamma+1)}{2\lambda^2} + (\gamma-\tfrac{1}{2})\mu + \tfrac{1}{2}\mu^2\lambda^2.$$

Moreover the energy spectrum of $\mathcal{h}$ can be determined completely. The energy levels $\hat{\mu}_\ell$, $\ell=0,1,2\ldots$, are discrete, two-fold degenerate ($\hat{\mu}_{2\ell}=\hat{\mu}_{2\ell+1}$) and are given by

$$\hat{\mu}_{2\ell} = \mu \, (2\ell+2\gamma+1). \qquad (4\text{-}5)$$

That is, apart from a γ-dependent minimum energy, the

energy levels are equally spaced reminiscent of the harmonic oscillator problem. These levels in the ultralocal model are interpreted as eigenlevels of an idealized, "pseudo-free" case -- not free in the sense of the Fock representation -- in which the interaction potential V is "zero" but its representation changing effects have been, in part, taken into account. The uniform ladder of levels $\hat{\mu}_\ell$ represent the excitation spectrum of $\mathcal{h}$, which is heuristically the spectrum of localized energy levels. In the harmonic oscillator an equal spacing of energy levels is equated with an absence of interaction, independently of the magnitude of the zero-point energy. By analogy, we interpret the equal spacing of the levels $\hat{\mu}_\ell$ as characterizing a sequence of "excitation types" which are declared, by fiat, to be "interaction free." Such a view holds whatever value we choose for γ.

As we change the model function from $\hat{c}(\lambda)$ to a general $c(\lambda)$, the potential changes from $\hat{v}(\lambda)$ to $v(\lambda)$ and the energy levels change accordingly. Deviations from the standard provided by (4-5) describe positive, or negative, interaction energies induced among the excitation types by the new terms in the potential. So long as $y(\lambda)\geq a+b\lambda^2$, $b>0$, for example, such Hamiltonians likewise have discrete, doubly degenerate, energy levels μ_ℓ, and corresponding energy eigenstates $u_\ell(\lambda)$. It is evident that we can rewrite the Hamiltonian in the form

$$\mathcal{H} = \Sigma\mu_\ell \int A^\dagger_\ell(\underline{x})\ A_\ell(\underline{x})d\underline{x} = \Sigma\ \mu_\ell\ N_\ell,$$

where

$$A_\ell(\underline{x}) \equiv \int u^*_\ell(\lambda)A(\underline{x},\lambda)d\lambda.$$

Each N_ℓ is a conventional number operator (for excitation type ℓ), and the spectrum of $\mathcal{H}$ is clearly given by $\Sigma\ \mu_\ell n_\ell$, where $n_\ell=0,1,2\ldots$, and $\Sigma n_\ell<\infty$. Observe that the entire spectrum of $\mathcal{h}$ is displayed in the "one particle" subspace where $\Sigma n_\ell=1$.

Although we have argued the case for a discrete spectrum of $\mathcal{h}$ the essential ideas are identical whatever the spectrum of $\mathcal{h}$ is.

Absence of Canonical Conjugate Field

Consider the expression

$$\dot{\varphi}(\underline{x}) \equiv i[\mathcal{H},\varphi(\underline{x})]$$
$$= \int B^{\dagger}(\underline{x},\lambda)(-i\partial/\partial\lambda)\ B(\underline{x},\lambda)d\lambda$$

as formally computed from the bilinear form of both $\mathcal{H}$ and $\varphi(\underline{x})$. This expression is only a form, and not a local operator, as readily seen by setting up a sequence of the type envisaged in Chapter 2. Nevertheless, we could formally imagine the hypothetical construct

$$\dot{\varphi}_r^2(\underline{x}) = \int B^{\dagger}(\underline{x},\lambda)(-\partial^2/\partial\lambda^2)\ B(\underline{x},\lambda)d\lambda$$

which is seen to be an ingredient in $\mathcal{H}(\underline{x})$; namely, that ingredient formally corresponding to $\pi^2(\underline{x})$. Although such a term is not a local operator it is greatly aided by a renormalization term

$$\gamma(\gamma+1)\varphi_r^{-2}(\underline{x}) = \int B^{\dagger}(\underline{x},\lambda)\ \frac{\gamma(\gamma+1)}{\lambda^2}\ B(\underline{x},\lambda)d\lambda$$

which by itself is also ill defined. We interpret this latter term as part of a necessary operator renormalization to supplement the ill defined term $\dot{\varphi}_r^2(\underline{x})$ in the construction of the Hamiltonian.

If we ignore the fact that $\dot{\varphi}(\underline{x})$ is only a form, we would formally compute that

$$[\varphi(\underline{x}),\dot{\varphi}(\underline{y})] = i\ \delta(\underline{x}-\underline{y}) \int B^{\dagger}(\underline{x},\lambda)B(\underline{x},\lambda)d\lambda$$

which has a divergent c-number term,

$$i\delta(\underline{x}-\underline{y}) \int c^2(\lambda)d\lambda .$$

Such a divergence is usually indicative of an infinite field strength renormalization.

Time-Dependent Field

The simple relation

$$e^{i\mathcal{H}t}\ A(\underline{x},\lambda)\ e^{-i\mathcal{H}t} = e^{-i\hbar t}A(\underline{x},\lambda)$$

allows us to express the time-dependent field operator in the form

$$\begin{aligned}\varphi(\underset{\sim}{x},t) &= e^{i\mathcal{H}t}\varphi(\underset{\sim}{x})\ e^{-i\mathcal{H}t}\\ &= \int A^{\dagger}(\underset{\sim}{x},\lambda)\ e^{i\hbar t}\lambda e^{-i\hbar t}\ A(\underset{\sim}{x},\lambda)d\lambda\\ &\quad + \int c(\lambda)\lambda\ e^{-i\hbar t}\ A(\underset{\sim}{x},\lambda)d\lambda\\ &\quad + \int A^{\dagger}(\underset{\sim}{x},\lambda)\ e^{i\hbar t}\ \lambda c(\lambda)d\lambda .\end{aligned}$$

For the field to become an operator with space smearing only it is necessary that $\lambda c(\lambda)$ be square integrable near $\lambda=0$, and this property carries over to the field at any sharp time t. However, if we permit space-time smearing to define our operators, then we can choose model functions for which $\lambda c(\lambda)\notin L^2$ near $\lambda=0$ and consider singularity parameters γ for arbitrary values $\gamma\geq\frac{1}{2}$. We note, first, that the definition of $\hbar$ (and thus of $\mathcal{H}$) is valid as it stands for any $\gamma\geq\frac{1}{2}$. We next argue, in the case of a discrete spectrum for $\hbar$, that for suitable e(t),

$$\int e(t)\ e^{i\hbar t}\ \lambda c(\lambda)dt \tag{4-6}$$

is square integrable for $\gamma\geq\frac{1}{2}$. To see this we need only consider the sum $\Sigma|\varkappa_{\ell}|^2$ where

$$\begin{aligned}\varkappa_{\ell} &\equiv \iint e(t)\ u_{\ell}^{*}(\lambda)e^{i\hbar t}\ \lambda c(\lambda)d\lambda dt\\ &= k_{\ell}\int e(t)\ e^{i\mu_{\ell}t}dt\\ &= \tilde{e}(\mu_{\ell})k_{\ell},\end{aligned}$$

and

$$k_{\ell} \equiv \int u_{\ell}^{*}(\lambda)\lambda c(\lambda)d\lambda .$$

Since $|u_{\ell}(\lambda)| \sim |\lambda|^{1+\gamma}$ near $\lambda=0$, k_{ℓ} exists for all ℓ, and it may be estimated that $k_{\ell}\propto \ell^{-\alpha}$ with $\alpha\leq\frac{1}{2}$ provided $\gamma\geq\frac{3}{2}$ consonant with the fact that $\lambda c(\lambda)$ is square integrable provided that $\gamma<\frac{3}{2}$. But, if $\tilde{e}(\mu_{\ell})$ falls sufficiently fast for large argument (equivalent to large ℓ), the sequence $\varkappa_{\ell}$ becomes square summable. This demonstrates that

(4-6) is square integrable, which is sufficient to show that $\varphi(\underset{\sim}{x},t)$ becomes an operator with space-time smearing. A similar argument applies in the cases where the spectrum of $\mathcal{h}$ is not purely discrete.

However, even if $\gamma \geq \frac{3}{2}$ so that time smearing is required, we can always consider the renormalized fields

$$\varphi_r^\theta(\underset{\sim}{x}) = \int B^\dagger(\underset{\sim}{x},\lambda)\ \lambda^\theta\ B(\underset{\sim}{x},\lambda)d\lambda$$

which for sufficiently large θ again become operators with space smearing alone. For present purposes let us adopt an "odd" definition of λ^θ, i.e.,

$$\lambda^\theta \equiv (\text{sign } \lambda)\ |\lambda|^\theta .$$

Then, it suffices that $\theta > \gamma - \frac{1}{2}$, and it follows that

$$E^\theta(f) \equiv \langle 0|e^{i\varphi_r^\theta(f)}|0\rangle$$

$$= e^{-\int d\underset{\sim}{x}\int[1-\cos[\lambda^\theta f(\underset{\sim}{x})]]\ c^2(\lambda)d\lambda} .$$

Clearly by a change of variables this expression becomes identical to that given in (4-2) for some transformed model function. This means that the renormalized θ field power (appropriate to a model function $c(\lambda)$ and a singularity parameter $\gamma<\theta+\frac{1}{2}$) is unitarily equivalent to one of the fields covered by Eq. (4-2)(where necessarily $\gamma< \frac{3}{2}$).

Summary

We are now in a position to recapitulate the basic solution for the ultralocal models. For each real, even, nowhere vanishing, twice differentiable (save at $\lambda=0$) model function of the form

$$c(\lambda) = \frac{1}{|\lambda|^\gamma}\ e^{-y(\lambda)},$$

for which $\gamma \geq \frac{1}{2}$ and for $\theta > \gamma - \frac{1}{2}$,

$$\int \lambda^{2\theta}/(1+\lambda^{2\theta})\ c^2(\lambda)d\lambda < \infty,$$

we associate the Hamiltonian

$$\mathcal{H} = \int d\underline{x} \int A^{\dagger}(\underline{x},\lambda)\, \mathcal{h}\, A(\underline{x},\lambda)\, d\lambda ,$$

where

$$\mathcal{h} = -\tfrac{1}{2}\frac{\partial^2}{\partial\lambda^2} + \frac{c''(\lambda)}{c(\lambda)} .$$

The (space-time distribution) local field operator is given by

$$\varphi(\underline{x},t) = \int A^{\dagger}(\underline{x},\lambda)\, e^{i\mathcal{h}t}\lambda e^{-i\mathcal{h}t}\, A(\underline{x},\lambda)\, d\lambda$$

$$+ \int A^{\dagger}(\underline{x},\lambda)\, e^{i\mathcal{h}t}\lambda c(\lambda)\, d\lambda$$

$$+ \int c(\lambda)\, \lambda e^{-i\mathcal{h}t}\, A(\underline{x},\lambda)\, d\lambda .$$

As before the potential terms in the Hamiltonian are formally given by renormalized field powers such as

$$\varphi_r^{2n}(\underline{x}) = \int B^{\dagger}(\underline{x},\lambda)\lambda^{2n} B(\underline{x},\lambda)\, d\lambda ,$$

which links the representation of $\varphi(\underline{x},t)$ to the Hamiltonian.

The connection to the motivating classical problem seems to be best understood by temporarily reinserting the dependence on $\hbar$. In this case

$$\mathcal{h} \equiv -\tfrac{1}{2}\hbar^2\frac{\partial^2}{\partial\lambda^2} + \frac{\hbar^2\gamma(\gamma+1)}{2\lambda^2} + \hbar^2 e + v_o(\lambda)$$

where e is a constant, $e=(\gamma-\frac{1}{2})y''(0)$, and $y(\lambda)$ $[=y(\lambda,\hbar)]$ is chosen so that

$$v_o(\lambda) \equiv \tfrac{1}{2} m_0^2\lambda^2 + V[\lambda]$$

is $\hbar$ independent as usual. For the pseudo-free case, $y(\lambda) = \frac{1}{2}\mu\lambda^2/\hbar$ so that $\hbar^2 e = \hbar(\gamma-\frac{1}{2})\mu$ and $v_o(\lambda) = \frac{1}{2}\mu^2\lambda^2$. Observe, in the general case, that the singular term λ^{-2}

makes no classical contribution as $\hbar \to 0$ [as is already familiar from the analogous problem of a radial Schrödinger equation for a spherically symmetric potential in which, in the classical limit, the angular momentum vanishes whenever the angular momentum quantum number (the analogue of γ) is held fixed].

As is evident from the preceding discussion, the characterization of the solution is complete apart from specification of γ which cannot be deduced from classical arguments alone. To determine γ we shall appeal to a very different argument and to a very special class of models, the scale-invariant models.

4.3 IMPLICATIONS OF SCALE INVARIANCE

Scale Invariant Models

In order to determine whether any of the ultralocal models qualify as scale invariant we invoke the t=0 invariance criterion given in (3-2). This is most simply employed if we specialize to multiples of a characteristic function, $f(\underline{x}) = r\chi_{\Delta}(\underline{x})$. The expectation functional then becomes

$$E^{\theta}(f) \equiv C(r) = e^{-\Delta L[r]},$$

and invariance under scaling $[E^{\theta}(f)=E^{\theta}(f_{(S)})$, where $f_{(S)}(\underline{x}) = S^{d_\theta - s} f(S^{-1}\underline{x})]$ takes on the form

$$e^{-S^{s}\Delta L[S^{d_\theta - s} r]} = e^{-\Delta L[r]}.$$

It follows directly that $L[r] = k|r|^{\alpha}$, where $\alpha = s/(s-d_\theta)$, and $k \equiv L[1]$. These functions lead to the (symmetric) stable distributions discussed in Chapter 2. We noted there that α is restricted such that $0<\alpha\leq 2$. The case $\alpha=2$ corresponded to the Gaussian, while $0<\alpha<2$ were all non-Gaussian with $c(\lambda)$ being a homogeneous function.

A more direct analysis is useful as well. Scale invariance of the expectation functional, $E^{\theta}(f)=E^{\theta}(f_{(S)})$, requires that the expression

$$\int d\underline{x} \int \{1-\cos[\lambda^{\theta} S^{d_\theta - s} f(S^{-1}\underline{x})]\} c^2(\lambda) d\lambda$$

be independent of S. This can only be satisfied by a homogeneous $c(\lambda)$, namely $c(\lambda) \propto |\lambda|^{-\gamma}$, in which case it follows (by a simple change of variables) that

$$d_\theta = (1 - \frac{\theta}{2\gamma-1})s, \tag{4-7}$$

and thus the exponent $\alpha = (2\gamma-1)/\theta$.

We note that the only model functions describing scale invariant theories are given by $c(\lambda) = |\lambda|^{-\gamma}$ (or multiples thereof), where $\gamma > \frac{1}{2}$. The functional form for stable distributions implies that for scale invariant theories

$$<0|e^{i\varphi_r^\theta(f)}|0> = e^{-k\int|f(\underset{\sim}{x})|^\alpha d\underset{\sim}{x}},$$

where $(\alpha \neq 2)$

$$k = \int [1-\cos(\lambda^\theta)]\ |\lambda|^{-2\gamma} d\lambda .$$

It follows that $||\varphi_r^\theta(f)|0>|| = \infty$, i.e. $|0>$ is not in the domain of $\varphi_r^\theta(f)$. This is only a technical difficulty and does not invalidate their existence (a suitable time smeared field can be applied to the vacuum).

Determination of Scale Dimension and Singularity Parameter

As a next step in exploring the scale invariant models we invoke scaling with respect to time. In analogy to (3-2), full space-time scale invariance, i.e.,

$$E^\theta(g) = E^\theta(g_{(S)}),$$

where

$$g_{(S)}(\underset{\sim}{x},t) = S^{d_\theta - s - 1} g(S^{-1}\underset{\sim}{x}, S^{-1}t),$$

leads to the easily established requirement that

$$\int d\underset{\sim}{x} \int c(\lambda)\{1-\cos[\int (e^{i\hbar t\lambda^\theta} e^{-i\hbar t}) S^{d_\theta - s - 1} g(S^{-1}\underset{\sim}{x}, S^{-1}t)dt]\}c(\lambda)d\lambda$$

be independent of S. Observe for $c(\lambda) = |\lambda|^{-\gamma}$ that

$$\mathcal{h} = -\tfrac{1}{2}\frac{\partial^2}{\partial\lambda^2} + \frac{\gamma(\gamma+1)}{2\lambda^2}$$

which is homogeneous in λ^{-2}. To secure full independence as far as the variable t is concerned requires (simply using the homogenity of $\mathcal{h}$) that

$$d_\theta = s-\tfrac{1}{2}\theta,$$

$$\gamma = s+\tfrac{1}{2};$$

namely, two conditions rather than the single (compatible) condition (4-7) obtained for t=0. Observe that by this argument the singularity parameter is "½ plus the number of space dimensions".

We assume that this evaluation holds for $\gamma(\lambda)\neq 0$, i.e., even for nonscale invariant theories. For example, with this choice the spectrum of the pseudo-free model functions

$$\hat{c}(\lambda) = \frac{e^{-\frac{1}{2}\mu\lambda^2}}{|\lambda|^{s+\frac{1}{2}}}$$

is given [on reference to (4-5)] by

$$\hat{\mu}_{2\ell} = 2\mu(\ell+s+1)$$

which leads to a ladder of excitation levels with a starting level $[(2\mu)(s+1)]$ to spacing (2μ) ratio of s+1, the number of space-time dimensions. (It is amusing to imagine that such a property almost constitutes an "experimental prediction" under the hypothetical conditions for which it applies.)

Dilation Operator

The explicit construction of the dilation operator V(S) is straightforward in the ultralocal models. It follows that

$$V(S) = \exp\{(\ln S)\int\int A^\dagger(\underline{x},\lambda)D\,A(\underline{x},\lambda)\,d\underline{x}d\lambda\} \qquad (4\text{-}8)$$

where

$$D \equiv \tfrac{1}{2}(\underset{\sim}{x}\cdot\frac{\partial}{\partial \underset{\sim}{x}} + \frac{\partial}{\partial \underset{\sim}{x}}\cdot\underset{\sim}{x}) + \tfrac{1}{4}(\lambda\,\frac{\partial}{\partial\lambda} + \frac{\partial}{\partial\lambda}\,\lambda).$$

It is clear the $V(S)|0\rangle = |0\rangle$, and a straightforward computation shows that

$$V^{-1}(S)\varphi^{\theta}_{r}(\underset{\sim}{x},t)V(S) = S^{d_\theta}\,\varphi^{\theta}_{r}(S\underset{\sim}{x},St)$$

provided $c(\lambda) \propto |\lambda|^{-\gamma}$.

Once given the form of $V(S)$ we may investigate the scale transformation of a general (nonscale invariant) field $\varphi^{\theta}_{r}(\underset{\sim}{x},t)$ described by $c(\lambda)=|\lambda|^{-\gamma}\exp[-y(\lambda)]$. in that case,

$$V^{-1}(S)\varphi^{\theta}_{r}(\underset{\sim}{x},t)V(S) \equiv S^{d_\theta}\,\tilde{\varphi}^{\theta}_{r}(S\underset{\sim}{x},St),$$

where $\tilde{\varphi}^{\theta}_{r}$ is characterized by

$$\tilde{c}(\lambda) \equiv |\lambda|^{-\gamma}\exp[-y(S^{-\frac{1}{2}}\lambda)]$$

and by the corresponding $\tilde{\mathcal{H}}$ ($\tilde{\mathcal{H}}\tilde{c}=0$). For the pseudo-free case $\hat{c}(\lambda)$, where $y(\lambda)=\frac{1}{2}\mu\lambda^2$, the transformed field corresponds to a transformed mass $\tilde{\mu}=S^{-1}\mu$, a transformation which is consistent with the conventional viewpoint lending credence to the choice $\gamma=s+\frac{1}{2}$ even for nonscale invariant theories.

Alternative Scale Transformations

Although we have fixed γ on the basis of scale invariance arguments it is instructive to understand the significance of the remaining γ values. We have argued that the ultralocal models may be interpreted as the limit of covariant Hamiltonians in which the coefficient (α say) of $\frac{1}{2}(\underset{\sim}{\nabla}\varphi_{c\ell})^2$ vanishes. It is such a term which dictates the equality of the $\underset{\sim}{x}$ and t scaling for covariant theories as is evident, for example, in the wave equation

$$\frac{\partial^2\theta}{\partial t^2} = \alpha\,\frac{\partial^2\theta}{\partial \underset{\sim}{x}^2}$$

The appropriate scaling, $\underset{\sim}{x} \to S\underset{\sim}{x}$ and $t \to St$, holds for all $\alpha > 0$, and it applies equally well in the limit $\alpha = 0$. However the limiting equation may be obtained in a number of different ways. Consider the wave equation defined by

$$\frac{\partial^2 \theta}{\partial t^2} = \alpha \left(\frac{\partial^2}{\partial \underset{\sim}{x}^2}\right)^{\beta} \theta, \qquad \beta > 0, \tag{4-9}$$

which is covariant under the scaling

$$\underset{\sim}{x} \to S\underset{\sim}{x}; \quad t \to S^{\beta} t.$$

Only $\beta = 1$ characterizes the relativistic case. When $\alpha = 0$ all these wave equations are equal and the whole class (i.e., all β) of scaling transformations apply.

In the case of the ultralocal models, where $\alpha = 0$, the whole class of scaling transformations should also apply and not merely the special case $\beta = 1$. If we repeat our previous analysis based on the assumptions that $V_\beta(S)|0\rangle = |0\rangle$ and

$$V_\beta^{-1}(S)\varphi_r^\theta(\underset{\sim}{x},t)\ V_\beta(S) = S^{d_{\theta,\beta}}\varphi_r^\theta(S\underset{\sim}{x}, S^\beta t),$$

we would discover that

$$d_{\theta,\beta} = s - \tfrac{1}{2}\theta\beta,$$

$$\gamma = (s/\beta) + \tfrac{1}{2},$$

which connects γ to the number of space dimensions and the presumed underlying wave equation (4-9), i.e., to β. In this calculation it would follow that $V_\beta(S)$ has the same form as (4-8) except that the coefficient of $\lambda \frac{\partial}{\partial \lambda}$, etc., in D now reads $\frac{1}{4}\beta$.

On the basis of these arguments the singularity parameter γ can be determined, at least for $\gamma > \frac{1}{2}$. The case $\gamma = \frac{1}{2}$ is anomalous in this regard (although it can be considered as a limiting case as $\beta \to \infty$). We remind the reader that the relativistically determined γ values are $\gamma = s + \frac{1}{2}$, i.e., $\gamma = \frac{3}{2}, \frac{5}{2}, \frac{7}{2}$, etc.

4.4 PASSAGE TO THE FREE THEORY

An important issue remaining to be discussed is what might be called "turning off the coupling constant," namely, the passage from the interacting models to a non-interacting, free, Fock solution for a particular mass m. This transformation has been discussed elsewhere for restricted singularity parameters $\frac{3}{2}>\gamma\geq\frac{1}{2}$. Here we discuss the general case $\gamma\geq\frac{1}{2}$.

Field Operator

Let us assume that $c(\lambda)$ is such that $\lambda^{\theta}c(\lambda)\in L^2$ for some $\theta>\gamma-\frac{1}{2}$, all of which are sufficient to overcome the singularity at the origin. For future reference we shall also need an upper bound on θ, namely $\gamma+\frac{1}{2}\geq\theta$. With the "odd" definition of λ^{θ} we have already noted that

$$E^{\theta}(f) = e^{-\int d\underset{\sim}{x}\int\{1-\cos[\lambda^{\theta}f(\underset{\sim}{x})]\}c^2(\lambda)d\lambda},$$

while the free, Fock theory of mass m is characterized by the functional

$$E_F(f) = \langle 0|e^{i\varphi_F(f)}|0\rangle = e^{-\frac{1}{4m}\int f^2(\underset{\sim}{x})d\underset{\sim}{x}}.$$

We "turn off the coupling constant" by choosing a sequence of models -- hence a sequence of model functions $c_\eta(\lambda)$ -- such that, as $\eta\to\infty$,

$$\int[1-\cos(\lambda^{\theta}f)]c_\eta^2(\lambda)d\lambda \to \frac{1}{4m}f^2.$$

This is accomplished by the formal requirement that

$$\lambda^{2\theta}c_\eta^2(\lambda) \to \frac{1}{2m}\delta(\lambda);$$

for example, if we adopt

$$c_\eta^2(\lambda) = \eta^{2\theta+1}c^2(\eta\lambda)$$

where $c(\lambda)$ has been prescaled (without change of $\hbar$ or $\mathcal{K}$) so that

$$\int \lambda^{2\theta}c^2(\lambda)d\lambda = (2m)^{-1}.$$

Observe that any preassigned mass m can be obtained by such a procedure.

According to the above conditions, the sequence of expectation functionals converges to the expectation functional of the free theory. Thus we may say that

$$\varphi_r^\theta(\underline{x}) \to \varphi_F(\underline{x})$$

in the sense of expectation functionals. This convergence does not occur as operator convergence in the Hilbert space $\mathfrak{H}$. However, the free field $\varphi_F(\underline{x})$ has its own well known operator realization in an appropriate Hilbert space $\mathfrak{H}_F$.

Hamiltonian

The convergence of the Hamiltonian $\mathcal{H}$ to the free (ultralocal) Hamiltonian $\mathcal{H}_F$ of mass m is more complicated. Although $\varphi_r^\theta(\underline{x})$ is a local operator (since $\theta>\gamma-\frac{1}{2}$), $\dot{\varphi}_r^\theta(\underline{x})$ is only a form (since we impose $\gamma+\frac{1}{2}\geq\theta$). It follows, moreover, in a formal way, that

$$[\varphi_r^\theta(\underline{x}),\ \dot{\varphi}_r^\theta(\underline{y})] = i\delta(\underline{x}-\underline{y}) \int B^\dagger(\underline{x},\lambda)\ \theta^2\lambda^{2(\theta-1)}\ B(\underline{x},\lambda)d\lambda$$

which by hypothesis has a leading c-number singularity

$$i\delta(\underline{x}-\underline{y}) \int \theta^2\lambda^{2\theta-2}c^2(\lambda)d\lambda$$

that diverges because we suppose that $\theta\leq\gamma+\frac{1}{2}$.

To remove the infinite multiple in the commutation relation let us consider the properties that follow from the scaled and modified Hamiltonians

$$\mathcal{H}_\varepsilon \equiv \int d\underline{x} \int A^\dagger(\underline{x},\lambda)(b^\dagger M_\varepsilon b)A(\underline{x},\lambda)d\lambda$$

where, as before, $b=2^{-\frac{1}{2}}c(\lambda)\ (\partial/\partial\lambda)c^{-1}(\lambda)$, and

$$M_\varepsilon \equiv \frac{e^{-\varepsilon^2/\lambda^2}}{\int e^{-\varepsilon^2/\lambda^2}\theta^2\lambda^{2\theta-2}c^2(\lambda)d\lambda}.$$

As $\varepsilon\to 0$, the operators $\mathcal{H}_\varepsilon$ converge to a form and not an operator. We may see this most easily by taking coherent

state matrix elements with wave functions $\psi(\underset{\sim}{x},\lambda) \equiv \lambda^{\theta}c(\lambda)h(\underset{\sim}{x},\lambda)$, where $h(\underset{\sim}{x},\lambda)$ is C^2 save at $\lambda=0$ where C^1 suffices, and has compact support in R^s. If $|\psi>$ and $|\psi'>$ denote two such coherent states, it follows that

$$\lim_{\varepsilon\to 0} <\psi|\mathcal{H}_{\varepsilon}|\psi'> = \tfrac{1}{2}\int h^*(\underset{\sim}{x},0)h'(\underset{\sim}{x},0)d\underset{\sim}{x}<\psi|\psi'>,$$

which is evidently not continuous in the bra $<\psi|$. The special states

$$|f> \equiv e^{i\varphi_r^{\theta}(f)}|0>$$

are coherent states of the usual type for which

$$\psi(\underset{\sim}{x},\lambda) = \{e^{i\lambda^{\theta}f(\underset{\sim}{x})}-1\}c(\lambda) \equiv \lambda^{\theta}c(\lambda)h(\underset{\sim}{x},\lambda),$$

and therefore $h(\underset{\sim}{x},0)=if(\underset{\sim}{x})$. Hence we learn that

$$\lim_{\varepsilon\to 0} <f|\mathcal{H}_{\varepsilon}|f'> = \tfrac{1}{2}\int f(\underset{\sim}{x})f'(\underset{\sim}{x})d\underset{\sim}{x}<f|f'>,$$

in which, although the set of states $|f>$ are total, the lack of continuity in the bra $<f|$ is not so obvious. Finally, if we let $\eta\to\infty$ in the model functions $c_{\eta}(\lambda)$ we learn that*

$$\lim_{\eta\to 0}\lim_{\varepsilon\to 0} <f|\mathcal{H}_{\varepsilon}|f'> = \tfrac{1}{2}\int f(\underset{\sim}{x})f'(\underset{\sim}{x})d\underset{\sim}{x}_F<f|f'>_F$$

which is well known to characterize the matrix elements ${}_F<f|\mathcal{H}_F|f'>_F$ of the free Fock Hamiltonian of mass m in the Hilbert space $\mathfrak{H}_F$. In the sense of the above analysis, $\mathcal{H}\to\mathcal{H}_F$.

For the space translation generator we immediately find as $\eta\to\infty$ that

$$<f|e^{-i\underset{\sim}{a}\cdot\underset{\sim}{P}}|f'> \equiv <f|f'_a> \to {}_F<f|f'_a>_F \equiv {}_F<f|e^{-i\underset{\sim}{a}\cdot\underset{\sim}{P}_F}|f'>_F .$$

In this sense, $\underset{\sim}{P} \to \underset{\sim}{P}_F$.

---

* Here we introduce $<0|f>_F \equiv E_F(f)$ and $|f>_F \equiv e^{i\varphi_F(f)}|0>$, which form a total set for $\mathfrak{H}_F$.

Conclusion

It is noteworthy that a transition to the free, Fock theory exists for all $\gamma \geq \frac{1}{2}$, and that each theory can lead to a free theory with any preassigned mass value.

It is remarkable how complete an understanding can be found for the ultralocal models on the basis of simple symmetry arguments. The power of these arguments will again become apparent in the models discussed in the next chapter.

5. DIASTROPHIC FIELD THEORIES

5.1 HEURISTIC, CLASSICAL MOTIVATION

The basic extension of one problem to another which distinguishes the ultralocal fields can be applied more generally. By way of illustration consider the classical Hamiltonian of a covariant, φ^4, theory

$$H_1[\pi_{c1},\varphi_{c1}] = \int\{\tfrac{1}{2}[\pi^2_{c1}(\underline{x})+(\underline{\nabla}\varphi_{c1}(\underline{x}))^2+m_o^2\varphi^2_{c1}(\underline{x})] + \tfrac{1}{4}g\,\varphi^4_{c1}(\underline{x})\}\,d\underline{x} \tag{5-1}$$

and the diastrophic field theory it engenders. We enlarge the configuration space so that $\pi_{c1}(\underline{x}) \to \pi_{c1}(\underline{x},w)$ and $\varphi_{c1}(\underline{x}) \to \varphi_{c1}(\underline{x},w)$, where $w\in R$, and adopt the classical model

$$\begin{aligned} H(\pi_{c1},\varphi_{c1}) &= \int H_1[\pi_{c1}(w),\,\varphi_{c1}(w)]\,dw \\ &\equiv \int\{\tfrac{1}{2}[\pi^2_{c1}(\underline{x},w)+(\underline{\nabla}\varphi_{c1}(\underline{x},w))^2+m_o^2\varphi^2_{c1}(\underline{x},w)] \\ &\quad + \tfrac{1}{4}g\varphi^4_{c1}(\underline{x},w)\}\,d\underline{x}dw. \end{aligned} \tag{5-2}$$

The essential point to observe is that there is no mechanism for communication of field values from one w value to another -- distinct w values label statistically independent fields for all time. This feature is also clear in the field equation that follows from (5-2), namely $(x=\underline{x},t)$

$$(\Box + m_o^2)\,\varphi_{c1}(x,w) = -g\varphi^3_{c1}(x,w). \tag{5-3}$$

In fact, if we choose as initial values

$$\varphi_{c1}(\underset{\sim}{x},w) = \varphi_{c1}(\underset{\sim}{x})\ \chi(w),$$

$$\dot{\varphi}_{c1}(\underset{\sim}{x},w) = \dot{\varphi}_{c1}(\underset{\sim}{x})\ \chi(w),$$

where $\chi(w)$ is a characteristic function (say for $|w| \leq \frac{1}{2}$), then the solution to the diastrophic model is just

$$\varphi_{c1}(x,w) = \varphi_{c1}(x)\ \chi(w)$$

where $\varphi_{c1}(x)$ is the corresponding solution of the base theory given in (5-1). That is, all the solutions of the base theory are contained in the solutions of the diastrophic theory. (This argument is of course independent of the covariance of the base theory.)

Not only can we have the solution $\varphi_{c1}(x)$ in the "stratum" $|w| \leq \frac{1}{2}$, we can simultaneously have two quite independent solutions in disjoint strata. To show this we need only note that

$$\varphi_{c1}(x,w) = \varphi_{c1}(x)\ \chi(w) + \hat{\varphi}_{c1}(x)\ \hat{\chi}(w)$$

is a solution also, provided that $\varphi_{c1}(x)$ and $\hat{\varphi}_{c1}(x)$ are solutions of the base theory and that $\chi(w)\ \hat{\chi}(w) \equiv 0$. This picture obviously extends to any number of solutions in disjoint strata, and all solutions are suitable limits of such multiple-strata solutions.

Suppose the diastrophic classical model (5-2) could be quantized. For one thing it should correspond to a covariant theory as is evident from the equation of motion (5-3). Moreover, since within one stratum (say $|w| \leq \frac{1}{2}$), we can recover the entire set of classical solutions of the base theory, it is not unreasonable that a scattering theory could be set up, built from asymptotic states having characteristic functions in w for test functions. We also have another fact at our disposal, namely that if, in the quantum theory based on (5-2), the coefficient of the spacial gradient went to zero we must recover an ultralocal model of the preceding chapter based on the configuration space R^{s+1}, the latter R coming from the variable w.

Superficially, on the basis of perturbation theory for example, the diastrophic theory is no less singular than the base theory. We approach their solution by non-perturbative techniques, and, although we are only able to present a certain "super structure" of the solution, it is noteworthy how much structural information can be found. Our discussion makes heavy use of analogies, both explicit and implicit, with the treatment of the ultralocal models.

5.2 SUPER STRUCTURE OF OPERATOR SOLUTION

As in the previous chapter our assumptions are minimal and reasonable. For a large class of symmetric potentials we assume there is a positive, self-adjoint Hamiltonian $\mathcal{H}$, and a unique ground state $|0\rangle$ satisfying $\mathcal{H}|0\rangle = 0$. Also we assume there is a space-translation generator $\underset{\sim}{\mathcal{P}}$ satisfying $[\underset{\sim}{\mathcal{P}},\mathcal{H}] = 0$, and $|0\rangle$ is a nondegenerate eigenstate of $\underset{\sim}{\mathcal{P}}$, $\underset{\sim}{\mathcal{P}}|0\rangle = 0$. Likewise, we postulate the generators of the Lorentz group for a covariant diastrophic model.

With regard to the field operator $\varphi(x,w) = \varphi(\underset{\sim}{x},t,w)$ we assume it transforms conventionally, and is self adjoint when smeared with appropriate space (or space-time) plus w-dependent test functions. For purposes of illustration we shall assume that space plus w-dependent test functions suffice. (The appropriate generalization to include time smearing will be noted subsequently.)

Determination of Expectation Functional

We study the expectation functional for the field operator

$$\varphi(f) = \int\int \varphi(\underset{\sim}{x},w)\ f(\underset{\sim}{x},w)d\underset{\sim}{x}dw.$$

The "ultralocal" nature of the dynamics with respect to the variable w implies that

$$E(f) = \langle 0|\ e^{i\varphi(f)}|0\rangle = e^{-\int dwL\{f(\cdot,w)\}}. \qquad (5\text{-}4)$$

Here, the notation for L denotes, formally,

$$L\{f(\cdot,w)\} = \sum_{n=o}^{\infty} \int\ldots\int M_n(\underset{\sim}{x}_1,\ldots\underset{\sim}{x}_n)f(\underset{\sim}{x}_1,w)\ldots f(\underset{\sim}{x}_n,w)\ d\underset{\sim}{x}_1\ldots d\underset{\sim}{x}_n.$$

Symmetry of the potential implies that $E(-f) = E(f)$. Let us assume that $E(f)$ is defined for all $f \in C_0^\infty$ (minimally ray continuous) and is positive definite. It follows from the functional form in (5-4) that $E(f)$ is infinitely divisible. Clearly

$$[E(f)]^{1/m} = e^{-(1/m)\int dw L\{f(\cdot,w)\}}$$

$$= e^{-\int dw' L\{f(\cdot,mw')\}} \equiv E(f_{[m]}),$$

where $f_{[m]}(\underset{\sim}{x},w) \equiv f(\underset{\sim}{x},mw) \in C_0^\infty$. It follows that $[E(f)]^{1/m}$ is an expectation functional for all m and is therefore infinitely divisible. This property, which holds for <u>all</u> $f(\underset{\sim}{x},w)$, is a consequence of the ultralocality of just <u>one</u> of the variables, namely w.

According to Eq. (3-9), $E(f)$ never vanishes and thus $L\{f(\cdot,w)\}$ is defined for all f. Continuity of $E(f)$ extends to product test functions of the form $f(\underset{\sim}{x},w) = \hat{f}(\underset{\sim}{x})\,\chi_\Delta(w)$, where χ_Δ is a characteristic function and $\hat{f}(\underset{\sim}{x})$ is smooth. For these cases

$$E(\hat{f}) = e^{-\Delta L\{\hat{f}(\cdot)\}}$$

which since it must be positive definite for all $\Delta>0$ implies that $L\{f(\cdot)\}$ can be determined from Eq. (3-9). As in the ultralocal models we associate the Gaussian terms with the free (Fock representation) theory and focus instead on the non-Gaussian part. Thus we have determined that

$$E(f) = e^{-\int dw \int [1-\cos(\Lambda,f(w))]\, d\sigma(\Lambda)},$$

where, formally, the notation means

$$(\Lambda,f(w)) = \int \Lambda(\underset{\sim}{x})\, f(\underset{\sim}{x},w) d\underset{\sim}{x}.$$

In the present theory, the "model function" is contained in the measure σ. In analogy with the ultralocal case, we shall require that $\int d\sigma(\Lambda) = \infty$.

Field Operator

An operator realization for the field can be given in terms of the formulation presented in Section 3.4. For this purpose we assume that

$$\mathfrak{h} \equiv \mathfrak{L}^2(\mathcal{V}',\sigma) \times L^2(R) \equiv \hat{\mathfrak{L}}^2$$

composed of elements $\psi(\Lambda,w)$ such that

$$(\psi,\psi) \equiv \int dw \int |\psi(\Lambda,w)|^2 d\sigma(\Lambda) < \infty.$$

It is also convenient to employ elements of $\mathfrak{L}^2(\mathcal{V}',\sigma)$ alone, say $\psi(\Lambda)$, for which

$$\langle\psi,\psi\rangle \equiv \int |\psi(\Lambda)|^2 d\sigma(\Lambda).$$

The round or angular brackets distinguish the two cases. In an obvious embedding we can imagine that

$$\psi(\Lambda,w) \in \mathfrak{L}^2(\mathcal{V}',\sigma)$$

for almost all w, in which case we set

$$\langle\psi(w),\psi(w)\rangle \equiv \int |\psi(\Lambda,w)|^2 d\sigma(\Lambda).$$

Observe that the function "one" is not an element of $\mathfrak{L}^2$ [since $\int d\sigma(\Lambda) = \infty$], but that $\int \Lambda(\underline{x}) g(\underline{x}) d\underline{x}$, for suitable $g(\underline{x})$, is σ-square integrable "near the origin".

In the notation of Section 3.4 we may represent the field operator $\varphi(\underline{x},w)$ as

$$\varphi(\underline{x},w) = \langle B(w), \Lambda(\underline{x})\, B(w)\rangle$$

[cf., Eq.(3-5) and also Eq. (4-3)]. Here we have chosen

$$B(\varphi) = A(\varphi) + \int \varphi^*(\Lambda,w)\, d\sigma(\Lambda) dw,$$

which holds for $\varphi \in \hat{\mathfrak{L}}^2 \cap \hat{\mathfrak{L}}^1$, and which can be viewed as the limit of unitarily equivalent operators for a non-Cauchy sequence $\xi_n(\Lambda,w) \to 1$ in the sense of Chapter 3. In terms of coherent state matrix elements we have

$$<\psi|\varphi(\underline{x},w)|\psi'> / <\psi|\psi'>$$

$$= \int[\psi^*(\Lambda,w)+1]\,\Lambda(\underline{x})\,[\psi'(\Lambda,w)+1]\,d\sigma(\Lambda),$$

where $|\psi> = N \exp[A(\psi)^\dagger]|0>$, $\psi \in \mathfrak{h}$, as usual.* In spite of the abstract notation, the fundamental idea to appreciate is simplicity itself: The representation of the field operator is bilinear in conventional (Fock representation) annihilation and creation operators!

Functions of the Field and Cyclicity of the Representation

It is clear that we may write

$$\varphi(\underline{x},w)\varphi(\underline{x}',w') = \delta(w-w')\,<B(w),\Lambda(\underline{x})\Lambda(\underline{x}')B(w)> + :\varphi(\underline{x},w)\varphi(\underline{x}',w'):,$$

where : : means normal order with respect to $B^\dagger$ and B. As in the preceding chapter the singular coefficient of the first term (for w=w') permits a test function sequence to pick out the partially renormalized field product

$$[\varphi(\underline{x},w)\varphi(\underline{x}',w)]_r \equiv <B(w),\ \Lambda(\underline{x})\Lambda(\underline{x}')B(w)>,$$

where multiplication occurs in "Λ-space". This is a partial renormalization because, as $\underline{x}' \to \underline{x}$, another, model dependent, renormalization may be involved; however, we do not discuss this problem.

By extension we can clearly build partially renormalized polynomials composed of terms such as

$$[\varphi(\underline{x}_1,w)\ldots\varphi(\underline{x}_n,w)]_r \equiv <B(w),\Lambda(\underline{x}_1)\ldots\Lambda(\underline{x}_n)B(w)>$$

all of which are functions of the field $\varphi(\underline{x},w)$. It is from such operators that we expect to build the nonlinear interaction terms in the Hamiltonian just as we did in the ultralocal models. All such terms are bilinear in the B's (or A's)!

---

* We adopt, as in Chapter 4, a principal value definition such that $\int\Lambda(x)d\sigma(\Lambda) = 0$.

Cyclicity of the representation can be established in much the same manner as for the ultralocal models. In a limiting fashion, let us build the operator

$$G \equiv \int \langle B(w), h(\Lambda,w)B(w)\rangle dw$$

where $h(\Lambda,w)$ is a suitable functional, e.g.,

$$h(\Lambda,w) \in \hat{\mathfrak{L}}^1 \cap \hat{\mathfrak{L}}^\infty .$$

The operator e^G essentially takes the vacuum into a coherent state since

$$\begin{aligned} e^G|0\rangle &= :\, e^{\int \langle B(w), \{e^{h(\Lambda,w)}-1\}B(w)\rangle dw} :\, |0\rangle \\ &= e^{g+A(\psi)^\dagger}|0\rangle \end{aligned}$$

where g is a numerical factor, and where

$$\psi(\Lambda,w) \equiv \{e^{h(\Lambda,w)}-1\} \in \hat{\mathfrak{L}}^2$$

in virtue of our assumption on h. As $h(\Lambda,w)$ varies in its allowed domain it clearly covers a dense set of elements in $\hat{\mathfrak{L}}^2$, which thus ensures the desired cyclicity. It is note-worthy that this remarkable cyclicity has come about by the fact that $w \in R$. One should never underestimate the power of the continuum!

A Few Generators

The simplest generators to deal with generate translations in the configuration space. For example, the operator

$$\mathbb{b} = \int dw \langle A(w), (-i\partial/\partial w)A(w)\rangle$$

induces translations in the stratum variable w. The space-translation generator $\underset{\sim}{\mathcal{P}}$ is likewise bilinear in $A^\dagger$ and A,

$$\underset{\sim}{\mathcal{P}} = \int dw \langle A(w), \not{\underset{\sim}{p}}\, A(w)\rangle ,$$

where $\not{\underset{\sim}{p}}$ is a self-adjoint operator on $\mathfrak{L}^2$ with the property that

$$i[\underset{\sim}{\mathcal{P}},\Lambda(\underset{\sim}{x})] = \underset{\sim}{\nabla}\,\Lambda(\underset{\sim}{x}).$$

Note that $\underset{\sim}{\mathcal{P}}|0\rangle=0$; but uniqueness demands there exist no invariant state in $\mathcal{L}^2$.

The form of these operators is already completely dictated by the field operator representation since a total set of matrix elements is given by a relation such as

$$\langle f|e^{i\underset{\sim}{a}\cdot\underset{\sim}{\mathcal{P}}}|f'\rangle = \langle f_a|f'\rangle = E(f'-f_a)$$

where

$$|f\rangle \equiv e^{i\varphi(f)}|0\rangle.$$

Structure of Hamiltonian

The operator form of the Hamiltonian is as conceptually simple as a collection of harmonic oscillators! As in the ultralocal models we expect that $\mathcal{H} = \int\mathcal{H}(w)dw$ and that $\mathcal{H}(w)$ is constructed from A and $A^\dagger$, but only at the point w. We are led by the same reasoning as in the ultralocal case to the form

$$\mathcal{H} = \int dw\langle A(w), \mathcal{k}\, A(w)\rangle$$

where $\mathcal{k}$ is a self-adjoint operator on $\mathcal{L}^2$. Again $\mathcal{H}|0\rangle = 0$, and to have uniqueness of the ground state and a non-negative spectrum we need $\mathcal{k}>0$. That is, there can be no time-translation invariant state in $\mathcal{L}^2$, nor, as we have already noted, any space-translation invariant state. Thus, with regard to the base quantum theory -- that defined in $\mathcal{L}^2$ -- we deal with an unconventional field theory, one without a normalizable, invariant state.

For the Lorentz group (when this group applies) identical reasoning demands that

$$\underset{\sim}{\mathcal{K}} = \int dw\,\langle A(w), \underset{\sim}{\mathcal{k}}\, A(w)\rangle,$$

$$\underset{\sim}{\mathcal{J}} = \int dw\,\langle A(w), \underset{\sim}{\mathcal{j}}\, A(w)\rangle,$$

for the boosts and rotations, respectively. Fulfillment of the Poincare Lie algebra by $\mathcal{H},\underset{\sim}{\mathcal{P}},\underset{\sim}{\mathcal{K}}$ and $\underset{\sim}{\mathcal{J}}$ is assured provided $\mathcal{k},\mathcal{p}$, $\underset{\sim}{\mathcal{k}}$ and $\underset{\sim}{\mathcal{j}}$ already fulfill the Lie algebra on $\mathcal{L}^2$.

However, in this latter case we deal with an unusual representation since there is no normalizable state annihilated by $\mathcal{K}$ or $\mathcal{L}$.

Stratons and Excitations

We may suggestively refer to

$$\mathcal{N} \equiv \int dw \, < A(w), A(w)>$$

as the "straton" number operator, and the subspace where $\mathcal{N} = n$ as the n-straton subspace. Evidently $[\mathcal{N},\mathcal{K}]=0$, etc., so that all the generators we have introduced conserve straton number.

Straton conservation implies that we can restrict $\mathcal{K}$ (and the other generators) to any particular n-straton subspace. It is clear from the construction that the entire spectrum of $\mathcal{K}$ is contained already in the one straton subspace! Hence, among the one straton states we expect to find the one particle states of the base theory, two particle scattering states, three particle scattering states, etc. Among the one straton states should be found a two particle "in" state and a (different) two particle "out" state, etc., which implies, as well, that the conventional scattering theory of the base theory is already contained in the one straton subspace. Thus, unlike the usual one particle S-matrix (which is trivial), we expect nontrivial scattering to exist among initial and final one straton states.

Two straton states would describe two independent, uncoupled scattering events conveniently pictured as taking place at disjoint strata. Multi-straton states have a corresponding interpretation.

Time-Dependent Field

The form of the time-dependent field operator can be readily found using the relation

$$e^{i\mathcal{K}t}<\varphi,A(w)> e^{-i\mathcal{K}t} = <\varphi,e^{-i\mathcal{K}t}A(w)> = <e^{i\mathcal{K}t}\varphi,A(w)>$$

where $<\varphi,A(w)>$ is the annihilation operator which maps the

coherent state $|\psi>$, $\psi(\Lambda,w) \in \hat{\mathfrak{L}}^2$ into $<\varphi,\psi(w)>|\psi>$.
If we note that

$$\varphi(\underline{x},w) = <A(w),\Lambda(\underline{x})A(w)> + <\Lambda(\underline{x}),A(w)> + <A(w),\Lambda(\underline{x})>,$$

then we find $(x=\underline{x},t)$

$$\varphi(x,w) \equiv e^{i\mathcal{H}t}\varphi(\underline{x},w)e^{-i\mathcal{H}t} = <A(w),\ e^{i\hbar t}\Lambda(\underline{x})e^{-i\hbar t}A(w)> + <e^{i\hbar t}\Lambda(\underline{x}),\ A(w)> + <A(w),\ e^{i\hbar t}\Lambda(\underline{x})>.$$

It is convenient to define this as

$$\varphi(x,w) \equiv <A(w),\ \Lambda(x)\ A(w)> + <\overline{\Lambda}(x),\ A(w)> + <A(w),\ \overline{\Lambda}(x)>,$$

which introduces the notation $\Lambda(x)$, an operator on $\mathfrak{L}^2$, and $\overline{\Lambda}(x)$ an element of $\mathfrak{L}^2$ (of course when smeared). We emphasize that the field operator, for all times, is simply a bilinear expression in A and $A^{\dagger}$*.

---

* The space-time field $\varphi(x,w)$, and the given generators for $\mathcal{H}$ and $\underline{\mathcal{P}}$, would serve to characterize the theory in the event that space-time smearing is required. This permits, as in Chapter 4, the generalization to additional measures $\sigma(\Lambda)$. These may be different in that only $[\int\Lambda(\underline{x})g(\underline{x})d\underline{x}]^{\theta}$, $\theta>1$, may be square integrable near zero (rather than $\theta=1$). These measures may also be different if $\hbar$ cannot be constructed to act on $\mathfrak{L}^2$ but requires an enlarged space; i.e., the field at fixed time of the base theory is not cyclic. Even if this is the case, the given expressions for $\varphi(x,w)$, $\mathcal{H}$ and $\underline{\mathcal{P}}$ apply with the new measure.

Asymptotic Fields

The bilinear form for $\varphi(x,w)$ is in fact preserved in the asymptotic limit yielding "in" and "out" operators. Let "ex" denote either of these. Then we must have

$$\varphi_{ex}(x,w) = \langle A(w),\ \Lambda_{ex}(x)\ A(w)\rangle + \langle \bar{\Lambda}_{ex}(x),\ A(w)\rangle + \langle A(w),\ \bar{\Lambda}_{ex}(x)\rangle. \qquad (5\text{-}5)$$

In the simplest form of asymptotic theory we would associate the one particle states (say of mass m) with the one straton states formed from $\varphi_{ex}(x,w)\,|0\rangle$. The only term of (5-5) which contributes is the last and we find

$$\varphi_{ex}(x,w)\,|0\rangle = \langle A(w),\ \bar{\Lambda}_{ex}(x)\rangle\ |0\rangle \equiv \int \bar{\Lambda}_{ex}(x)\ d\sigma(\Lambda)\,|\Lambda,w\rangle.$$

Stability of the one particle state requires that

$$\bar{\Lambda}_{in}(x) = \bar{\Lambda}_{out}(x) \equiv \bar{\Lambda}_{o}(x).$$

The overlap of two such states is just

$$\langle 0|\varphi_{ex}(x,w)\varphi_{ex}(x',w')\,|0\rangle = \delta(w-w') \times \langle \bar{\Lambda}_{o}(x),\ \bar{\Lambda}_{o}(x')\rangle,$$

from which it follows that we should have

$$\int \bar{\Lambda}^{*}_{o}(x)\bar{\Lambda}_{o}(x')d\sigma(\Lambda) = -i\Delta_{+}(x-x';m^{2}).$$

Consider next the product $\varphi_{ex}(x,w)\varphi_{ex}(x',w')$ which contains a one straton creation operator term proportional to $\delta(w-w')$. To emphasize this term we appeal to our partial renormalization of the product. The vector

$$[\varphi_{ex}(x,w)\varphi_{ex}(x',w)]_{r}\,|0\rangle$$

has the one straton component

$$\langle A(w), \Lambda_{ex}(x)\bar{\Lambda}_o(x')\rangle \; |0\rangle$$

$$\equiv \int \Lambda_{ex}(x)\bar{\Lambda}_o(x')d\sigma(\Lambda)\,|\Lambda, w\rangle,$$

which we interpret as a two particle state. These states can differ for "in" and "out" and their overlap would correspond to a two-particle elastic scattering amplitude. From these examples one sees the way in which scattering states can in fact lie in the one straton subspace.

It should be noted that the asymptotic fields $\varphi_{ex}(x,w)$ are not canonical (c-number commutator) unless $\Lambda_{ex}(x) = 0$. This behavior is mandatory in order to have non-trivial scattering for otherwise $\varphi_{in}(x,w) = \varphi_{out}(x,w)$ and no scattering arises.

Indeed, the form of the S-matrix must be given by

$$S = e^{i\int dw \langle A(w),\, \eta A(w)\rangle},$$

which, as usual, must fulfill

$$S^{-1}\varphi_{in}(x,w)S = \varphi_{out}(x,w).$$

This is satisfied provided we have

$$e^{-i\eta}\Lambda_{in}(x)e^{i\eta} = \Lambda_{out}(x)$$

and

$$e^{i\eta}\bar{\Lambda}_o(x) = \bar{\Lambda}_o(x),$$

which are quite plausible.

Truncated Vacuum Expectation Values

On the basis of the bilinear form for the field operator, the truncated vacuum expectation values for the field may be readily calculated. It follows directly that

$$<0|\varphi(x_1,w_1)\ldots\varphi(x_n,w_n)|0>^T$$

$$= \delta(w_1-w_2)\ldots\delta(w_{n-1}-w_n)$$

$$\times <\bar{\Lambda}(x_1),\ \Lambda(x_2)\ldots\Lambda(x_{n-1})\bar{\Lambda}(x_n)>.$$

Moreover, if we introduce

$$\varphi(x) \equiv \int_{-\frac{1}{2}}^{\frac{1}{2}} \varphi(x,w)\,dw,$$

an operator appropriate to the stratum $|w|\leq\frac{1}{2}$, we find simply

$$<0|\varphi(x_1)\ldots\varphi(x_n)|0>^T$$

$$= <\bar{\Lambda}(x_1),\Lambda(x_2)\ldots\Lambda(x_{n-1})\bar{\Lambda}(x_n)>$$

$$\equiv \int \overline{\Lambda^*}(x_1)\Lambda(x_2)\ldots\Lambda(x_{n-1})\bar{\Lambda}(x_n)\,d\sigma(\Lambda). \qquad (5\text{-}6)$$

It is noteworthy that only one straton intermediate states appear in the construction of the truncated functions. If there are any asymptotic states (possibly requiring composite field operators), or even if there is no scattering theory at all, these facts can, in principle, be deduced from the truncated functions (5-6).

5.3 SPECULATION ON ADDITIONAL PROPERTIES

A number of additional properties of diastrophic quantum fields can be postulated on a heuristic basis by analogy with similar results for ultralocal models. We may imagine that the dynamics and field representation are linked through a functional differential equation "$\mathcal{K}1 = 0$" [similar to $\mathcal{K}c(\lambda) = 0$]. This would lead to joint expressions

$$\mathcal{H} = \int dw\ <A(w),\mathcal{K}A(w)>$$

$$= \int dw\ <B(w),\mathcal{K}B(w)>.$$

In the latter form we could expect to identify renormalized powers of the field characterizing the interaction potential appearing as monomials in $\Lambda(\underline{x})$ within $\mathcal{K}$ along with functional differential operators, $\delta/\delta\Lambda(\underline{x})$. If we suppose

that the highest-order functional differential operator in $\mathcal{H}'$ contributes the formal term

$$-\tfrac{1}{2}\int[\delta^2/\delta\Lambda^2(\underset{\sim}{x})]\,d\underset{\sim}{x},$$

then it would follow that

$$[\varphi(\underset{\sim}{x},w),\ \dot{\varphi}(\underset{\sim}{x}',w')]$$

$$= i\delta(\underset{\sim}{x}-\underset{\sim}{x}')\delta(w-w')\ \langle B(w),\ B(w)\rangle,$$

which has a leading c-number singularity

$$i\ \delta(\underset{\sim}{x}-\underset{\sim}{x}')\ \delta(w-w')\int d\sigma(\Lambda).$$

On this basis we also expect that $\dot{\varphi}(\underset{\sim}{x}',w')$ is only a form and not an operator when smeared with a test function $f(\underset{\sim}{x}',w')$.

The passage to the free, Fock theory of mass m_o could presumably be arranged along lines similar to those appropriate to the ultralocal models. For the field operator, for example, we could obtain convergence of the expectation functionals basically by a suitable scaling. Consider a hypothetical sequence of covariant measures $\sigma_\eta(\Lambda)$ with the property that

$$\int \Lambda(\underset{\sim}{x})\Lambda(\underset{\sim}{y})\,d\sigma_\eta(\Lambda) = -i\int \Delta_+(\underset{\sim}{x}-\underset{\sim}{y};\ m^2)\ \rho_\eta(m^2)\,dm^2$$

[$\rho_\eta(m^2)$ is the two-point spectral weight] converges to

$$\lim \int \Lambda(\underset{\sim}{x})\Lambda(\underset{\sim}{y})\,d\sigma_\eta(\Lambda) = -i\Delta_+(\underset{\sim}{x}-\underset{\sim}{y};m_o^2)$$

$$\equiv 2b(\underset{\sim}{x},\underset{\sim}{y}). \qquad (5\text{-}7)$$

Then it follows that the sequence

$$E_\eta(f) \equiv e^{-\int dw\int[1-\cos(\Lambda,f(w))]\eta^2 d\sigma_\eta(\eta\Lambda)}$$

$$= e^{-\int dw\int\{1-\cos[\eta^{-1}(\Lambda,f(w))]\}\eta^2 d\sigma_\eta(\Lambda)}$$

converges to

$$E_F(f) = e^{-\int dw(f(w),\ bf(w))}$$

where b is the kernel defined by (5-7). This is the appropriate expectation functional for the covariant diastrophic free field of mass m_o.

Pseudo-Free Models

The general formalism outlined above applies, in principle, to a number of covariant diastrophic theories* such as the φ^4 theory appearing in (5-2), or for other symmetric potentials. The difficulties in computation are not necessarily simpler in such a study -- namely determination of a suitable $\sigma(\Lambda)$ -- but they may be different ones. However, among all such models that can be studied two types stand out as prime candidates for early attack. In the ultralocal models, the pseudo-free cases were among the simplest possible. The analogue of the pseudo-free cases in the present context should yield a very different way to view a conventional free theory, one in which the interaction is vanishingly small but one in which its representation changing properties have, in large measure, been taken into account. Presumably such a pseudo-free theory should have no nontrivial scattering, but rather it serves to "prepare the ground." Adding interaction to the pseudo-free theory seems to be conceptually and practically easier than adding it to the free (Fock) theory, as we comment on in the next chapter. Indeed, in favorable cases, the field representations of the two (i.e. pseudo-free and interacting) are locally equivalent if the behavior of the ultralocal models carries over. It would seem that a study of pseudo-free, covariant, diastrophic quantum fields would be a useful enterprise.

A second class of interesting models could be the scale-invariant ones, which proved so useful in the study of the ultralocal fields. By analogy, these models would have a number of delicate technical questions, but again, if analogies hold true, the scale-invariant theories should correspond to the zero-mass, pseudo-free covariant diastrophic fields and may most likely be obtained from the pseudo-free models as a limit in which their only parameter -- the mass -- vanishes.

* And by implication to noncovariant diastrophic theories of many varieties although we have not discussed any as an illustration. We trust the reader's imagination to make the necessary changes.

6. THE CLASSIC FUNCTIONAL TECHNIQUES REVISTED

What is the relationship, if any, of the field theory solutions presented in these notes to the classic functional formalisms as presented in Chapter 1? For the sake of illustration we take the most understood of our models, the ultralocal models of Chapter 4. The usual functional formulations purport to "solve the problem"; but what is the true state of affairs? The explicit solutions we have obtained can shed some light on this situation.

Diagonalization of the Field

We take the simplest ultralocal models for which

$$E(f) = e^{-\int d\underline{x}\int\{1-\cos[\lambda f(\underline{x})]\}c^2(\lambda)d\lambda}$$

with a singularity parameter γ satisfying $\frac{1}{2} \leq \gamma < \frac{3}{2}$. According to the theorems presented in Chapter 3, it should be possible to realize E(f) in the form

$$E(f) = \int e^{i(\Lambda,f)} d\mu(\Lambda),$$

where $\Lambda(\underline{x}) \in \mathcal{V}'$. A heuristic, physical picture of the kind of elements in $\mathcal{V}'$ can be given on the basis of results abstracted from the theory of "shot noise." We suppose that we formally may write

$$\Lambda(\underline{x}) \equiv \sum_{j=1}^{\infty} \lambda_j \, \delta(\underline{x}-\underline{y}_j), \tag{6-1}$$

where λ_j, $\underline{y}_j$ are independent, and identical random variables for all j. The characteristic function for the distribution-valued stochastic variable $\Lambda(\underline{x})$ may be given by

$$\langle e^{i(\Lambda,f)} \rangle = \prod_{j=1}^{\infty} \langle e^{i\lambda_j f(\underline{y}_j)} \rangle_j,$$

which we interpret to mean

$$\lim_{N\to\infty} \{ \int e^{i\lambda f(\underline{y})} d\mu_N(\lambda,\underline{y}) \}^N$$

where μ_N is a sequence of normalized measures. The appearance of the N^{th} power is a consequence of the assumed equivalence of distributions for each j. Under appropriate conditions (which we adopt) it follows that

$$\lim_{N\to\infty} \{1 + \frac{N}{N} \int [e^{i\lambda f(y)} - 1] d\mu_N(\lambda, y)\}^N$$

$$= e^{\int [e^{i\lambda f(y)} - 1] d\sigma(\lambda, y)}$$

where $\sigma = \lim N\mu_N$. Independence of λ and y, uniformity of the distribution in y, and symmetry and absolute continuity of the λ distribution finally determine that

$$<e^{i(\Lambda, f)}> = e^{-\int dy \int \{1-\cos[\lambda f(y)]\} c^2(\lambda) d\lambda}$$

as desired [cf. Eq. (4-2)]. Consequently, we may formally regard

$$\Lambda(x) = \Sigma_j \lambda_j \delta(x - y_j)$$

as characteristic of the elements in $\mathcal{V}'$.

To make this realization more plausible we note that

$$\Lambda(x)\Lambda(x') = \sum_{j,k} \lambda_j \lambda_k \delta(x-y_j)\delta(x'-y_k)$$

$$= \sum_j \lambda_j^2 \delta(x-y_j)\delta(x'-y_j)$$

$$+ \sum_{j\neq k} \lambda_j \lambda_k \delta(x-y_j)\delta(x'-y_k).$$

Observe that the first term may also be written as

$$\delta(x-x') \Sigma_j \lambda_j^2 \delta(x-y_j)$$

which suggests that

$$\Lambda_r^2(x) = \Sigma_j \lambda_j^2 \delta(x-y_j)$$

corresponds to $\varphi_r^2(x)$. Higher order renormalized products are defined analogously.

Consider next the quantity

$$\dot{\Lambda}(x) \equiv \Sigma_j (-i\partial/\partial\lambda_j) \delta(x-y_j) \qquad (6\text{-}2)$$

which is formally a differential operator on the space $\mathcal{V}'$. It follows that

$$[\Lambda(\underline{x}),\dot{\Lambda}(\underline{x}')] = i\delta(\underline{x}-\underline{x}')\Sigma_j \delta(\underline{x}-\underline{y}_j).$$

However the quantity $H(\underline{x}) = \Sigma_j \delta(\underline{x}-\underline{y}_j)$ is not a meaningful random variable since (if $f \neq 0$)

$$<e^{i(H,f)}> = e^{-\int d\underline{x}\int\{1-\cos[f(\underline{x})]\}c^2(\lambda)d\lambda} = 0$$

due to the divergence of $\int c^2(\lambda)d\lambda$.

Within this framework, the (functional) differential representation of the Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H} = \int\{\tfrac{1}{2}[\dot{\Lambda}_r^2(\underline{x}) + \gamma(\gamma+1)\Lambda_r^{-2}(\underline{x}) + 2e + m_0^2\Lambda_r^2(\underline{x})] + (V[\Lambda(\underline{x})])_r\}d\underline{x}. \qquad (6\text{-}3)$$

Expressed in terms of the realization (6-1) and (6-2) we determine that

$$\mathcal{H} = \int\{\Sigma_j[-\tfrac{1}{2}\frac{\partial^2}{\partial\lambda_j^2} + v(\lambda_j)]\ \delta(\underline{x}-\underline{y}_j)\}d\underline{x}$$

$$= \Sigma_j[-\tfrac{1}{2}\frac{\partial^2}{\partial\lambda_j^2} + v(\lambda_j)].$$

The similarity to the differential operator for $\mathcal{H}$ is clear, but, of course, the interpretation is quite different.

The expressions given above represent the best interpretation of a diagonalization of the field (at t=0), and a realization of the Hamiltonian as a functional differential operator. It should be noted that the class of elements composing $\mathcal{V}'$ is of basic importance for it determines the form of the renormalized product. Moreover, the operator $-i\delta/\delta\Lambda(\underline{x})$ actually does not appear in $\mathcal{H}$, but its role is assumed by the formally similar quantity $\dot{\Lambda}(\underline{x})$. Adding these features to the need for the unusual and unexpected field renormalization term $\frac{1}{2}\gamma(\gamma+1)\Lambda_r^{-2}(\underline{x})$ drastically reduces the value of the straightforward, commonly assumed functional form like (1-3).

T-Product Generating Functional

Consider another of the functional formulations of Chapter 1. For the ultralocal models ($\frac{1}{2} \leq \gamma < \frac{3}{2}$) let us construct

$$\Omega\{j\} = \langle 0|T\, e^{i\int \varphi(\underline{x},t) j(\underline{x},t) d\underline{x}dt}|0\rangle \qquad (6\text{-}4)$$

where $j(\underline{x},t)$ is a smooth function. This expression can be evaluated for the ultralocal models, and one finds

$$\Omega\{j\} = \exp\left(\int d\underline{x} \int c(\lambda)\{T\, e^{i\int \lambda(t) j(\underline{x},t)dt} - 1\} c(\lambda) d\lambda\right),$$

where $\lambda(t) = e^{i\hbar t \lambda} e^{-i\hbar t}$. Strictly speaking, this expression is not precise and should be supplemented by the rule that $c(\lambda)$ is a formal left and right eigenvector for $e^{i\hbar t}$ with eigenvalue one. Thus a term like

$$\int c(\lambda)\lambda(t_1)\lambda(t_2)\ldots\lambda(t_{n-1})\lambda(t_n)c(\lambda)d\lambda$$

is understood to mean

$$\int c(\lambda)\lambda e^{-i\hbar t_1}\lambda(t_2)\ldots\lambda(t_{n-1})e^{i\hbar t_n}\lambda c(\lambda)d\lambda .$$

We discuss only a few questions based on the formula for $\Omega\{j\}$. Suppose one had a particular solution corresponding to one model and wished to consider the perturbation to another model with a different potential. For example, assume that we knew $\hat{\Omega}\{j\}$ based on the pseudo-free model function $\hat{c}(\lambda) = |\lambda|^{-\gamma} \exp(-\frac{1}{2}\mu\lambda^2)$. Considering the interaction term $g\lambda^4$ to be a perturbation in the potential we could expect that a relation like (1-4) holds, namely

$$\Omega\{j\} = N\, e^{-ig\int Z^3\delta^4/\delta j(x)^4 dx}\, \hat{\Omega}\{j\}$$

where $Z^{-1} = \delta(0)$ denotes the factor associated with renormalized field powers. To be more precise we anticipate that $\Omega\{j\}$ is the limit as $g(x) \to g$, where $g(x) \in C_0^\infty$, and $g(x) \geq 0$, of

$$\Omega_g\{j\} = N_g\, e^{-i\int g(x) Z^3\delta^4\delta/j(x)^4 dx}\hat{\Omega}\{j\} . \qquad (6\text{-}5)$$

This latter expression can be made meaningful provided the field representations for $\hat{c}(\lambda)$ and $c(\lambda)$ are <u>locally</u>

equivalent, and the condition for this to hold is just $\{c(\lambda) - \hat{c}(\lambda)\} \in L^2$. In that case N_g is neither 0 nor ∞ and is chosen so that $\Omega_g\{0\} = 1$.

We define the operator

$$\mathcal{G}_g \equiv \int g(x)\, Z^3 \delta^4 / \delta j(x)^4 dx$$

by its action on functionals of the form

$$\mathcal{F} \equiv \Sigma_n (n!)^{-1} \int\int A_n(y; t_1, \ldots, t_n) j(y,t_1) \ldots j(y,t_n) dy dt_1 \ldots dt_n \tag{6-6}$$

to be

$$\mathcal{G}_g \mathcal{F} \equiv \Sigma_n (n!)^{-1} \int \ldots \int A_{n+4}(y; t, t, t, t, t_1, \ldots, t_n) g(y,t) \times j(y,t_1) \ldots j(y,t_n) dy dt_1 \ldots dt_n. \tag{6-7}$$

The inclusion of the space-time cutoff $g(x)$ removes any support requirements on the coefficients A_n; in fact, in the case of interest A_n, $n>0$, is independent of y. Certain of the functionals $\mathcal{F}$ would be analytic functionals for $\mathcal{G}_g$ such that $e^{-i\mathcal{G}_g}\mathcal{F} \equiv \Sigma (m!)^{-1} (-i)^m \mathcal{G}_g^m \mathcal{F}$. The operation $e^{-i\mathcal{G}_g}$ can be extended by linearity and continuity to additional functionals which need no longer be analytic for $\mathcal{G}_g$. It seems reasonable that $\hat{\Omega}\{j\}$ is a functional of this type, and that (6-5) holds true. Finally $\Omega\{j\} = \lim \Omega_g\{j\}$ as $g(x) \to g$.

One advantage of this picture is that the unusual renormalization λ^{-2} in $\hat{\mathcal{K}}$ is already included in $\hat{\Omega}$, and only the genuine interaction term λ^4 need be explicitly displayed in constructing Ω. Yet we know this to be true only because we already have the operator solution to draw on. What exact form the renormalized interaction assumes cannot be foretold in the general case.

It is noteworthy that there are many functionals on which $e^{-i\mathcal{G}_g}$ acts in a trivial fashion. For example, on the basis of our earlier definition

$$e^{-i\mathcal{G}_g} e^{i(j,f)} = e^{i(j,f)}$$

where

$$(j,f) = \int j(\underset{\sim}{x},t)\ f(\underset{\sim}{x},t)d\underset{\sim}{x}dt$$

and $f\in C_0^\infty$, say. One may be tempted to extend this trivial action to other functionals by linearity and continuity -- but what is the appropriate choice of topology? For example, an extension of the trivial action, consistent with (6-7), can be made to the functional

$$\Omega_o\{j\} = e^{-\int j(\underset{\sim}{x},t_1)b(t_1,t_2)j(\underset{\sim}{x},t_2)dt_1dt_2d\underset{\sim}{x}},$$

which is the form of (6-4) appropriate to the free Fock representation. Yet we would not like to extend, by continuity, a trivial action to the functional $\hat{\Omega}\{j\}$. What is the distinction? In the Hilbert space characterized by $\Omega_o\{j\}$ [e.g., in the functional Hilbert space constructed (at sharp time) in the fashion of Section 3.2] the action of $\mathcal{O}_g=0$. On the other hand, $\mathcal{O}_g$ is nonzero in other functional Hilbert spaces such as that characterized by $\hat{\Omega}\{j\}$. Any superficial "extension by continuity" would tend to gloss over the delicate distinction of different functional Hilbert spaces. Clearly, considerable care must be exercised in the general problem in finding the proper definition of an interaction operator.

Speculation on Covariant Models

Let us once again consider $\Omega\{j\}$ as given in (6-4). In analogy with (1-4) we wish to explore the possibility of restoring the (spacial gradient)^2 to the ultralocal models thereby giving a covariant theory. Formally, we might anticipate that the covariant functional $\Omega_c\{j\}$ would be the limit of the expressions

$$\Omega_{cg}\{j\} = N_{cg}\ e^{i\frac{1}{2}\int g(x)[\underset{\sim}{\nabla}\delta/\delta j(x)]_r^2\,dx}\Omega\{j\} \qquad (6\text{-}8)$$

as $g(x)\to 1$, where r stands for some unknown renormalization. For this expression to be useful we must somehow be able to define

$$\overline{\mathcal{O}}_g \equiv -\tfrac{1}{2}\int g(x)[\underset{\sim}{\nabla}\delta/\delta j(x)]_r^2 dx$$

on functionals of the form (6-6). If this could be accomplished we sould have the ingredients of a covariant theory, or so it would seem. Unfortunately, no natural, straightforward definition has been found, and one is reluctant to introduce momentum-space cutoffs blindly just to get a definition. After all, momentum-space cutoffs make the ultralocal interaction terms compatible with the Fock representation; but there is no hope for such a scheme because one is ignorant of the unexpected yet necessary operator renormalization [i.e., $\Lambda_r^{-2}(x)$ as given in (6-3)]. It is not known whether or not unexpected renormalization terms would be needed in restoring the spacial gradient term, and if so what form they would take. On the other hand, the inadequacy of the scheme represented by (6-8) is by no means certain, and perhaps it should not be dismissed so lightly.

Conclusion

It is hoped that the reader may have found in the material presented in these notes further evidence for the continuing appeal of functional techniques. Few can deny the potential power and diversity of both the old and the ever-enlarging new functional techniques in attacking the fundamental problems of quantum field theory. So numerous are the unsolved problems and so challenging are their solutions that one feels a kindred relation with the Colorado miner of yester-year in his tireless search for the tiny deposits of precious ore deeply hidden the vast wilderness. One can only hope, as he did then, that "Surely, there must be gold in them thar hills."

REFERENCES

Chapter 1: There are a number of classic sources for the traditional functional methods commonly used in quantum field theory. Among these we might note the textbooks

Berezin, F.A., The Method of Second Quantization, translated by N. Mugibayashi and A. Jeffrey (Academic Press, New York, 1966).

Bogoliubov, N.N., and D. V. Shirkov, Introduction to the Theory of Quantized Fields, translated by G. M. Volkoff (Interscience, New York, 1959).

Novozhilov, Iu. V. and A. V. Tulub, Method of Functionals in the Quantum Theory of Fields, translated from the Russian (Gordon and Breach, New York, 1961).

Rzewuski, J., Field Theory, Vol. 2, translated from the Polish (Iliffe, London, 1968).

Schweber, S.S., An Introduction to Relativistic Quantum Field Theory (Harper & Row, New York, 1962).

The model field theories with classical Hamiltonian given in Eq. (1-5) with $N=\infty$ are formulated and solved in

Klauder, J.R., "Rotationally-Symmetric Model Field Theories," J. Math. Phys. 6, 1666 (1965).

The demonstration that these models cannot be solved by taking the limit of the cutoff models as defined by Eq. (1-5) for each N, as $N\to\infty$, is discussed in

Abarbanel, H.D.I., J. R. Klauder and J. G. Taylor, "Green's Functions for Rotationally Symmetric Models," Phys. Rev. 152, 1198 (1966).

Chapter 2: A number of texts on functional analysis discuss Hilbert space techniques, and among these we cite

Kato, T., Perturbation Theory for Linear Operators, (Springer-Verlag, New York, 1966).

Riesz, F. and B. Sz.-Nagy, Functional Analysis (F. Ungar, New York, 1955).

Yosida, K., Functional Analysis (Springer-Verlag, New York, 1968), Second Edition.

Our discussion on probability theory and characteristic functions draws heavily on the text

Lukacs, E., Characteristic Functions (Hafner, New York, 1970), Second Edition.

Coherent states and their various properties are extensively discussed in

Klauder, J.R. and E.C.G. Sudarshan, Fundamentals of Quantum Optics (Benjamin, New York, 1968).

Chapter 3: The characterization of group representations by expectation functionals is extensively discussed in the text

Naimark, M.A., Normed Rings, translated by L.F. Boron (Noordhoff, Groningen, 1964).

The classic text for the functional Fourier representations of positive-definite functionals, nuclear spaces and the like is

Gel'fand, I.M. and N. Ya. Vilenkin, Generalized Functions Vol. 4: Applications of Harmonic Analysis, translated by A. Feinstein (Academic Press, New York, 1964).

Our discussion of infinitely divisible representations of field algebras is based on the article

Klauder, J.R., "Exponential Hilbert Space: Fock Space Revisited," J. Math. Phys. 11, 609 (1970).

The canonical representation quoted in Eq. (3-8) appears in the article by

Tortrat, A., "Sur la structure des lois indefiniment divisibles (class $\mathcal{J}(X)$) dans les espaces vectoriels X(sur le corps reel)," Z. Wahrscheinlichkeitstheorie Verw. Geb. 11, 311 (1969).

Chapter 4: The idea to study covariant models in which the spacial gradient is dropped from the Hamiltonian is not new. Other, earlier attempts to solve these models are unsatisfactory in one or another respect. Our discussion follows, and, with regard to scale transformations extends, the work presented in

Klauder, J.R., "Ultralocal Scalar Field Models," Commun. Math. Phys. 18, 307 (1970);

Klauder, J.R., "Ultralocal Quantum Field Theory," Acta Phys. Austr., Suppl. VIII, 227 (1971);

Klauder, J.R., "Heisenberg Operator Equations of Motion," Proceedings of the Symposium on Basic Questions in Elementary Particle Physics, Munich, June 8-18, 1971.

References to earlier treatments of ultralocal models appear in the first two references cited above.

* * * * * * * *

Since the preparation of these lecture notes, several additional articles have appeared that pertain to the subjects treated. For ultralocal models (Chapter 4) the fact that there is no self-adjoint local momentum operator canonically conjugate to the field operator (independent of any dynamical consideration whatsoever) is proved in

Hegerfeldt, G. C. and J. R. Klauder, "Fields without Partners," Nuovo Cimento, 10A, 723 (1972).

A simplified discussion of diastrophic field theories (Chapter 5) that leads rather directly to many of the results presented in these lectures appears in

Klauder, J. R., "Covariant Diastrophic Quantum Field Theory," Phys. Rev. Letters 28, 769 (1972).

This short note also outlines the essential operator structure for diastrophic spinor (i.e., fermion) fields and shows that such fields require the coexistence of boson fields for their construction. A fuller account of ultralocal fermion fields appears in

Klauder, J. R., "Ultralocal Spinor Field Models," (submitted for publication).

The argument presented in Chapter 4 that the fundamental field operators of interest are bilinear expressions in the basic annihilation and creation operators is rigorously established for both boson and fermion fields in

Ruskai, M. B. and J. R. Klauder, "Local Products as Operators," (submitted for publication).

LECTURES ON THE THERMODYNAMIC LIMIT FOR COULOMB SYSTEMS

Elliott H. Lieb*
Massachusetts Institute of Technology
Cambridge, Massachusetts

and

Joel L. Lebowitz+
Yeshiva University
New York, N.Y.

I. Introduction

These lecture notes present an outline of the proof of the existence of the thermodynamic limit for Coulomb systems. A brief statement of the main results has appeared previously, Lebowitz and Lieb (1969), and the full work will appear shortly, Lieb and Lebowitz (1971). What we have tried to do in these notes is to present the ideas and methods used in constructing this proof while leaving out most of the details of the analysis. In some places, such as section III, we treat only the simplest kind of Coulomb system: two species of charged particles (one positive and one negative) whose only interaction is through the Coulomb potential. In other places we simply state various lemmas and theorems without proof.

The basic pre-requirement for the existence of a thermodynamic limit for Coulomb systems in the Dyson-Lenard Theorem, Dyson-Lenard (1965), which gives a lower bound to the energy of a system of charged particles. It is therefore very fortunate that the proof of this theorem is presented in a particularly nice form, in Professor Lenard's lectures which are included in this volume.

*Supported in part by National Science Foundation Grant GP26526
+Supported in part by AFOSR #44620-71-C-0013

Statement of the Problem

Statistical Mechanics as developed by Gibbs and others rests on the hypothesis that equilibrium properties of matter can be completely described in terms of a phase-space average, or partition function, $Z = \mathrm{Tr}\{\exp(-\beta H)\}$, with H the Hamiltonian and β the reciprocal temperature. It was realized early that there were grave difficulties in justifying this assumption in terms of basic microscopic dynamics. These questions, which involve the time evolution of macroscopic systems, have still not been satisfactorily resolved, but the great success of equilibrium statistical mechanics in offering qualitative and quantitative equilibrium explanations for such varied phenomena as superconductivity, specific heats of crystals, chemical equilibrium constants, etc. have left little doubt about the essential correctness of the partition function method. However, since Z cannot be evaluated explicitly for any reasonable physical Hamiltonian H, comparison with experiment always involves some uncontrolled approximations. Hence, the following problem deserves attention: Is it true that the thermal properties of matter obtained from an exact evaluation of the partition function would be extensive and otherwise have the same form as those postulated in the science of thermodynamics? In particular, does the thermodynamic, or bulk, limit exist for the Helmholtz free energy/unit volume derived from the partition function, and if so, does it have the appropriate convexity, i.e., stability properties?

To be more precise: Let $\{\Lambda_j\}$ be a sequence of bounded open sets (domains) in R^d with Λ_j becoming infinitely large as $j \to \infty$ in some 'reasonable way' which will be specified later. (We shall be concerned primarily with $d = 3$ but many of our results are valid for all d). The volume (Lebesgue measure) of Λ_j will be denoted by $V(\Lambda_j)$ and $V(\Lambda_j) \to \infty$ as $j \to \infty$. Consider now a sequence of systems consisting of S species of particles in the domains $\{\Lambda_j\}$. Let $\underline{N}_j = (N_j^1, \ldots, N_j^S)$ be the particle number vector specifying the system in Λ_j, i.e. N_j^i is a non-negative integer and is the number of particles of species i contained in Λ_j. The canonical partition function of the jth system at reciprocal temperature β is then given by

$$\text{by} \qquad Z(\beta, \underset{\sim}{N}_j ; \Lambda_j) = \sum_{\alpha=0}^{\infty} \exp\left[-\beta E_\alpha (\underset{\sim}{N}_j ; \Lambda_j)\right]$$
$$\equiv \exp\left[V(\Lambda_j)\, g(\beta, \underset{\sim}{\rho}_j ; \Lambda_j)\right] \tag{1.1}$$

where $E_\alpha (\underset{\sim}{N}_j ; \Lambda_j)$ are the energy levels of the jth system, $\underset{\sim}{\rho}_j \equiv \underset{\sim}{N}_j / V(\Lambda_j)$ is the particle density vector, and $-\beta g (\beta, \underset{\sim}{\rho}_j ; \Lambda_j)$ is the Helmholtz free energy/unit volume of the jth system. According to statistical mechanics, knowledge of g determines all the equilibrium properties of this system. The question to be studied is the following: Given a sequence of particle density vectors $\{\underset{\sim}{\rho}_j\}$ which approach a limit $\underset{\sim}{\rho}$ as $j \to \infty$, does $g(\beta, \underset{\sim}{\rho} ; \Lambda_j)$ approach a limit, $g(\beta, \underset{\sim}{\rho})$, as $j \to \infty$ and is this limit in some sense of the particular sequence of domains $\{\Lambda_j\}$ and density vectors $\{\underset{\sim}{\rho}_j\}$ used in going to the limit? If so, does the limiting free energy density have, as a function of $\underset{\sim}{\rho}$ and β the convexity properties required for thermodynamic stability, i.e. is $g(\beta, \underset{\sim}{\rho})$ convex in β and concave in $\underset{\sim}{\rho}$? (With regard to β, we see from (1.1) that each $g(\beta, \underset{\sim}{\rho}_j ; \Lambda_j)$ is convex in β. Therefore, if this limit $g(\beta, \underset{\sim}{\rho})$ exists it will automatically be convex in β. Consequently we can set $\beta = 1$ and omit mention of β, and shall do so henceforth.)

The proof of the above for the free energy obtained from the canonical ensemble and the proof that the 'same' results are obtained, in the thermodynamic limit, from the microcanonical and grand canonical ensembles as well, have come to he recognized (by some people) as one of the basic goals of statistical mechanics and is referred to as proving the existence of the thermodynamic limit.

Background: Tempering and the Coulomb Potential

Various authors have evolved a technique for proving the existence of the thermodynamic limit for systems whose Hamiltonians satisfy certain conditions. (The different names associated with this development are: Van Hove, Lee and Yang, van Kampen, Wills, Mazur and van der Linden, Griffiths, and in particular Ruelle and Fisher. The reader is referred to Fisher (1964) and Ruelle (1969) for an exposition and references. For a synopsis and more

references see also Lebowitz (1968) and Griffiths (1971).) In particular it was necessary to assume that the interaction between the particles constituting the microscopic units of macroscopic matter were short range or 'tempered'. This means that there exists a fixed distance $r_o \geq 0$ and constants $C \geq 0$ and $\epsilon > 0$ such that the inter-domain interaction potential energy between N_1 particles in a domain Λ_1, N_2 particles in a domain Λ_2, ..., and N_K particles in a domain Λ_K, has a bound in terms of the minimum distances r_{ij} between Λ_i and Λ_j,

$$I(N_1, \ldots, N_K) \equiv U(N_1 \oplus \ldots \oplus N_K) - \sum_{i=1}^{K} U(N_i) \leq C \sum_{i \neq j}^{K} r_{ij}^{-(d+\epsilon)} N_i N_j \tag{1.2}$$

whenever $r_{ij} \geq r_o$ for all $i \neq j$. We have written here $U(N) \equiv U(\underset{\sim}{x}_1, \ldots \underset{\sim}{x}_N)$ for the total potential energy of N particles at positions $\underset{\sim}{x}_\ell \in R^d$. (We shall generally not indicate that the particles belong to different species when this is not essential and shall denote $\sum_1^S N_\alpha$ by N).

The requirement of tempering unfortunately excludes the Coulomb potential which is the true potential relevant for real matter. That a nice thermodynamic limit exists for systems with Coulomb forces is a fact of common experience, but the proof that it does so is a much more subtle matter than for short range forces. It is screening, brought about by the long range nature of the Coulomb force itself, that causes the Coulomb force to behave as if it were short range. This has the consequence, as we shall prove in these notes, that when the sequence of systems are overall neutral then the approach of $g(\underset{\sim}{\rho}_j ; \Lambda_j)$ to its limit $g(\rho)$ and the properties of $g(\rho)$ are the same as those obtained for systems with tempered interactions (except that the ρ^i, $i = 1, \ldots, S$ are constrained by the neutrality requirement). In particular $g(\rho)$ is the same for different 'shapes' of the domains in $\{\Lambda_j\}$. This shape independence disappears when the constraint of charge neutrality is lifted and systems with a 'non-negligible' amount of net charge are considered. The true long range nature of the Coulomb force now becomes manifest, leading in some cases to a shape

dependent limit of the free energy density and in other cases (when the excess charge is large) to an infinite limit (cf section IV).

Background: H-stability and the Dyson-Lenard Theorem

The basic condition on the N body Hamiltonian

$$H(N) = \sum_{\ell=1}^{N} (P_\ell^2) + U(\underset{\sim}{x}_1, \ldots \underset{\sim}{x}_N) , \tag{1.3}$$

where m_ℓ is the mass and $\underset{\sim}{p}_\ell$ the momentum (momentum operator in quantum mechanics) of the ℓth particle, required for the existence of thermodynamics is that there exists a constant $B < \infty$, such that for all N

$$E_o(N) \geq -BN . \tag{1.4}$$

Here $E_o(N)$ is the ground state energy of the N particle system in infinite space, $\underset{\sim}{x}_i \in R^d$, defined by

$$E_o(N) = \inf_{\Psi} [(\Psi, H(N) \Psi)/(\Psi , \Psi)] \tag{1.5}$$

with the $\Psi (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N)$ elements of a properly constructed Hilbert space in which H(N) is a self-adjoint operator. The functions $\Psi(\underset{\sim}{x}_1 \ldots, \underset{\sim}{x}_N)$ have to satisfy the proper symmetry relations whenever the coordinates of two particles belonging to the same species are interchanged: $\Psi \to \Psi$ or $\Psi \to -\Psi$ for bosons or fermions respectively. (Since the spin does not appear directly in the Hamiltonian we can, and do, treat particles of the same type having different values of their spins in the z-direction as belonging to different species.)

We shall refer to condition (1.4) as H-stability. Heuristically, H-stability insures against collapse of the system. Mathematically it provides an upper bound to the sequence $\{g(\underset{\sim}{\rho}_j ; \Lambda_j\}$ and this bound plays an essential role in the proof. It should be emphasized however that H-stability does not in itself imply a thermodynamic limit. As an example, it is trivial to prove H-stability for charged particles all of one sign, and it is equally obvious that the thermodynamic limit does not exist in that case.

We shall refer to condition (1.4) as H-stability. Heuristically, H-stability insures against collapse of the system. Mathematically it provides an upper bound to the sequence $\{g(\rho_j ; \Lambda_j)\}$ and this bound plays an essential role in the proof. It should be emphasized however that H-stability does not in itself imply a thermodynamic limit. As an example, it is trivial to prove H-stability for charged particles all of one sign, and it is equally obvious that the thermodynamic limit does not exist in that case.

To satisfy (1.4) it is clearly sufficient that the potential energy U (N) by itself have a lower bound of the same form:

$$\text{Inf } U(\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N) \geq -NB, \text{ all } \underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N . \quad (1.6)$$

Indeed for a classical system (1.6) is also necessary for (1.4). There are a large variety of interaction potentials for which the existence of the lower bound (1.6) can be verified explicitly. The simplest of these is the case when U(N) can be written as the sum of a positive term and a term consisting of a sum of pair potentials $v(\underset{\sim}{x}_i - \underset{\sim}{x}_j)$ which is bounded below and has the asymptotic behaviors,

$$\lim_{|\underset{\sim}{r}| \to 0+} |\underset{\sim}{r}|^{d+\delta} v(\underset{\sim}{r}) \to + \text{ and } \lim_{|\underset{\sim}{r}| \to \infty} |\underset{\sim}{r}|^{d+\epsilon} v(r) \geq 0 \quad (1.7)$$

for some $\delta > 0$ and $\epsilon > 0$. (This result is due to Morrey (1955) who appears to have been the first to consider bounds of the form (1.6) for non-Coulomb potentials.) More general types of potentials satisfying (1.6) have been considered by other authors [3], [17] .

It is clear, however, that (1.6) will not be satisfied by a system of point charges with charges q_i of different signs, i = 1,..., N . The interparticle Coulomb potential has the form, for d = 4,

$$U_c(\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N) = \frac{1}{2} \sum_{i \neq j} q_i q_j |\underset{\sim}{x}_i - \underset{\sim}{x}_j|^{-1}, \quad (1.8)$$

and the potential energy of even a single pair of oppositely charged particles has no lower bound. Interestingly though, if the particles have har cores, i.e. U(N)

contains in addition to its Coulomb part (1.8) a term which is $+\infty$ if $|\underset{\sim}{x}_i - \underset{\sim}{x}_j| < R$, then Onsager (1939) showed that (1.6) is satisfied. Onsager's proof is so simple that we shall present it here (in a form communicated to us by Penrose*): Since the particles cannot approach each other any closer than a distance R, the effect of the Coulomb interaction between the particles will be the same, Newton (1687), if the charge of each particle is distributed in any spherically symmetric way within a ball of radius $\frac{1}{2}R$ centered on the position of that particle, e.g. a uniform charge density. Now, as is well-known from electrostatics, (Kellog 1929),

$$\tfrac{1}{2} \sum_{i \neq j} q_i q_j |\underset{\sim}{x}_i - \underset{\sim}{x}_j|^{-1} = \tfrac{1}{2} \int_{R^d} \underset{\sim}{E}^2(\underset{\sim}{x})\, d\underset{\sim}{x} - \sum_{i=1}^{N} (\text{self energy of the } i\text{th particle}) \geq -NB , \tag{1.9}$$

where $\underset{\sim}{E}(\underset{\sim}{x})$ is the electrostatic field, and B is the maximum self energy of any of the balls.

Onsager's results were generalized somewhat by Fisher and Ruelle (1966). This work, however, still left open the question of whether a system of point Coulomb charges, which may be taken as the building blocks of real matter, has a lower bound of the form (1.4). Now when dealing with a quantum system of charges the non-existence of a lower bound to $-|\underset{\sim}{x}_i - \underset{\sim}{x}_j|^{-1}$ might appear not as serious as in the classical case since we expect that the Heisenberg uncertainty principle, which prevents particles from having their positions 'close to each other' without also having a large kinetic energy, will insure the existence of a lower bound to the ground state energy. This is indeed the case for any finite system, (-13.5 electron volts for a system composed of one electron and one proton), and generally $E_o(N) > -\infty$, for any N, Kato (1966). We need however a bound proportional to N and this, it turns out, the uncertainty principle alone cannot provide. The required result was proven by Dyson and Lenard (1967-8), who showed that (1.4) holds for a system of point Coulomb charges when all species with negative and/or positive charges are fermions. This

*See also Penrose's comments [16] on using electromagnetic energy considerations to establish the thermodynamic limit for charged and magnetic systems.

is happily the case in nature where the electrons are fermions. (When neither of the charges are fermions, Dyson (1967) found an upper bound to the ground state energy that is proportional $-N^{7/5}$; hence such a system will not be thermodynamically stable.)

We note here that the Dyson-Lenard lower bound is valid whenever the masses of the fermion particles are finite (the masses only affect the numerical value of B which is of no interest here). Hence it remains valid if the kinetic energy term in the Hamiltonian is multiplied by some δ , $0 < \delta < 1$, e.g. $\delta = \frac{1}{2}$.

Basic Inequalities and Outline of the Notes

Let us consider a system of $\underset{\sim}{N} = (N^1, \ldots, N^s)$ particles in a domain Λ with a Hamiltonian $H(N ; \Lambda)$

$$H(N ; \Lambda) = -\tfrac{1}{2}\hbar^2 \sum_{i=1}^{N} (m_1)^{-1}\Delta_i + U_c (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N) + U_T (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N) . \tag{1.10}$$

Here $N \equiv \sum_{\alpha=1}^{s} N^\alpha$, and $\underset{\sim}{x}_i \in \Lambda$ is the coordinate of a particle of species one for $1 \le i \le N^1$, and of a particle of species two for $N^1 < i \le N^1 + N^2$, etc.. $U_c (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N)$ is the Coulomb potential defined in (1.8), so that $m_i = m_1$, $q_i = e_1$ for $1 \le i \le N^1$, etc. with m_α and e_α , $\alpha = 1, \ldots, S$, the mass and charge of a particle of the αth species. $U_T (\underset{\sim}{x}_1, \ldots, \underset{\sim}{x}_N)$ is a tempered and stable potential satisfying (1.2) and (1.6) (which is also translationally and rotationally invariant). It is not altogether useless to include tempered potentials along with the true Coulomb potentials because one might wish to consider model systems in which ionized molecules are the elementary particles. Although we shall omit U_T in most of these notes, it should be understood that all the stated theorem are valid for the full Hamiltonian (1.10). $H(N ; \Lambda)$ is a self-adjoint operator, defined via the Friedrichs extension. (In the physicists language this corresponds to using a Hilbert space in which the wave functions vanish on the boundary of Λ .) When the statistics of the particles satisfy the

conditions of the Dyson-Lenard theorem then $H(N ; \Lambda)$ will satisfy the inequality (note the factor ½)

$$H(N ; \Lambda) \geq -\frac{\hbar^2}{4} \sum_{i=1}^{N} (m_i)^{-1} \Delta_i - N \Phi \tag{1.11}$$

with Φ some constant, $\Phi < \infty$. The canonical partition function $Z(\underset{\sim}{N} ; \Lambda)$ and the corresponding $g(\underset{\sim}{\rho} ; \Lambda)$ will have the bounds

$$Z(\underset{\sim}{N} ; \Lambda) \leq \exp [N\Phi] \prod_{\alpha=1}^{s} Z_{o,\alpha} (N^{\alpha} ; \Lambda) \tag{1.12a}$$

$$g(\underset{\sim}{\rho} ; \Lambda) \leq \rho \Phi + \sum_{\alpha=1}^{s} g_{o,\alpha} (\rho^{\alpha} ; \Lambda) \tag{1.12b}$$

where $\rho \equiv \sum_{\alpha=1}^{s} \rho^{\alpha}$ and $z_{o,\alpha}$ (resp. $g_{o,\alpha}$) is the partition function (resp.-free energy/unit volume) of an ideal gas (fermion or boson according to the statistic of species α) of particles with masses $m_{\alpha}' = 2m_{\alpha}$. The inequality (1.13) readily yields a uniform bound on any sequence $\{g(\underset{\sim}{\rho}_j ; \Lambda_j)\}$ whenever the $\underset{\sim}{\rho}_j$ are in a compact subset of R^s, (with $\rho_j^{\alpha} \geq 0$) .

We now give a sketch of the method used in our proof. As usual, one first proves the existence of the limit for a standard sequence of domains. The limit for an arbitrary domain is then easily arrived at by packing that domain with the standard ones. The basic inequality that is needed is that if a domain Λ contains K disjoint sub-domains $\Lambda_1, \Lambda_2, \ldots, \Lambda_K$ and if the inter-domain interaction be neglected then

$$Z(\sum_{i=1}^{K} \underset{\sim}{N}_i ; \Lambda) \geq \prod_{i=1}^{K} Z (\underset{\sim}{N}_i ; \Lambda_i) \tag{1.13a}$$

or

$$g (\sum_{i=1}^{K} f_i \underset{\sim}{\beta}_i ; \Lambda) \geq \sum_{i=1}^{K} f_i \, g(\underset{\sim}{\rho}_i ; \Lambda_i) \tag{1.13b}$$

where $f_i \equiv V(\Lambda_i)/V(\Lambda)$ is the fraction of the volume of Λ occupied by Λ_i. If the distance between every pair of sub-domains is not less than r_o , one can use (1.2) to obtain a useful bound on the tempered part of the omitted interdomain interaction energy, $I_T(N_1, \ldots, N_K)$.

The normal choice for the standard domains are cubes Γ_j , with Γ_{j+1} being composed of 2^d copies of Γ_j, together with 'corridors'. One chooses $\underset{\sim}{N}_{j+1} = 2^d \underset{\sim}{N}_j$. Neglecting I_T one would have from (1.14) that $g(\underset{\sim}{\rho}_{j+1}\ ;\ \Gamma_{j+1}) \geq g(\underset{\sim}{\rho}_j\ ;\ \Gamma_j)$ which, since $g(\underset{\sim}{\rho}_\ell\ ;\ \Gamma_\ell)$ is bounded above implies the existence of a limit. To justify neglect of I_T one makes the corridors increase in thickness with increasing j ; although V_j^c , the corridor volume, approaches ∞ one makes $V_j^c/V_j \to 0$ in order that $\Sigma_1^\infty f_j \to 1$ as $j \to \infty$. The positive ϵ of (1.2) allows one to accomplish these desiderata.

Obviously, such a strategy will fail with Coulomb forces, but fortunately there is another way to bound the inter-domain energy. The essential point is that it is not necessary to bound this energy for all possible states of the systems in the sub-domains; it is only necessary to bound the 'average' interaction between domains which is much easier. This is expressed mathematically by using the Peierls-Bogoliubov inequality (Jensen's inequality in the classical case) to show that for $N = \Sigma_1^K \underset{\sim}{N}_i$

$$Z(\underset{\sim}{N}\ ;\ \Lambda) \geq \exp\left[-\langle I(\underset{\sim}{N}_1,\ldots\underset{\sim}{N}_K)\rangle\right] \prod_{i=1}^{K} A(\underset{\sim}{N}_i\ ;\ \Lambda_i) \quad (1.14)$$

where $\langle I\rangle$ is the <u>average</u> inter-domain energy in an ensemble where each sub-domain is <u>independent</u>. Where $\langle I\rangle$ vanishes, (1.14) reduces to (1.13a) and, in general, there is a corresponding equation for $g(\underset{\sim}{\rho}\ ;\ \Lambda)$ as in (1.13b).

To prove (1.14) consider the case K = 2 and let $\{\Psi_j\}$, j = 1,2,..., be a set of functions consisting of all properly symmetrized and normalized functions of the form

$$\Psi_j \equiv \Psi_{n,m} \equiv \Phi_n(\underset{\sim}{x}_1,\ldots \underset{\sim}{x}_{N_1}\ ;\ \chi_m(\underset{\sim}{x}_{N_1+1},\ldots \underset{\sim}{x}_{N_1+N_2}\ ;\ \Lambda_2) \quad (1.15)$$

where the $\{\Phi_n\}$ and $\{\chi_m\}$ are a <u>complete</u> orthonormal set of eigen functions in the Hilbert spaces of $H(N_1\ ;\ \Lambda_1)$ and $H(N_2\ ;\ \Lambda_2)$. The $\{\Psi_j\}$ are clearly an orthonormal set (possible incomplete) in the Hilbert space of $H(N_1 + N_3\ ;\ \Lambda)$. Hence

$$Z(N_1+N_2 ; \Lambda) \geq \sum_j (\Psi_j, \{\exp[-H(N_1+N_2 ; \Lambda)]\} \Psi_j) \tag{1.16}$$

$$= \sum_j (\Psi_j, \{\exp[-H(N_1 ; \Lambda_1) - H(N_2;\Lambda_2) - I(N_1,N_2)]\} \Psi_j)$$

where the last equality follows from the fact that the support of the $\{\Phi_n\}$ is in Λ_1 and the support of the $\{\chi_m\}$ is in Λ_2. The convexity of the exponential function implies (Peierls-Bogoliubov inequality) for any pair of self-adjoint operators $\underset{\sim}{A}$ and $\underset{\sim}{B}$ and any set of normalized vectors $\{f_\ell\}$ in the domain of $\underset{\sim}{A}$ and $\underset{\sim}{B}$ that

$$\sum_\ell (f_\ell, [\exp(\underset{\sim}{A}+\underset{\sim}{B})] f_\ell) \geq \sum_\ell \exp \{f_\ell, [\underset{\sim}{A}+\underset{\sim}{B}] f_\ell)\} \tag{1.17}$$

$$\geq \{\sum_\ell \exp (f_\ell, \underset{\sim}{A} f_\ell)\} \exp [\langle \underset{\sim}{B} \rangle_{\underset{\sim}{A}}]$$

where

$$\langle \underset{\sim}{B} \rangle_{\underset{\sim}{A}} \equiv \{\sum_\ell [\exp(f_\ell, \underset{\sim}{A} f_\ell)] (f_\ell, \underset{\sim}{B} f_\ell)\}/\{\sum_\ell \exp (f_\ell, \underset{\sim}{A} f_\ell)\}. \tag{1.18}$$

Applying (1.17) to (1.16) and remembering that $\{\Phi_n\}$ and $\{\chi_m\}$ are complete in the Hilbert spaces of $H(N_1 ; \Lambda_1)$ and $H(N_2 ; \Lambda_2)$ respectively yields the desired inequalities,

$$Z(\underset{\sim}{N}_1 + \underset{\sim}{N}_2 ; \Lambda) \geq Z(\underset{\sim}{N}_1 ; \Lambda_1) Z(\underset{\sim}{N}_2 ; \Lambda_2 ; I_1) \tag{1.19}$$

$$\geq Z(\underset{\sim}{N}_1 ; \Lambda_1) Z(\underset{\sim}{N}_2 ; \Lambda_2) \exp [-\langle I(N_1 , N_2)\rangle] .$$

Here $Z(N_2 ; \Lambda_2 : I_1)$ is the partition function of N_2 particles in Λ_2 with a Hamiltonian

$$H(N_2 ; \Lambda_2 : I_1) = H(N_2 ; \Lambda_2) + I_1 (x_{N_1+1},\ldots,\underset{\sim}{x}_{N_1+N_2}) , \tag{1.20}$$

with

$$I_1 \equiv \mathrm{Tr}_1 \{I(N_1,N_2) \exp [-H(N_1 ; \Lambda_1]\}/Z(\underset{\sim}{N}_1; \Lambda_1), \tag{1.21}$$

the subscript 1 indicating that the trace is taken with

respect to the $\{\Phi_m\}$. Hence I_1 is the value of the interdomain interaction energy for a given configuration of the N_2 particles in Λ_2 averaged over the canonical ensemble of N_1 particles in Λ_1 which are unaffected by the presence of the particles in Λ_2. Similarly,

$$\langle I(N_1,N_2)\rangle \equiv \frac{\mathrm{Tr}_{1,2}\{I(N_1,N_2)\exp[-H(N_1;\Lambda_1)]\exp[-H(N_2;\Lambda_2)]\}}{Z(\underset{\sim}{N}_1;\Lambda_1)\,Z(\underset{\sim}{N}_2;\Lambda_2)}, \tag{1.22}$$

the trace now being taken over a complete set of functions corresponding to a system consisting of a specified set of N_1 particles in Λ_1 and another set of N_2 particles in Λ_2. The quantity $\langle I(N_1,N_2)\rangle$ thus corresponds to the average of the interaction between the particles in Λ_1 and those in Λ_2 when the states and the probabilities of different states in each box are completely unaffected by the presence of the other box. This corresponds to taking the average of $I(N_1,N_2)$ with a density matrix which is a direct product of the unperturbed density matrices in Λ_1 and Λ_2.

We now make the observation, which is one of the crucial steps in our proof, that if Λ_1 and/or Λ_2 are balls then, because of the rotational (and translational) symmetry of the Hamiltonian the unperturbed density matrices (corresponding to no interaction between Λ_1 and Λ_2) are spherically symmetric about the centers of Λ_1 and/or Λ_2. This implies in particular that the average unperturbed charge density in Λ_1 and/or Λ_2 is spherically symmetric and hence by Newton's theorem the Coulomb contribution to $\langle I(N_1,N_2)\rangle$ in (1.19) is the same as would be obtained if all the charges in the ball domain were concentrated at its center and would vanish when the ball is overall neutral.

This clearly generalizes to the Coulomb part of $\langle I(N_1,\ldots,N_K)\rangle$ in (1.14) and this leads us to choose balls, rather than cubes, for our standard domains. There is of course a price to be paid for this since balls do not pack into each other as nicely as cubes do and necessitates our packing the standard ball domain B_j not only with balls of type B_{j-1} but with balls of types B_0,

$B_1, \ldots, B_{j-2}$ as well. The geometrical problem involved in this is handled in sec. II.

We mention here that the use of (1.14) permits us to prove the existence of the thermodynamic limit for systems containing electric or magnetic dipoles which interact with long range, i.e. non-tempered, potentials (falling off only as $|x_i - x_j|^{-3}$). The average interaction between domains will vanish since the expectation value of the dipole moment of any particle will be zero in the absence of an external electric or magnetic field, Griffiths (1968)*.

For such systems it is essential, however, that the particles have hard cores since otherwise they will not satisfy the H-stability condition. Indeed, $E_0(N)$ will not be bounded below. For this reason we cannot include spin-spin couplings between the elementary charges in our analysis. (These couplings are intrinsically of a relativistic nature and present entirely new problems; Dyson- private communication.)

Needless to say we do not deal with the strong (nuclear) and weak interactions. As pointed out by Dyson (1967), the magnitude of the nuclear forces is so large that they would give completely different binding energies for molecules and for crystals if they played any role in the thermal properties of ordinary matter. We are also neglecting, of course, gravitational forces which certainly are important for large aggregates of matter and thus might be thought important in the 'thermodynamic limit'. To quote Onsager (1967), however, 'The common concept of a homogeneous phase implies dimensions that are large

*Griffiths' proof for dipoles does not use (1.19) but relies on the complete symmetry between 'up and down' orientations of the dipoles. Using such symmetry Griffiths (unpublished) was able to prove the existence of the thermodynamic limit for a system of charged particles in which the positive and negative particles are identical under charge conjugation, e.g. positrons and electrons. When such an additional symmetry is present the rotational invariance of the Hamiltonian becomes unimportant and it is not necessary to use balls as we do. Unfortunately such symmetries are not present in real systems.

compared to the molecules and small compared to the moon. When we speak of the thermodynamic limit which is mathematically the infinite system limit we have in mind its physical application to systems containing, $10^{22} \sim 10^{28}$ particles, i.e. systems which are large enough for surface effects to be negligible and yet small enough for internal gravitational effects also to be completely negligible. (An external gravitational field will of course have some effect but does not present any fundamental problem.)

II. On Packing A Domain With Balls

In this section we address ourselves to a geometric construction which is fundamental to our proof of the existence of the thermodynamic limit, namely the possibility of packing a ball or a cube by smaller balls such that the packing is both complete and rapid.

We shall always use the word domain to mean a bounded, open set in R^d. If Λ is a domain and $B = \{B_i\}$ is a denumerable family of domains such that $B_i \subset \Lambda$ for all i we shall say that B is packed in Λ if the $\{B_i\}$ are all disjoint. The packing is complete if $\Sigma_i V(B_i) = V(\Lambda)$ where $V(\Lambda)$ is the volume (lebesgue measure of Λ .

Definition: For a domain $\Lambda \subset R^d$ and a real number h we define

$$\begin{aligned} \Lambda_h &= \{r: r \in \Lambda, d(r; \sim \Lambda) < h\} \quad \text{for } h > 0 \\ &= \{r: r \in \sim \Lambda, d(r;\Lambda) \leq -h\} \quad \text{for } h \leq 0 \ , \end{aligned} \tag{2.1}$$

where $d(;)$ is the distance function and $\sim$ denotes complement. We also define $V(h;\Lambda)$ to be the volume of Λ_h.

We shall frequently make use of the fact (Lemma 2 of Section 8 in Fisher (1964) that the number, N_y, of cubes of side $2y$ that can be packed in Λ satisfies the inequality

$$N_y \geq (2y)^{-d}[V(\Lambda) - V(2y\sqrt{d};\Lambda)] \tag{2.2}$$

Definition: Let σ_d be the volume of a ball of unit radius in R^d. $g_d \equiv 2^{-d}\sigma_d$ is the fraction of the volume of a cube of side $2y$ filled by a ball of radius y when the ball is packed in the cube. We also define $\alpha_d \equiv (2^d - 1)2\sqrt{d}$.

Clearly, for a ball B of radius $r \geq 2y\sqrt{d} \geq 0$

$$V(2y\sqrt{d};B) \leq V(-2y\sqrt{d};B) \leq \alpha_d \sigma_x r^{(d-1)} y \ . \tag{2.3}$$

The main theorem we wish to prove is that we can find a sequence of balls of decreasing radius, of which the j^{th} type has radius $r_j = \delta^j$ (with $\delta < 1$), such that we can completely pack a unit d-dimensional ball ($r = 1$) with these and, moreover, we can do this rapidly.

Theorem 2.1: Let p be a positive integer and, for all positive integers j, define radii $r_j = (1 + p)^{-j}$ and integers $n_j = p^{j-1}(1 + p)^{j(d-1)}$. Then if $1 + p \geq \alpha_d + g_d^{-d}$ it is possible to pack $\bigcup_{j=1}^{\infty}$ (n_j balls of radius r_j) in a unit d-dimensional ball. The packing is complete since $\sum_{j=1}^{\infty} n_j r_j^d = 1$.

Proof: We shall give an explicit construction for accomplishing the packing stated in the theorem by using (2.2) and (2.3). First cover the unit ball by a cubic array of cubes of side $2r_1$. We shall show that there are n_1 of these cubes which are contained in the unit ball. We can place a ball of radius r_1 at the center of each of these cubes. We then cover the unit ball by a cubic array of cubes of side $2r_2$ and show that there are n_2 of these cubes which are contained in the unit ball and which do not intersect the first n_1 balls. The argument is repeated inductively. Thus, we have to show that after placing all balls up to and including those of radius r_j we can pack n_{j+1} in a cubic array into Ω_j, which is the interior of the unfilled portion of the unit ball. (We must prove this for $j \geq 0$, with $r_o \equiv 0$.) For $j \geq 0$,

$$V(\Omega_j) = \sigma_d - \sigma_d \sum_{k=o}^{j} n_k r_k^d = \sigma_d \left(\frac{p}{p+1}\right)^j .$$

Clearly, $V(2\sqrt{d}r_{j+1};\Omega_j)$ is bounded above by M_j which is the sum of the $V(-2\sqrt{d}r_{j+1};B)$ for each ball of $\bigcup_{\ell=0}^{j}$ (n_ℓ balls of radius r_ℓ) separately, plus $V(2\sqrt{d}r_{j+1};B)$ for the unit ball. Thus, by (2.3), if $2\sqrt{d}r_{j+1} < r_j$ (which is true when p satisfies the hypothesis)

$$V(2\sqrt{d}r_{j+1};\Omega_j) \leq M_j \leq \alpha_d \sigma_d r_{j+1}^{d-1}\{1 + \sum_{k=o}^{j} n_k r_k^{d-1}\}$$
$$= (p^j + p - 2)(p - 1)^{-1}(1 + p)^{-(j+1)}\alpha_d\sigma_d \equiv \hat{M}_j . \qquad (2.4)$$

Using (2.2) it is sufficient to show that

$$(2r_{j+1})^d n_{j+1} \leq [V_j - \hat{M}_j] \leq [V(\Omega_j) - V(2\sqrt{d}r_{j+1};\Omega_j)]$$

Inserting the relevant quantities, we require that

$$1 \leq g_d[p + 1 - \alpha_d \frac{1 + p^{-j}(p - 2)}{p - 1}]$$

for all $j \geq 0$. By the hypothesis $p \geq 2$. Then $p^{-j}(p - 2) \leq (p - 2)$ and hence it is sufficient that

$$1 \leq g_d[p + 1 - \alpha_d] \quad ,$$

which agrees with the hypothesis.

The minimum ratio of successive radii, $1 + p$, required by this construction is 27 for $d = 3$. We note that the fraction of volume of the unit ball occupied by all the balls of radius r_j is

$$f_j = n_j r_j^d = p^{-1}\lambda^j \quad , \tag{2.5}$$

where

$$\lambda = p(1 + p)^{-1} < 1 \quad . \tag{2.6}$$

Moreover, the fraction of volume left unfilled after the balls of type j have been packed is λ^j. This implies that the packing is "exponentially fast".

It can be shown that Theorem 2.1 is also true if "unit d-dimensional ball" is replaced by "d-dimensional cube of volume σ_d .

III. Thermodynamic Limit for Spherical Domains

In this section we shall prove the existence of the thermodynamic limit for a two component system of charges $e_1 > 0$, $e_2 < 0$ and $|e_1/e_2|$ is rational and that the particles interact via the Coulomb potential alone. To do so we shall define a sequence of standard balls $\{B_j\}$ of increasing radii $\{R_j\}$, $j = 0, 1, \ldots$.

Definition: Let $1 + p$ satisfy the condition of Theorem 2.1 and be even. (The fact that $1 + p$ is even will not be used until later.) Choose an $R_o > 0$. The balls, B_o, B_1, ..., forming the standard sequence, are chosen to have radii

$$R_j = R_o(1 + p)^j \tag{3.1}$$

The volume of B_j will be denoted by V_j.

The packing described in Theorem 2.1 will be referred to as the standard packing of the ball B_K with balls $\{B_j\}$, $j = 0, 1, \ldots, K - 1$.

Filling of Balls with Particles

In the following we shall fill the standard balls with particles in various ways. However, we shall always observe the following convention: Each ball will have charge neutrality. We take q particles of type 1 and ℓ particles of type 2 such that $qe_1 + \ell e_2 = 0$, and such that q and ℓ have no common divisor, as the fundamental unit and this will be referred to simply as a multiplet. Densities and (multiplet) numbers will be in terms of this unit.

We define

$$g_j(\rho) = (V_j)^{-1} \ell n Z(N = \rho V_j; B_j) \quad , \tag{3.2}$$

where N is the number of multiplets and where we have set $\beta = 1$ for convenience.

Since N must be an integer, an obvious restriction is thereby placed on ρ. However, following Fisher (1964) we can define g for all ρ by linear interpolation as follows:

Definition: Let $f(N)$ be a function from the integers to the reals. If $n = N + \eta$, with N an integer and $0 \le \eta \le 1$, we extend $f(\cdot)$ to the reals by $f(n) \equiv f(N) + \eta[f(N+1) - f(N)]$.

The usefulness of this definition is made manifest by the following lemma.

Lemma 3.1: Let $\mathbb{Z}^+$ be the non-negative integers, $\mathbb{R}^+$ the non-negative reals, and $\mathbb{R}$ the reals. Let $f, h_1, h_2, \ldots, h_M$ be functions from $\mathbb{Z}^+$ to $\mathbb{R}$ and let $f, h_1, \ldots, h_M$ be the extended functions from $\mathbb{R}^+$ to $\mathbb{R}$ as in the above definition. Let $N_j \in \mathbb{Z}^+$ and $n_j \in \mathbb{R}^+$. If $f(\Sigma_1^M N_j) \ge \Sigma_1^M h_j(N_j)$ for all $\{N_j\}$ then $f(\Sigma_1^M n_j) \ge \Sigma_1^M h_j(n_j)$ for all $\{n_j\}$.

The proof follows by induction on M. The case $M = 1$ is obvious and $M = 2$ is proved in Fisher (1964), footnote 25.

Let us now consider a standard packing of B_K and place N_j multiplets in all balls of type j, $j = 0, 1, \ldots, K-1$. The total number of multiplets in B_K is then

$$N = \Sigma_0^{K-1} N_j n_{K-j} , \tag{3.3}$$

so that

$$\rho = N/V_K = p^{-1} \Sigma_0^{K-1} \rho_j \gamma^{K-j} \tag{3.4}$$

Our fundamental inequality on the partition function of a subdivided domain, together with the vanishing of the average Coulomb interaction for neutral balls implies:

Theorem 3.1: Let $\rho_1, \ldots, \rho_{K-1}$ be non-negative reals and let $\rho = p^{-1} \Sigma_0^{K-1} \rho_j \gamma^{k-j}$. Then

$$g_K(\rho) \geq \frac{1}{p} \sum_{j=0}^{K-1} \gamma^{K-j} g_j(\rho_j) \tag{3.5}$$

Limit of $g_k(\rho)$ as $k \to \infty$

Our next task is to use Theorem 3.1 to establish the thermodynamic limit of $g_k(\rho)$ for the standard sequence of balls. To accomplish this we define, for each $j \geq 0$, a standard density sequence (depending on ρ) as follows:

$$\rho_j = \rho \text{ for } j > 0; \quad \rho_j = \rho(1-\gamma)^{-1} \text{ for } j = 0$$

It is understood that when $\rho = 0$, $g_j(0) = 0$.

With ρ held fixed, let us denote $g_j(\rho_j)$ simply by g_j. Then, from Theorem 3.1,

$$g_k = \frac{1}{p} \sum_{j=0}^{k-1} \gamma^{k-j} g_j + c_k \tag{3.6}$$

for $k > 0$, where c_k is a non-negative real number.

The solution of (3.6), valid for $k > 0$, is easily found to be

$$g_k = \gamma c_k + (1-\gamma) \sum_{j=1}^{k} c_j + (1-\gamma) g_0 \tag{3.7}$$

Equation (3.7) establishes a limit for g_k because: (a) g_0 is finite; (b) As each $c_j \geq 0$, and as we know that g_k has an upper bound by H-stability, the sum involving the c's must converge. This implies that $c_k \to 0$ and hence (3.7) must have a limit. We shall call this limit $g(\rho)$.

Further examination of (3.7) leads to a lower bound for g which is proportional to ρ for sufficiently small ρ.

Our analysis of (3.6) thus yields

Theorem 3.2: Let ρ be a fixed multiplet density. Then

$$g(\rho) = \lim_{k \to \infty} g_k(\rho)$$

exists and is finite. Furthermore, there exists a $\rho_1 > 0$ such that for ρ in the closed-open interval $[0,\rho_1)$, $g(\rho)$ is bounded below by $\alpha\rho$ with α finite and independent of ρ.

Convexity of the Free Energy

With the limit, $g(\rho)$, in hand we can next establish convexity. It is here that we use the fact that $1 + p$ was chosen to be even. This permits us to place densities ρ'_j corresponding to a final density ρ' (with $\rho'_0=\rho'(1-\gamma)^{-1}$, $\rho'_j=\rho'$, $j > 0$) in half the balls of each type and densities ρ''_j in the other half. Taking the limit $j \to \infty$ yields

$$g(\tfrac{1}{2}\rho' + \tfrac{1}{2}\rho'') \geq \tfrac{1}{2} g(\rho') + \tfrac{1}{2} g(\rho''). \qquad (3.8)$$

We can now follow the standard arguments used for non-Coulomb systems to establish the concavity and hence continuity of $g(\rho)$. Similarly, the approach of $g_j(\rho)$ to $g(\rho)$ can be shown (by means of Dini's theorem) to be uniform on any closed interval $[0,\rho']$, $\rho' < \infty$.

Neutral Multicomponent Systems with Coulomb and Tempered Interactions in General Domains

Thus far we have established the limit and the convexity of the free energy/unit volume for an overall neutral system composed of two species of charged particles interacting with Coulomb forces only and confined to the standard sequence of balls. This permitted us to deal with a neutral multiplet as though it were a single particle.

We shall now state the general theorem on the properties of the free energy/unit volume for an overall system composed of S species of particles with charges $e_1,\ldots,e_s$. We suppose these charges to be rational fractions of each other so that, in appropriate units, the e_i may be taken to be integers. In nature all elementary charges are in fact integral multiples of the electron charge. The e_i may not be all of one sign, but we do allow some of them to be zero. We shall represent particle numbers by a vector $\underset{\sim}{N} = (N^1,\ldots,N^S)$, so that charge neutrality is represented by $\underset{\sim}{N} \cdot \underset{\sim}{E} = 0$ with $\underset{\sim}{E} = (e_1,\ldots,e_s)$. In a like manner we shall represent particle densities by a vector $\underset{\sim}{\rho}$.

The particles comprising our system may have, in addition to their Coulomb interactions, other kinds of interaction potentials as long as those interactions are tempered and the full Hamiltonian is H-stable (this will always be true when the additional interactions are themselves stable). When these tempered interactions include hard cores, there will be some convex domain in $\mathbb{R}^s$ in which a vector $\underset{\sim}{\rho}$ must lie in order for the density to be less than the close packing density. We shall denote the fact that $\underset{\sim}{\rho}$ is in this domain by writing $|\underset{\sim}{\rho}| < \rho_c$.

We consider a general sequence of domains $\{\Lambda_j\}$ tending to infinity in a reasonable way. To define reasonable we introduce the following conditions on a sequence of domains in $\mathbb{R}^d$:

A. A sequence of domains $\{\Lambda_j\}$ tends to infinity in the sense of Van Hove if $V(\Lambda_j) \to \infty$ and $V(h;\Lambda_j)/V(\Lambda_j) \to 0$ as $j \to \infty$ for each fixed h. (For definitions see (2.1)).

B. A sequence of domains $\{\Lambda_j\}$ satisfies the ball condition if there exists a $\delta > 0$ such that

$$V(\Lambda_j)/V(\hat{B}_j) \geq \delta \ , \tag{3.9}$$

where $\hat{B}_j$ is the ball of smallest radius containing Λ_j.

C. A sequence of domains $\{\Lambda_j\}$ tends to infinity in the sense of Fisher if $V(\Lambda_j) \to \infty$ and if there exists a continuous function $\pi: \mathbb{R}^1 \to \mathbb{R}^1$, with $\pi(0) = 0$ such that

$$V(\alpha[V(\Lambda_j)]^{1/d};\Lambda_j)/V(\Lambda_j) \leq \pi(\alpha) \tag{3.10}$$

for all α and all j.

Obviously, condition C implies A. It also implies condition B as shown in Fisher (1964). On the other hand, neither condition A nor B implies the other, nor do conditions A and B together imply C.

Definition: A regular sequence of domains, $\{\Lambda_j\}$, in $\mathbb{R}^d$ is one satisfying conditions A and B if only strongly tempered potentials (in addition to the Coulomb potential) are present. If weakly tempered potentials are also present then the stronger condtion C must be satisfied.

Our final result for neutral systems, which we state here without proof, is

Theorem 3.3: Let $\{\Lambda_j\}$ be a regular sequence of domains. Let $\{\underset{\sim}{N}_j\}$ be a sequence of non-negative, integer valued particle number vectors satisfying the neutrality conditions, $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$, and let $\varrho_j = V(\Lambda_j)^{-1}\underset{\sim}{N}_j$. If $\lim_{j\to\infty} \varrho_j = \varrho$, with $|\varrho| < \rho_c$ then

(i) $\lim_{j\to\infty} g(\varrho_j;\Lambda_j) = g(\varrho)$ exists and is independent of the sequence of domains or particle numbers.

(ii) $g(\varrho)$ is continuous and concave in the convex domain $D=\{\varrho:|\varrho| < \rho_c\} \quad \{\varrho:\varrho \cdot \underset{\sim}{E}=0\}$ and $g(0) = 0$.

(iii) Let K be a compact subset of D. Suppose that for each $\varrho \in K$ we have a sequence $\{\underset{\sim}{N}_j(\varrho)\}$ and the corresponding sequence $\{\varrho_j(\varrho)\}$ with the additional hypothesis that $\varrho_j(\varrho) \to \varrho$ uniformly on K. Then $g_j(\varrho_j(\varrho)) \to g(\varrho)$ uniformly on K.

IV. Systems With Net Charge

In the last section we showed that a sequence of systems of charged particles has a thermodynamic limit when the finite systems in the sequence have no net charge, that is $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$. The free energy density in this limit, $-g(\varrho)$, is independent of the shape of the domains Λ_j and depends only on the limit of the particle density vector $\underset{\sim}{N}_j/V(\Lambda_j)$.

It is intuitively clear that this condition of strict charge neutrality, $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$, is unnecessarily restrictive. We expect that a 'small' amount of uncompensated charge will have no effect on the free energy density in the thermodynamic limit while a 'large' amount of uncompensated charge will lead to a divergent free energy density in that limit. The dividing line between 'small' and 'large' should be when the excess charge Q_j, in a domain Λ_j, increases in proportion to the 'surface area' of Λ_j as $j \to \infty$. In this case we expect the thermodynamic limit of the free energy density to exist but that its value depends also on the limiting shape of the domains Λ_j.

These expectations come from macroscopic electrostatic theory (Kellog (1929)) which shows that the lowest energy configuration for any net charge Q confined to a domain Λ is obtained when Q is concentrated at the boundary of Λ. This configuration of the charge is described in electrostatics by a two dimensional charge density $\sigma(\underset{\sim}{x})$, $\underset{\sim}{x} \in S_\Lambda$, where S_Λ is the surface of Λ. (We shall only consider three dimensional systems here, that is $\Lambda_j \subset R^3$). This surface charge density will be such as to make the electrostatic potential constant in the interior of Λ, i.e., there will be no electric field in Λ. The electrostatic energy of this surface layer is equal to $\frac{1}{2}Q^2/C(\Lambda)$ where $C(\Lambda)$ is the capacitance of Λ.

For a given domain shape, $C(\Lambda)$ is proportional to $[V(\Lambda)]^{\frac{1}{3}}$ and the electrostatic energy per unit volume will thus be proportional to $[Q/V^{\frac{1}{3}}]^2$, the square of the 'average surface charge density'. Hence for a sequence of domains $\{\Lambda_j\}$ with volumes $\{V_j\}$ and capacitances $\{C_j\}$ each containing a net charge Q_j such that as $j \to \infty$, $V_j \to \infty$, $C_j/V_j^{\frac{1}{3}} \to c$ and $Q_j/V_j^{\frac{1}{3}} \to \sigma$, the minimum electrostatic energy per unit volume $\tilde{e}_j$ will also approach a limit,

$$\tilde{e} = \lim_{j\to\infty} \tilde{e}_j = \frac{1}{2}\sigma^2/c \tag{4.1}$$

Note that (4.1) refers solely to the macroscopic electrostatic energy per unit volume of the charge Q in the domain Λ or on the surface S_Λ. We shall now state a theorem which shows that in the thermodynamic limit the difference between the free energy densities of a neutral system, obtained in Section III and of a system containing some extra charged particles is given precisely by (4.1). For technical reasons the theorem is proved only for a sequence of domains whose shapes approach ellipsoids in the sense defined below. This is more restrictive than is desirable or (probably) necessary as will be clear from the derivation of the theorem.

Definition: Let E be an open ellipsoid of unit volume and capacity c_E. A sequence of domains $\{\Lambda_j\}$, $j = 1,2,\ldots$, will be called asymptotically similar to E if $V(\Lambda_j) \to \infty$ and if there exist ellipsoids $\{E_j'\}$ and $\{E''\}$ similar to E such that $E_j' \subset \Lambda_j \subset E_j''$ and $V(E_j'')/V(E_j') \to 1$ as $j \to \infty$. The capacity of Λ_j will clearly lie between the capacities of

E_j' and E_j''. These latter capacities are $c_E[V(E_j')]^{\frac{1}{3}}$ and $c_E[V(E_j'')]^{\frac{1}{3}}$ respectively.

Theorem 4.1: Let $\{\Lambda_j\}$ be a sequence of domains asymptotically similar to an ellipsoid E, and let $\{\underset{\sim}{N}_j\}$, and $\{\underset{\sim}{n}_j\}$ be sequences of integer particle number vectors such that $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$, $\underset{\sim}{n}_j \cdot E = Q_j$, and

$$\lim_{j\to\infty} \underset{\sim}{N}_j/V(\Lambda_j) = \underset{\sim}{\rho}, \quad \lim_{j\to\infty} \underset{\sim}{n}_j/V(\Lambda_j) = 0, \quad \lim_{j\to\infty} Q_j[V(\Lambda_j)]^{-\frac{2}{3}} = \sigma.$$

Then if $|\underset{\sim}{\rho}| < \rho_c$,

$$\lim_{j\to\infty} g([\underset{\sim}{N}_j + \underset{\sim}{n}_j]/V(\Lambda_j;\Lambda_j) = g(\underset{\sim}{\rho}) - \frac{1}{2}\sigma^2/c_E.$$

Remarks: (i) Since $E_j' \subset \Lambda_j \subset E_j''$, it follows from the basic inequality that $Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j;E_j'') \geq Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j;\Lambda_j) \geq Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j; E_j')$. Moreover, since $V(E_j'')/V(E_j') \to 1$ as $j\to\infty$ it is sufficient to prove the theorem for a sequence of ellipsoids $\{E_j\}$ similar to E, whose volumes are the same as that of the $\{\Lambda_j\}$. With each E_j we associate a pair of homothetic ellipsoids, E_j^- and E_j^+ similar to E_j such that $E_j^- \subset E_j \subset E_j^+$ and $V(E_j^+)/V(E_j^-) \to 1$ as $J \to \infty$. The volumes and capacities of E_j^-, E_j, and E_j^+ will be denoted by $(L_j^-)^3$, L_j^3, $(L_j^+)^3$ and C_j^-, C_j, C_j^+ respectively. Clearly $C_j^{\pm}=c_E L_j^{\pm}$ and $C_j=c_E L_j$. The interiors of the ellipsoidal shells $E_j^+\backslash E_j$ and $E_j\backslash E_j^-$ will be called D_j^+ and D_j^- respectivcly.

(ii) The reason for the introduction of ellipsoidal domains, is their well known electrostatic property (Kellog (1929)) that a uniform three dimensional charge density τ in an ellipsoidal shell such as D_j^+ (defined above) has a self energy $\frac{1}{2}\tau^2 V(D_j^+)^2/C_j'$ and produces a constant potential $\tau V(D_j^+)/\bar{C}_j$ in the interior of E_j, with $c_E L_j \leq \bar{C}_j \leq C_j' \leq c_E L_j^+$. This fact will enable us to obtain bounds on the partition functions for the domains $\{E_j\}$ in a simple manner. Identical methods would work also for any other sequence of domains for which there

are shell domains surrounding each Λ_j with the above mentioned properties of the shells $D_j^{\pm}$.

(iii) The proof of Theorem (4.1) will proceed by establishing bounds on the free energy of these systems. For this we shall need the free energies of two kinds of neutral systems: the first kind consists of $\underset{\sim}{N}_j$ particles in E_j^-; the second kind is of system in E_j^+ which contains an additional species of particles so that it has altogether S + 1 species. The new specie, which, following Aristotle, we shall call hyle will be labeled by the index zero. Its charge e_o will be ± 1 (in units in which all e_i, $i = 1,\ldots, S$, are integers). The sign of e_o will be chosen as the opposite of the sign (which we shall take to be independent of j) of the excess charge Q_j, that is $e_o Q_j < 0$. The new neutral system will have an S + 1 component particle number vector $N_j + n_j + n_j^o = (n_j^o, N_j^1 + n_j^1, \ldots, N_j^S + n_j^S)$ with $n_j^o = |Q_j|$, $n_j^o e_o = -Q_j$ so that the system is overall neutral. The hyle particles will only have Coulomb interactions and will be fermions in order to comply with the Dyson-Lenard theorem.

Lower Bound on the Partition Functions of Charged Ellipsoids

We consider a packing of E_j^- with balls and we distribute the $\underset{\sim}{N}_j$ particles, $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$, among the balls such that each ball is neutral and call the resulting partition function $Z(\underset{\sim}{N}_j; B(E_j^-))$. The remaining n_j particles we place in D_j^-. It then follows from our basic inequality and the fact that each ball is neutral that

$$Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j; E_j) \geq Z(\underset{\sim}{N}_j; B(E_j^-)) Z(\underset{\sim}{n}_j; D_j^-). \qquad (4.2)$$

It can be shown, using Theorem 3.1, that the packing for each j can be chosen so that upon taking the logarithm of (4.2) and dividing by $V(E_j)$ one obtains

$$\lim_{j\to\infty} \inf\{g([\underset{\sim}{N}_j + \underset{\sim}{n}_j]/V(E_j); E_j) - [V(E_j)]^{-1} \ell n Z(n_j; D_j^-)\} \geq g(\varrho). \qquad (4.3)$$

Since $n_j/V(E_j) \to 0$, the only contribution from $\ell n Z(\underset{\sim}{n}_j; D_j^-)/V(E_j)$ which survives when $j \to \infty$ is the Coulomb self energy of the charges in D_j^-. We now use the

following general inequality for the partition function of a system of N particles in a somain Λ, with a Hamiltonian H;

$$\ell nZ(N;\Lambda) \geq -J^{-1} \sum_{\alpha=1}^{J} (\psi_\alpha, H\psi_\alpha) , \qquad (4.4)$$

where $\{\psi_\alpha\}$, $\alpha = 1,\ldots, J$ is any properly symmetrized and normalized set of functions of the N particle coordinates $\underset{\sim}{x}_i$, $i = 1,\ldots, N$ and spins, which vanish unless $\underset{\sim}{x}_i \in \Lambda$. Applying (4.4) to $Z(\underset{\sim}{n}_j;D_j^-)$ with a choice of ψ_α which corresponds to the $\underset{\sim}{n}_j$ particles being situated in little balls centered on the vertices of a cubical lattice covering D_j^-, we obtain a lower bound on this self energy corresponding to a uniform distribution of the charge Q_j in D_j^-,

$$\lim_{j\to\infty} \sup[V(E_j)]^{-1} \ell nZ(\underset{\sim}{n}_j;D_j^-) \geq -\frac{1}{2}\sigma^2/c_E. \qquad (4.5)$$

This yields

$$\lim_{j\to\infty} \inf\, g([\underset{\sim}{N}_j + \underset{\sim}{n}_j]/V(E_j);E_j) \geq g(\varrho) - \frac{1}{2}\sigma^2/c_E \qquad (4.6)$$

Upper Bound on the Partition Functions of Charged Ellipsoids

Let $Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j + \underset{\sim}{n}_j^o;E_j^+)$ be the partition function of a system in the domain E_j^+ having S + 1 species with $n_j^o = |Q_j|$ hyle particles of charge $e_o = -Q_j/|Q_j|$, as in remark (iii) after Theorem 4.1. The masses m_o of the hyle particles may be chosen arbitrarily. We then have

$$Z(\underset{\sim}{N}_j + \underset{\sim}{n}_j^o;E_j^+) \geq Z(\underset{\sim}{N} + \underset{\sim}{n}_j;E_j)Z(n_j^o;D_j^+:W_j). \qquad (4.7)$$

Here $Z(n_j^o;D_j^+:W_j)$ is the partition function of n_j^o particles of species zero whose Hamiltonian consists of a kinetic energy term, a Coulomb pair interaction term, and an external one-body electrostatic potential $w_j(\underset{\sim}{x}_i)$, $i=1,..,n_j^o$, produced by the (canonical ensemble) average charge density of the $\underset{\sim}{N}_j + \underset{\sim}{n}_j$ particles in E_j. Taking logarithms in (4.7) and dividing by $V(E_j^+)$ gives the upper bound

$$\lim_{j\to\infty} \sup\{g([\underset{\sim}{N}_j+\underset{\sim}{n}_j]/V(E_j^+);E_j^+)+[V(E_j^+)]^{-1}\ell nZ(n_j^o;D_j^+:W_j)\} \leq g(\varrho). \qquad (4.8)$$

Here, too, the only contribution from $\ell nZ(n_j^o;D_j^+:W_j)$ which survives in the limit is the Coulomb energy which now consists of two parts: the self energy of the charges in D_j^+ and the mutual electrostatic energy between the charge $-Q_j$ in D_j^+ and Q_j in E_j. Now if the charge $-Q_j$ were smeared out uniformly in D_j^+ then, because of the properties of the ellipsoidal shells mentioned in remark (ii), the sume of these two energies would be $\frac{1}{2}Q_j^2/C_j' - Q_j^2/\bar{C}_j$ with C_j' and $\bar{C}_j$ both approaching C_j as $j \to \infty$. It can be shown indeed by using inequality (4.4) with a suitable choice of $\{\psi_\alpha\}$, that

$$\lim_{j\to\infty} \inf[V(E_j)]^{-1}\ell nZ(n_j^o;D_j^+:W_j) \geq \frac{1}{2}\sigma^2/c_E. \qquad (4.9)$$

Combining this with (4.8) and (4.6) yields Theorem 4.1.

When the magnitude of the charge contained in Λ_j, $Q_j = \underset{\sim}{M}_j \cdot \underset{\sim}{E}$, (where $\underset{\sim}{M}_j$ is an integer particle number vector), increases faster than $V(\Lambda_j)^{\frac{2}{3}}$, i.e., $|Q_j|V(\Lambda_j)^{-\frac{2}{3}} \to \infty$, then it is possible to show that $g(\underset{\sim}{M}_j/V(\Lambda_j);\Lambda_j) \to -\infty$ for any regular sequence of domains $\{\Lambda_j\}$.

V. Grand Canonical Ensemble

The grand canonical partition function for a system of S species in a domain Λ_j with chemical potentials μ_i, $i = 1,\ldots,S$, is defined as

$$\Xi(\underset{\sim}{\mu};\Lambda_j) = \sum_{N^1=o}^{\infty} \cdots \sum_{N^S=o}^{\infty} \exp[\underset{\sim}{\mu} \cdot \underset{\sim}{N}]Z(\underset{\sim}{N};\Lambda_j), \qquad (5.1)$$

where $\underset{\sim}{\mu} = (\mu_1,\ldots,\mu_S)$, and we have set $\beta = 1$. The grand canonical pressure is defined as

$$\pi(\underset{\sim}{\mu};\Lambda_j) = V(\Lambda_j)^{-1}\ell n\Xi(\underset{\sim}{\mu};\Lambda_j). \qquad (5.2)$$

We also define the neutral grand canonical partition function Ξ', by restricting the summations in the right side of (5.1) to neutral systems for which $\underset{\sim}{N} \cdot \underset{\sim}{E} = 0$. The The function Ξ' will clearly depend only on that part of the vector $\underset{\sim}{\mu}$ which is perpendicular to $\underset{\sim}{E}$, i.e., on $\underset{\sim}{\mu} = \underset{\sim}{\mu} - (\underset{\sim}{\mu} \cdot \underset{\sim}{E})\underset{\sim}{E}/(\underset{\sim}{E} \cdot \underset{\sim}{E})$, and will thus be a function of only S - 1 independent variables,

$$\Xi'(\underset{\sim}{\mu};\Lambda_j) = \Xi'(\underset{\sim}{\mu}';\Lambda_j) = \sum_{(\underset{\sim}{N}\cdot\underset{\sim}{E}=o)} \exp[\underset{\sim}{\mu}'\cdot\underset{\sim}{N}]Z(\underset{\sim}{N};\Lambda_j). \qquad (5.3)$$

Similarly,

$$\pi'(\underset{\sim}{\mu};\Lambda_j) = \pi(\underset{\sim}{\mu}';\Lambda_j) = V(\Lambda_j)^{-1} \ell n \Xi'(\underset{\sim}{\mu}';\Lambda_j) \qquad (5.4)$$

is the neutral grand canonical pressure.

As in section IV, we shall convine our attention here to domains $\Lambda_j \subset \mathbb{R}^3$.

Remark: As is well known, if $\mu > o$, the grand canonical partition function Bose gas is infinite for large j (Bose-Einstein condensation). One can prove, Ruelle (1969), that if the particles interact with a tempered super-stable potential then this pressure does exist for all μ, while for a tempered potential which is only stable the pressure exists only for small values of μ (depending on β), i.e., $\mu < f(\beta)$.

For Coulomb systems to be H-stable the Dyson-Lenard theorem requires that all charged bosons have charges of the same sign. We can show that if the only bosons present are charged ones then $\lim_{j\to\infty} \pi(\underset{\sim}{\mu};\Lambda_j)$ exists for all values of the μ_i, $(-\infty \leq \mu_i < \infty, i=1,\ldots,S)$; see Lemma 5.3. If, however, our systems contains some species of neutral bosons, say $e_1 = e_2 = \ldots = e_\ell = 0$, $\ell \leq S - 2$, then the corresponding μ_i, $i = 1,\ldots, \ell$ will have to be appropriately small unless the tempered potentials involving these uncharged particles satisfy some super-stability condition. Since the part of the proof which involves the uncharged components does not differ from the standard ones we shall assume from now on that all the species are charged with $e_1,\ldots,e_a > 0$ and $e_{a+1},\ldots,e_S < 0$. We shall assume that species $a + 1,\ldots,S$ are fermions and that some or all of species $1,\ldots,a$ may be bosons.

We shall now state the main theorem of this section.

Theorem 5.1

For any regular sequence of domains $\{\Lambda_j\}$, $\pi(\underset{\sim}{\mu}) = \lim_{j\to\infty} \pi(\underset{\sim}{\mu};\Lambda_j) = \lim_{j\to\infty} \pi'(\underset{\sim}{\mu}';\Lambda_j) = \pi'(\underset{\sim}{\mu}')$ exists and is related to the Helmholtz free energy density by

$$\pi'(\underset{\sim}{\mu}') = \sup_{|\underset{\sim}{\varrho}|<\rho_c} [\underset{\sim}{\varrho} \cdot \underset{\sim}{\mu}' + g(\underset{\sim}{\varrho})] , \qquad (5.5)$$

the supremum being taken only over values of $\underset{\sim}{\varrho}$ for which $\underset{\sim}{\varrho} \cdot \underset{\sim}{E} = 0$.

Proof: The proof that $\lim \pi'(\mu'; \Lambda_j)$ exists and is given by (5.5) is analogous to Fisher's (1964) proof of a similar result for one component systems interacting only with tempered potentials with the additional result that the μ_i are arbitrary even if some of the components are bosons. (The reason for this is that if the boson density is large then the fermion density must also be large to insure charge neutrality. See Lemma 5.3.) The new element entering Theorem 5.1 is the equality of $\pi(\underset{\sim}{\mu};\Lambda_j)$ and $\pi'(\underset{\sim}{\mu}';\Lambda_j)$ in the thermodynamic limit. This means in essence that the terms in the grand partition function for which $\underset{\sim}{N}_j \cdot \underset{\sim}{E} \neq 0$ do not contribute to the pressure in this limit and hence $\lim_{j\to\infty} \pi(\underset{\sim}{\mu};\Lambda_j)$ depends only on S - 1 variables. Now since $\pi(\underset{\sim}{\mu};\Lambda_j) \geq \pi'(\underset{\sim}{\mu}';\Lambda_j)$, Theorem 5.1 will be established if we can prove that $\pi(\mu;\Lambda_j) \leq \pi(\mu';\Lambda_j) + \delta_j$ with $\delta_j \to 0$ as $j \to \infty$. This is accomplished with the help of the following three lemmas which we shall give here without proof (assuming for simplicity that there are no hard cores).

Lemma 5.1: Let $\underset{\sim}{M} = (M^1,\ldots,M^s)$ be an integer particle number vector such that $\underset{\sim}{M} \cdot \underset{\sim}{E} = Q$. It is then possible to decompose $\underset{\sim}{M}$ into a "neutral" part $\underset{\sim}{N}$ and a "charged" part $\underset{\sim}{n}$, $\underset{\sim}{M} = \underset{\sim}{N} + \underset{\sim}{n}$ such that (i) $\underset{\sim}{N}$ and $\underset{\sim}{n}$ are both integer particle number vectors; (ii) $\underset{\sim}{N} \cdot \underset{\sim}{E} = 0$, $\underset{\sim}{n} \cdot \underset{\sim}{E} = Q$; (iii) it is impossible to decompose $\underset{\sim}{n}$ into a non-zero neutral part and a charged part; (iv) $|n| \equiv \sum_{i=1}^{s} n^i \leq \lambda|Q|$ with λ a constant.

Lemma 5.2: Let $\{\Lambda_j\}$ be a regular sequence of domains with $V(\Lambda_j) = V_j$ and let K be a compact subset of ρ. Let $\underset{\sim}{\mu}$ be a fixed chemical potential. Then there exists a sequence of numbers $\{\varepsilon_j\}$ (depending on K and $\underset{\sim}{\mu}$), tending to zero as $j \to \infty$, such that

$$\underset{\sim}{\mu} \cdot \underset{\sim}{n}V_j^{-1} + g(\underset{\sim}{M}V_j^{-1};\Lambda_j) - g(\underset{\sim}{N};\Lambda_j) \leq \varepsilon_j , \qquad (5.6)$$

whenever $\underset{\sim}{M}V_j^{-1} \in K$ and $\underset{\sim}{M} = \underset{\sim}{N} + \underset{\sim}{n}$ as in Lemma 5.1.

Lemma 5.3: Let $\{\Lambda_j\}$ be a sequence of regular domains with volumes $\{V_j\}$. Then there exists some fixed, strictly positive constants k and α independent of j such that

$$Z(\underset{\sim}{M};\Lambda_j) \leq \{\prod_{i=1}^{a} Z^{+}_{o,i}(M_i;\Lambda_j)\}\{\prod_{i=a+1}^{s}(M_i!)^{-1}[V_j]^{M_i}\}$$

$$\cdot \exp\{k\sum_{i=1}^{s} M_i - \alpha V_j^{-\frac{1}{3}}(\mu \cdot E)^2\}$$

for j sufficiently large. Here, $Z^{+}_{o,i}(M_i;\Lambda_j)$ is the ideal Bose gas partition function of M_i bosons of species i in the domain Λ_j.

The proof of Theorem 5.1 now proceeds as follows: Using Lemma 5.3 we establish that

$$\Xi(\underset{\sim}{\mu};\Lambda_j) \leq 2^{\sum_{M^1=o}^{[V_j/v_o]} \cdots \sum_{M^s=o}^{[V_j/v_o]}} \exp[\underset{\sim}{\mu} \cdot \underset{\sim}{M}] Z(\underset{\sim}{M};\Lambda_j) \quad (5.7)$$

for j sufficiently large, where v_o is some fixed small volume. The inequality (5.7) is easily obtained for non-Coulomb systems when the interactions among the bosons is superstable. The physical content of Lemma 5.3 is that the Coulomb energy is as efficacious as a superstable interaction in this respect; the Coulomb energy discourages a large excess of bosons over fermions. The number of terms in (5.7) is at most $(1 + V_j/v_o)^s$. If we now write $\underset{\sim}{M} = \underset{\sim}{N} + \underset{\sim}{n}$ as in Lemma 5.1 and use (5.6) we readily find that,

$$\Xi'(\underset{\sim}{\mu};\Lambda_j) \leq \Xi(\underset{\sim}{\mu};\Lambda_j) \leq$$

$$\leq 2 \sum_{(\underset{\sim}{N}\cdot\underset{\sim}{E}=o)} \cdots\cdots \sum e^{\underset{\sim}{\mu}\cdot\underset{\sim}{N}} Z(\underset{\sim}{N};\Lambda_j) \sum_{n^1=o}^{[V_j/v_o]} \cdots \sum_{n^s=o}^{[V_j/v_o]} e^{\underset{\sim}{\mu}\cdot\underset{\sim}{n}} \frac{Z(\underset{\sim}{N}+\underset{\sim}{n};\Lambda_j)}{Z(\underset{\sim}{N};\Lambda_j)}$$

$$\leq 2(2 + V_j/v_o)^s [\exp(\epsilon_j V_j)]\Xi'(\underset{\sim}{\mu};\Lambda_j), \quad (5.8)$$

so that

$$\pi'(\underset{\sim}{\mu};\Lambda_j) \leq \pi(\underset{\sim}{\mu};\Lambda_j) \leq \pi'(\underset{\sim}{\mu};\Lambda_j) + \delta_j \ , \quad (5.9)$$

and $\delta_j \to 0$ as $j \to \infty$. Equation (5.9) proves the equivalence of $\pi'(\underset{\sim}{\mu};\Lambda_j)$ and $\pi(\underset{\sim}{\mu};\Lambda_j)$ in the thermodynamic limit. The proof of the existence of $\pi'(\underset{\sim}{\mu})$ and (5.5) is identical to that for systems with tempered potentials.

Remark: Theorem 5.1 shows in a striking way the special nature of the Coulomb potential. In the absence of the Coulomb potential, but for any tempered potential, one can, by properly choosing the various chemical potentials μ_i, induce essentially any desired ratio of the densities ρ_i of the various species. For Coulomb potentials, on the other hand, only neutral densities are permitted in the thermodynamic limit. To be more specific, it can be readily shown that $\langle Q \rangle_j V_j^{-1} \to \infty$, where $\langle Q \rangle_j$ is the expectation value of the charge in Λ_j, for an arbitrary choice of the chemical potentials μ_i, $i = 1,\ldots,S$.

An interesting question arises about the behavior of the charge fluctuations $\langle [Q - \langle Q \rangle_j]^2 \rangle_j V_j^{-1}$ as $j \to \infty$. It seems certain on the basis of our previous results that this will approach zero (probably as $V_j^{-\frac{1}{3}}$) when $j \to \infty$, but we have not established this rigorously.

VI. The Microcanonical Ensemble For Neutral Systems

In the foregoing pages we discussed the existence and properties of the canonical and grand canonical free energies per unit volume. The microcanonical ensemble is and ensemble of even more physical and historical importance. From it the requisite thermodynamic properties of the canonical and grandcanonical ensembles may be deduced directly on general grounds, but the converse is not true. The microcanonical partition, function $\Omega(E,\underset{\sim}{N};\Lambda)$, is a function of energy, E, the domain, Λ, and the particle number vector $\underset{\sim}{N}$. There are many ways to define Ω, but in any case one defines an entropy/unit volume, σ, as a function of density, ρ, and energy/unit volume, ε, by

$$\sigma(\varepsilon,\underset{\sim}{\rho};\Lambda) \equiv V^{-1} \ell n \Omega(\varepsilon V, \underset{\sim}{\rho} V;\Lambda), \tag{6.1}$$

where $V = V(\Lambda)$. In addition to showing that σ has a thermodynamic limit which is concave in $(\varepsilon,\underset{\sim}{\rho})$, one also has to show that the various definitions of Ω yield the same limiting σ function. [See Ruelle (1969) and references quoted therein.]

Instead of following the usual route of first defining σ and then its inverse function $\epsilon(\sigma,\rho;\Lambda)$, we define ϵ directly to suit our purposes. We then show that it has all the requisite thermodynamic properties for neutral systems in general domains as we did in Section III for the canonical free energy. It can also be shown that our definitions of ϵ and σ (which is defined to be the inverse of our ϵ function) agree with the usual definitions in the thermodynamic limit. The "equivalence" of the microcanonical ensemble to the canonical and grand-canonical ensembles in this limit is a consequence of the general arguments developed for non-Coulomb systems (cf. Ruelle, 1969).

The Microcanonical Energy Function, ϵ

Definition: Consider a quantum system in a domain Λ (of volume V) with particle density ρ. Let $E_1 \le E_2 \le \ldots$ be the eigenvalues of the Hamiltonian arranged in increasing order (including multiplicity). Let $\sigma \in \mathbb{R}^1$ and let $\ell \ge 1$ be the smallest integer $\ge \exp(\sigma V)$. Then the energy function is defined by

$$\epsilon(\sigma,\rho;\Lambda) \equiv (V\ell)^{-1} \sum_{i=1}^{\ell} E_i \ . \tag{6 2}$$

Remarks: (i) H-stability provides the lower bound

$$\epsilon(\sigma,\rho;\Lambda) \ge |\rho|\Phi \ , \tag{6.3}$$

for some constant, Φ.

(ii) The range of $V\epsilon(\sigma,\rho;\Lambda)$ is $[E_1,\infty]$ since the Hamiltonian is unbounded above.

(iii) It is clear from the definitions that ϵ is non-decreasing in σ. Hence, the energy function has a pseudo-inverse called the entropy function which will be denoted by $\sigma(\epsilon,\rho;\Lambda)$. It is given explicitly by

$$\sigma(\epsilon,\rho,\Lambda) = \sup\{\sigma: \epsilon(\sigma,\rho;\Lambda) \le \epsilon\} \ . \tag{6.4}$$

Implicit in Eq. (6.2) is the notion that each E_i is defined for all ρ by linear interpolation. Thus, the definition, (6.4), of σ is not the same as one would obtain if one defined σ for non-integral particle numbers

by linear interpolation of σ. In other words, we have given priority to the energy function. It is also to be noted that while the domain of ϵ (in σ) is $(-\infty,\infty)$, the domain of σ (in ϵ) is $[E_1/V,\infty]$.

We now use the minimax principle which states that if $\{\psi_i\}$, $i = 1,\ldots,\ell$ is a set of ℓ orthonormal functions (called variational functions) in the domain of the Hamiltonian, H, and that if we form the ℓ-square Hermitian matrix A whose elements are $A_{ij} = (\psi_i,H\psi_j)$, and label the eigenvalues of A as $\lambda_1 \le \lambda_2 \le \ldots \le \lambda_\ell$ then $\lambda_i \ge E_i$ for $i = 1,\ldots,\ell$. In particular, for integral particle numbers,

$$\epsilon(\sigma,\underset{\sim}{\rho};\Lambda) \le (V\ell)^{-1}\mathrm{Tr}A, \tag{6.5}$$

where $\exp(\sigma V) = \ell$. This formula shows the advantage of our definition of ϵ because all we need to know are the diagonal elements of A.

To apply this principle, let $\Lambda \supset \Lambda_1 \cup \Lambda_2$, with Λ_1 and Λ_2 disjoint, and let $\underset{\sim}{N} = \underset{\sim}{N}_1 + \underset{\sim}{N}_2$ be the respective particle number in the various domains. If $\{\psi_i^1,E_i^1\}$, $i = 1,\ldots,n_1$ (resp. $\{\psi_i^2,E_i^2\}$, $i = 1,\ldots,n_2$) are the first n_1 (resp. n_2) eigenfunctions and eigenvalues in Λ_1 (resp. Λ_2), we can form the set of n_1n_2 variational functions in Λ by $\psi_{ij} = \psi_i^1 \otimes \psi_j^2$. To evaluate the right hand side of (6.5) we need consider only $A_{ij,ij}$ and this is given by

$$A_{ij,ij} = E_i^1 + E_j^2 + U_{ij}\ , \tag{6.6}$$

where U_{ij} is the expectation value of the inter-domain part of the potential energy. Obviously, (6.6) generalizes in a trivial way when Λ contains more than two disjoint subdomains.

The average interaction, U_{ij}, consists of a non-Coulomb, but tempered part and a Coulomb part. The former can be easily bounded and we shall ignore it in these notes. Bounding the Coulomb part U_{ij}^C is slightly more complicated.

Suppose that Λ_1 in the previous discussion is a ball, B. Each index i denoting the eigenfunctions and eigenvalues of the Hamiltonian in B can best be written as a pair (α,m) where α denotes the principal quantum numbers,

including the angular momentum, $L(\alpha)$ (irreducible representation of the rotation group), and m denotes the magnetic quantum number (row of the representation). The energy E_i depends only on α and not on m. Suppose further that n_1 is such that for every α all the levels (α,m) with $-L(\alpha) \le m \le L(\alpha)$ appear in the list $1,\ldots,n_1$ if any one (α,m') does. In that case we shall say that n_1 is perfect. $\sum_{m=-L}^{L} U^C_{(\alpha,m),j}$, which is part of the sum in (6.5), we have to evaluate an average charge density in Λ_1 which involves integrals over all but one of the N_1 particle coordinates in B, such as

$$I_\alpha(\underset{\sim}{r}) = \sum_{m=-L}^{L} \int_{B^{N_1-1}} |\psi^1_{(\alpha,m)}(\underset{\sim}{r},\underset{\sim}{r}_2,\ldots,\underset{\sim}{r}_{N_1})|^2 d\underset{\sim}{r}_2 \ldots d\underset{\sim}{r}_{N_1}$$

Clearly I_α depends only on the distance of $\underset{\sim}{r}$ from the center of B. If, in addition, we postulate that $\underset{\sim}{N}_1 \cdot \underset{\sim}{E} = 0$, i.e., that Λ_1 contains a neutral mixture of particles, then the average Coulomb potential outside of Λ_1 will vanish by Newton's theorem. That is

$$\sum_{i=1}^{n_1} U^C_{ij} = 0 \text{ for all } j \,, \tag{6.7}$$

regardless of the shape of Λ_2 and of its constituent particles. If n_1 is not perfect, it lies between two perfect numbers μ and ν, $\mu < n_1 < \nu$, $\nu - \mu = 2L(\alpha) + 1 \equiv t$, where α is the last principle quantum number appearing in the first n_1 levels. The sum $\sum_{i=1}^{\mu} U^C_{ij} = 0$ and can be ignored.

We are then left with $U = \sum_{i=\mu+1}^{n_1} U^C_i$ where $U^C_i \equiv \sum_{j=1}^{n_2} U^C_{ij}$. The key fact is that we can relabel the last t levels in ν such that $U \le 0$. This is so because $\sum_{i=\mu+1}^{\nu} U^C_i = 0$.

Writing, for $i = 1, 2$, $x_i = V_i/V$ and $\exp(\sigma_i V_i) = n_i$, then if $\exp(\sigma V) = n_1 n_2$ we have $\sigma = x_1\sigma_1 + x_2\sigma_2$. If we now denote the energy function of Λ_i by ε_i and if Λ_1 is a ball, then the preceding discussion shows that

$$\varepsilon(x_1\sigma_1+x_2\sigma_2, x_1\ell_1+x_2\ell_2;\Lambda) \le x_1\varepsilon_1(\sigma_1,\ell_1;\Lambda_1)+x_2\varepsilon_2(\sigma_2,\ell_2;\Lambda_2). \tag{6.8}$$

It can be shown that (6.8) is true even when n_1, n_2, $\underset{\sim}{N}_1$ and $\underset{\sim}{N}_2$ are not integral and that it generalizes to more than two subdomains provided all but one of them is a ball. Thus, we have established precisely the analogue of the inequalities on the g function of Section III. Therefore, the same analysis as that given in Section III will lead to the same conclusions for the energy function.

Our results are summarized in the following theorem:

Theorem 6.1: (i) Let $\{\Lambda_j, \underset{\sim}{N}_j\}$ be a sequence of regular domains and integer valued particle number vectors satisfying the neutrality condition $\underset{\sim}{N}_j \cdot \underset{\sim}{E} = 0$ and such that $\underset{\sim}{\rho}_j = V(\Lambda_j)^{-1}\underset{\sim}{N}_j$ satisfies $|\underset{\sim}{\rho}| < \rho_c$. Let a sequence of entropies $\{\sigma_j\}$ also be given and suppose that $\underset{\sim}{\rho} \to \underset{\sim}{\rho}$ with $|\underset{\sim}{\rho}| < \rho_c$ and $\sigma_j \to \sigma$. Then, the energy functions $\epsilon(\sigma_j, \underset{\sim}{\rho}_j; \Lambda_j)$ converge to a function $\epsilon(\sigma, \underset{\sim}{\rho})$ which is independent of the particular sequence.

(ii) $\epsilon(\sigma, \underset{\sim}{\rho})$ is continuous and convex in $(\sigma, \underset{\sim}{\rho})$ in the domain

$$D = \{(\sigma, \underset{\sim}{\rho}) : |\underset{\sim}{\rho}| < \rho_c, \underset{\sim}{\rho} \cdot \underset{\sim}{E} = 0, -\infty < \sigma < \infty\} .$$

It is also non-decreasing in σ.

(iii) $\epsilon(\sigma, \underset{\sim}{0}) = 0$.

(iv) Let K be a compact subset of D. Suppose that for each $(\sigma, \underset{\sim}{\rho}) \in K$ we have a sequence $\{\sigma_j(\sigma, \underset{\sim}{\rho}), \underset{\sim}{\rho}_j(\sigma, \underset{\sim}{\rho})\}$ which approaches $(\sigma, \underset{\sim}{\rho})$ uniformly on K. Then $\epsilon(\sigma_j, \underset{\sim}{\rho}_j; \Lambda_j)$ approaches $\epsilon(\sigma, \underset{\sim}{\rho})$ uniformly on K.

(v) The entropy function, $\sigma(\epsilon, \underset{\sim}{\rho}, \Lambda)$, also approaches a limit $\sigma(\epsilon, \underset{\sim}{\rho})$ uniformly on compacts.

(vi) $\sigma(\epsilon, \underset{\sim}{\rho})$ is continuous, and concave in the domain $D = \{(\epsilon, \underset{\sim}{\rho}) : |\underset{\sim}{\rho}| < \rho_c, \underset{\sim}{\rho} \cdot \underset{\sim}{E} = 0, \epsilon_1(\underset{\sim}{\rho})\}$, where $\epsilon_1(\underset{\sim}{\rho}) = \lim_{j\to\infty} E_1(\underset{\sim}{\rho}_j; \Lambda_j) V(\Lambda_j)^{-1}$ and $E_1(\underset{\sim}{\rho}; \Lambda)$ is the lowest eigenvalue of the Hamiltonian in Λ. It is also non-decreasing in ϵ and its range is not bounded above.

(viii) $\sigma(\epsilon, \underset{\sim}{\rho})$ and $\epsilon(\sigma, \underset{\sim}{\rho})$ are inverse functions.

REFERENCES

1. F. J. Dyson and A. Lenard, J. Math. Phys. 8, 423 (1967).

2. F. J. Dyson, J. Math. Phys. 8, 1538 (1967).

3. M. E. Fisher, Arch. Ratl. Mech. Anal. 17, 377 (1964).

4. M. E. Fisher and D. Ruelle, J. Math. Phys. 7, 260 (1966).

5. R. B. Griffiths, Rigorous Results and Theorems, to appear in Phase Transitions and Critical Points, C. Domb and M. S. Green, editors (Academic Press).

6. T. Kato, Perturbation Theory for Linear Operators, (Springer, 1966).

7. O. D. Kellog, Foundations of Potential Theory, (Dover Publication, 1953).

8. J. L. Lebowitz, Ann. Rev. Phys. Chem. 19, 389 (1968).

9. J. L. Lebowitz and E. Lieb, Phys. Rev. Lett. 22, 631 (1969).

10. A. Lenard and F. J. Dyson, J. Math. Phys. 9, 698 (1968).

11. E. H. Lieb and J. L. Lebowitz, The Constitution of Matter, to appear in Advances in Math.

12. C. B. Morrey, Comm. Pure. Appl. Math. 8, 279 (1955).

13. I. Newton, Philosophia Naturalis Principia Mathematica (1687): translated by A. Motte, revised by F. Cajori (University of California Press, Berkeley, California, 1934), Book 1, p. 193, propositions 71, 76.

14. L. Onsager, J. Phys. Chem. 43, 189 (1939).

15. L.Onsager, The Neurosciences, p. 75, G. C. Quarton, T. Melinchick and F. O. Schmitt, editors (Rockefeller University Press, 1967).

16. O. Penrose, in Statistical Mechanics, Foundations and Applications, edited by T. Bak (W. A. Benjamin, Inc., New York, 1967), p. 98.

17. D. Ruelle, Statistical Mechanics (W. A. Benjamin, N.Y. 1969).

DISSIPATIVE SYSTEMS AND DIFFERENTIAL ANALYSIS*

David Ruelle**
Institute for Advanced Study
Princeton, New Jersey

Introduction

In these notes we make a sketchy description of the application to dissipative systems of the qualitative theory of differential equations. One might say that the subject to be described is largely non-existent. Indeed, much remains to be understood about the qualitative behaviour of solutions of differential equations in finite dimension. Furthermore the differential equations for the time-evolution of dissipative systems are in infinite dimension, and an existence and uniqueness theorem for their solutions is usually not known. On the other hand, dissipative systems exhibit some of the most fascinating of natural phenomena (think for instance of the flow of liquids), and obviously deserve the growing interest devoted to them (see Prigogine[15], Thom[24]).

Chapter 1. The physical principles.

1. Nature of dissipative systems.

Dissipative systems are macroscopic systems described in terms of macroscopic variables in such a manner that time evolution leads to a continual dissipation of microscopic information (entropy production). More precisely, we consider a dissipative system as extended in physical space, and such that its state in small regions of space may be approximated by a thermodynamic equilibrium state. This

*This is a short version of lectures given in spring and summer 1971 in Brandeis and Boulder respectively. I hope to publish later an extended treatment of the same subject.
**Permanent address: IHES, 91. Bures-sur-Yvette, France.

state (described by a finite number of parameters) varies in general in space and time, and the system as a whole is not in equilibrium.

The time evolution of a dissipative system is given by a differential equation

$$\frac{d\xi}{dt} = X(\xi) \tag{1.1}$$

on a suitable manifold M. Here ξ is a vector field over physical space, the components of which are, at each point, the parameters describing the local thermodynamic equilibrium state. The r.h.s. of (1.1) is a vector field X on the infinite dimensional manifold M. The vector field X is locally of a rather unrestricted character, and therefore there is usually no useful variational principle associated with dissipative problems. On the other hand, the unrestricted nature of X makes the qualitative behaviour of dissipative systems less complicated in general than that of conservative (Hamiltonian) systems: the Kolmogorov-Arnold-Moser phenomena do not occur here (see for instance Arnold and Avez[2]).

2. Evolution equations.

Consider a continuous system consisting of a fluid mixture of a finite number of constituents which may participate in chemical reactions. The following processes lead to entropy production:
(i) internal friction due to viscosity
(ii) heat transfer
(iii) diffusion of matter
(iv) chemical reactions
Processes (i), (ii), and (iii) are called transfer processes. The evolution equations for the system under consideration are obtained by expressing the conservation of the masses of the various constituents and the conservation of the fluid momentum and energy.

We shall write the evolution equations in the special case of a system with only one constituent: the processes (iii) and (iv) are then not present.

We let ρ be the density, p the pressure, T the temperature, ϵ resp. s the energy resp. entropy per unit mass of fluid. All these variables satisfy the equilibrium thermodynamic relations. The velocity of the fluid is denoted by $v = (v_i)$. The following quantities are also associated with the dissipative processes (i) and (ii):

(i) viscosity stress tensor $\sigma = (\sigma_{ij})$

(ii) heat flow $q = (q_i)$

The evolution equations are then as follows

(a) Mass conservation (continuity equation)

$$\frac{\partial \rho}{\partial t} = - \frac{\partial}{\partial x_i} (\rho v_i) + E$$

(b) Momentum conservation

$$\frac{\partial}{\partial t} (\rho v_i) = - \frac{\partial}{\partial x_j} [\rho v_i v_j + (p\delta_{ij} - \sigma_{ij})] + F_i$$

(c) Energy conservation

$$\frac{\partial}{\partial t} (\tfrac{1}{2} \rho v^2 + \rho \epsilon)$$

$$= - \frac{\partial}{\partial x_i} [(\tfrac{1}{2} \rho v^2 + \rho\epsilon) v_i + (p\delta_{ij} - \sigma_{ij}) v_j + q_i] + G$$

The source terms $E, F=(F_i), G$ have been introduced for greater generality. They are functions of the position x and of the thermodynamic variables and the velocity at x. Notice that we may rewrite (c) as

$$\frac{\partial}{\partial t}(\rho \epsilon) = - \frac{\partial}{\partial x_i}(\rho \epsilon v_i + q_i) - (p \delta_{ij} - \sigma_{ij}) \frac{\partial v_i}{\partial x_j}$$

$$- \tfrac{1}{2} E v^2 + F_i v_i + G$$

The entropy s satisfies

$$T \, ds = d\epsilon - \frac{p}{\rho^2} d\rho$$

hence, omitting the source terms in (a), (b), (c),

$$T\left[\frac{\partial}{\partial t}(\rho s) + \frac{\partial}{\partial x_i}(\rho s v_i)\right] = \rho T\left(\frac{\partial s}{\partial t} + v_i \frac{\partial s}{\partial x_i}\right)$$

$$= \rho\left[\frac{\partial \varepsilon}{\partial t} - \frac{p}{\rho^2}\frac{\partial \rho}{\partial t} + v_i \frac{\partial \varepsilon}{\partial x_i} - \frac{p}{\rho^2} v_i \frac{\partial \rho}{\partial x_i}\right]$$

$$= \frac{\partial}{\partial t}(\rho\varepsilon) + \frac{\partial}{\partial x_i}(\rho\varepsilon v_i) + p\frac{\partial v_i}{\partial x_i} = -\frac{\partial q_i}{\partial x_i} + \sigma_{ij}\frac{\partial v_i}{\partial x_j}$$

Therefore

$$\left[\frac{\partial}{\partial t}(\rho s) + \frac{\partial}{\partial x_i}(\rho s v_i)\right] + \frac{\partial}{\partial x_i}\left(\frac{q_i}{T}\right) = -\frac{1}{T^2} q_i \frac{\partial T}{\partial x_i} + \frac{1}{T}\sigma_{ij}\frac{\partial v_i}{\partial x_j}$$

The r.h.s. is the rate of internal entropy production. It is positive and may be rewritten as

$$\sum_m J_m X_m \geq 0$$

where the X_k are <u>forces</u> $\frac{\partial T}{\partial x_i}$, $\frac{\partial v_i}{\partial x_j}$ and the J_k are <u>fluxes</u> or <u>currents</u>.

The fluxes J_m depend on the forces X_n and vanish when the latter vanish (at equilibrium). In the linear approximation we have thus <u>phenomenological laws</u>

$$J_m = \sum_n L_{mn} X_n$$

where the coefficients depend on the thermodynamic variables. From the reversibility of the microscopic laws of motion one derives the <u>Onsager reciprocity relations</u> of the type

$$L_{mn} = L_{nm}$$

Furthermore some coefficients L_{mn} vanish identically as a consequence of physical invariance laws.

In our case σ_{ij} and q_i are given by the phenomenological laws

$$\sigma_{ij} = \eta\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\,\delta_{ij}\,\frac{\partial v_k}{\partial x_k}\right) + \zeta\,\delta_{ij}\,\frac{\partial v_k}{\partial x_k}$$

$$q_i = \lambda\,\frac{\partial T}{\partial x_i} \qquad \text{(Fourier's law)}$$

The entropy production is

$$-\frac{1}{T^2}\,q_i\,\frac{\partial T}{\partial x_i} + \frac{1}{T}\,\sigma_{ij}\,\frac{\partial v_i}{\partial x_j}$$

$$= \frac{\lambda}{T^2}\left(\frac{\partial T}{\partial x_i}\right)^2 + \frac{\eta}{2T}\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}\,\delta_{ij}\,\frac{\partial v_k}{\partial x_k}\right)^2 + \frac{\zeta}{T}\left(\frac{\partial v_k}{\partial x_k}\right)^2$$

hence $\lambda,\ \eta,\ \zeta \geq 0$; λ is the thermal conductivity and $\eta,\ \zeta$ are the viscosity coefficients.

For an incompressible fluid, the evolution equations are

$$0 = \frac{\partial}{\partial x_i}\,v_i$$

$$\frac{\partial v_i}{\partial t} = -v_j\,\frac{\partial v_i}{\partial x_j} - \frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j}\,\sigma_{ij} = -v_j\,\frac{\partial v_i}{\partial x_j} - \frac{\partial p}{\partial x_i} + \frac{\eta}{\rho}\,\Delta\,v_i$$

(Navier-Stokes equation)

$$\frac{\partial \varepsilon}{\partial t} = -v_i\,\frac{\partial \varepsilon}{\partial x_i} - \frac{\partial q_i}{\partial x_i} + \sigma_{ij}\,\frac{\partial v_i}{\partial x_j} = -v_i\,\frac{\partial \varepsilon}{\partial x_i} + \lambda\,\Delta\,T + \frac{1}{2}\,\eta$$

$$\times\left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i}\right)^2$$

where η and λ have been assumed constant, and the source terms have been omitted.

3. Remarks on the time evolution of dissipative systems

Because of entropy production, an isolated dissipative system will "die out" i.e. tend to equilibrium. A more interesting situation is that in which the system receives a steady negative entropy flow from the outside; in that case non equilibrium regimes will be established, which may be stationary or non stationary.

The evolution equations have to be supplemented by boundary conditions. In particular, the velocity v of a viscous fluid at a solid boundary is that of the boundary (the fluid sticks). The boundary conditions and the source terms in the evolution equations both contribute to the entropy balance of a dissipative system. A stationary non equilibrium state can thus be achieved in spite of a strictly positive internal entropy production.

A mathematical theory of solutions of the evolution equations has been developed only in very special cases. In general, compressibility leads to characteristic physical phenomena: sound waves and shock waves, and the existence of shock waves necessitates the introduction of special prescriptions to supplement the evolution equations. The situations is better for an incompressible fluid but the non linear terms - in particular $v_j \frac{\partial v_i}{\partial x_j}$ - are not of a type treated by standard methods. In the well-studied case of the Navier-Stokes equation for 3-dimensional flows, no general existence and uniqueness theorem is known.

The idealization process which leads to the mathematical description of dissipative systems uses various approximations. In particular, in specific problems some dissipative processes are considered, and others disregarded. Apart from that the following approximations play an important role
(i) Macroscopic approximation. This implies in particular that the "mean free path" of molecules is

small in comparison with macroscopic lengths (a condition usually not satisfied in plasma physics).
(ii) Linearity of phenomenological laws. This is generally assumed for transfer processes but not for chemical reactions (where use is made of the nonlinear "mass action law"). Nonlinear phenomenological laws have been considered for incompressible viscous fluids in order to make the problem of existence and uniqueness for the evolution equations more amenable to study (see Ladyzhenskaja[12], supplement p. 193).
(iii) Constant phenomenological coefficients and thermodynamic functions. This is assumed for mathematical convenience.

4. Examples of dissipative systems.

We can only sketch the wealth of intriguing phenomena occuring in dissipative systems. In the next section we consider the viscous flow between rotating cylinders. For the flow of a viscous fluid past an obstacle see Feynman[7] II Section 41-4. For the onset of convexion in a fluid layer heated below (Bénard problem) see Chandrasekhar[4].

A simple and interesting example of dissipative system is provided by a homogeneous solution of chemical reactants. Here the evolution equation

$$\frac{dc}{dt} = X(c) \tag{4.1}$$

describes the change in time of a vector c in a finite-dimensional concentration space.

A stationary state (or steady state) is a constant solution of (4.1): $c(t) = c_0$ i.e. $X(c_0) = 0$. By maintaining the concentrations of some of the reactants at non equilibrium values, a non equilibrium stationary state is usually obtained. In some cases a periodic state with period t_0 appears, i.e. a non constant solution of (4.1) such that $c(t+t_0) = c(t)$, we observe then chemical oscillations. An example of chemical oscillations is easily realized at room temperature in the Beloussov reaction; the following mixture being alternatively yellow and colorless:

Malonic acid	0.3 Molar
Cerous nitrate	0.005 - 0.01 M
H_2SO_4	3.0 M
Na Br O_3	0.05 - 0.01 M

More colors are obtained by adding a little Ferroin.

5. Viscous flow between rotating cylinders.

Two coaxial circular cylinders with radii R_1 and R_2 rotate with constant angular velocities Ω_1 and Ω_2 (Fig. 1). The space between them is filled with a viscous incompressible fluid, and the flow of the fluid is investigated for various choices of Ω_1, Ω_2.

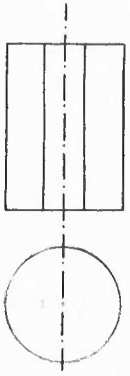


Fig. 1

The evolution equations are the Navier-Stokes equation

$$\frac{\partial v_i}{\partial t} = - v_j \frac{\partial v_i}{\partial x_j} - \frac{\partial p}{\partial x_i} + \nu \Delta v_i$$

(where $\nu = \eta/\rho$ is the kinematic viscosity) supplemented by the incompressibility relation

$$\frac{\partial v_i}{\partial x_i} = 0$$

We introduce cylindrical coordinates r, φ, z, the boundary conditions are then

$$v_r = v_z = 0, \quad v_\varphi = \Omega_1 R_1 \quad \text{for } r = R_1$$

$$v_r = v_z = 0, \quad v_\varphi = \Omega_2 R_2 \quad \text{for } r = R_2$$

It is also required that there be no net flow of the fluid along the vertical direction. This problem has a simple steady (time-independent) solution, the rotating Couette flow given by

$$v_\varphi = \frac{\Omega_2 R_2^2 - \Omega_1 R_1^2}{R_2^2 - R_1^2} r + \frac{(\Omega_1 - \Omega_2) R_1^2 R_2^2}{R_2^2 - R_1^2} \frac{1}{r} \qquad (5.1)$$

The flow between rotating cylinders has been extensively studied experimentally, the fluid being air or water (with suspended particles for visibility). A variety of flow patterns are obtained. We describe some of them, but warn the reader that new experiments might change details in the following descriptions (see Fig. 2).

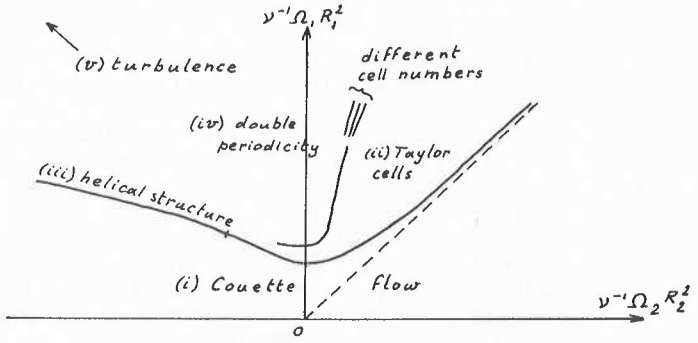


Fig. 2.

(i) The Couette flow (5.1) is observed for small $|\Omega_1|, |\Omega_2|$, and also when $\Omega_2 R_2^2 > \Omega_1 R_1^2 > 0$

(ii) For higher values of Ω_1 the Couette flow is replaced by another steady flow, the structure of which has been theoretically and experimentally investigated by G. I. Taylor[23]. Here the fluid column is vertically divided in cells - Taylor cells - and a vertical and radial motion is superimposed on the (horizontal) Couette flow (Fig. 3). The new flow is still time independent and rotationally symmetric, but has lost the invariance under vertical translations (translations along the z-axis). We have here a beautiful

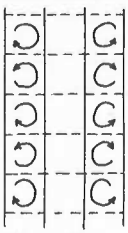


Fig. 3.

example of symmetry breakdown in physics.
(iii) when the cylinders turn in opposite directions, instead of a transition Couette flow → Taylor cells, one may observe a transition Couette flow → helical wave form. The helical wave form consists in that the Taylor cells, instead of being annular, are joined to form a structure which winds helically between the two cylinders (see Snyder[22]). Here the time invariance and the rotational invariance of the flow have both been destroyed.
(iv) In other situations the Taylor cells retain their general shape but loose their rotational invariance, which is replaced by an angular periodicity with period $\frac{2\pi}{m}$ ($m = 1,2,\ldots$) - the boundaries of the cells undulate with this period - and time invariance is lost. These flows are called doubly periodic.
Notice that the experiments are made with cylinders of finite height L (one can idealize this situation by restricting one's attention to solutions of the Navier-Stokes equation which are periodic with period L with respect to the z ordinate). Let n be the total number of cells. It has been shown experimentally (Coles[5]) that, for each choice of Ω_1, Ω_2 in a certain region, there exist several possible flows characterized by different values of m and n and stable under small disturbances.
(v) There is a domain of values of the pair (Ω_1, Ω_2) which leads to turbulent flow. Here is an experimental (!) description (extracted from Coles[5]) of the onset of turbulence: "...The ensuing motion then remains laminar and doubly periodic for a time, as successively higher-order harmonics of the basic frequencies are excited. At sufficiently high speeds however, the flow becomes noticeably no longer quite laminar, in the sense that irregularities have begun to appear, especially in the motions of smallest scale. Further increases in speed then increase the degree of irregularity until finally the flow can only be described as fully turbulent."

Chapter 2. Differential analysis.

6. Qualitative theory of differential equations.

In the next chapter we propose to use the qualitative theory of differential equations to study the evolution equation

$$\frac{dx}{dt} = X(x)$$

of a dissipative system. It is not possible here to reproduce all the necessary background of differential analysis - ancient and modern - and we shall mostly limit ourselves to some bibliography.

Calculus in Banach space, in particular the implicit function theorem, is described in Dieudonné[6]. See Lang[13] for manifolds, vector fields, and the basic existence and uniqueness theorems for differential equations. The C^k vector fields on a manifold have (usually) a natural topology of Banach or Fréchet space $\mathfrak{X}^k$; a property of vector fields is called C^k generic if it is satisfied on a residual set of $\mathfrak{X}^k$. This is discussed in Abraham[1], which is also referred to for the study of closed orbits of a vector field and their associated Poincaré maps.

Remarkable progress has been made recently in understanding the structure of orbits of vector fields or diffeomorphisms. Smale[21] is the basic reference in this area; it discusses the main results on the nonwandering set of a diffeomorphism and gives in particular examples of strange attractors. In section 7 below we give some indications on the work of Hirsch, Pugh, and Shub concerning the center manifold theorem, and invariant manifolds with normally hyperbolic diffeomorphism.

7. Some results on invariant manifolds.

Let 0 be a fixed point of a C^k diffeomorphism f of the Banach space E. We suppose that the spectrum of the derivative Df(0) is disjoint from the unit circle, i.e. 0 is a hyperbolic fixed point. The

points x such that $\lim_{n \to \infty} f^n s = 0$ form a C^k manifold called stable manifold and tangent at 0 to the linear space corresponding to the part of the spectrum of Df(0) which is inside the unit circle. Similarly the unstable manifold consists of those x such that $\lim_{n \to -\infty} f^n x = 0$ and corresponds to the part of the spectrum of Df(0) which is outside the unit circle. The situation for vector fields is analogous.

Clearly, the stable and unstable manifolds are invariant under f. If f is not hyperbolic but has a finite number of isolated eigenvalues of finite multiplicity on the unit circle, various invariant manifolds can still be defined under suitable conditions: stable, centerstable, center, centerunstable, and unstable. The center manifold will interest us most: it is tangent to the subspace of E corresponding to the part of the spectrum of Df(0) which is on the unit circle. The existence of a center manifold which is C^k, $k < +\infty$, is asserted by the center manifold theorem, for which see Hirsch, Pugh, and Shub[8]. There is no uniqueness: there may be several center manifolds.

We shall use the center manifold to study a diffeomorphism f_μ, depending on the real parameter μ, when one or several eigenvalues of $Df_\mu(0)$ cross the unit circle, say for $\mu=0$. It is then convenient to add one dimension to E and to consider the diffeomorphism $(x,\mu) \to (f_\mu x, \mu)$ of $E \times \mathbb{R}$ at the fixed point (0,0). The new diffeomorphism has just one more eigenvalue (1) on the unit circle.

Similar results hold if 0 is a critical point of a vector field X : X(0) = 0. The role of Df(0) is taken by the Jacobian and that of the unit circle by the imaginary axis.

Let f be a diffeomorphism of a manifold, which leaves invariant a compact submanifold V. It is sometimes possible to say that a diffeomorphism f' close to f has a compact invariant manifold V' close to V. This is the case if f is normally hyperbolic to V. Roughly speaking, this means that f is more

hyperbolic (i.e. contracting or expanding) in directions normal to V than tangent to V. For a precise statement see Hirsch, Pugh, and Shub[8] or[9].

8. Normal forms.

If we want to study a vector field near a critical point, or a diffeomorphism near a fixed point, it is generally useful to choose the coordinates so that the lower order terms in the expression for the vector field or diffeomorphism take a simple form. Following Jost and Zehnder[11] we shall make use of such normal forms in discussing bifurcation problems. Here we state a typical result.

Theorem. Let $X_\mu = (X_{\mu i})$ be a vector field in $\mathbb{R}^{2q}$ depending on the parameter $\mu \in \mathbb{R}$ and such that $X_\mu(0) = 0$, $(x, \mu) \to X_\mu(x)$ is C^k while $x \to X_\mu(x)$ is C^ℓ for all μ, $3 \le k \le \ell \le +\infty$. Let $A_\mu = (A_{\mu ij})$ be the Jacobian of X_μ at 0:

$$A_{\mu ij} = \frac{\partial X_{\mu i}}{\partial x_j}(0)$$

We let A_μ have $2q$ complex eigenvalues $\lambda_1, \ldots, \lambda_q$, $\bar{\lambda}_1, \ldots, \bar{\lambda}_q$. We assume that there is no linear relation of the form

$$\sum_{n=1}^{q} (s_n \lambda_n + t_n \bar{\lambda}_n) = \lambda_i$$

$(1 \le i \le q)$ for any choice of integers $s_n, t_n \ge 0$ such that $\sum_n (s_n + t_n) = 2$ or 3, except relations of the form $\lambda_i + \lambda_j + \bar{\lambda}_j = \lambda_i$ (this exception permits the λ_j to become pure imaginary). There is then a change of coordinates $\mathbb{R}^{2q} \to \mathbb{C}^q$ such that $(x,\mu) \to \psi_\mu x$ is C^{k-3} and, for each μ, $x \to \psi_\mu x$ is C^∞.

In the new coordinates $z_1,\dots,z_q$ *the vector field* X_μ *takes the form* $(Z_1,\dots,Z_q)$:

$$Z_i = (\lambda_i + \sum_{n=1}^{q} b_{in}|z_n|^2)z_i + o_3$$

where the λ_i *are* C^{k-1} *and the* b_{in} *are* C^{k-3} *functions of* μ; $(x, \mu) \to Z_{\mu i}$ *is* C^{k-3} *and, for each* μ, $x \to Z_{\mu j}$ *is* C^{ℓ}.

Chapter 3. Bifurcation theory of dissipative systems.

9. General principles.

Let again the evolution equation of a dissipative system be

$$\frac{dx}{dt} = X(x) \tag{9.1}$$

For most dissipative systems this is a differential equation on an infinite dimensional manifold. Furthermore the vector field X is often not well behaved so that the standard theory does not apply. Under these circumstances it will be convenient to think of (9.1) as a differential equation on a finite-dimensional manifold. This will be sufficient for the intuitive picture we want to give here - some references will be given to precise results.

We consider a dissipative system subjected to a time-independent external action described by a real parameter μ. Equation (9.1) then becomes

$$\frac{dx}{dt} = X_\mu(x) \tag{9.2}$$

where we assume that $(x,\mu) \to X_\mu(x)$ is C^k $(1 \le k \le \infty)$. The study of changes in the structure of orbits of a differential equation as a parameter is varied is the object of *bifurcation theory*. We are thus interested in the bifurcation theory of dissipative systems.

For $\mu = 0$, we suppose that the external action on the system vanishes. (9.2) has then a solution $x = \xi_0$ corresponding to equilibrium, and all solutions $x(t)$ of (9.2) tend to ξ_0 when $t \to +\infty$. We assume therefore that ξ_0 is an attracting critical point of the vector field X_0 in the sense that the eigenvalues of the Jacobian have strictly negative real parts. The Jacobian $A = (A_{ij})$ defined by

$$A_{ij} = \frac{\partial X_{oi}}{\partial x_j}(0)$$

is then an invertible matrix, and therefore the implicit function theorem shows the existence of a smooth function $\mu \to \xi_\mu$ such that $X_\mu(\xi_\mu) = 0$ for μ in a neighbourhood of 0. For sufficiently small μ the eigenvalues of the Jacobian of X_μ at ξ_μ have strictly negative real parts (by continuity) and ξ_μ is thus an attracting critical point for X_μ. We may interpret ξ_μ as a steady state of the dissipative system described by the vector field X_μ; it is stable in the sense that small perturbations are damped off (exponentially) as time tends to $+\infty$. As μ increases, one or more eigenvalues of $A(\mu)$ may cross the imaginary axis, ξ_μ then "looses its stability". In hydrodynamics the operator corresponding to A is unbounded and its study is the basis of linear stability theory (see Lin[14], Sattinger[18]).

10. Loss of stability of a steady state and Hopf bifurcation.

As above let A_μ be the Jacobian of the vector field A_μ at the critical point ξ_μ. For small μ, ξ_μ is attracting. As μ increases a real eigenvalue $\lambda(\mu)$ may cross 0, or a pair of complex conjugate eigenvalues $\lambda(\mu)$, $\bar{\lambda}(\mu)$ may cross the imaginary axis (it is intuitively clear that other possibilities are not generic).

If the real eigenvalue $\lambda(\mu)$ crosses 0 for $\mu = \mu_1$, A_{μ_1} is clearly not invertible. Therefore the implicit function theorem which we used to prove the existence of ξ_μ does not apply. What happens generically

is that, for $\mu=\mu_1$, the attracting critical point ξ_μ coalesces with a saddle type critical point ξ'_μ and, for $\mu > \mu_1$, there is no attracting set close to ξ_{μ_1} (Fig. 4). This possibility is relatively uninteresting. We shall see in a moment that, when two complex conjugate eigenvalues of A cross the imaginary axis, the steady solution ξ_μ is replaced by a periodic solution. It remains thus to understand how, in the flow between rotating cyclinders for instance (see Section 5) a steady flow (Couette) can loose its stability and be replaced by another steady flow (Taylor cells). What happens here is that, because of the symmetry of the system (invariance under the group of vertical translations and reflections) the vector field X_μ has a non generic behaviour. The study of systems with an invariance group has therefore to be made separately; it may lead to bifurcations from one steady flow to other steady flows.

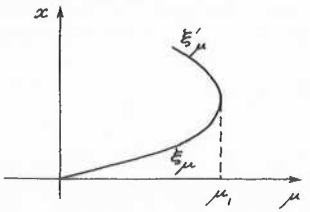


Fig. 4

Suppose now that two complex conjugate eigenvalues $\lambda(\mu)$, $\bar{\lambda}(\mu)$ of A_μ cross the imaginary axis for $\mu=\mu_1$. There exists then, in a neighbourhood of (ξ_{μ_1}, μ_1) a one-parameter family of closed orbits of X_μ. This fact was first proved by E. Hopf[10], and the associated transition from critical point to periodic orbit is known as Hopf bifurcation.

It is easy to understand how the Hopf bifurcation takes place in 2 dimensions. For $\mu < \mu_1$ (resp. $\mu > \mu_1$) the integral curves of X_μ spiral inward towards ξ_μ (resp. outward from ξ_μ). This is true very close to ξ_μ. Some distance away from ξ_μ, the transition from spiraling inward to outward may be delayed (see Fig. 5) or advanced, and this leads to a closed orbit.

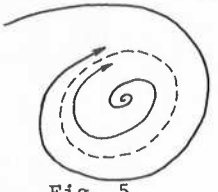


Fig. 5

The original proof of Hopf assumes that $(x,\mu) \to X_\mu(x)$ is real analytic. A proof in the C^k case $(3 \leq k < +\infty)$ is obtained as follows. First use the center manifold theorem to reduce the problem to 2 dimensions. Then put the vector field in the normal form

$$Z = (\lambda(\mu) + b(\mu)|z|^2)|z| + o_3$$

The vector field $Z - o_3$ has the closed orbit

$$|z| = (-\text{Re } \lambda(\mu)/\text{Re } b(\mu))^{\frac{1}{2}}$$

Using the theorem of Hirsch, Pugh and Shub on invariant manifolds with normally hyperbolic diffeomorphism (or vector field) one can show that this closed orbit is not destroyed by the perturbation o_3.

The Hopf bifurcation explains the occurence of periodic oscillation in hydrodynamic systems after loss of stability of a steady solution (see Bruš-linskaja[3], Sattinger[19]). It explains also chemical oscillations (see Sel'kov[20], I am indebted to J. P. Eckmann for explaining this reference to me).

11. Other bifurcations.

The bifurcations of a closed orbit can be studied by investigating the corresponding Poincaré map Φ. The closed orbit is attracting if the eigenvalues of $D\Phi(0)$ are in the open unit disk. Loss of stability occurs when the unit circle is crossed either by a pair of complex conjugate eigenvalues or by a real eigenvalue at ± 1. Crossing at $+1$ generically leads to a destruction of the closed orbit. The other possibilities may lead to the replacement of the original attracting closed orbit by one or several attracting closed orbits the periods of which are an integral multiple of the original period. In particular in the case of a real eigenvalues crossing at -1, one has a doubling of the period. It is interesting to remark that such a doubling of period is apparently observed in some chemical oscillations (see Fig. 6 in Pye and Chance[16]).

It is difficult to have a complete picture of the structure of orbits of (9.2) when μ becomes large. It is however possible to obtain some idea of what happens by using the normal forms of section 8. We assume thus that ξ_μ is a critical point of X_μ and that, as μ increases, successive pairs of complex conjugate eigenvalues $\lambda_1, \bar{\lambda}_1, \ldots, \lambda_q, \bar{\lambda}_q$ cross the imaginary axis from left to right. By using somehow the center manifold theorem and then introducing suitable complex coordinates $z_1, \ldots, z_q$ we may assume that the vector field X_μ has the form

$$(Z_1, \ldots, Z_q) = (Z_1', \ldots, Z_q') + o_3$$

$$Z_i' = (\lambda_i + \sum_{n=1}^{q} b_{in} |z_n|^2) z_i$$

We consider first the equations

$$\frac{dz_i}{dt} = Z_i'$$

They imply

$$\frac{d}{dt} |z_i|^2 = 2(\mathrm{Re}\ \lambda_i + \sum_{n=1}^{q} \mathrm{Re}\ b_{in} |z_n|^2) |z_i|^2$$

Therefore the manifolds defined by the equations

$$\mathrm{Re}\ \lambda_i + \sum_{n=1}^{q} \mathrm{Re}\ b_{in} |z_n|^2 = 0 \quad \text{or} \quad |z_i|^2 = 0$$

for each i are invariant for Z'. It is easy in specific cases to discuss the attracting, saddle-type, or repulsive character of each one of these manifolds. For instance if q=2 the following possibilities exist (among others which are less interesting)

(a) The following two one-dimensional invariant manifolds (closed orbits) are attracting:

$$\mathrm{Re}\ \lambda_1 + \sum_{n=1}^{2} \mathrm{Re}\ b_{1n} |z_n|^2 = 0, \quad |z_2|^2 = 0$$

and

$$|z_2|^2 = 0,\ \mathrm{Re}\ \lambda_2 + \sum_{n=1}^{2} \mathrm{Re}\ b_{2n}|z_n|^2 = 0$$

(b) The following two-dimensional invariant manifold (torus) is attracting

$$\mathrm{Re}\ \lambda_1 + \sum_{n=1}^{2} \mathrm{Re}\ b_{1n}|z_n|^2 = 0,$$

$$\mathrm{Re}\ \lambda_2 + \sum_{n=1}^{2} \mathrm{Re}\ b_{2n}|z_n|^2 = 0$$

If the perturbation o_3 is sufficiently small, one finds (see section 7) that Z (and hence X_μ) has again two attracting closed orbits in case (a) or one attracting 2-torus in case (b).

The interest of case (a) is to show that a dissipative system may conceivably have several different attractors. Which attractor is chosen by the system depends then on initial conditions. An experimental example of this situation was mentioned in Section 5 where several "doubly periodic" flows may coexist.

One can see that case (b) corresponds to the bifurcation of a closed orbit (as discussed above) when two complex conjugate eigenvalues of $D\Phi(0)$ cross the unit circle. Using Peixoto's theorem (see for instance Abraham[1] one finds that the attracting invariant 2-torus carries a finite number of attracting closed orbits.

We consider now the analog of case (b) when $q>2$. We have here an attracting q-dimensional torus, and the problem is to understand the nature of the orbits on this torus. As it turns out very complicated orbit structures may occur; in particular strange attractors arise for $q \geq 4$ (see Ruelle and Takens[17]). The integral curve of a vector field near a strange attractor has an apparently irregular and chaotic appearance, and depends very sensitively on initial conditions. It was proposed by Ruelle and Takens[17]

that the phenomenon of turbulent flow in a viscous liquid is described by an integral curve of the equations of motion which is asymptotic to a strange attractor. This proposal is in agreement with the qualitative properties of turbulent flow and, as we have just shown, is not unreasonable from the point of view of bifurcation theory.

REFERENCES

1. R. Abraham and J. Marsden, Foundations of Mechanics, New York, Benjamin, 1967.

2. V. I. Arnold et A. Avez, Problèmes ergodiques de la mécanique classique, Gauthier-Villars, Paris, 1967.

3. N. N. Brušlinskaja, The behaviour of solutions of the equations of hydrodynamics when the Reynolds number passes through a critical value, Doklady A.N.S.S.R. 162, No. 4 (1965).

4. S. Chandrasekhar, Hydrodynamic and hydromagnetic stability, Clarendon, Oxford, 1961.

5. D. Coles, Transition in circular Couette flow, J. Fluid Mech. 21, 385-425 (1965).

6. J. Dieudonné, Foundations of modern analysis, Academic Press, New York, 1960.

7. R. P. Feynman, R. B. Leighton and M. Sands, The Feynman lectures on Physics, Vol II. Addison-Wesley, Reading, Mass., 1964.

8. M. Hirsch, C. C. Pugh, and M. Shub, Invariant manifolds, To appear

9. M. Hirsch, C. C. Pugh, and M. Shub, Invariant manifolds, Bull. A.M.S. 76, 1015-1019 (1970).

10. E. Hopf. Abzweigung einer periodischen Lösung von einer stationären Lösung eines Differentialsystems, Ber. Mat.-Phys. Kl. Sächs. Akad.Wiss. Leipzig 94, 1-22 (1942).

11. R. Jost and E. Zehnder, A generalization of the Hopf bifurcation theorem. Helv. Phys. Acta. To appear.

12. O. A. Ladyzhenskaja, The mathematical theory of viscous incompressible flow, Second English edition, Gordon and Breach, New York, 1969.

13. S. Lang, Introduction to differentiable manifolds, Second Edition, Interscience, New York, 1966.

14. C. C. Lin, The theory of hydrodynamic stability, Cambridge University Press, 1966.

15. I. Prigogine, Introduction à la thermodynamique des processus irréversibles, Dunod, Paris, 1968.

16. K. Pye, and B. Chance, Sustained sinusoidal oscillations of reduced pyridine nucleotide in a cell-free extract of Sacchoromyces carlsbergensis, Proc. Nat. Acad. Sci. U.S.A. 55, 888-894 (1966).

17. D. Ruelle and F. Takens, On the nature of turbulence, Commun. math. Phys. 20, 167-192 (1971).

18. D. H. Sattinger, The mathematical problem of hydrodynamic stability, J. Math. Mech. 19, 797-817 (1970).

19. D. H. Sattinger, Bifurcation of periodic solutions of the Navier-Stokes equations, Arch. Rat. Mech. Anal. 41, 66-80 (1971).

20. E. E. Sel'kov, Self-oscillations in glycolysis, 1. A simple kinetic model, European J. Biochem. 4, 79-86 (1968).

21. S. Smale, Differentiable dynamical systems, Bull. A.M.S. 73, 747-817 (1967).

22. H. A. Snyder, Stability of rotating Couette flow, I. Asymetric waveforms. Phys. Fluids, 11, 728-734 (1968).

23. G. I. Taylor, Stability of a viscous liquid contained between two rotating cylinders, Phil, Tran. Roy. Soc. A 223, 289-343 (1923).

24. R. Thom. Stabilité structurelle et morphogénèse, Benjamin, New York.

COMPOSITE PARTICLES IN MANY-BODY SYSTEMS

Arthur Y. Sakakura
Colorado School of Mines
Golden, Colorado

I. Introduction

In these lectures a second quantized formalism for composite particles in many-body systems will be presented. In order to explicitly define our terms we consider a system of N protons and N electrons interacting through the coulomb potential. At non-relativistic energies the proton and the electron are elementary particles, whereas the H-atom, H_2-molecule, H^--ion, etc. are the composite particles, which we will indiscriminately call the atoms.

It is well known[1] that when composite particles are far apart, they may be considered as integral entities with internal structures and obeying definite statistics, i.e., H-atoms and H_2-molecules are bosons, and H^--ions are fermions. When they are close together, the indistinguishability of the constituent elementary particles negates the notion of individual atoms. Yet, the concept of atoms as entities has been most useful, so we wish to express the hamiltonian and other observables in terms of atomic, protonic, and electronic creation-annihilation operators from first principles.

The emphasis is important, for one can always write the atom-atom interaction with a phenomenological potential. First principles means: given a system of protons and electrons interacting through the coulomb potential and taking into account that these particles are fermions, derive systematically the atom-atom and other composite particle interactions.

The important problem of deriving the interactions aside, it is an advantage to introduce the composite particles in the zero order. This means that the hamiltonian

will explicitly describe processes such as

$$H^+ + H \quad H^+ + H^+ + e^-$$

$$H^+ + e^- \quad H \qquad \text{(I-1)}$$

etc., so that the equilibrium and non-equilibrium properties relating to various chemical constituents can be calculated in the zero order without recourse to many-particle green functions.

II. Statement of the Problem

To further simplify the admittedly involved formalism, the only atom to be considered will be the H-atom. This choice does not alter the problem, but merely precludes the explicit appearance of terms involving H_2, H^-, etc. How these are to be incorporated will be indicated later. The proton coordinates including the spins will be denoted by $x_1 \ldots . x_N$, the electron coordinates by $y_1 \ldots . y_N$, and the entire set by $\underline{xy}$. The hamiltonian in the Schrödinger representation is

$$H_S = \sum_{i=1}^{N} (T_+(i)+T_-(i)) + \sum_{i>j=1}^{N} [v(|\vec{x}_j-\vec{x}_j|)+v(|\vec{y}_i-\vec{y}_j|)]$$

$$- \sum_{i,j=1}^{N} v(|\vec{x}_j-\vec{y}_i|) \qquad \text{(II.1)}$$

where v is the repulsive coulomb potential, and T_+ and T_- are the kinetic energy operators of the proton and the electron, respectively.

The hilbert space in which H_S operates is spanned by

$$|\underline{xy}> \equiv |x_1> \ldots |xy> |y_1> \ldots |y_N> \qquad \text{(II.2)}$$

whereas the physically admissible states ("physical states") form a subspace defined by

$$A_N|\Phi> = |\Phi> \qquad \text{(II.3)}$$

where

$$A_N|\underline{xy}> = (N!)^{-2} \sum_P \delta_P |x_P >\ldots|x_{PN}> |y_P >\ldots|y_{PN}>$$

$$\equiv (N!)^{-2} \sum_P \delta_P |P\underline{xy}> \qquad \text{(II.4)}$$

and where P is the product, (permutation of the N proton coordinates) x (permutation of the N electron coordinates). δ_P is the signature of P. A_N is then the operator which skew symmetrizes the state with respect to the N proton coordinates and the N electron coordinates simultaneously. It is hermetian and idempotent, and commutes with all the operators of the system observeables.

The standard representation, which we will call the elementary particle representation, is the second quantized one, wherein

$$H_E = \int \psi(x)^+ T_+ \psi(x)dx + \int \psi(y)^+ T_- \psi(y)dy$$

$$+\tfrac{1}{2}\int \psi^+(x_1)\psi^+(x_2)v(|\vec{x}_1-\vec{x}_2|)\psi(x_2)\psi(x_1)dx_1dx_2$$

$$- \int \psi^+(x)\psi^+(y)v(|\vec{x}-\vec{y}|)\psi(y)\psi(x)dxdy$$

$$+\tfrac{1}{2}\int \psi^+(y_1)\psi^+(y_2)v(|\vec{y}_1-\vec{y}_2|)\psi(y_2)\psi(y_1)dy_1dy_2 \qquad \text{(II.5)}$$

and

$$A_N = \int \psi(x_1)^+\ldots\psi^+(x_N)\psi(y_1)^+\ldots\psi(y_N)^+|0>(N!)^{-2}$$

$$<0|\psi(y_N)\ldots\psi(y_1)\psi(x_N)\ldots\psi(x_1)\ d\underline{xy} \qquad \text{(II.6)}$$

where $\psi(x)$ and $\psi(y)$ are the usual wave function operators for the proton and the electron, respectively, and $|0>$ is the normalized vacuum state. The operation, $\int dx$, means integration over the space coordinates and summation over the spin coordinates, whereas the operation, $\int \underline{dxy}$, means the integration and summation over the coordinates of all the particles. The ψ obey the normal commutation relations:

$$[\psi(x),\psi(x')^+]_+=\delta(x,x') \qquad [\psi(y),\psi^+(y')]_+=\delta(y,y')$$

$$0=[\psi(x),\psi(x')]_+=[\psi^+(x),\psi^+(x')]_+=[\psi^+(y),\psi^+(y')]_+=[\psi(y),\psi(y)']_+$$

$$0=[\psi(x),\psi(y)]_+=[\psi(x),\psi(y)^+]_+=[\psi(x)^+,\psi(y)^+]_+=[\psi(x)^+,\psi(y)]_+ \tag{II.7}$$

The last set of relations enables the product, $\psi(y)\psi(x)$, to commute with both $\psi(x')$ and $\psi(y')$, thus partaking a useful characteristic of an independent bose operator.

Two points need emphasizing. The hamiltonian, H_E, contains no explicit references to composite particles, and A_N is the identity operator for the sub space containing N particles each of two different fermions.

Girardeau[2] introduced atoms via the basis

$$|\alpha_1\ldots\alpha_N\rangle \equiv (N!)^{-1}\sum_R \int d\underline{xy}\,|\underline{xy}\rangle\, \varphi_{\alpha_{R1}}(x_1y_1)\ldots\varphi_{\alpha_{RN}}(x_Ny_N) \tag{II.8}$$

where φ_α is the complete set of atomic wave functions, and R is a permutation of N objects. One now introduces the creation and annihilation operators for the state, α, obeying the bose commutation relations:

$$[a_\alpha,a_\alpha^+,] = \delta(\alpha,\alpha'),\ [a_\alpha,a_\alpha,]=[a_\alpha^+,a_\alpha^+,]=0 \tag{II.9}$$

Girardeau obtained

$$H_A=\sum_\alpha E_\alpha a_\alpha^+ a_\alpha+\tfrac{1}{2}\sum_{\alpha_1\alpha_2\alpha_1'\alpha_2'} a_{\alpha_1}^+ a_{\alpha_2}^+ \langle\alpha_2\alpha_1|V_{AA}|\alpha_1'\alpha_2'\rangle a_{\alpha_2'}a_{\alpha_2'} \tag{II.10}$$

where E_α is the energy of α including the center of mass energy, and V_{AA} is the coulomb interaction energy of two atoms:

$$V_{AA} = v(|\vec{x}-\vec{x}_2|)+v(|\vec{y}_1-\vec{y}_2|)-v(|\vec{x}_1-\vec{y}_2|)-v(|\vec{y}_1-\vec{x}_2|) \tag{II.11}$$

We will call this the atomic representation.

A_N, however, is not the identity for the subspace of N atoms, so the physical states must obey the subsidiary

condition (II.3) or its equivalent, the satisfaction of which presents a formidable problem.

Another source of difficulty is the necessary inclusion of the continuum atomic states for completeness. By treating these states as atoms, recombination and ionization terms do not appear in H_A. A mixed representation with a basis of bound atomic states and free electron and free proton states will explicitly display these desired terms. This, however, opens a new Pandora's box of problems associated with orthogonality and completeness. The lectures by Brittin[3], Girardeau[4], and myself are all concerned with different approaches to the problems of orthogonality and completeness and of the elimination of the subsidiary condition in the mixed representation.

In these lectures, a complete orthonormal set in the mixed representation will be shown with one curious feature: every orthonormal subspace, except one, contains a continuum boson, the remainder of the particles being bound atoms and/or free electrons and protons. The elimination of the continuum boson is equivalent to the strong orthogonality of Girardeau[4,5].

We avoid the subsidiary conditions by utilizing Giradeau's[6] observation about the projected hamiltonian, $\tilde{H}$,

$$\tilde{H} \equiv A_N H = H A_N = A_N H A_N \qquad \text{(II.12)}$$

The eigenstates of $\tilde{H}$ are the physical states of H. So instead of examining H (and other observables) in the physical subspace, we consider $\tilde{H}$ in the entire hilbert space. In other words, the burden of symmetry is to be carried by $\tilde{H}$, rather than by the state.

Unfortunately, the eigenstates belonging to the zero eigenvalue of H include all the unphysical states so that the solution of the Schrödinger equation,

$$i\hbar \, |\dot{\Phi}> = \tilde{H} \, | \, \Phi > \qquad \text{(II.13)}$$

may contain a large time-independent component of the unphysical states. But the physically interesting quantities are thermal averages of observables, O, which can be written,

$$\begin{aligned}\langle 0\rangle &= \text{Trace over physical states } \rho\ 0 \\ &= \text{Trace over all states } A_N \rho\ 0 \\ &= \text{Trace over all states } \tilde{\rho}\ \tilde{0} \end{aligned} \qquad \text{(II.14)}$$

where ρ is the density matrix with the normalization

$$\text{Trace over all states } \tilde{\rho} = 1 \qquad \text{(II.15)}$$

and where $\tilde{\rho}$ and $\tilde{0}$ are related to ρ and 0 by equation (II.12).

We will determine $\tilde{H}$ and A_N, and demonstrate a systematic procedure for determining $\tilde{0}$.

III. The Mixed Representation

The representation of the projected operator, $\tilde{0}$, is carried out in the space, S_c, which is the union of the orthogonal subspaces, S_M, where the number of bound atoms, M, ranges from zero to N. The subspace, S_N, contains N bound bosons. The remaining subspaces, S_M, with M less than N, contain M bound bosons, one continuum boson, N-M-1 protons and N-M-1 electrons. The space is a subspace of the original hilbert space but is "large enough" to contain the physical states (see, III.23). The presence of the single continuum boson is necessitated by the requirement of completeness and orthogonality in the original hilbert space. (see, III.5 et seq).

We will first derive the expression for the projected operator, $\tilde{0}$, in S_c (III.1). Then we will isolate the factors containing the continuum boson operators and express $\tilde{0}$ in terms of $\tilde{\tilde{0}}$, the corresponding operator in the space containing no continuum bosons. Finally, we will exhibit the relationship between Girardeau's "strong orthogonality" and the present formulation.

We will first show that the operator $\tilde{0}$ can be written in the form

$$0 = \sum_{M,M'=0}^{N} \sum_{\alpha_1 \cdots \alpha_M}^{\text{Bound States}} \sum_{\alpha'_1 \cdots \alpha'_{M'}}^{\text{Bound States}} \sum_{\beta_{M+1},\beta'_{M'+1}}^{\text{Continuum States}}$$

$$\int dx_{M+2} dy_{M+2} \cdots dx_N dy_N dx'_{M'+2} dy'_{M'+2} \cdots dx'_N dy'_N$$

$$a^+_{\alpha_1} \cdots a^+_{\alpha_M} a^+_{\beta_{M+1}} \psi^+(x_{M+2})\psi^+(y_{M+2}) \cdots \psi^+(x_N)\psi^+(y_N)|0\rangle [M!M_+!M_-!]^{-\frac{1}{2}}$$

$$\times \langle 0|\psi(y_N)\psi(x_N)\cdots\psi(y_{M+2})\psi(x_{M+2})A_{\beta_{M+1}} A_{\alpha_M} \cdots A_{\alpha_1} 0_E (N!)^{-2}$$

$$\times A^+_{\alpha_1} \cdots A^+_{\alpha'_{M'}} A^+_{\beta_{M'+1}} \psi^+(x'_{M'+2})\psi^+(y'_{M'+2}) \cdots \psi^+(x'_N)\psi^+(y'_N)|0\rangle$$

$$\times [M'!M'_+!M'_-!]^{-\frac{1}{2}} \langle 0|\psi(y'_N)\psi(x'_N)\cdots\psi(y'_{M'+2})\psi(x'_{M'+2})$$

$$\times a_{\beta'_{M'+1}} a_{\alpha'_{M'}} \cdots a_{\alpha'_1} \qquad \text{(III.1)}$$

The a and the ψ have already been introduced and obey the commutation rules, (II.7) and (II.9). In addition the bose operators, a^+_α, a_α, a^+_β, a_β, commute with both ψ and ψ^+. Moreover, both a_α and a^+_α commute with both a_β and a^+_β. That is, the continuum bosons and the bound bosons are treated as kinematically independent. We will always employ the convention that α will run over the bound states and β over the continuum states so that the summation ranges will not be explicitly stated hereafter. 0_E is the second quantized form of 0 in the elementary particle representation. A_α and A_β are defined as follows,

$$A_\alpha = \int \varphi^*_\alpha(xy)\psi(y)\psi(x)dxdy, \quad A_\beta = \int \varphi^*_\beta(xy)\psi(y)\psi(x)dxdy \qquad \text{(III.2)}$$

whereas A^+_α and A^+_β are their hermetian conjugates.

M_+, the number of proton operators appearing as factors on the left hand side of a given term, and M_-, the number of electron operators appearing as factors on the left hand side of a given term, are both equal to N-M-1 when $M \neq N$, and are equal to zero when N = M. The primed quantities

refer to the right hand factors (annihilation operators) with the same restrictions.

Note that there is a single continuum boson in all states except when M or M′ are equal to N. Then there is no continuum boson.

The kets on the left, including the factor, $(M!M_+!M_-!)^{-\frac{1}{2}}$, are the normalized basis vectors for states containing M bound bosons, one or zero continuum boson, M_+ heavy fermions, and M_- electrons. These basis vectors are orthogonal for different values of M, and together, they span the space in which

$$M + M_+ + (0 \text{ or } 1) = N = M + M_- + (0 \text{ or } 1). \qquad \text{(III.3)}$$

where the number of continuum bosons is zero when N = M and one otherwise.

Note carefully that the operators are "reflected" across $|0\rangle\langle 0|$ at both ends of the expression. This simplifies sign manipulation when the summation is actually carried out.

The starting point is really a definition owing to the completeness of $|\underline{xy}\rangle$.

$$\tilde{0} = \int d\underline{xy}\; d\underline{x'y'} |\underline{xy}\rangle\langle\underline{xy}| A_N 0\, A_N |\underline{x'y'}\rangle\langle\underline{x'y'}| \qquad \text{(III.3)}$$

The trick is to rewrite the identity operator, incorporating atoms and taking advantage of some properties of A_N. Consider the one atom projection operators,

$$I_B(1,1') = \sum_\alpha \varphi_\alpha(x_1,y_1)\varphi^*_\alpha(x_1',y_1') \qquad \text{(III.5)}$$

$$I_C(1,1') = \sum_\beta \varphi_\beta(x_1,y_1)\varphi^*_\beta(x_1',y_1') \qquad \text{(III.6)}$$

$$I(1,1') = I_B(1,1') + I_C(1,1') \equiv \delta(x_1,x_1')\delta(y_1,y_1') \qquad \text{(III.7)}$$

with the properties

$$\int I_B(1,1')I_B(1',1'')dx'dy' = I_B(1,1'') \qquad \text{(III.8)}$$

$$\int I_B(1,1')I_C(1',1'')dx'dy' = 0 \qquad \text{(III.9)}$$

$$\int I_C(1,1')I_C(1',1'')dx'dy' = I_C(1,1'') \qquad \text{(III.10)}$$

Now construct the projection operator, I_M, whose matrix elements are

$$\langle \underline{xy}|\, I_N\,|\underline{x'y'}\rangle = I_B(1,1')\ldots I_B(N,N')$$

$$\langle \underline{xy}|\, I_M\,|\underline{x'y'}\rangle = I_B(1,1')\ldots I_B(M,M')I_C(M+1,M+1')$$

$$\times\, I(M+2,M+2')\ldots I(N,N'), \quad 1 \le M \le N-1$$

$$\langle \underline{xy}|\, I_o\,|\underline{x'y'}\rangle = I_C(1,1')\ldots I(N,N') \qquad \text{(III.11)}$$

From the above, one can show

$$I_M I_{M'} = \delta(M,M') \qquad \text{(III.12)}$$

and

$$\sum_{M=0}^{N} I_M = \int d\underline{xy}|\,\underline{xy}\rangle\langle\underline{xy}| \qquad \text{(III.13)}$$

which is the unit operator for the entire hilbert space. The orthogonality, (III.12) is a weak one since it depends upon integration over the entire configuration space. It depends critically upon the position of the single I_C in (III.11). We now rewrite

$$\tilde{0} = \sum_{M,M'=0}^{N} I_M A_N 0_S A_N I_{M'} \qquad \text{(III.14)}$$

and introduce the projection operator, P_M, with the matrix elements

$$\langle \underline{xy}|\, P_M\,|\underline{x'y'}\rangle = (M!M_+!M_-!)^{-\frac{1}{2}} \sum_S \delta_S \langle \underline{xy}|\, S\underline{x'y'}\rangle \qquad \text{(III.15)}$$

Where S is the product, (permutation of the M couples $x_1y_1,\ldots x_My_M$) $\times$ (permutation of the proton coordinates, $x_{M+2},\ldots x_N$) $\times$ (permutation of the electron coordinates, $y_{M+2},\ldots y_N$), and δ_S is the product of the signatures of the electron and the proton permutations. It can be shown that

$$P_M^2 = P_M, \quad P_M^+ = P_M \tag{III.16}$$

$$P_M A_N = A_N P_M = A_N \tag{III.17}$$

$$P_M I_M = I_M P_M \tag{III.18}$$

The last property results only because the single continuum boson is treated as a different particle and is not mixed in the permutation, S. We now write

$$\begin{aligned}\tilde{0} &= \sum_{M,M'=0}^{N} I_M P_M A_N 0_S A_N P_{M'} I_{M'} \\ &= \sum_{M,M'=0}^{N} P_M I_M A_N 0_S A_N \\ &= \sum_{M,M'=0}^{N} \int |M\underline{xy}\rangle\langle \underline{xy}| A_N 0_S A_N |\underline{x'y'}\rangle\langle M'\underline{x'y'}| d\underline{xy}\, d\underline{x'y'}\end{aligned} \tag{III.19}$$

where

$$|M\underline{xy}\rangle \equiv P_M I_M |\underline{xy}\rangle \tag{III.20}$$

The orthogonality and completeness relations are,

$$\begin{aligned}\langle M'\underline{x'y'}|M\underline{xy}\rangle &= \langle \underline{x'y'}| I_{M'} P_{M'} P_M I_M |\underline{xy}\rangle \\ &= \langle \underline{x'y'}| P_{M'} I_{M'} I_M P_M |\underline{xy}\rangle \\ &= \langle \underline{x'y'}| P_M I_M P_M |\underline{xy}\rangle\ \delta(M',M)\end{aligned} \tag{III.21}$$

and

$$\begin{aligned}\int |M\underline{xy}\rangle\langle \underline{xy}M'| d\underline{xy} &= \int P_M I_M |\underline{xy}\rangle\langle \underline{xy}| I_{M'} P_{M'} d\underline{xdy} \\ &= P_M I_M I_{M'} P_{M'} \\ &= P_M I_M P_M \delta(M,M')\end{aligned} \tag{III.22}$$

Clearly, $P_M I_M P_M$ is the idempotent projection operator onto the subspace M, so the entire space, S_c, is the union of

orthogonal subspaces. The identity for S_c, which is "smaller" than the original hilbert space, is given by $\sum_{M=0}^{N} I_M P_M I_M$. The space contains all the physical states, for

$$A_N \sum_{M=0}^{N} I_M P_M I_M = A_N \sum_{M=0}^{N} P_M I_M = A_N \sum_{M=0}^{N} I_M = A_N \quad \text{(III.23)}$$

as one might have expected.

In second quantization,

$$|Mxy\rangle = (M!M_+!M_-!)^{-\frac{1}{2}} \int \psi(x_1,y_1)^+ \ldots \psi(x_M,y_M)^+ \psi_c(x_{M+1},y_{M+1})^+ \times \psi^+(x_{M+2})\psi^+(y_{M+2}) \ldots \psi^+(x_N)\psi^+(y_N)\, d\underline{xy}\, |0\rangle \quad \text{(III.24)}$$

where

$$\psi(x,y)^+ = \sum_\alpha \varphi_\alpha^*(x,y) a_\alpha^+, \qquad \psi_C(x,y)^+ = \sum \varphi_\beta^*(x,y) a_\beta^+ \quad \text{(III.25)}$$

are the wave function operators of the bound and continuum bosons respectively. The order, or the pairing, of the free proton and electron operators is due to the paired form of $I(j,j')$.

Finally, the matrix element,

$$\begin{aligned}\langle xy|A_N O_S A_N|x'y'\rangle &= (N!)^{-2}\langle 0|\psi(x_1)\ldots\psi(x_N)\psi(y_1\ldots\psi(y_N)O_E \\ &\quad \times \psi^+(y'_N)\ldots\psi^+(y'_1)\psi^+(x'_N)\ldots\psi^+(x'_1)|0\rangle \\ &= (N!)^{-2}\langle 0|\psi(y_N)\psi(x_N)\ldots\psi(y_1)\psi(x_1)O_E \\ &\quad \times \psi^+(x'_1)\psi^+(y'_1)\ldots\psi^+(x'_N)\psi^+(y'_N)|0\rangle \end{aligned} \quad \text{(III.26)}$$

allows pairing of the particles without a change in sign, since the primed and unprimed operators are moved symmetrically with respect to the "reflection" across O_E. Here, we have made a transition from the Schrödinger O_S to the O_E in the elementary particle representation.

Upon combining (III.2, 19, 24, 25, 26), we obtain (III.1).

We will now achieve simplification by isolating the effect of the continuum bosons.[5] We will show that (III.1) can then be written as

$$\tilde{O} = \Lambda^{+} \tilde{\tilde{O}} \Lambda \tag{III.27}$$

where

$$\Lambda^{+} = P_{N,0,0} + \sum_{\beta} a_{\beta}^{+} A_{\beta} (\hat{M}_{+}\hat{M}_{-})^{-\frac{1}{2}} \sum_{M=0}^{N-1} P_{M,N-M,N-M}$$

$$\Lambda = P_{N,0,0} + \sum_{M=0}^{N-1} P_{M,N-M,N-M} (\hat{M}_{+}\hat{M}_{-})^{-\frac{1}{2}} \sum_{\beta} A_{\beta}^{+} a_{\beta} \tag{III.28}$$

which are complex conjugate quantities. The very important operator, $\tilde{\tilde{O}}$, is given by (III.1) except that there are no continuum boson operators. That is, instead of

$$\tilde{O} = \sum_{\beta,\beta'\ldots} a_{\beta}^{+} \ldots |0\rangle\langle 0| \ldots A_{\beta} \ldots A_{\beta'}^{+} \ldots |0\rangle\langle 0| \ldots a_{\beta'} \ldots \tag{III.1}$$

there is

$$\tilde{\tilde{O}} = \int dx_{M+1} dy_{M+1} dx'_{M'+1} dy'_{M'+1} \ldots \psi^{+}(x_{M+1}) \psi^{+}(y_{M+1}) \ldots |0\rangle$$

$$\times \langle 0| \ldots \psi(y_{M+1}) \psi(x_{M+1}) \ldots \psi^{+}(x'_{M'+1}) \psi^{+}(y'_{M'+1}) \ldots |0\rangle$$

$$\cdot \langle 0| \ldots \psi(y'_{M'+1}) \psi(x'_{M'+1}) \ldots \tag{III.29}$$

where all the irrelevant (for the moment!) debris have been suppressed. Hence $\tilde{\tilde{O}}$ is defined in the subspaces with M bound bosons, N-M each of electrons and protons, and no continuum bosons. The union of these orthogonal subspaces from M=0 to M=N is the space, S_B.

In (III.28), $\hat{M}_{+}$ and $\hat{M}_{-}$ are number operators for the proton and the electron, respectively. P_{rst} is the identity operator for the subspace containing r bound bosons, s protons, t electrons, and no continuum bosons. Explicitly,

$$P_{rst} = \sum_{\alpha_1 \ldots \alpha_r} \int dx_1 \ldots dy_s dy_1 \ldots dy_t a^+_{\alpha_1} \ldots a^+_{\alpha_r}$$

$$\psi^+(x_1) \ldots \psi^+(x_s) \psi^+(y_1) \ldots \psi^+(y_t) |0\rangle$$

$$[r!s!t!]^{-1} \langle 0| \psi(y_t) \ldots \psi(y_1) \psi(x_s) \ldots \psi(x_1) a_{\alpha_r} \ldots a_{\alpha_1}$$

(III.30)

with the properties

$$P_{rst} P_{r's't'} = \delta(r,r')\delta(s,s')\delta(t,t') P_{rst}$$

$$a_\alpha P_{rst} = P_{r-1st} a_\alpha$$

$$\psi(x) P_{rst} = P_{rs-1t} \psi(x)$$

$$\psi(y) P_{rst} = P_{rst-1} \psi(y) \qquad \text{(III.31)}$$

We now derive (III.27). Consider the left hand portion of (III.1) in the neighborhood of the vacuum state projection operator, $|0\rangle\langle 0|$.

$$\sum_\beta \int dx_{M+2} dy_{M+2} \ldots dx_N dy_N \ldots a^+_\beta \psi^+(x_{M+2}) \psi^+(y_{M+2}) \ldots \psi^+(x_N) \psi^+(y_N) |0\rangle$$

$$\times [M!(N-M-1)!(N-M-1)!]^{-\frac{1}{2}} \langle 0 | \psi(y_N) \psi(x_N) \ldots \psi(y_{M+2}) \psi(x_{M+2}) A_\beta \ldots$$

$$= \sum_\beta \ldots a^+_\beta P_{0,N-M-1,N-M-1} A_\beta \ldots [(M!)^{-1}(N-M-1)!(N-M-1)!]^{\frac{1}{2}} \ldots$$

$$= \sum_\beta \ldots a^+_\beta A_\beta P_{o,N-M,N-M} [(M!)^{-1}(N-M-1)!(N-M-1)!]^{\frac{1}{2}} \ldots$$

$$= \sum_\beta \ldots a^+_\beta A_\beta (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} P_{0,N-M,N-M} [(M!)^{-1}(N-M)!(N-M)!]^{\frac{1}{2}} \ldots$$

$$= \sum_\beta \ldots a^+_\beta A_\beta (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} P^2_{0,N-M,N-M} [(M!)^{-1}(N-M)!(N-M)!]^{\frac{1}{2}} \qquad \text{(III.32)}$$

Upon restoring the bound state operators in the above,

$$\sum_{\alpha_1 \cdots \alpha_M} \sum_{\beta} \int dx_{M+2} dy_{M+2} \cdots dx_N dy_N a^+_{\alpha_1} \cdots a^+_{\alpha_M} a^+_{\beta}$$

$$\times \psi^+(x_{M+2}) \psi^+(y_{M+2}) \cdots \psi^+(x_N) \psi^+(y_N) |0\rangle$$

$$\times [M!(N-M-1)!(N-M-1)!]^{-\frac{1}{2}}$$

$$\times \langle 0| \psi(y_N) \psi(x_N) \cdots \psi(y_{M+2}) \psi(x_{M+2}) A_\beta A_{\alpha_M} \cdots A_{\alpha_1} \cdots$$

$$= \sum_{\beta} a^+_{\beta} A_{\beta} (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} \sum_{\alpha_1 \cdots \alpha_M} a^+_{\alpha_1} \cdots a^+_{\alpha_M} P^2_{0,N-M,N-M}$$

$$\times [(M!)^{-1}(N-M)!(N-M)!]^{+\frac{1}{2}}$$

$$= \sum_{\beta} a^+_{\beta} A_{\beta} (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} P_{M,N-M,N-M} \sum_{\alpha_1 \cdots \alpha_M} a^+_{\alpha_1} \cdots a^+_{\alpha_M}$$

$$\times P_{0,N-M,N-M} [(M!)^{-1}(N-M)!(N-M)!]^{\frac{1}{2}}$$

$$= \sum_{\beta} a^+_{\beta} A_{\beta} (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} P_{M,N-M,N-M}$$

$$\times \sum_{\alpha_1 \cdots \alpha_M} \int dx_{M+1} dy_{M+1} \cdots dx_N dy_N$$

$$\times a^+_{\alpha_1} \cdots a^+_{\alpha_M} \psi^+(x_{M+1}) \psi^+(y_{M+1}) \cdots \psi^+(x_N) \psi^+(y_N) |0\rangle$$

$$\times [(M!)(N-M)!(N-M)!]^{-\frac{1}{2}}$$

$$\times \langle 0| \psi(y_N) \psi(x_N) \cdots \psi(y_{M+1}) \psi(x_{M+1}) A_{\alpha_M} \cdots A_{\alpha_1} \cdots$$

(III.33)

Proceeding similarly with the factors clustered about the vacuum state projection operator on the right hand side of (III.1), one obtains

$$\sum_{\alpha'_1 \cdots \alpha'_{M'}} \sum_{\beta'} \int dx'_{M'+2} dy'_{M'+2} \cdots dx'_N dy'_N \cdots A^+_{\alpha'_1} \cdots A^+_{\alpha'_{M'}} A^+_{\beta'}$$

$$\times \psi^+(x'_{M'+2}) \psi^+(y'_{M'+2}) \cdots \psi^+(x'_N) \psi^+(y'_N) [M'!(N-M'-1)!(N-M'-1)!]^{-\frac{1}{2}}$$

$$\times \langle 0| \psi(y'_N) \psi(x'_N) \cdots \psi(y'_{M'+2}) \psi(x'_{M'+2}) a_{\beta'} a_{\alpha'_N} \cdots a_{\alpha'_1}$$

$$= \sum_{\alpha'_\beta \ldots \alpha'_M} \int dx'_{M'+1} dy'_{M'+1} \ldots dx'_N dy'_N \ldots A^+_{\alpha'_1} \ldots A^+_{\alpha'_{M'}}$$

$$\times \psi^+(x'_{M'+1}) \psi^+(y'_{M'+1}) \ldots \psi^+(x'_N) \psi^+(y'_N) |0\rangle$$

$$\times [M'!(N-M')!(N-M')!]^{-\frac{1}{2}}$$

$$\times \langle 0 | \psi(y'_N) \psi(x'_N) \ldots \psi(y'_{M'+1}) \psi(x'_{M'+1}) a_{\alpha'_{M'}} \ldots a_{\alpha'_1}$$

$$\times \sum_{\beta'} P_{M',N-M',N-M'} (\hat{M}_+ \hat{M}_-)^{-\frac{1}{2}} A^+_{\beta'} a_{\beta'} \qquad \text{(III.34)}$$

Sandwiching O_E between the two and summing over M and M' yields (III.27,28). We finally write

$$\tilde{O} = \Lambda^+ \tilde{O} \Lambda$$

$$= \Lambda^+ \sum_{M,M'=0} \sum_{\substack{\alpha_1 \ldots \alpha_M \\ \alpha'_1 \ldots \alpha'_{M'}}} \int dx_{M+1} dy_{M+1} \ldots dx_N dy_N$$

$$\times dx'_{M'+1} dy'_{M'+1} \ldots dx'_N dy'_N a^+_{\alpha_1} \ldots a^+_{\alpha_M}$$

$$\times \psi^+(x_{M+1}) \psi^+(y_{M+1}) \ldots \psi^+(x_N) \psi^+(y_N) |0\rangle [M!(N-M)!(N-M)!]^{-\frac{1}{2}}$$

$$\times \langle 0 | \psi(y_N) \psi(x_N) \ldots \psi(y_{M+1}) \psi(x_{M+1}) A_{\alpha_M} \ldots A_{\alpha_1} O_E (N!)^{-2}$$

$$\times A^+_{\alpha'_1} \ldots A^+_{\alpha'_{M'}} \psi^+(x'_{M'+1}) \psi^+(y'_{M'+1}) \ldots \psi^+(x'_N) \psi^+(y'_N) |0\rangle$$

$$\times [M'!(N-M')!(N-M')!]^{-\frac{1}{2}}$$

$$\times \langle 0 | \psi(y'_N) \psi(x'_N) \ldots \psi(y'_{M'+1}) \psi(x'_{M+1}) a_{\alpha'_{M'}} \ldots a_{\alpha'_1} \Lambda \qquad \text{(III.27)}.$$

The problem now reduces to the determination of $\tilde{O}$.

Before doing so, we will exhibit the relationship between the present formulation and Girardeau's[4,5]; wherein the states are in S_B, the sub space containing no continuum bosons. Only the physical states are considered by using A_N as the metric operator. In addition, the physical states

satisfy the "strong orthogonality" condition,

$$A_\alpha |\psi\ ) = |\psi) \tag{III.35}$$

Note that the operator, Λ, maps any state of S_C into a subspace of S_B, which we will call S_C'. Similarly, Λ^+ maps any state of S_B into a subspace of S_C, which we will call S_B'. In both mappings, distinct vectors map into distinct vectors, but the two mappings are not inverses of each other in general. But for physical states, one can define a one-one correspondence, a consequence of which is (III.35).

From the identity

$$A_N = A_N^{\ 2} \tag{III.36}$$

and the definition,

$$A_N = \Lambda\tilde{\tilde{A}}\Lambda^+ \tag{III.37}$$

it follows that

$$0 = \Lambda\Lambda^+\tilde{\tilde{A}}(1-\Lambda\Lambda^+\tilde{\tilde{A}})\Lambda, \qquad 0 = \Lambda^+(1-\tilde{\tilde{A}}\Lambda\Lambda^+)\tilde{\tilde{A}}\Lambda\Lambda^+ \tag{III.38}$$

That is, both S_C' and S_B' are composed of two mutually orthogonal subspaces. The case of S_B', this is nothing more than the decomposition of A_N. Let $|\psi\rangle$ be a vector in S_C and $|\psi)$ be its image in S_C' defined by

$$|\psi) = \Lambda|\psi\rangle \tag{III.39}$$

We will always used the angular kets for vectors in S_C and the rounded ket for vectors in S_B.

The physical states satisfy

$$A_N|\psi\rangle = |\psi\rangle \tag{III.40}$$

so that

$$(\Lambda^+\tilde{\tilde{A}})|\psi) = |\psi\rangle \tag{III.41}$$

and

$$(\Lambda\Lambda^+\tilde{\tilde{A}})|\psi) = |\psi) \tag{III.42}$$

The last equation is the necessary condition that $|\psi)$ be an image of a physical state.

Consider now the states in S_C' satisfying

$$\tilde{\tilde{A}}|\psi) = |\psi) \tag{III.43}$$

These will have images in the physical states given by

$$|\psi\rangle = A_N \Lambda^+ \tilde{\tilde{A}}|\psi) \tag{III.44}$$

which in turn will have images in S_C' given by

$$|\psi') = \Lambda A_N \Lambda^+ \tilde{\tilde{A}}|\psi) \tag{III.45}$$

$$= (\Lambda\Lambda^+ \tilde{\tilde{A}})(\Lambda\Lambda^+ \tilde{\tilde{A}})|\psi)$$

$$= (\Lambda\Lambda^+ \tilde{\tilde{A}})|\psi)$$

by the idempotency of $\Lambda\Lambda^+\tilde{\tilde{A}}$ in S_C'. Consequently, if we now require that

$$(\Lambda\Lambda^+ \tilde{\tilde{A}})|\psi) = |\psi) \tag{III.46}$$

the transformation becomes one-one. Together with (III.43), then

$$\Lambda\Lambda^+|\psi) = |\psi) \tag{III.47}.$$

In S_C', the above equation is a consequence of condition (III.43). In the larger space S_B, both conditions (III.43,46) serve to define the physical state. From the identity

$$\hat{M}_+\hat{M}_- = \sum_\beta A_\beta^+ A_\beta + \sum_\alpha A_\alpha^+ A_\alpha \tag{III.48}$$

we write

$$\Lambda\Lambda^+ = 1 - \sum_\alpha A_\alpha^+ (\hat{M}_+ + 1)^{-\frac{1}{2}} (\hat{M}_- + 1)^{-\frac{1}{2}} A_\alpha \tag{III.49}$$

so that the condition (III.47) becomes

$$A_\alpha|\psi) = 0 \tag{III.50}$$

which is the strong orthogonality condition.

To recapitulate, solving the following problem in S_C

$$i\hbar \frac{d}{dt} |\psi\rangle = \tilde{H}|\psi\rangle \tag{III.51}$$

is equivalent to solving the following problem in S_B.

$$i\hbar \frac{d}{dt} |\varphi) = \tilde{\tilde{H}}|\varphi), \; A_\alpha|\varphi) = 0 \tag{III.52}$$

IV. The Determination of $\tilde{\tilde{U}}$.

Previously[7], the direct summation of (III.27) was given wherein the matrix elements appearing in it were expanded in analogy with the Ursell expansion in statistical mechanics. Here we will give a simpler derivation yielding the same forms for the low order terms, but will not yield without much manipulation, the normal ordered exponential form of the previous work.

One observes from (III.27) that the matrix elements of $\tilde{\tilde{U}}$ are given by

$$\begin{aligned}
&[M!(N_+-M)!(N_--M)!M'!(N_+-M')!(N_--M')!]^{-\frac{1}{2}} \\
&\times \langle 0|\psi(y_N)\ldots\psi(y_{N_--M})\psi(x_{N_+})\ldots\psi(x_{N_+-M})a_{\alpha_M}\ldots a_{\alpha_1}\tilde{\tilde{U}} \\
&\times a^+_{\alpha_1},\ldots a^+_{\alpha_{M'}},\psi(x'_{N_+-M'})^+\ldots\psi(x'_{N_+})^+\psi(y'_{N_--M'})^+\ldots\psi(y'_{N_-})|0\rangle \\
&\equiv (N_+!N_-!)^{-1}\langle 0|\psi(y_{N_-})\ldots\psi(y_{N_--M})\psi(x_{N_+})\ldots\psi(x_{N_+-M})A_{\alpha_M}\ldots A_{\alpha_1}0_E \\
&\times A^+_{\alpha_1},\ldots A^+_{\alpha_{M'}},\psi(x'_{N_+-M'})^+\ldots\psi(x'_{N_+})^+\psi(y'_{N_--M'})^+\ldots\psi(y'_{N_-})|0\rangle
\end{aligned} \tag{IV.1}$$

where we have generalized by allowing the number of protons, N_+, and the number of electrons, N_-, to take different values.

We baldly assume the existence of the normal ordered expansion

$$0 = \sum_{r's't'} \sum_{\substack{\alpha_1\ldots\alpha_r \\ \alpha'_1\ldots\alpha'_{r'}}} \int dx_1\ldots dx_s dx'_1\ldots dx'_s, dy_1\ldots dy_t dy'_1\ldots dy'_t,$$

$$\times a^+_{\alpha_1}\ldots a^+_{\alpha_r}\psi^+(x_1)\ldots\psi^+(x_s)\psi^+(y_1)\ldots\psi^+(y_t)$$

$$\times \langle y_t\ldots y_1 x_s\ldots x_1\alpha_r\ldots\alpha_1|0|\alpha'_1\ldots\alpha'_r,x'_1\ldots x'_s,y'_1\ldots y'_t,\rangle_c$$

$$\times \psi(y'_t,)\ldots\psi(y'_1)\psi(x'_1)a_{\alpha'_r}\ldots a_{\alpha'_1}$$

$$\times (r!s!t!r'!s'!t'!)^{-1} \qquad \text{(IV.2)}$$

where the "connected" matrix elements $\langle\rangle_c$, are skew symmetric in the interchange of coordinates, x, x′, y, y′, and symmetric in the interchange of α,α'. These are evaluated by considering (IV.1) for the few particle states. Observe that the terms containing no bose operators ($r,r'=0$, or $M,M'=0$) merely yield 0_E, and need not be explicitly evaluated. The first few non-trivial terms are

$$\langle\alpha|0|\alpha'\rangle_c = \langle 0|A_\alpha 0_E A^+_{\alpha'}|0\rangle$$

$$N_+=N_-=M=M'=1 \qquad \text{(IV.3)}$$

$$\langle yx|0|\alpha'\rangle_c = \langle 0|\psi(y)\psi(x)0_E A^+_{\alpha'}|0\rangle$$

$$N_+=N_-=M'=1,M=0 \qquad \text{(IV.4)}$$

$$\langle y|0|y'\rangle_c\langle\alpha|\alpha'\rangle + \langle y|y'\rangle\langle\alpha|0|\alpha'\rangle_c + \langle y\alpha|0|\alpha'y'\rangle_c$$

$$= 2^{-1}\langle 0|\psi(y)A_\alpha 0_E A^+_{\alpha'},\psi(y')^+|0\rangle$$

$$N_+=M=M'=1,N_-=2 \qquad \text{(IV.5)}$$

$$\langle x|0|x'\rangle_c\langle\alpha|\alpha'\rangle + \langle x'|x'\rangle\langle\alpha|0|\alpha'\rangle_c + \langle x\alpha|0|\alpha'x'\rangle_c$$

$$= 2^{-1}\langle 0|\psi(x)A_\alpha 0_E A^+_{\alpha'},\psi(x')^+|0\rangle$$

$$N_-=M=M'=1,N_+=2 \qquad \text{(IV.6)}$$

$$2^{-1}\{\langle\alpha_2|0|\alpha_2'\rangle_c\langle\alpha_1|\alpha_1'\rangle + \langle\alpha_2|0|\alpha_1'\rangle_c\langle\alpha_1|\alpha_2'\rangle$$
$$+ \langle\alpha_2|\alpha_2'\rangle\langle\alpha_1|0|\alpha_1'\rangle_c + \langle\alpha_2|\alpha_1'\rangle\langle\alpha_1|0|\alpha_1'\rangle_c$$
$$+ \langle\alpha_2\alpha_1|0|\alpha_1'\alpha_2'\rangle_c\}$$
$$= 2^{-2}\langle 0|A_{\alpha_2}A_{\alpha_1}0_E A_{\alpha_1'}A_{\alpha_2'}|0\rangle$$
$$N_+=N_-=M=M'=2 \qquad \text{(IV.7)}$$

Clearly, every connected matrix element can be expressed in terms of the corresponding and lower order ordinary matrix elements.

For the case of $\tilde{H}$, we have,

$$\langle 0|A_\alpha H_E A^+_{\alpha'}|0\rangle = \langle\alpha|H_A|\alpha'\rangle = E_\alpha\delta(\alpha,\alpha') \qquad \text{(IV.8)}$$

where H_A is the atomic hamiltonian, and the bracket is the matrix element in our original hilbert space of distinguishable particles. Moreover,

$$\langle 0|\psi(y)\psi(x)H_E A_{\alpha'}|0\rangle = \langle xy|H_A|\alpha'\rangle = E_{\alpha'}\varphi_{\alpha'}(xy) \qquad \text{(IV.9)}$$

and

$$\langle 0|\psi(y)A_\alpha H_E A^+_{\alpha'}\psi(y')|0\rangle = \langle y\alpha|(1-I_-)(H_A+T_-+V_{A-})|\alpha' y'\rangle \qquad \text{(IV.10)}$$

where T_- is the kinetic energy of the extra electron, V_{A-} is the coulomb interaction between the atom and the electron, and I_- is the electron exchange operator.

$$\langle 0|\psi(x)A_\alpha H_E A^+_{\alpha'}\psi(x')|0\rangle = \langle x\alpha|(1-I_+)(H_A+T_++V_{A+})|\alpha' x'\rangle \qquad \text{(IV.11)}$$

where T_+ is the kinetic energy of the extra proton, V_{A+} is the coulomb interaction between the proton and the atom, and I_+ is the proton exchange operator. Finally,

$$\langle 0|A_{\alpha_2}A_{\alpha_1}H_E A^+_{\alpha_1'}A^+_{\alpha_2'}|0\rangle$$

$$= \langle\alpha_2\alpha_1|(1-I_-)(1-I_+)(H_A(1)+H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle$$

$$= \langle\alpha_2\alpha_1|(1-I_+I_-)(1-I_-)(H_A(1)+H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle$$

$$= \langle\alpha_2\alpha_1|(1-I_-)(H_A(1)+H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle$$

$$+ \langle\alpha_1\alpha_2|(1-I_-)(H_A(1)+H_A(2)+V_{AA}|\alpha_1'\alpha_2'\rangle \qquad \text{(IV.12)}$$

where $H_A(1)$ and $H_A(2)$ are the hamiltonians of the first and second atoms, respectively, and V_{AA} is the coulomb interaction between the two of them. We have used the idempotency of the exchange operators and the fact that I_-I_+ has the effect of exchanging an electron-proton pair.

Thus,

$$\langle\alpha|H|\alpha'\rangle = E_\alpha\delta(\alpha,\alpha') \qquad \text{(IV.3')}$$

$$\langle yx|H|\alpha'\rangle = E_{\alpha'}\varphi_{\alpha'}(xy) \qquad \text{(IV.4')}$$

$$\langle y\alpha|1+|\alpha'y'\rangle_c = 2^{-1}\langle y\alpha|(1-I_-)(H_A+T_-+V_{A-})|\alpha'y'\rangle$$

$$-\langle y\alpha|T_-|\alpha'y'\rangle-\langle y\alpha|H_A|\alpha'y'\rangle$$

$$= \langle y\alpha|V_{A-}|\alpha'y'\rangle$$

$$+\langle y\alpha|\left(\frac{-1-I_-}{2}\right)(H_A+T_-+V_{AA}|\alpha'y\alpha\rangle$$

$$\equiv \langle y\alpha|V_{A-}|\alpha'y'\rangle+\langle y\alpha|I'H_{A-}|\alpha'y'\rangle \qquad \text{(IV.5')}$$

where we have introduced some obvious notations. Similarly,

$$\langle x\alpha|H|\alpha'x'\rangle_c = \langle x_\alpha|V_{A+}|\alpha'x'\rangle+\langle x\alpha|\left(\frac{-1-I_+}{2}\right)(H_A+T_++V_{A+})|\alpha'x'\rangle$$

$$\equiv \langle x\alpha|V_{A+}|\alpha'x'\rangle+\langle x\alpha|I_+'H_{A+}|\alpha'x'\rangle \qquad \text{(IV.6')}$$

Finally,

$$\begin{aligned}
\langle\alpha_2\alpha_1|H|\alpha_1'\alpha_2'\rangle_c &= 2^{-1}\langle\alpha_2\alpha_1|(1-I_-)(H_A(1)+H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle \\
&+ 2^{-1}\langle\alpha_1\alpha_2|(1-I_-)(H_A(1)+(H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle \\
&- \langle\alpha_1\alpha_2|(H_A(1)+H_A(2))|\alpha_2'\alpha_1'\rangle \\
&- \langle\alpha_1\alpha_2|H_A(1)+H_A(2)|\alpha_1'\alpha_2'\rangle \\
&= \langle\alpha_2\alpha_1|V_{AA}|\alpha_1'\alpha_2'\rangle \\
&+ \langle\alpha_2\alpha_1|\left(\frac{-1-I_-}{2}\right)(H_A(1)H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle \\
&+ \langle\alpha_1\alpha_2|V_{AA}|\alpha_1'\alpha_2'\rangle \\
&+ \langle\alpha_1\alpha_2|\left(\frac{-1-I_-}{2}\right)(H_A(1)+H_A(2)+V_{AA})|\alpha_1'\alpha_2'\rangle \\
&\equiv \langle\alpha_2\alpha_1|V_{AA}|\alpha_1'\alpha_2'\rangle+\langle\alpha_2\alpha_1|I'H_{AA}|\alpha_1'\alpha_2'\rangle \\
&+ \langle\alpha_1\alpha_2|V_{AA}|\alpha_1'\alpha_2'\rangle+\langle\alpha_1\alpha_2|I'H_{AA}|\alpha_1'\alpha_2'\rangle
\end{aligned} \tag{IV.7'}$$

We thus obtain,

$$\begin{aligned}
\tilde{\tilde{H}} &= H_E + \sum_\alpha E_\alpha a_\alpha^+ a_\alpha \\
&+ \sum_\alpha E_\alpha (a_\alpha^+ A_\alpha + A_\alpha^+ a_\alpha) \\
&+ \sum_{\alpha,\alpha'} \int dx dx' \psi^+(x) a_\alpha^+ \langle x\alpha|V_{A+}+I_+'H_{A+}|\alpha' x'\rangle \psi(x') a_{\alpha'} \\
&+ \sum_{\alpha,\alpha'} \int dy dy' a_\alpha^+ \psi^+(y') \langle y\alpha|V_{A-}+I_-'H_{A-}|\alpha' y'\rangle \psi(y') a_{\alpha'}
\end{aligned}$$

$$+ \tfrac{1}{2} \sum_{\substack{\alpha_1\alpha_2 \\ \alpha_1''\alpha_2''}} a^+_{\alpha_1} a^+_{\alpha_2} \langle \alpha_2\alpha_1 | (V_{AA}+I'_-H_{AA}) | \alpha_1''\alpha_2'' \rangle a_{\alpha_2''} a_{\alpha_1''} \qquad \text{(IV.13)}$$

The first term is the ion-electron hamiltonian (II.5). The second term is the free atom hamiltonian. The third term describes the process

$$H \leftrightharpoons p + e^- \qquad \text{(IV.14)}$$

which was first obtained by Stolt and Brittin[8].

The fourth term describes the process

$$H + p \leftrightharpoons H + p \qquad \text{(IV.15)}$$

and consists of two parts, the direct coulomb interaction V_{A+}, and I'_+H_{A+}, which does not agree with the $\tilde{E}_{pa}$ of Stolt and Brittin, who have the operator I_+ rather than I'_+. Our result is in partial agreement with Girardeau[9] and is physically reasonable for it corrects the total hamiltonian by eliminating the part of the atom-proton hamiltonian which is symmetric in the exchange of the two protons. The fifth term is the analogous one for the electron-atom scattering.

The last term describes the process,

$$H + H \leftrightharpoons H + H \qquad \text{(IV.16)}$$

which again consists of the direct part, V_{AA}, and the correction, I'_-H_{AA}. These terms were first obtained by Girardeau[6] except that the atomic states included the continuum states-

Terms corresponding to processes such as $p+H \leftrightharpoons p+p+e^-$ were omitted as we were only concerned with two body processes. The inclusion would be trivial, for one only needs the connected matrix, $\langle x_1\alpha | H | y'x_2'x_1' \rangle$.

The expansion of $\tilde{A}$ is a bit more subtle. The $M=M'=0$ term of (III.27) with 0_E set equal to one is nothing more than $P_{0,N,N}$, the identity in the subspace with no bound atoms. Thus, instead of the expansion (IV.2), one needs

$$A=1+\sum_{\substack{rst\\ r's't'}}\sum_{\substack{\alpha\ldots\alpha_t\\ \alpha'_1\ldots\alpha'_r}}\int dx_1\ldots dx_s dx'_1\ldots dx'_s, dy_1\ldots dy_t dy'_1\ldots dy'_t,$$

$$\times\ a^+_{\alpha_1}\ldots a^+_{\alpha_r}\psi^+(x_1)\ldots\psi^+(x_s)\psi^+(y_1)\ldots\psi^+(y_t)$$

$$\times\ \langle y_t\ldots y_1 x_s\ldots x_1\alpha_r\ldots\alpha_1|A|\alpha'_1\ldots\alpha'_r, x'_1\ldots x'_t, y'_1\ldots y'_t,\rangle_c$$

$$\times\ \psi(y'_t,)\ldots\psi(y'_1)\psi(x'_s,)\ldots\psi(x'_1)a_{\alpha'_r,}\ldots a_{\alpha'_1,} \qquad \text{(IV.17)}$$

which differs from (IV.2) by the presence of the identity. The definition of the connected matrix elements are then

$$\langle 0|a_\alpha a_{\alpha,}|0\rangle+\langle\alpha|A|\alpha'\rangle_c = \langle 0|A_\alpha A^+_{\alpha,}|0\rangle = \delta(\alpha,\alpha')$$

$$N_+=N_-=M'=M=1 \qquad \text{(IV.18)}$$

$$\langle 0|\psi(y)\psi(x)a^+_{\alpha,}|0\rangle+\langle yx|A|x'\rangle_c = \langle 0|\psi(y)\psi(x)A^+_{\alpha,}|0\rangle = \varphi_{\alpha,}(xy)$$

$$N_+=N_-=M'=1, M=0 \qquad \text{(IV.19)}$$

$$\langle 0|\psi(y)a_\alpha a^+_{\alpha,}\psi^+(y')|0\rangle+\langle\alpha|A|\alpha'\rangle_c\langle y|y'\rangle+\langle\alpha|\alpha'\rangle\langle y|A|y'\rangle_c$$

$$+\ \langle y\alpha|A|\alpha' y'\rangle_c$$

$$= 2^{-1}\langle 0|\psi(y)A_\alpha A^+_{\alpha,}\psi^+(y')|0\rangle$$

$$= 2^{-1}\langle y\alpha|(1-I_-)|\alpha' y'\rangle$$

$$N_+=M=M'=1, N_-=2 \qquad \text{(IV-20)}$$

$$\langle 0|\psi(x)a_\alpha a^+_{\alpha,}\psi^+(x')|0\rangle+\langle\alpha|A|\alpha'\rangle_c\langle x|x'\rangle+\langle\alpha|\alpha'\rangle\langle x|A|x'\rangle_c$$

$$+\ \langle x\alpha|A|\alpha' x'\rangle_c$$

$$= 2^{-1}\langle 0|\psi(x)A_\alpha A^+_{\alpha,}\psi^+(y')|0\rangle$$

$$= 2^{-1}\langle x\alpha|(1-I_+)|\alpha' x'\rangle$$

$$N_-=M=M'=1, N_+=2 \qquad \text{(IV.21)}$$

$$2^{-1}\{\langle 0|a_{\alpha_2}a_{\alpha_1}a^+_{\alpha_1'}a^+_{\alpha_2'}|0\rangle+\langle\alpha_2|A|\alpha_2'\rangle_c\langle\alpha_1|\alpha_1'\rangle$$

$$+\langle\alpha_2|A|\alpha_1'\rangle_c\langle\alpha_1|\alpha_2'\rangle+\langle\alpha_2|\alpha_2'\rangle_c\langle\alpha_1|A|\alpha_1'\rangle_c+\langle\alpha_2|\alpha_1'\rangle\langle\alpha_1|A|\alpha_2'\rangle_c$$

$$+\langle\alpha_2\alpha_1|A|\alpha_1'\alpha_2'\rangle_c\}$$

$$=2^{-2}\langle 0|A_{\alpha_2}A_{\alpha_1}A^+_{\alpha_2'}A^+_{\alpha_2'}|0\rangle$$

$$=2^{-2}\langle\alpha_2\alpha_1|(1-I_-)(1-I_+)|\alpha_1'\alpha_2'\rangle$$

$$=2^{-2}\langle\alpha_2\alpha_1|1-I_-|\alpha_1'\alpha_2'\rangle+2^{-2}\langle\alpha_1\alpha_2|1-I_-|\alpha_1'\alpha_2'\rangle$$

$$N_+=N_-=M=M'=2 \qquad \text{(IV.22)}$$

We then have

$$\langle\alpha|A|\alpha'\rangle_c = 0 \qquad \text{(IV.18')}$$

$$\langle yx|A|\alpha'\rangle_c = \varphi_{\alpha'}(xy) \qquad \text{(IV.19')}$$

$$\langle y\alpha|A|\alpha'y'\rangle_c = \langle y\alpha|I_-'|\alpha'y'\rangle \qquad \text{(IV.20')}$$

$$\langle x\alpha|A|\alpha'x'\rangle_c = \langle x\alpha|I_+'|\alpha'x'\rangle \qquad \text{(IV.21')}$$

$$\langle\alpha_2\alpha_1|A|\alpha_1'\alpha_2'\rangle_c = \langle\alpha_2\alpha_1|I_-'|\alpha_1'\alpha_2'\rangle+\langle\alpha_1\alpha_2|I_-'|\alpha_1'\alpha_2'\rangle \qquad \text{(IV.22')}$$

In the derivation of (III.20′) and (III.21′), we used the fact that $\langle x|A|x'\rangle_c$ and $\langle y|A|y'\rangle_c$ were zero respectively, since the projection of A into the subspace with no bound atoms is the identity.

Thus, we have

$$\approx{A} = 1 + \sum_\alpha (a^+_\alpha A_\alpha + A^+_\alpha a_\alpha)$$

$$+ \sum_{\alpha,\alpha'}\int dxdx' a^+_\alpha \psi^+(x)\langle x\alpha|I_+'|\alpha'x'\rangle\psi(x')a_{\alpha'}$$

$$+ \sum_{\alpha,\alpha'} \int dy dy' a_{\alpha}^{+} \psi^{+}(y) \langle y\alpha | I_{-}' | \alpha' y' \rangle \psi(y') a_{\alpha'}$$

$$+ \tfrac{1}{2} \sum_{\substack{\alpha_1\alpha_2 \\ \alpha_1'\alpha_2'}} a_{\alpha_1}^{+} a_{\alpha_2}^{+} \langle \alpha_2\alpha_1 | I_{-}' | \alpha_1'\alpha_2' \rangle a_{\alpha_2'} a_{\alpha_1'} \qquad \text{(IV.22)}$$

V. Concluding Remarks

We have obtained expressions for the projected hamiltonian, $A_N H$, and A_N itself in the mixed representation, which explicitly display terms corresponding to the ionization, recombination, scattering, etc., of elementary and composite particles. Two representations were shown. The one involves the solution of the Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi\rangle = \tilde{H} |\psi\rangle \qquad \text{(V.1)}$$

where

$$\tilde{H} = \Lambda^{+} \tilde{\tilde{H}} \Lambda \qquad \text{(V.2)}$$

operating in S_C, and the other involving the solution of the Schrödinger equation,

$$i\hbar \frac{d}{dt} |\psi) = \tilde{\tilde{H}} |\psi) \qquad \text{(V.3)}$$

with the subsidiary condition

$$A_{\alpha} |\psi) = 0 \qquad \text{(V.4)}$$

operating in S_B. In either representation, the hamiltonian serves to project the wave function into the physical space except for the subspace belonging to the zero eigenvalues of the respective hamiltonians. These must be supplemented by the relations

$$\tilde{A} |\psi\rangle = |\psi\rangle \qquad \text{(V.5)}$$

or

$$\tilde{\tilde{A}} |\psi) = |\psi) \qquad \text{(V.6)}$$

The form (V.2) is not as formidable as it appears, for its effect is to replace a proton-electron pair operator with a continuum boson operator, i.e.,

$$\Lambda^+ \tilde{\mathcal{O}} \Lambda \equiv \tilde{\mathcal{O}} \Lambda^+ \Lambda \qquad (V.7)$$

where $\tilde{\mathcal{O}}$ is simply related $\mathcal{O}$ by the above mentioned replacement and $\Lambda^+\Lambda$ yields one when operating upon a physical state. Explicit form has been reported earlier[7].

Since one must eventually calculate either the thermal averages or expectation values, it may be more profitable to directly evaluate Trace $A_N \rho \mathcal{O}$, rather than explicitly finding $A_N \mathcal{O}$ as we have done here. In fact, the arguments leading to (III.27) results in some interesting expressions for A_N, which will facilitate the trace operations.

Finally the presence of more than one composite specie can be taken into account by the generalization of (III.11) et seq. One introduces the bound and continuum projection operators for H_2 molecules, say, and decompose the product of the free electron and proton operators in the same way that the entire hilbert space was decomposed by means of the H atom projection operators. This and the above matter will be reported elsewhere.

VI. Acknowledgement

The author is grateful to Professor Wesley E. Brittin for encouragement and stimulating discussions. He also wishes to acknowledge Professor T. Matsubara of Kyoto University, where a portion of this manuscript was written.

REFERENCES

1. P. Ehrenfest and J.R. Oppenheimer, "Note on the Statistics of Nuclei", Phys. Rev. 37, 333-338 (1931).

2. M. Girardeau, "Formulation of the Many-Body Problem for Composite Particles", J. Math. Phys. 4, 1096-1115 (1963).

3. W.E. Brittin, This volume, 55.

4. M. Girardeau, This volume, 147.

5. This equivalence was not shown in the lectures but will be included in the notes for completeness.

6. M. Girardeau, "Formulation of the Many-Body Problem for Composite Particles, II, The Projected Hamiltonian", J. Math. Phys. 12, 1799-1810 (1971).

7. A.Y. Sakakura, "Composite Particles in Many-Body Systems", Phys. Rev. Lett. 27, 822-826 (1971).

8. R.H. Stolt and W.E. Brittin, "New Method for Treating Systems Containing Elementary and Composite Particles (Non-Relativistic Theory)", Phys. Rev. Lett. 27, 616-619 (1971).

9. M. Girardeau, "Second Quantization Representation for Systems of Atoms, Nuclei, and Electrons", Phys. Rev. Lett. 27, 1416 (1971).

ERGODIC THEORY IN ALGEBRAIC STATISTICAL MECHANICS

G. L. Sewell
Queen Mary College
London E1 4NS England

1. Introduction

The C*-algebraic formulation of statistical mechanics provides a framework for the study of properties of assemblies of particles in the thermodynamical limit. The use of this limit constitutes an idealization which permits sharp mathematical characterizations of properties that would otherwise be masked by finite-size effects.

The object of the present course is to formulate a unified approach, within the algebraic framework, to the following statistical mechanical problems:- (a) the characterization of pure thermodynamical phases; (b) the characterization of phase transition, especially with regard to symmetry changes; and (c) theory of the approach to equilibrium. Our treatment of these problems will be centered on a non-commutative generalization of classical ergodic theory and also on the celebrated KMS conditions.

The subject matter will be presented as follows. In section 2, we shall present, in summarized form, the mathematical equipment we need: this will consist of definitions and standard results concerning C*-algebras, abstract ergodic theory and KMS conditions. In section 3, we shall outline the algebraic formulation of states, observables and space-translational ergodic theory for "infinite volume" physical systems. In section 4, we shall adapt the formalism of section 3 to statistical mechanics. In particular, we shall formulate the Gibbs state and the time translations in the islands of ‡ those states according to

‡ By the island of a state ϕ on an algebra $\mathcal{A}$, we mean the set of states of the form $\Psi\circ\pi$, where Ψ is a normal state on the image of $\mathcal{A}$ under the ϕ-induced GNS representation π.

the scheme proposed by Dubin and myself[1] (DS). This will lead us to a characterization of the pure phases and of symmetry breakdown in equilibrium states (section 4.7). In section 5, we shall apply the formalism of section 4 to the theory of the approach to equilibrium in the island of a Gibbs state. I think that the results of this section, though of a familiar form, are new at least as consequences of the DS scheme. In section 6, we shall apply the same formalism to the study of phase transitions, with particular reference to symmetry breakdown and, in the case of transitions of the second kind, the divergence of a correlation length (appropriately defined) at the critical point. Our study of phase transitions will be largely centered on a new algebraic treatment, due to Maria Marinaro and myself[2], of a class of Ising models which, in the specific case of the soluble two-dimensional one, leads to the required characterizations of the phase transitions.

2. Mathematical Equipment

2.1 Σ-systems

We define a Σ-system to be a triple $(\mathfrak{A},S,\alpha(G))$, where $\mathfrak{A}$ is a C^*-algebra[3], S the set of all states on $\mathfrak{A}$ and α a homomorphism of a group G into Aut $\mathfrak{A}$, the automorphisms of $\mathfrak{A}$. The set of all pure states on $\mathfrak{A}$ will be denoted by S_p. We define α^* to be the representation of G induced in S by α:- $(\alpha^*(g)\phi)(A) = \phi(\alpha(g)A)$, $\forall A\in\mathfrak{A}$, $\phi\in S$, $g\in G$.

Let $\phi\in S$. We denote by $(\mathcal{H}_\phi, \pi_\phi, \Omega_\phi)$ the GNS triple (representation space, representation, cyclical vector) induced by the action of ϕ on $\mathfrak{A}$. Correspondingly, we define $\mathcal{I}(\phi)$, the island of states associated with ϕ, to be the set of states of the form $\psi = \Psi\circ\pi_\phi$ where Ψ is a normal state on $\pi_\phi(\mathfrak{A})$: thus $\mathcal{I}(\phi)$ corresponds to the set of density matrices in $\mathcal{H}_\phi$.

We denote by C_G the set of all G-invariant states on $\mathfrak{A}$, i.e. $C_G = \{\phi\,|\,\phi\in\mathfrak{A}^*,\ \alpha^*(g)\phi = \phi,\ \forall g\in G\}$. For $\phi\in C_G$, we define U_ϕ to be the representation of G induced in $\mathcal{H}_\phi$ by α:-

$$U_\phi(g)\Omega_\phi = \Omega_\phi;\ \pi_\phi(\alpha(g)A) = U_\phi(g)\pi_\phi(A)U_\phi(g^{-1});\quad \forall A\in\mathfrak{A},\ g\in G$$

We define P_ϕ to be the projection operator for the maximal subspace of $\mathcal{H}_\phi$ that is stable under $U_\phi(G)$.

The set C_G is convex and w*-compact. We denote by E_G the extremal elements of this set. Adopting a usual terminology, whose significance will become manifest in sections 2.4, 2.5, we refer to the elements of E_G as the G-ergodic states.

2.2 Classical Σ-systems

We term a Σ-system $(\mathcal{G}, S, \alpha(G))$ classical if $\mathcal{G}$ is abelian and possesses an identity element. On the other hand, we define a C-system to be a pair $(K, \tau(G))$, where K is a compact space and τ a homomorphism of the group G into the automorphisms of K.

It follows from these definitions that, corresponding to a given C-system $(K, \tau(G))$, there exists a classical Σ-system $(\mathcal{G}, S, \alpha(G))$ such that $\mathcal{G} = \mathcal{B}(K)$, the set of all bounded, continuous, complex-valued functions on K, with supremum norm; and $(\alpha(g)A)(k) = A(\tau(g^{-1})K)$, $\forall A\in\mathcal{B}(K)$, $k\in K$, $g\in G$. Further, the state space S corresponds to the set $\mathcal{P}(K)$ of all probability measures (positive measures of total mass 1) on K; i.e. if $\phi\in S, \exists$ a unique probability measure μ_ϕ on K such that $\phi(A) = \int_K A(k)\,d\mu_\phi(k)$, $\forall A\in\mathcal{B}(K)$.

Conversely, it follows from the Gelfand isomorphism that every classical Σ-system $(\mathcal{G}, S, \alpha(G))$ may be constructed from a C-system in this way, with $K = S_p$ (equipped with the w*-topology) and $\tau(g) = \alpha^*(g)$, $\forall g\in G$. Hence there is a one-to-one correspondence between C-systems and classical Σ-systems. We note that in this correspondence between $(K, \tau(G))$ and $(\mathcal{G}, S, \alpha(G))$, $\mathcal{G}$ is separable if and only if K is metrisable.

2.3 Amenable groups[4]

Let G be a locally compact group; and let $\mathcal{B}(G)$ be the C*-algebra of all bounded, continuous, complex-valued functions on G, with supremum norm. Let a be the homomorphism of G into Aut $\mathcal{B}(G)$, defined by:-$(a(g)A)(g') = A(g^{-1}g')$, $\forall A\in\mathcal{B}(G)$; $g,g'\in G$. Then if $A\in\mathcal{B}(G)$, we define O(A), the orbit of A, to be $\{a(g)A;\ g\in G\}$. The set W(G), of weakly

almost periodic functions on G, is defined[5] to be the set of elements A, of $\mathfrak{B}(G)$, for which O(A) is relatively compact in the weak topology of $\mathfrak{B}(G)$. W(G) is then a closed subspace of the Banach space $\mathfrak{B}(G)$.

The group G is said to be amenable if there exists at least one state η_G on $\mathfrak{B}(G)$, such that η_G is invariant under $a^*(G)$. η_G is then termed an invariant mean (more precisely a left-invariant mean) on $\mathfrak{B}(G)$. In general, an amenable group G will possess more than one invariant mean. However, these means all coincide on W(G).

Examples of amenable groups are:- compact groups, Z(the integers), R(the reals), finite-dimensional Euclidean groups; but <u>not</u> the Lorentz group. In the particular case of R, all invariant means η_R reduce, on W(R), to the form:-

$$\eta_R(A) = \lim_{T\to\infty} \frac{1}{T} \int_0^T dt\ A(t).$$

<u>Note</u> If G is an amenable group and U a strongly-continuous unitary representation of G in a Hilbert space $\mathcal{H}$, then[5] the function $g \to (\psi_1, U(g)\psi_2)$ belongs to W(G), $\forall \psi_1, \psi_2 \in \mathcal{H}$. Further, the application to this function of an invariant mean η_G yields the value $(\psi_1, P\psi_2)$, where P is the projector for the subspace of $\mathcal{H}$ that is stable under U(G): this is the mean L^2 ergodic theorem. Hence in the notation of section 2.1, if $\phi \in C_G$, then the function $g \to \phi((\alpha(g)A)B)\ (\equiv (U_\phi(g)\pi_\phi(A^*)\Omega_\phi, \pi_\phi(B)\Omega_\phi))$ belongs to W(G), and the action of η_G on this function yields the value $(\pi_\phi(A^*)\Omega_\phi, P_\phi\pi_\phi(B)\Omega_\phi)$. This result is the essential constituent of the mean ergodic theorems of sections 2.4, 2.5.

2.4 Classical mean ergodic theory

Let $(K, \tau(G))$ be a C-system, for which G is amenable, K metrisable and τ a continuous representation of G in K. We define $\mathcal{C}_G$ to be the set of all G-invariant probability measures on K, i.e. $\mathcal{C}_G = \{\mu \mid \mu \in \mathcal{P}(K);\ \mu = \mu \circ \tau(g), \forall g \in G\}$. We define $\mathcal{E}_G$ to be the subset of elements μ of $\mathcal{C}_G$ for which K is indecomposable into subspaces of non-zero μ-measure that are stable under $\tau(G)$. $\mathcal{E}_G$ is termed the set of G-ergodic measures on K, and corresponds to the set of extremal elements of $\mathcal{C}_G$ (c.f. Ref. 6, Ch. 10).

We now note the following results of classical ergodic theory:-

(a) (Ref. 6, Ch. 10). There is a unique integral representation of $\mathcal{C}_G$ in $\mathcal{E}_G$; i.e., given $\mu\in\mathcal{C}_G$, $\exists$ a unique measure m_μ on $\mathcal{C}_G$ such that $m_\mu(\mathcal{C}_G\backslash\mathcal{E}_G = 0$ and

$$\mu(A)(\equiv \int A\, d\mu) = \int_{\mathcal{C}_G} \nu(A)\, d\, m_\mu(\nu),\ \forall A\in\mathcal{B}(K).$$

(b) (consequence of mean $L^2(K,\mu)$ ergodic theorem). Let $\mu\in\mathcal{C}_G$ and let η_G be an invariant mean on $\mathcal{B}(G)$. Then the following statements are equivalent:-

(i) $\mu\in\mathcal{E}_G$

(ii) $\eta_G\mu(\alpha(\cdot)A)B) = \mu(A)\mu(B)$, $A, B\in\mathcal{B}(K)$,

where $\alpha(G)$ is the group induced by $\tau(G)$ in Aut $\mathcal{A}$, according to the prescription of section 2.3.

In view of the one-to-one correspondence between C-systems and classical Σ-systems (c.f. section 2.2), we may translate the above results of classical ergodic theory into the following form. Let $(\mathcal{A}, S, \alpha(G))$ be a classical Σ-system for which $\mathcal{A}$ is separable, G amenable and α a continuous representation of G in $\mathcal{A}$. Let C_G, E_G be defined as in section 2.1. Then:-

(a)' There is a unique integral representation of C_G in E_G; i.e. given $\phi\in C_G$, $\exists$ a unique measure μ_ϕ on C_G, such that $\mu_\phi(C_G\backslash E_G) = 0$ and $\phi = \int_{C_G} \sigma\, d\, \mu_\phi(\sigma)$.

(b)' Let $\phi\in C_G$ and let η_G be an invariant mean on $\mathcal{B}(G)$. Then the following statements are equivalent:-

(i) $\phi\in E_G$

(ii) $\eta_G\phi((\alpha(\cdot)A)B) = \phi(A)\phi(B)$, $\forall A, B\in\mathcal{A}$,

i.e. ϕ is <u>weakly clustering</u> with respect to G.

2.5 Non-commutative mean ergodic theory

The reformulation of classical mean ergodic theory in C*-algebraic terms has the advantage of being generalizable

to non-commutative algebras. We now formulate the non-commutative generalization of the theory of section 2.5.

We introduce the following standard definitions of "quasi-commutative" properties which may serve to replace the commutativity of $\mathcal{A}$ in classical ergodic theory.

(a) We say that $\mathcal{A}$ is G-asymptotically abelian if $\lim_{g\to\infty} [\alpha(g)A,B]_- = 0 \quad \forall A,B\in\mathcal{A}$. By $\lim_{g\to\infty}$ we mean the following. Let f be a function from G into C. We say that $\lim_{g\to\infty} f(g) = 0$ if, given $\epsilon > 0$, $\exists$ a compact subset K_ϵ of C such that $|f(g)| < \epsilon \ \forall g\in G\setminus K_\epsilon$.

(b) We say that $\mathcal{A}$ is η_G-abelian[7] if α is a strongly-continuous representation of G in $\mathcal{A}$ and η_G is an invariant mean on $\mathcal{B}(G)$ such that

$$\eta_G(\psi_1,\pi_\phi[\alpha(\cdot)A,B]_-\psi_2) = 0, \ \forall\psi_1,\psi_2\in\mathcal{H}_\phi; \ A, \ B\in\mathcal{A}, \ \phi\in S.$$

(c) We say that $\mathcal{A}$ is G-abelian[8] (resp. G-abelian in the representation π_ϕ) if $P_\phi\pi_\phi(\mathcal{A})P_\phi$ is abelian $\forall\phi\in C_G$ (resp. for the particular state $\phi\in C_G$).

For cases where $\mathcal{A}$ is a C*-algebra in a Hilbert space $\mathcal{H}$, we introduce the following definitions:-

(a)' We say that $\mathcal{A}$ is weakly G-asymptotically abelian in $\mathcal{H}$ if $\lim_{g\to\infty} (\psi_1,[\alpha(g)A,B]_-\psi_2) = 0$; $\forall A, \ B\in\mathcal{A}$; $\psi_1,\psi_2\in\mathcal{H}$.

(b)' We say that $\mathcal{A}$ is η_G-abelian in $\mathcal{H}$ if, for all $\psi_1,\psi_2\in\mathcal{H}$ and $A_1,A_2\in\mathcal{A}$, the function $g\to(\psi_1[\alpha(g)A,B]_-\psi_2)$ is continuous; and if the application of the invariant mean η_G to this function yields zero.

It folllows from these definitions that, if $\mathcal{A}$ is G-asymptotically abelian and α is a strongly continuous representation of G in $\mathcal{A}$, then $\mathcal{A}$ is η_G-abelian. Further, if $\mathcal{A}$ is η_G-abelian or G-asymptotically abelian, then it is necessarily G-abelian (c.f. Ref. 9, Corollary 6.2.10 and Prop. 6.2.16).

The next two theorems generalize classical mean ergodic theory to the case where $\mathcal{A}$ is G-abelian, though not necessarily abelian.

Theorem 2.5.1 Let $\mathfrak{A}$ be G-abelian and norm-separable. Then $\exists$ a unique integral representation of C_G in E_G; i.e., given $\phi \in C_G$, $\exists$ a unique measure μ_ϕ on C_G such that $\mu_\phi(C_G \backslash E_G) = 0$ and $\phi = \int_{C_G} \sigma \, d\mu_\phi(\sigma)$.

Proof c.f. Lanford and Ruelle[8] (Theorem 3.1 and Corollary 3.2).

Theorem 2.5.2 Let $\phi \in C_G$. Then:-

(i) If P_ϕ is one-dimensional, then $\phi \in E_G$.

(ii) If $\mathfrak{A}$ is G-abelian in the representation π_ϕ and if $\phi \in E_G$, then P_ϕ is one-dimensional.

(iii) If U_ϕ is a strongly-continuous representation of G in $\mathcal{H}_\phi$, then ϕ is weakly clustering (i.e. $\eta_G^\phi((\alpha(\cdot)A)B) = \phi(A)\phi(B)$, $\forall A, B \in \mathfrak{A}$) if and only if P_ϕ is one-dimensional..

Proof c.f. Ruelle[9] (Theorem 6.3.3 and Proposition 6.3.5).

Thus it follows from the above theorems that the following conditions suffice for the generalization of classical ergodic theory to the non-commutative case:- (a) $\mathfrak{A}$ is G-abelian; (b) $\mathfrak{A}$ is norm-separable; and (c) U_ϕ is a strongly continuous representation of G in $\mathcal{H}_\phi$. In fact, it will be seen in section 3 that, in cases of physical interest, where $\mathfrak{A}$ has a "quasi-local" structure, one may obtain the essential content pertinent to physics of the above theorems without recourse to the separability assumption.

2.6 The KMS conditions

Let $(\mathfrak{A}, S, \gamma(R))$ be a Σ-system, with γ a homomorphism of the real line, R, into Aut $\mathfrak{A}$. Given $\beta(\in R) > 0$, we say that $\phi(\in S)$ satisfies the KMS conditions[10,11,12] corresponding to $(\gamma(R), \beta)$ if, for each pair A, $B \in \mathfrak{A}$, $\exists$ functions f_{AB}, g_{AB} on the complex plane C, such that:-

(i) $f_{AB}(t) = \phi((\alpha(t)A)B)$; $g_{AB}(t) = \phi(B(\alpha(t)A))$; $\forall A, B \in \mathfrak{A}$; $t \in R$.

(ii) f_{AB} (resp. g_{AB}) is analytic in the strip $\mathrm{Im} z \in (-\beta, 0)$ (resp. $(0, \beta)$) and continuous on its boundaries.

(iii) $f_{AB}(z) = g_{AB}(z+i\beta)$, $\forall z \in C$.

If ϕ satisfies these conditions, then[11]:- (1) ϕ is invariant under $\gamma^*(R)$; and (2) Ω_ϕ is both cyclical and separating with respect to $\pi_\phi(\mathcal{A})''$, the weak closure of $\pi_\phi(\mathcal{A})$.

Conversely to (2), we have the following theorem, due to Tomita.

Theorem 2.6.1 Let ϕ be a state on a C*-algebra $\mathcal{A}$, and let Ω_ϕ be cyclical and separating with respect to $\pi_\phi(\mathcal{A})''$. Then, given β ($\in R$) > 0, $\exists$ a unique homomorphism $\gamma_\phi : R \to$ Aut $\pi_\phi(\mathcal{A})''$ such that, if $\tilde{\phi}$ denotes the state on $\pi_\phi(\mathcal{A})''$ defined by $\phi(\cdot) = (\Omega_\phi, (\cdot)\Omega_\phi)$, then $\tilde{\phi}$ satisfies the KMS conditions with respect to $(\gamma_\phi(R), \beta)$.

Proof c.f. Takesaki[12], Theorems 13.1 and 13.2.

3. States of Physical Systems

3.1 The quasi-local algebra

Let X be a locally compact, non-compact space corresponding to that occupied by the physical system under consideration. We assume that X is either a finite-dimensional Euclidean space or a lattice of points (with discrete topology) in such a space: in either case X corresponds to an amenable group. Let L be the set $\{Y\}$ of all bounded subsets of X. By a standard construction (Ref. 9, Ch. 7), we assign to each $Y \in L$ a C*-algebra $\mathcal{A}_Y$ (whose self-adjoint elements correspond to the observables for the region Y) such that:-

(i) $\mathcal{A}_Y$ is isotonic with respect to Y;

(ii) $\mathcal{A}_Y$ commutes[‡] with $\mathcal{A}_{Y'}$ if Y and Y' are mutually disjoint; and

(iii) denoting $\bigcup_{Y \in L} \mathcal{A}_Y$ by $\mathcal{A}_L$, $\exists$ a homomorphism α of X into Aut $\mathcal{A}_L$, such that $\alpha(x)\mathcal{A}_Y \equiv \mathcal{A}_{xY}$.

‡ In the case of a system of fermions; one has to define $\mathcal{A}_Y$ as the even subalgebra for the region Y in order to obtain this commutativity condition.

We define $\mathcal{A}$ to be the norm closure of $\mathcal{A}_L$, and extend $\alpha(X)$ by continuity from $\mathcal{A}_L$ to $\mathcal{A}$. In a usual way, we refer to $\mathcal{A}$ as the C*-algebra of quasi-local bounded observables. The group $\alpha(X)$ corresponds to space translations.

It follows easily from (i)-(iii) and our definitions of and $\alpha(X)(\in \mathrm{Aut}\,\mathcal{A})$ that $\mathcal{A}$ is X-asymptotically abelian.

We shall henceforth restrict ourselves to the following standard cases:-

Lattice systems (classical or quantal). Here, $\mathcal{A}$ is norm-separable, and α is a strongly continuous representation of X in $\mathcal{A}$.

Quantal continuous systems. The $\mathcal{A}_Y$'s are W*-algebras in a Fock-Hilbert space $\mathcal{H}_F$, and $\alpha(X)$ is unitarily implemented in $\mathcal{H}_F$ by a strongly continuous representation U_F of X.

3.2 Locally normal states

Consider a quantal continuous system $(\mathcal{A}, S, \alpha(X))$, equipped with the above-described quasi-local structure. We denote by $\mathfrak{L}$ the set of all locally normal states, i.e. the states whose restrictions to all the local algebras $\{\mathcal{A}_Y | Y \in L\}$ are normal. These are presumably the physically significant elements of S, since they correspond to the states for which there is zero probability of finding an infinity of particles in a bounded region of space[13]. Thus, when considering states of continuous quantal systems, we shall henceforth restrict our analysis to the class $\mathfrak{L}$. (In the case of lattice systems, there is no need to make such a restriction, since the number of particles (spins) in a bounded region is a fortiori finite.)

Let us now note the following properties of $\mathfrak{L}$:-

(i) If $\phi \in \mathfrak{L}$, then $\mathcal{H}_\phi$ is separable (Ref. 8, Prop. 4.3).

(ii) If $\phi \in C_X \cap \mathfrak{L}$, then the unitary representation U_ϕ of X induced by α in $\mathcal{H}_\phi$ is strongly continuous: this follows from the strong continuity of U_F in $\mathcal{H}_F$ and the strong continuity of ϕ on the unit ball of each local algebra $\mathcal{A}_Y$ (i.e. the local normality of ϕ).

3.3 X-ergodicity

Theorem 3.3.1 With the above definitions and assumptions, the set C_X has the following properties:-

(1) In the case of a lattice system,

(a) $\exists$ a unique integral representation of C_X in E_X: and

(b) $\phi(\in C_X)$ is weakly clustering with respect to X if and only if $\phi \in E_X$.

(2) In the case of quantal continuous system,

(a) $\exists$ a unique integral representation of $C_X \cap \mathfrak{L}$ in $E_X \cap \mathfrak{L}$; and

(b) $\phi(\in C_X \cap \mathfrak{L})$ is weakly clustering with respect to S if and only if $\phi \in E_X \cap \mathfrak{L}$.

Proof Since $\mathfrak{A}$ is X-abelian in all cases, we see that (1) follows directly from Theorems 2.5.1-2 and our above specifications concerning lattice systems (Sec. 3.1): while (2) follows from the strong continuity of U_ϕ, together with Theorems 2.5.2 and Ref. 8, Theorem 4.3.

Thus, in case of physical interest, there is a unique decomposition of any element ϕ, of C_X, into X-ergodic components; and there is an identification of E_X with the weakly clustering elements of X. Further, we note that weak spatial clustering is arguably a property of a pure X-invariant thermodynamical phase (c.f. Ref. 9, Ch. 6.5), and thus the X-ergodic decomposition law, given by Theorem 3.3.1, has a physical significance: the characterization of pure thermodynamical phases will be further discussed in Secs. 4.6 and 4.7.

4. Formulation of Statistical Mechanics

4.1 The Σ-systems σ_n

We shall now outline a procedure for the formulation of equilibrium states and of dynamics of an 'infinite' system, as limits of the corresponding quantities for finite systems. For this purpose, we introduce an increasing sequence $L' = \{Y_n\}$ of elements of the above- described set L, such that $\bigcup_n Y_n = X$. Corresponding to each

$Y_n \in L'$, we employ the standard methods of finite-volume statistical mechanics to construct a triple $(\mathcal{G}^{(n)}, S^{(n)}, \gamma^{(n)}(R))$, where $\mathcal{G}^{(n)} \equiv \mathcal{G}_{Y_n}$ is the algebra of observables of a system of the particles concerned, confined to the region Y_n; $S^{(n)}$ is the set of all states on $\mathcal{G}^{(n)}$; and $\gamma^{(n)}$ is a homomorphism of R into Aut $\mathcal{G}^{(n)}$, corresponding to time-translations of the system. Thus the triple $(\mathcal{G}^{(n)}, S^{(n)}, \gamma^{(n)}(R))$ is a Σ-system, which we denote by σ_n. Likewise, by standard methods, we construct the Gibbs (canonical or grand canonical) state $\phi_\beta^{(n)}$ on $\mathcal{G}^{(n)}$ corresponding to the inverse temperature $\beta(>0)$.

Note In the quantum-mechanical case[10], $\phi_\beta^{(n)}$ satisfies the KMS conditions corresponding to $(\gamma^{(n)}(R), \beta)$. Also, in the case of a continuous quantal system, $\phi_\beta^{(n)}$ is a normal state on $\mathcal{G}^{(n)}$.

4.2 Gibbs states on $\mathcal{G}$.

Let $A \in \mathcal{G}_L$. Then it follows from the constructions of Secs. 3.1, 4.1 that, for n sufficiently large, $A \in \mathcal{G}^{(n)}$ and thus $\phi_\beta^{(n)}(A)$ is well-defined. We now introduce the following standard postulate (c.f. Ref. 10):-

(I) $\lim_{n\to\infty} \phi_\beta^{(n)}(A)$ exists, $\forall A \in \mathcal{G}_L$.

Since $\mathcal{G}_L$ is norm-dense in $\mathcal{G}$, it follows from (I) that there exists a state ϕ_β on $\mathcal{G}$ that is uniquely defined by the prescription:-

$$\phi_\beta(A) = \lim_{n\to\infty} \phi_\beta^{(n)}(A), \quad \forall A \in \mathcal{G}_L.$$

We term ϕ_β the Gibbs state of the "infinite" system, whose observables are $\mathcal{G}$.

4.3 Time-translations: the HHW scheme

The Haag-Hugenholtz-Winnink[10] treatment of infinite-volume statistical mechanics was based on (I), together with the following postulate concerning time-translations:-

(II) Given $A \in \mathcal{G}_L$ and $t \in R$, $\gamma^{(n)}(t)A$ converges normwise in the limit $n\to\infty$. As a consequence of this postulate, one obtains the result that there exists a homomorphism γ of R into Aut $\mathcal{G}$, uniquely defined by the prescription:-

$$\gamma(t)A = \text{norm-}\lim_{n\to\infty} \gamma^{(n)}(t)A, \quad \forall A\in\mathcal{G}_L, \ t\in R.$$

Thus, in this scheme, time-translations of the infinite system are taken to correspond to the group $\gamma(R)$ of automorphisms of $\mathcal{G}$.

On the basis of assumptions (I) and (II), HHW obtained the powerful result that, in the quantum-mechanical cases, ϕ_β satisfies the KMS conditions with respect to $(\gamma(R), \beta)$.

4.4 Critique of the HHW scheme

The HHW postulates (I) and (II) are known to be valid for free fermions and also for a wide class of lattice systems[14]. On the other hand, as shown by Dubin and myself (Appendices of Ref. 1), postulate (II) (though not (I)) fails in the cases of the BCS model and the ideal Bose gas. For this reason, we proposed the following scheme, based on weaker axioms than those of HHW.

4.5 The DS scheme[1]

This scheme is based on the following two postulates, which are shown to be weaker (as a pair) than (I) and (II).

(I)' Given $k\ (\in Z_+) <\infty$; $A_1,\ldots,A_k\in\mathcal{G}_L$; and $t_1,\ldots,t_k\in R$, $\lim_{n\to\infty} \phi_\beta^{(n)}((\gamma^{(n)}(t_1)A_1)\ldots(\gamma^{(n)}(t_k)A_k))$ exists.

(II)' Given k, $1(\in Z_+)$; $A_1,\ldots,A_{k+1}\in\mathcal{G}_L$; and $t_1,\ldots,t_{k+1}\in R$, $\lim_{m\to\infty}\lim_{n\to\infty} \phi_\beta^{(n)}((\gamma^{(n)}(t_1)A_1)\ldots(\gamma^{(n)}(t_k)A_k) \times (\gamma^{(m)}(t_{k+1})A_{k+1})\ldots(\gamma^{(m)}(t_{k+1})A_{k+1}))$ exists, and is equal to $\lim_{n\to\infty} \phi_\beta^{(n)}((\gamma^{(n)}(t_1)A_1)\ldots(\gamma^{(n)}(t_{k+1})A_{k+1}))$.

These assumptions were shown to be valid for all the cases considered, including the BCS model and the ideal Bose gas. From a physical standpoint, they may be considered to have the advantage of expressing properties of correlation functions, rather than of abstract algebraic entities.

The principle results of the DS scheme are as follows: (1) Trivially, the HHW postulate (I) is satisfied; and therefore (Ref. 15, Prop. 1) the resultant Gibbs state ϕ_β is locally normal. We shall denote by $(\mathcal{K}_\beta, \pi_\beta, \Omega_\beta)$ the

GNS triple induced by ϕ_β. Also, we shall denote by $\tilde{\phi}_\beta$ the state on $\pi_\beta(\mathfrak{A})''$ defined by:- $\tilde{\phi}_\beta(\cdot) = (\Omega_\beta, (\cdot)\Omega_\beta)$.

(2) $\exists$ a unique homomorphism $\tilde{\gamma}_\beta$ of R into Aut $\pi_\beta(\mathfrak{A})''$, such that

$$\lim_{n\to\infty} \phi_\beta^{(n)}((\gamma^{(n)}(t_1)A_1)\ldots(\gamma^{(n)}(t_k)A_k)) = \tilde{\phi}_\beta(\pi_\beta^{t_1}(A_1)\ldots\pi_\beta^{t_k}(A_k)),$$

$$\forall\ A_1,\ldots,A_k \in \mathfrak{A}_L;\ t_1,\ldots,t_k \in R;\ k < \infty,$$

where $\pi_\beta^t(A) \equiv \tilde{\gamma}_\beta(t)\pi_\beta(A)$.

(3) $\tilde{\phi}_\beta$ satisfies the KMS conditions with respect to $(\tilde{\gamma}_\beta(R),\beta)$.

We interpret these results as signifying that, in the island $\mathcal{I}(\phi_\beta)$, time-translations correspond to the automorphisms $\tilde{\gamma}_\beta(R)$ of $\pi_\beta(\mathfrak{A})''$, though not necessarily to automorphisms of $\mathfrak{A}$ itself. With this interpretation of $\tilde{\gamma}_\beta(R)$, we see that the results (1)-(3) constitute a recovery of the essential conclusions of HHW regarding the island $\mathcal{I}(\phi_\beta)$.

The automorphisms $\tilde{\gamma}_\beta(R)$ are, in fact, precisely the Tomita automorphisms described in Theorem 2.6.1.

4.6 Structure of $\mathcal{I}(\phi_\beta)$

Let us now investigate the structure of the island $\mathcal{I}(\phi_\beta)$ on the basis of the DS scheme, supplemented by the postulate that $\phi_\beta \in C_X$. This last hypothesis is valid, for example, if $x \in X, \phi_\beta$ is is unchanged if L' is replaced by xL' $(\equiv\{xY_n \mid Y_n \in L'\})$ in the construction of Sec. 4.2. In fact, it may be verified that ϕ_β is invariant under these translations of L' in all the tractable models mentioned in Sec. 4.4.

Thus, we examine the structure of $\mathcal{I}(\phi_\beta)$ on the basis of the following "axioms":-

(1) $\phi_\beta \in C_X$.
(2) In the case of continuous systems, $\phi_\beta \in \mathfrak{L}$.
(3) $\exists$ a homomorphism $\tilde{\gamma}_\beta$ of R into Aut $\pi_\beta(\mathfrak{A})''$, such that $\tilde{\phi}_\beta$ satisfies the KMS conditions with respect to $(\tilde{\gamma}_\beta(R),\beta)$.
It follows from this last axiom that $\tilde{\phi}_\beta$ is invariant under $\tilde{\gamma}_\beta^*(R)$ and thus that $\tilde{\gamma}_\beta(R)$ is unitarily implemented in $\mathcal{H}_\beta$

by a representation $\tilde{V}_\beta$ of R, uniquely defined by the requirements that

$$\tilde{V}_\beta(t)\Omega_\beta=\Omega_\beta, \text{ and } \tilde{V}_\beta(t)\pi_\beta(A)\tilde{V}_\beta(-t)=\tilde{\gamma}_\beta(t)\pi_\beta(A);\ \forall\ t\in R, A\in\mathcal{A}.$$

Further it follows easily from the KMS conditions (specifically from the part of condition (ii) (Sec. 2.6) pertaining to continuity on the boundaries) that $\tilde{V}_\beta$ is a strongly continuous representation of R in $\mathcal{H}_\beta$.

It follows from the general definition of $\mathcal{S}(\phi)$ (Sec. 2.1) that $\mathcal{S}(\phi_\beta)$ is in one-to one correspondence with the normal states on $\pi_\beta(\mathcal{A})$ and thus (by continuity) on $\pi_\beta(\mathcal{A})''$. Specifically, each $\psi\in\mathcal{S}(\phi_\beta)$ corresponds to a unique normal state $\tilde{\psi}$ on $\pi_\beta(\mathcal{A})''$, such that $\psi=\tilde{\psi}\circ\pi_\beta$. Thus, the automorphisms $\tilde{\gamma}_\beta(R)$, of $\pi_\beta(\mathcal{A})''$, induce a representation γ^*_β of R in $\mathcal{S}(\phi_\beta)$, given by:-

$$\gamma^*_\beta(t)\psi = \tilde{\gamma}^*_\beta(t)\psi\circ\pi_\beta, \forall\ t\in R.$$

Hence, γ^*_β corresponds to time-translations in $\mathcal{S}(\phi_\beta)$.

Let C_β be the set of states $\psi=\tilde{\psi}\circ\pi_\beta(\in\mathcal{S}(\phi_\beta))$ such that $\tilde{\psi}$ satisfies the KMS conditions with respect to $(\tilde{\gamma}_\beta(R),\beta)$. C_β is thus a convex, w^*-compact set. We shall denote its extremal elements by E_β. As a natural generalization of traditional (finite-volume) statistical mechanics, we regard C_β as constituting the set of thermal equilibrium states of the system at the inverse temperature β. Thus, as has been cogently argued by Emch, Knops and Verboven[16, 17], it is natural to interpret E_β as the set of states corresponding to pure thermodynamical phases at that inverse temperature.

We define $C_{\beta R}(\supset C_\beta)$ to be the set of elements of $\mathcal{S}(\phi_\beta)$ that are stable under $\gamma^*_\beta(R)$; and we denote by $E_{\beta R}$ the set of extremals of $C_{\beta R}$.

Theorem 4.6.1 Let $\mathcal{S}_\beta=\{\psi\,|\,\psi\in\mathcal{S}(\phi_\beta);\psi<\lambda\phi_\beta \text{ for some } \lambda\in R_+\}$. Then:-
(i) $\mathcal{S}_\beta$ is uniformly dense in $\mathcal{S}(\phi_\beta)$; and
(ii) $\exists$ a unique map $f:\mathcal{S}_\beta\to\mathcal{H}_\beta$ such that

$$(\gamma^*(t)\psi)(A) = (f(\psi),\ \tilde{V}_\beta(t)\pi_\beta(A)\Omega_\beta),\forall A\in\mathcal{A},\ t\in R,\ \psi\in\mathcal{S}_\beta.$$

Proof Since $\tilde{\phi}_\beta$ satisfies the KMS conditions with respect to $(\tilde{\gamma}_\beta(R), \beta)$, the required result follows easily from Ref. 18 (Prop. 2 and Ex. 5).

Theorem 4.6.2 Let P be a central projector for $\pi_\beta(\mathfrak{A})''$, and let χ_P be the element of $\mathcal{J}(\phi_\beta)$ defined by $\chi_P(\cdot) = (P\Omega_\beta, \pi_\beta(\cdot)P\Omega_\beta)/||P\Omega_\beta||^2$. Then $\mathcal{J}(\chi_P)$ is stable under $\gamma_\beta^*(R)$.

Proof It follows from our definitions that $\mathcal{J}(\chi_P)$ is the subset of $\mathcal{J}(\phi_\beta)$ constituted by states $\tilde{\Psi}_P o P \pi_\beta$, where $\tilde{\Psi}_P$ is a normal state on $\pi_\beta(\mathfrak{A})''$. Further, since $\tilde{\phi}_\beta$ satisfies the KMS conditions with respect to $(\tilde{\gamma}_\beta(R), \beta)$, it follows that P is stable under $\tilde{\gamma}_\beta^*(R)$ (c.f. Ref. 19, Corollary 2.5). Consequently, $\mathcal{J}(\chi_P)$ is stable under $\gamma_\beta^*(R)$.

Theorem 4.6.3 (i) There is a unique integral representation of C_β in E_β; and further, the decomposition $C_\beta \to E_\beta$ is precisely the central decomposition of C_β into factors.
(ii) If $\pi_\beta(\mathfrak{A})''$ is η_R-abelian (w.r.t. $\tilde{\gamma}_\beta(R)$) in the space $\mathcal{H}_\beta$, then E_β coincides with $E_{\beta R}$.
(iii) If $\psi \in C_X \cap C_\beta$, then the measure μ_ψ on E_X, induced by its X-ergodic decomposition, has support in $E_X \cap C_\beta$. Thus, the KMS decomposition is finer than or coincident with the X-ergodic one.

Proof In view of our axioms (1)-(3) and subsequent definitions:-
(i) follows from Ref. 12 (Theorem 15.4 and subsequent remark), together with Ref. 16, Theorem 1.2 and the fact that $\phi_\beta \in \mathfrak{L}$ implies that $\mathcal{J}(\phi_\beta) \in \mathfrak{L}$ (c.f. Ref. 20, Lemma 3.3).
(ii) is proved in the first sentence of the proof of Theorem 3.6 in Ref. 21 (independently of the norm-separability assumption for $\mathfrak{A}$); and
(iii) follows from (i), together with Ref. 22 (Theorem 3.2.1 and subsequent remark).

Corollary The Gibbs state ϕ_β must satisfy one of the following mutually exclusive properties:-
(i) $\phi_\beta \in R_\beta$
(ii) $\phi_\beta \notin E_X$, and the X-ergodic decomposition of ϕ_β coincides with its KMS decomposition
(iii) $\phi_\beta \notin E_\beta$, and the KMS decomposition of ϕ_β is a refinement of its X-ergodic decomposition.

4.7 Symmetry breakdown

Let $\psi \in C_\beta \cap C_X$, and let μ_ψ (resp ν_ψ) be the (unique) measure on E_X (resp E_β) associated with its X-ergodic (resp KMS) decomposition. Let τ be a homomorphism of a group H into Aut $\mathcal{A}$, let K_H be the set of states in $\mathcal{S}(\phi_\beta)$ that are stable under $\tau^*(H)$ and let $K_H' = \mathcal{S}(\phi_\beta) \backslash K_H$. Then we say that ψ undergoes an H-symmetry breakdown associated with its X-ergodic (resp. KMS) decomposition if $\psi \in K_H$ but $\mu_\psi(K_H')$ (resp $\nu_\psi(K_H')$) $\neq 0$.

Thus, accepting that E_β corresponds to the set of pure thermodynamical phases at the inverse temperature β, it follows from the Corollary to Theorems 4.6.3 that:-

(a) if alternative (i) is valid, then ϕ_β is a pure phase and undergoes no symmetry breakdown;

(b) if alternative (ii) is valid, then ϕ_β is a mixture of X-invariant pure phases, and may undergo a symmetry breakdown associated with its X-ergodic decomposition;

(c) if alternative (iii) is valid, then ϕ_β undergoes an X-symmetry breakdown (as in the crystalline state[16]), and possibly other breakdowns also, associated with its KMS decomposition into pure phases.

The theory of symmetry breakdown in phase transitions will be treated in section 6.

5. Temporal ergodicity and the approach to equilibrium

Assuming the axioms and definitions of section 4.7, we obtain the following theorem concerning dynamics of states in $\mathcal{S}(\phi_\beta)$.

Theorem 5 (i) $\exists$ a map $F: \mathcal{S}(\phi_\beta) \to C_{\beta R}$ such that

$$\eta_R(\gamma_\beta^*(\cdot)\psi) \equiv w^*\text{-}\lim_{T\to\infty} \frac{1}{T} \int_0^T \gamma_\beta^*(t)\psi \, dt = F(\psi), \quad \forall \psi \in \mathcal{S}(\phi_\beta)$$

(ii) If $\pi_\beta(\mathcal{A})''$ is weakly assymptotically abelian with respect to $\tilde{\gamma}_\beta(R)$ in $\mathcal{H}$, then

$$w^*\text{-}\lim_{t\to\infty} \gamma_\beta^*(t)\psi = F(\psi), \quad \forall \psi \in \mathcal{S}(\phi_\beta).$$

(iii) If $\pi_\beta(\mathcal{A})''$ is η_R-abelian with respect to $\tilde{\gamma}_\beta(R)$ in $\mathcal{H}_\beta$ then $F(\mathcal{J}(\phi_\beta)) \subset \mathcal{C}_\beta$.

(iv) If $\phi_\beta \in E_\beta$ and if $\pi_\beta(\mathcal{A})''$ is η_R-abelian in $\mathcal{H}_\beta$, then $F(\mathcal{J}(\phi_\beta)) = \{\phi_\beta\}$.

Proof (i) Let $\tilde{P}_\beta$ be the projector for the subspace of $\mathcal{H}_\beta$ that is invariant under $\tilde{V}_\beta(R)$. By applying von Neumann's mean ergodic theorem 4.6.1 (ii), we see that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt(\gamma_\beta^*(t)\psi)(A) = (f(\psi),\ \tilde{P}_\beta \pi_\beta(A)\Omega_\beta),\ \forall A\in\mathcal{A},\ \psi\in\mathcal{S}_\beta$$

The required result follows easily from this formula, together with Theorem 4.6.1 (i)

(ii) By theorem 4.6.3(i), $\tilde{\phi}_\beta$ belongs to the w*-convex hull of the factor states on $\pi_\beta(\mathcal{A})''$. Hence, if $\pi_\beta(\mathcal{A})''$ is weakly asymptotically abelian with respect to $\tilde{\gamma}_\beta(R)$ in $\mathcal{H}_\beta$, it follows from an easy adaptation of Ref. 20, Lemma 3.2‡ that

$$(w, \mathcal{H}_\beta)\text{-}\lim_{t\to\infty} \tilde{V}_\beta(t) = \tilde{P}_\beta = (w,\mathcal{H}_\beta)\text{-}\lim_{T\to\infty} \int_0^T dt\ \tilde{V}_\beta(t)$$

Hence, by (i) and Theorem 4.6.3 (ii)

$$w^*\text{-}\lim_{t\to\infty} \gamma_\beta^*(t)\psi = \eta_R(\gamma_\beta^*(\cdot)\psi) = F(\psi), \forall\ \psi\in\mathcal{S}_\beta.$$

The required result follows from this equation, together with Theorem 4.6.1 (i).

(iii) follows from Theorem 4.6.3 (ii) and the fact that $F : \mathcal{J}(\phi_\beta) \to \mathcal{C}_{\beta R}$ (by (i)).

(iv) Assuming that $\phi_\beta \in E_\beta$ and that $\pi_\beta(\mathcal{A})''$ is η_R-abelian in $\mathcal{H}_\beta$, it follows from Theorem 4.6.3 (ii) that $\phi_\beta \in E_{\beta R}$ and hence, from Theorem 2.5.2, that $\tilde{P}_\beta$ is the projection for Ω_β. Thus, since $F(\psi)(A) \equiv (f(\psi),\ \pi_\beta(A)\Omega_\beta),\ \forall\ \psi\in\mathcal{S}_\beta$, as shown in the proof of (i), it follows easily that $F(\psi) = \phi_\beta,\ \forall\ \psi\in\mathcal{S}_\beta$. Thus, in view of Theorem 4.6.1 (i), the required result follows by continuity.

‡ That lemma, in its original form, referred to space translations. However, it carries through equally well for time-translations when the above weak asymptotic abelian condition is fulfilled.

Comment It follows from this theorem that, if $\pi_\beta(\mathcal{A})''$ is weakly asymptotically abelian with respect to $\tilde{\gamma}_\beta(R)$ in $\mathcal{H}_\beta$, then the time-translate $\gamma_\beta^*(t)\psi$ tends (w^*) to an equilibrium state $F(\psi)$ as $t\to\infty$.
Further, in the case where ϕ_β is a pure phase, all states in $\mathcal{J}(\phi_\beta)$ evolve asymptotically to ϕ_β. On the other hand, it follows from Theorem 4.6 2-3 that any state ψ in the sector $\mathcal{J}(\chi_P)$ will tend to an equilibrium state in the same sector, i.e. $F(\mathcal{J}(\chi_P))\subset\mathcal{J}(\chi_P)\cap\mathcal{C}_\beta$. Thus a system whose initial state lies in a sector $\mathcal{J}(\chi_P)$, for which χ_P lacks certain of the symmetries of ϕ_β, will in general evolve asymptotically to an equilibrium state that also lacks those symmetries. From a physical standpoint, this is a satisfactory result: for one would anticipate, for example, that a ferromagnet which was initially disturbed from a polarized equilibrium state, would eventually relax back into such a state.

As regards the condition that $\pi_\beta(\mathcal{A})''$ be weakly asymptotically abelian with respect to $\tilde{\gamma}_\beta(R)$ in $\mathcal{H}_\beta$, one knows that this condition is fulfilled in certain tractable dynamical models, e.g. the ideal Fermi gas, the ideal Bose gas, certain Fermion lattice models[23] and the XY model[24]. Thus, the condition is at least compatible with the general principles of statistical mechanics.

In short, the use of the DS conditions in the algebraic formalism leads to a satisfactory framework for the theory of the approach to equilibrium within the island $\mathcal{J}(\phi_\beta)$, subject to the above realizable conditions. Moreover, a similar theory could not be obtained within the framework of the traditional finite-volume quantum statistical mechanics: for there the discrete spectrum of the Hamiltonian renders all time-dependent expectation values of observables quasi-periodic in t; and further the possibility of super-selection rules between different sectors such as $\mathcal{J}(\chi_P)$ is precluded by the uniqueness of the Hilbert space representation of the observables.

However, despite the advantages gained by the algebraic theory I still think that the present formulation of the theory of the approach to equilibrium lacks at least one essential ingredient which may loosely be termed "friction". By that I mean that the above-mentioned tractable models, which satisfy the asymptotic-abelian

condition, exhibit an approach to equilibrium which corresponds to nothing more than the spreading of a wave-packet for a free particle or free field. On the other hand, the mechanism of the approach to equilibrium of real systems is presumably governed by some "friction" process which operates locally, and which acts prior to the eventual dispersion of local disturbances (c.f. frictional term in the Navier-Stokes equation). For this reason, I think that the asymptotic abelian condition should be supplemented by some requirement of "local mixing" (corresponding to friction) in order to describe the approach to equilibrium in a more realistic way.

6. Phase transitions

6.1 General considerations

In general, phase transitions are characterized by:-

(a) a singularity or discontinuity in a thermodynamical function (e.g. the specific heat) at a certain temperature β_c^{-1};

(b) a spontaneous symmetry change (breakdown) on passing through the same temperature; and, in the case of transitions of the second kind;

(c)[25] a divergence of an otherwise finite correlation length, appropriately defined, as $\beta \to \beta_c \pm 0$; in cases where $\lim_{x\to\infty} \phi_\beta((\alpha(x)A)B)$ exists ($= F_\beta^{AB}$, say), one may define this length as

$$\sup_{A,B\in\mathfrak{A}} \{r \mid r\in R; \lim_{x\to\infty} [\phi_\beta((\alpha(x)A)B) - F_\beta^{AB}] e^{|x|/r'} = 0, \ \forall r < r'\}.$$

The theory of the characterization ($\mathfrak{A}$) has been treated systematically within the framework of traditional statistical mechanics (c.f. Ref. 26). In this treatment one formulates the properties of an intensive thermodynamical potential $f_V(\beta)$ (via a calculation of the partition function) for a system occupying volume V; and then one proceeds to the thermodynamical limit ($V\to\infty$) and examines whether the resultant function $f_\infty(\beta)$ has any singularities or discontinuities. This approach can, in principal, lead to the characterization (a); and, indeed, it is known to do so in the cases of certain exactly soluble models, e.g.

the BCS model, the Bose gas and the two-dimensional Ising model.

On the other hand, it is difficult to see how the traditional approach can lead to a rigorous treatment of spontaneous symmetry breakdown, as distinct from a breakdown induced by the application of some (usually fictitious) external field. In fact, it seems evident that the algebraic formalism provides one with the most natural means of approach to the characterizations (b) and (c). For this formalism is explicitly concerned with the properties of the symmetries and the correlations of the pure phase components of the Gibbs states (cf. Section 4.7.).

Actually, the algebraic formalism has been successfully employed to yield a theory of the characterization (b) for the BCS model[27,28] and for the ideal Bose gas[29]. In both these models the Gibbs state corresponds to a pure phase for $\beta < \beta_c$, but undergoes a gauge-symmetry breakdown for $\beta > \beta_c$ (here β_c is the thermodynamical transition temperature).

However, it is easily verified that neither the BCS model nor the ideal Bose gas posses the property (c). The BCS model lacks this property because it corresponds to a classical "mean field" theory, and consequently its correlation functions $\phi((\alpha(x)A)B)$ factorize in a way that precludes the possibility of (c). On the other hand, the Bose gas has the pathology that its correlation length ℓ is infinite for all $\beta > \beta_c$, and thus the critical point is not identified as the unique temperature where ℓ diverges.

In a recent work[2], which I shall outline below (Sections 6.2-5) Maria Marinaro and I have formulated an algebraic treatment of a rather wide class of (non-mean-field theoretic) Ising models, including the exactly soluble two-dimensional one, with the view to obtaining a theory of (b) and (c). In the two-dimensional case we prove that the model does exhibit the characteristics (b) and (c); thus we supplement Onsager's[30] result that it exhibits the property (a). In the more general case, we expressed space-correlation functions in terms of a certain semi-group, Σ, of contractions of a certain Hilbert space, and obtained conditions on Σ that sufficed to ensure that the system had the properties (b), (c): these conditions are fulfilled

in the two-dimensional case. In the more general case, this formulation might be useful for the purpose of constructing an axiomatic approach to the theory of phase transitions.

6.2 Ising Models[2]: the algebras.

Let X, the set of sites for the spins, be of the form T x Z, where T is a countably infinite point set and Z is the set of integers. Thus, in the case where $T = Z^d$, where d is a positive integer, X corresponds to a rectangular lattice of dimension (d+1).

We denote points in T, Z, X by t, n, x = (t,n). Translations along the Z-component of X correspond to the group $g = \{u^n | n \in Z\}$, where

$$u^n(t,n') = (t,n+n'), \ \forall t \in T; \ n,n' \in Z.$$

Let $\sigma = \{-1,1\}$, with discrete topology. We define K to be the topological power σ^X. K is thus a compact space (by Tychonoff's theorem). Points in K will be denoted by σ_X. The projection of σ_X corresponding to $x(\in X)$ will be denoted by σ_x. Thus σ_x may be interpreted as the spin at x; and K then corresponds to the space of spin configurations in X. The group g induces a group $G = \{U^n | n \in Z\}$ of K:-

$$(U^n \sigma_X)_x = \sigma_{u^n x}, \ \forall x \in X, \ n \in Z.$$

Let C(K) be the C*-algebra of bounded, continuous, complex-valued functions on K, with supremum norm. We take C(K) to be the algebra of observables, $\mathcal{A}$, for the system. We define the homomorphism α, of Z into Aut $\mathcal{A}$, corresponding to translations along the Z-component of X, by

$$(\alpha(n)A)(\sigma_X) = A(U^{-n}\sigma_X), \ \forall n \in Z, \ A \in \mathcal{A},$$

where U was defined above.

Corresponding to each finite point subset, Y, of X we define $\mathcal{A}_Y$ to be subalgebra of $\mathcal{A}$ consisting of functions A on K, such that the value of $A(\sigma_X)$ is independent of $\{\sigma_x | x \in X/Y\}$. By the Stone-Weierstrass theorem, the union

$\mathcal{A}_L$, of such algebras is dense in $\mathcal{A}$. Thus, $\mathcal{A}$ has the quasi-local structure described in Section 3.1.

For each finite point subset Y of X we define $P_Y(\in\mathcal{A}_Y)$ by:-

$$P_Y(\sigma_X) = \prod_{x\in Y}\sigma_x.$$

Thus one easily proves that $\mathcal{A}_Y$ is the vector space of linear combinations of $\{P_{Y'}|Y'\subset Y\}$. Hence, defining $\tilde{\mathcal{A}}$ to be the set of all P_Y's, for finite point subsets Y of X, it follows easily from our constructions that $\tilde{\mathcal{A}}$ is a linear basis for $\mathcal{A}$; i.e. $\mathcal{A}$ corresponds to the norm closure of the set of finite linear combinations of elements of $\tilde{\mathcal{A}}$. We define $\tilde{\mathcal{A}}_0$ to be the subset of $\tilde{\mathcal{A}}$ consisting of elements P_Y for which Y consists of an odd number of points.

We now introduce an auxiliary (non-commutative) algebra, $\mathcal{B}$, of Pauli spins over T. Thus, we assign to each $t\in T$ a two-dimensional Hilbert space $\mathcal{H}_t$. Corresponding to each finite point subset M of T, we define $\mathcal{H}_M = \bigotimes_{t\in M}\mathcal{H}_t$. Thus $\mathcal{H}_{MUM'} = \mathcal{H}_M \otimes \mathcal{H}_{M'}$ if $M \cap M' = \emptyset$. We then define $\mathcal{B}_M$ to be the C*-algebra of bounded operators in $\mathcal{H}_M$, with uniform norm. We identify $B(\in\mathcal{B}_M)$ with $B \otimes I_{M'\setminus M}(\in\mathcal{B}_{M'})$ for $M \subset M'$. Thus $\mathcal{B}_M$ is isotonic with respect to M; and the union, over all finite point subsets $M(\in T)$ of the $\mathcal{B}_M$'s is a normed *-algebra. We define $\mathcal{B}$ to be the norm closure of this union.

6.3 The Interactions.

As usual the Gibbs states on $\mathcal{A}$ are constructed as limits of 'finite-volume' Gibbs states, which in turn are specified by finite-volume Hamiltonians. In the present case, we define these as follows.

Let $\mathfrak{m}$ (resp. $\mathfrak{n}$) be a system of finite point subsets of T (resp. Z) whose union covers T (resp. Z). We shall always take the elements N of $\mathfrak{n}$ to be of the form $[n_1,n_2](\equiv\{n|n\in Z, n_1\leq n\leq n_2\})$.

In the two-dimensional case, where T = Z, we take the elements, M, of $\mathfrak{m}$ to be of the same form.

We assign to the region M x N (with $M \in \mathbb{m}$, $N \in \mathbb{n}$) a Hamiltonian $H_{MN} \in \mathcal{G}_{M \times N}$. In the two-dimensional case, this will take the form

$$H_{MN}(\sigma_X) = -J \sum_{n \in N} \sum_{m \in M} (\sigma_{m,n}\sigma_{m,n+1} + \sigma_{m,n}\sigma_{m+1,n}),$$

where $J>0$; and where, for $M = [m_1, m_2]$ and $N = [n_1, n_2]$, we identify m_2+1 with m_1, and n_2+1 with n_1 in the above summation (cyclical bounded conditions).

In the more general case, take

$$H_{MN} = \sum_{n \in N} f_{M,n} + \sum_{n \in N} g_{M,n,n+1}$$

where

$$g_{M,n,n+1}(\sigma_X) = -J \sum_{t \in M} \sigma_{t,n}\sigma_{t,n+1}$$

and where $f_{M,n} \in \mathcal{G}_{Mx\{n\}}$. Further restrictions are imposed on $f_{M,n}$ to ensure that it corresponds to a potential energy in a "hyperplane" $Mx\{n\}$, in which each spin is coupled to only a finite number of other spins.

6.4 Gibbs States.

We define the finite volume Gibbs state ϕ_β^{MN} on $\mathcal{G}_{M \times N}$ by:-

$$\phi_\beta^{MN}(A) \equiv E_{MN}(A)/E_{MN}(I)$$

where

$$E_{MN}(A) = \sum_{\{\sigma_x \mid x \in M \times N\}} A(\sigma_X)\exp(-\beta H_{MN}(\sigma_X)), \quad \forall A \in \mathcal{G}_{M \times N}$$

We prove that for any $A \in \mathcal{G}_L$, $\lim_{M\to\infty} \lim_{N\to\infty} \phi_\beta^{MN}(A)$ exists, and thus defines a Gibbs state ϕ_β on $\mathcal{G}$, as in section 4.2. Further, this state is invariant under the group $\alpha^*(Z)$, induced in the state space of $\mathcal{G}$ by the group $\alpha(Z)$, defined in section 6.2.

6.5 Properties of ϕ_β.

Using an extension of Onsager's transfer matrix formalism, we prove the following properties for ϕ_β in terms of the basis set $\tilde{\mathcal{G}}$ and the auxiliary algebra $\mathcal{B}$, both defined in section 6.2:-

(1) The state ϕ_β induces a map $\theta_\beta:\tilde{\mathcal{G}}\to\mathcal{B}$ and a state ψ_β on $\mathcal{B}$ such that $\phi_\beta=\psi_\beta\circ\theta_\beta$ on $\tilde{\mathcal{G}}$. Further the properties of ψ_β, θ_β are completely determined by the infinite-volume limiting properties (appropriately defined) of the Onsager transfer matrix (for transfer along the Z-component of X).

(2) Let $(\tilde{\mathcal{H}}_\beta,\pi_\beta,\omega_\beta)$ be the GNS triple corresponding to ψ_β. Then $\exists$ a quadruple $(\mathcal{H}_\beta,\xi_\beta,\xi^*_\beta,v_\beta)$ such that:-
(i) $\mathcal{H}_\beta(\ni\omega_\beta)$ is a closed subspace of $\tilde{\mathcal{H}}_\beta$
(ii) $\xi_\beta(=\pi_\beta\circ\theta_\beta),\xi^*_\beta$ are maps of $\tilde{\mathcal{G}}$ into $\mathcal{H}_\beta$, such that $\xi_\beta(\tilde{\mathcal{G}})(\equiv\xi^*_\beta(\tilde{\mathcal{G}}))$ is a linear basis set for $\mathcal{H}_\beta$.
(iii) v_β is a positive self-adjoint contraction in $\mathcal{H}_\beta$ and $v_\beta\omega_\beta=\omega_\beta$.
(iv) $\phi_\beta(A) = (\omega_\beta,\xi_\beta(A)) = (\xi^*_\beta(A),\omega_\beta)$, $\forall A\in\tilde{\mathcal{G}}$
(v) Given A, $A'\in\tilde{\mathcal{G}}$, $\exists n_o = n_o(A',A)$.

$$\phi_\beta(A'\alpha(n)A) = (\xi^*_\beta(A'),v_\beta^{n-n_o}\xi_\beta(A)),\ \forall A,A'\in\tilde{\mathcal{G}},\ n>n_o$$

Thus the structure of the state are determined by the quadruple $(\mathcal{H}_\beta,\xi_\beta,\xi^*_\beta,v_\beta)$, which in turn is determined by certain limiting properties of the finite-volume transfer matrices. In particular, space-translations in the Z_+ (and likewise the Z_-) direction corresponds to the one-parameter semi-group $\{v_\beta^n|n\in Z_+\}$ of positive self-adjoint contractions of $\mathcal{H}_\beta$. The formulation of correlations in terms of such contractions has obvious advantages for purpose of obtaining the asymptotic (e.g. the cluster) properties of such functions.

For the case of the two-dimensional model, in the absence of any external field, we utilize the properties of the finite volume transfer matrix, as obtained by Schultz, Mattis and Lieb[31]; and thereby obtain the following results:-
(I) For $\beta\neq\beta_c$, there is a gap $\Delta(\beta)$ between the principal eigenvalue (unity) of v_β and the rest of its spectrum. This gap $\to 0$ as $\beta\to\beta_c\pm 0$.

(II) For $\beta<\beta_c$, the subspace of $\mathcal{H}_\beta$ that is invariant under v_β consists of the scalar multiples of ω_β.
(III) For $\beta>\beta_c$, the principal eigenvalue (unity) of v_β becomes degenerate. Further, the principle eigenprojector of v_β is not orthogonal to the subspace of $\mathcal{H}_\beta$ generated by $\xi_\beta(\tilde{G}_o)(\equiv\xi^*_\beta(\tilde{G}_o))$, with $\tilde{G}_o$ defined as in section 6.2.

It is a simple matter to infer the following results from (I) - (III):-
(A) For $\beta<\beta_c$, ϕ_β is Z-ergodic (and therefore X-ergodic); while, for $\beta>\beta_c$, ϕ_β undergoes a symmetry breakdown, corresponding to spin inversion ($\sigma\to-\sigma$), associated with its Z-ergodic decomposition
(B) $\lim \phi_\beta(A\alpha(n)B)$ exists ($=F^{AB}_\beta$, say), $\beta\neq\beta_c$. Further, defining the correlation length $\ell(\beta)$ as $\sup_{A,B\in\mathcal{G}}\{r\,|\,r\in R;$ $\lim_{n\to\infty}(\phi_\beta(A\alpha(n)B) - F^{AB}_\beta)e^{n/r'} = 0, \forall\, r < r'\}$, we see that $\ell(\beta) = [\ell n(1-\Delta(\beta))^{-1}]^{-1}$. Thus, $\ell(\beta)$ is finite for $\beta\neq\beta_c$, and tends to infinity as $\beta\to\beta_c\pm0$.

Thus, the model exhibits the characteristics (b, (c), as well as (a). Since (b) and (c) are consequences of (I)-(III), it is tempting to propose these latter statements as "axioms" not merely for the two-dimensional case but for the wider class of models that we consider. We have, in fact, formulated a framework for the "scaling laws", even in the presence of an external magnetic field, on the basis of (I)-(III), together with a supplementary assumption concerning the spectral projection for v_β.

REFERENCES

1. D. Dubin and G. L. Sewell, J. Math. Phys. 11, 2990 (1970).

2. M. Marinaro and G. L. Sewell, Common.Math.Phys.24,310(1972).

For a comprehensive treatment of C*-algebras see
3. J. Dixmier, "Les C*-algebras et leurs representations" (Gauthier-Villars, Paris, 1964).

For a systemmatic treatment of amenable groups see
4. F. P. Greenleaf, "Invariant means on topological groups" (Van Nostrand, N.Y., 1969).

5. W. F. Eberlein, Trans. American Math. Soc. 67, 217 (1949).

6. R. R. Phelps, "Choquet's Theorem" (Van Nostrand, N.Y., 1966).

7. S. Doplicher, R. Kadison, D. Kastler and D. Robinson, Commun. Math. Phys. 6, 101 (1967).

8. O. Lanford and D. Ruelle, J. Math. Phys. 8, 1460 (1967).

9. D. Ruelle, "Statistical Mechanics" (Benjamin, N.Y., 1969).

10. R. Haag, N. M. Hugenholtz and M. Winnink, Commun. Math. Phys. 6, 189 (1967).

11. M. Winnink, Cargese Lectures (ed. D. Kastler), 1969.

12. M. Takesaki, "Tomita's theory of modular Hilbert algebras and its applications" (Springer-Verlag, N.Y., 1970).

13. G. F. Dell'Antonio, S. Doplicher and D. Ruelle, Commun. Math. Phys. 2, 223 (1966).

14. D. Robinson, Commun. Math. Phys. 6, 151 (1967)
Commun. Math. Phys. 7, 337 (1968)

15. S. Sakai, Proc. Japan Acad. 33, 438 (1957).

16. G. Emch, H. J. Knops and E. J. Verboven, J. Math. Phys. 11, 1655 (1970).

17. G. Emch and H. J. Knops, J. Math. Phys. 11, 3008 (1970).

18. C. Radin, Commun. Math. Phys. 21, 291 (1971).

19. H. Araki, Publ. RIMS, Kyoto Univ. Ser. A, 4, 361 (1968).

20. G. L. Sewell, J. Math. Phys. 11, 1868 (1970).

21. H. Araki and H. Miyata, Publ. RIMS, Kyoto Univ. Ser. A, 4, 373 (1968).

22. D. Ruelle, Cargese Lectures (Ed. D. Kastler), 1968.

23. H. Narnhofer, Acta Physica Austriaca 31, 349 (1970).

24. G. Emch and C. Radin, J.Math.Phys.12, 2043 (1971).

25. L. Kadanoff et al., Rev. Mod. Phys. 39, 395 (1967).

26. M. Fisher, Boulder Lecture Notes, 1966.

27. G. Emch and M. Guenin, J. Math. Phys. 7, 915 (1966).

28. W. Thirring and A. Wehrl, Commun. Math. Phys. 4, 303 (1967).

29. H. Araki and E. J. Woods, J. Math. Phys. 4, 637 (1963).

30. L. Onsager, Phys. Rev. 65, 117 (1944).

31. T. D. Schultz, D. C. Mattis and E. Lieb, Rev. Mod. Phys. 36, 856 (1964).

$$p_0 = \pm\sqrt{\vec{p}^2 + m^2}$$

this is not so for $m^2 < 0$. As long as only real energies are considered, i.e. as long as

$$p_0^2 = \vec{p}^2 + m^2 \geq 0,$$

we are restricted to momenta for which

$$\vec{p}^2 \geq -m^2 > 0.$$

This spells trouble for locality since it takes an integral over all momenta to achieve spatial localization.

In view of these difficulties we shall base tachyon quantization on a manifestly covariant and local expression. It has been emphasized by Tanaka[3] and especially by Schroer[4] that such a starting point is provided by the commutator equation

$$[A(x), A(y)] = i\Delta(x-y, m).$$

Inspection of

$$\Delta(x,m) = -\frac{1}{2\pi}\,\epsilon(x_0)\left\{\delta(x^2) - \theta(x^2)\frac{m}{2\sqrt{x^2}}\,J_1(m\sqrt{x^2})\right\}$$

shows that there is a well defined continuation to imaginary m where

$$\Delta(x,\pm i|m|) = -\frac{1}{2\pi}\,\epsilon(x_0)\left\{\delta(x^2) - \theta(x^2)\,\frac{|m|}{2\sqrt{x^2}}\,I_1(|m|\sqrt{x^2})\right\}.$$

As an aside it is worth mentioning that any such theory will be canonical: for $\varphi(\vec{x}) \equiv A(x)$ and $\pi(\vec{y}) \equiv \frac{\partial}{\partial y_0} A(y)$ at some fixed time $x_0 = y_0$ we have

$$[\varphi(\vec{x}), \pi(\vec{y})] = i\,\delta^3(\vec{x}-\vec{y}).$$

While it is challenging to speculate about regions of strongly attractive potentials as a natural habitat for tachyonic excitations[5-8] we shall here concentrate on quantizations with a ground state invariant under space time translations.

It turns out that this requirement imposes an indefinite metric on the representation space (one example each of representations with positive and indefinite metric have been constructed by Schroer[4]).

The further requirement of rotation invariance rules out all irreducible representations. The condition of Lorentz invariance, finally, will leave us with a one parameter family of causal, Poincaré invariant quantizations.

2. The Two-Point Function

We are interested in free field representations which are characterized by a translation invariant two point function

$$W_2(x-y) = \langle 0|A(x)\ A(y)|0\rangle .$$

Hence we make the ansatz

$$W_2(x) = \left[\frac{1}{2\pi}\right]^3 \int d^3\vec{k} \left[e^{ikx}\ \rho(\vec{k}) + e^{-ikx}\ \sigma(\vec{k})\right]$$

with

$$k_0 \equiv \omega = \begin{cases} \sqrt{\vec{k}^2 + m^2} & \text{if } \vec{k}^2 + m^2 > 0 \\ i\sqrt{-\vec{k}^2 - m^2} & \text{if } \vec{k}^2 + m^2 < 0. \end{cases}$$

We note for comparison that for the tardyon Fock representation we would have

$$\rho_{\text{Fock}} \equiv 0 \qquad \sigma_{\text{Fock}} = (2\omega)^{-1}.$$

The first thing to do will be to impose the commutation relation

$$W_2(x-y) - W_2(y-x) = i\Delta(x-y,m)$$

on our ansatz. By comparing

$$W_2(x) - W_2(-x)$$

with

$$i\Delta(x,m) = \left[\frac{1}{2\pi}\right]^3 \int d^3\vec{k} \left\{ e^{ikx} \left[-\frac{1}{2\omega} \right] + e^{-ikx} \left[\frac{1}{2\omega} \right] \right\}$$

we find that

$$\sigma(\vec{k}) = \frac{1}{2\omega(\vec{k})} + \rho(\vec{k})$$

is necessary and sufficient for the commutation relation to hold in the space generated by W_2. This condition shows clearly that to obtain the correct local commutation relations, contributions from all momenta are required. In particular we may not simply set

$$\rho = \sigma = 0$$

in the region of imaginary energies.

Having thus eliminated σ from our ansatz we proceed similarly to invoke hermiticity

$$A^\dagger(x) = A(x)$$

for the field operator.

For the two-point function this implies

$$W_2^*(-x) = W_2(x)$$

or in terms of ρ

$$\rho^*(\vec{k}) = \begin{cases} \rho(\vec{k}) & \text{if } \omega^2 > 0 \\ \\ \dfrac{1}{2\omega} + \rho(-\vec{k}) & \text{if } \omega^2 < 0. \end{cases}$$

To explore the metric properties of the representation space we focus on

$$\| \varphi(f) | 0 > \|^2 = \int d^3\vec{k}\ \tilde{f}^*(\vec{k})\ \tilde{w}_{\varphi\varphi}(\vec{k})\ \tilde{f}\ (\vec{k})$$

and

$$\| \pi(f) | 0 > \|^2 = \int d^3\vec{k}\ \tilde{f}^*(\vec{k})\ \tilde{w}_{\pi\pi}(\vec{k})\ \tilde{f}\ (\vec{k}).$$

Here

$$\varphi(f) = \int d^3\vec{x}\ \varphi(\vec{x})\ f(\vec{x})$$

$$\tilde{f}(\vec{k}) = \left[\frac{1}{2\pi}\right]^{3/2} \int d^3\vec{x}\ e^{-i\vec{k}\vec{x}} f(\vec{x})$$

and consequently

$$\tilde{w}_{\varphi\varphi}(\vec{k}) = \sigma(\vec{k}) + \rho(-\vec{k}) = \frac{1}{2\omega} + \rho(\vec{k}) + \rho(-\vec{k})$$

and

$$\tilde{w}_{\pi\pi}(\vec{k}) = \omega^2\left[\sigma(\vec{k}) + \rho(-\vec{k})\right]$$

$$= \omega^2\left[\frac{1}{2\omega} + \rho(\vec{k}) + \rho(-\vec{k})\right]$$

Evidently the "norms"

$$\|\varphi(f)|0>\|^2 = \langle 0|\varphi(f)\ \varphi(f)|0\rangle$$

and

$$\|\pi(f)|0>\|^2 = \langle 0|\pi(f)\ \pi(f)|0\rangle$$

are real (positive, zero) if and only if the kernels $\tilde{w}_{\varphi\varphi}$ and $\tilde{w}_{\pi\pi}$ are real (positive, zero) almost everywhere.

While reality of the kernels, and hence of the norms is a consequence of the hermiticity condition imposed on ρ we observe that for $\omega^2 < 0$ the kernels cannot both be positive, i.e.

All translation invariant representations of the tachyon commutation relations are endowed with an indefinite metric.

Actually this is true for the larger class of theories which have a ground state invariant under spatial translations and for which the time translations are generated by operators $U(t)$ (while not necessarily $U(t)|0> = |0>$). We plan to discuss such representations elsewhere.

At this point and before restricting our ansatz further by requiring rotational or full Poincaré invariance of the

ground state we pause to give an explicit construction of the field algebra.

3. Field Algebra and Hamilton Operator

Being mainly interested in the peculiarities brought about by $\omega^2 < 0$ we shall consider only the Fock representation as long as $\omega^2 > 0$, i.e.

$$\rho(\vec{k}) = 0 \qquad \text{for all } |\vec{k}| > |m|.$$

One verifies straightforwardly that the ansatz

$$A(x) = \left[\frac{1}{2\pi}\right]^{3/2} \int_{\omega^2 > 0} \frac{d^3\vec{k}}{\sqrt{2\omega}} \left[a(\vec{k})\ e^{-ikx} + a^*(\vec{k})\ e^{ikx}\right]$$

$$+\left[\frac{1}{2\pi}\right]^{3/2} \int_{\omega^2 < 0} d^3\vec{k} \left\{\left[\rho(\vec{k})\ b_+(-\vec{k}) + \rho^*(-\vec{k})\ b_+^*(\vec{k})\right] e^{ikx}\right.$$

$$\left.+\left[\chi(-\vec{k})\ b_-(\vec{k}) + \chi(\vec{k})\ b_-^*(-\vec{k})\right] e^{-ikx}\right\}$$

with

$$\chi(\vec{k}) = \begin{cases} 1 & \text{if } \rho(\vec{k}) \neq 0 \\ 0 & \text{if } \rho(\vec{k}) = 0 \end{cases}$$

and

$$b_\pm(\vec{k})|0> \equiv 0$$

$$[b_\pm(\vec{k}),\ b_\pm(\vec{k}')] = 0 \qquad [b_\pm(\vec{k}),\ b_\pm^*(\vec{k}')] = 0$$

$$[b_\pm(\vec{k}),\ b_\mp(\vec{k}')] = 0 \qquad [b_\pm(\vec{k}),\ b_\mp^*(\vec{k}')] = \delta^3(\vec{k}-\vec{k}')$$

gives rise to the correct commutation relation and n-point functions.

To disentangle the algebra of $b_\pm$ it is useful to introduce

$$a_\pm(\vec{k}) = \frac{1}{\sqrt{2}} (b_+(\vec{k}) \pm b_-(\vec{k}))$$

so that

$$[a_\pm(\vec{k}), a^*_\pm(\vec{k}')] = \pm\delta^3(\vec{k}-\vec{k}').$$

Since furthermore a_+ and a^*_+ commute with a_-, a^*_-, the representation space is the direct product of three Fock spaces

$$H = H_{\omega^2>0} \otimes H_+ \otimes H_-$$

where, however, H_- is endowed with an indefinite metric:

$$\langle 0| e^{\int z^*(\vec{k})\ a_\pm(\vec{k}) - z(\vec{k})\ a^*_\pm(\vec{k})\ d^3\vec{k}} |0\rangle = e^{\mp \frac{1}{2}\int_M |z(\vec{k})|^2\ d^3\vec{k}}$$

M will be specified below.

To ensure that the representation thus constructed is the one that arises from the generating functional

$$\langle 0| e^{iA(f)} |0\rangle = e^{-\frac{1}{2}\int f(x)\ W_2(x-y)\ f(y)\ dxdy}$$

we have to verify that it is cyclic.

To this end we note that the field algebra as well as the canonical algebra of φ and π are generated by

$$a(\vec{k}),\ a^*(\vec{k}) \qquad \text{for } \vec{k}^2 > -m^2$$

and by

$$\rho(\vec{k})\ b_+(-\vec{k}) + \rho^*(-\vec{k})\ b^*_+(\vec{k}),\ \chi(-\vec{k})\ b_-(\vec{k}) + \chi(\vec{k})\ b^*_-(-\vec{k})$$

for $\vec{k}^2 < -m^2$.

Hence the field algebra generates states

$$a^*(\vec{k})|0\rangle \text{ and } a^*_\pm(\vec{k})|0\rangle$$

from the vacuum, the latter if $\vec{k} \in \text{supp}\ \rho(-\vec{k}) \equiv M$. By iteration of this procedure one obtains the higher order states.

With this choice of M, then,

$$H = H_{\omega^2>0} \otimes H_+ \otimes H_-$$

is the cyclic representation space for the field A(x).

If there is a set $G \subset R^3$ of non-zero measure on which

$$\chi(\vec{k}) = \chi(-\vec{k}) = 1$$

then any operator

$$\int_G f(\vec{k})(b_+(-\vec{k}) + b_+^*(\vec{k}))\ d^3\vec{k}$$

will commute with A(x).

On the other hand, if there is no such G, the operators given above as generators of the field algebra specialize to

$$a(\vec{k}),\ a^*(\vec{k})$$

and

$$b_+(-\vec{k}),\ b_-^*(-\vec{k}) \text{ or } b_+^*(\vec{k}),\ b_-(\vec{k})$$

depending on whether we choose $-\vec{k} \in M$ or $\vec{k} \in M$. From these then we can build the $a^{(*)}_{\pm}(\vec{k})$ for $\vec{k} \in M$. Hence: The canonical and field algebras are irreducible if and only if $\rho(\vec{k}) \cdot \rho(-\vec{k}) = 0$ for almost all $\vec{k}$ with $\vec{k}^2 < -m^2$. In particular, rotation invariant representations with $\rho(\vec{k}) = \rho(|\vec{k}|)$ are always reducible.

We note that

$$\rho(\vec{k}) = -\theta(k_3)(2\omega)^{-1} \quad \text{for } \omega^2<0$$

provides an example of an irreducible representation.

The Hamiltonian, for which

$$i[\mathcal{H},A(x)] = \frac{\partial}{\partial x_0} A(x)$$

and

$$\mathcal{H}|0\rangle = 0$$

is given by

$$\mathcal{H} = \int_{\omega^2>0} d^3\vec{k}\omega(k)\ a^*(\vec{k})\ a(\vec{k}) + \int_M d^3\vec{k}\omega(\vec{k})(b_+^*(\vec{k})\ b_-(\vec{k})$$

$$-b_-^*(\vec{k})\ b_+(\vec{k})).$$

We recognize the $b_\pm^*(\vec{k})$ as creation operators for energy eigenstates with eigenvalues $\pm\omega(\vec{k})$.

4. Lorentz Invariance

To obtain a Lorentz invariant two-point function one might think of proceeding in analogy with our discussion of the tachyon commutator $i\Delta(x,m)$. However

$$i\Delta(x,m) = \left[\frac{1}{2\pi}\right]^3 \int d^3\vec{k}\ e^{i\vec{k}\vec{x}}\ \frac{\sin \omega x_0}{\omega}$$

depends on ω^2 only and hence

$$\Delta(x,i|m|) = \Delta(x,-i|m|).$$

On the other hand this is not so for

$$i\Delta_+(x) = \frac{1}{2}\ (i\Delta(x) + \Delta_1(x)) = \frac{i}{2}\ \Delta(x) + \begin{cases} \frac{m^2}{8\pi}\ \frac{N_1(m\sqrt{x^2})}{m\sqrt{x^2}} & x^2 > 0 \\ \frac{m^2}{4\pi}\ \frac{K_1(m\sqrt{-x^2})}{m\sqrt{-x^2}} & x^2 < 0 \end{cases}$$

as one sees quite easily from

$$\Delta_1(x) = \left[\frac{1}{2\pi}\right]^3 \int d^3\vec{k}\, e^{i\vec{k}\vec{x}} \frac{\cos \omega x_0}{\omega} .$$

As a result we have two Δ_+ - functions from which to construct a Lorentz invariant two-point function: The ansatz

$$W_2(x) = \alpha\, \Delta_+(x, i|m|) + \beta\, \Delta_+(x, -i|m|)$$

amounts to

$$\sigma(\vec{k}) = \frac{\alpha}{2\omega} \qquad \rho(\vec{k}) = -\frac{\beta}{2\omega} \quad \text{for } \omega^2 < 0$$

on which we have to impose the restrictions from the commutation relation

$$\sigma(k) - \rho(k) = \frac{1}{2\omega} \quad \text{i.e.} \quad \alpha + \beta = 1 .$$

and further from hermiticity

$$\rho^*(\vec{k}) - \rho(-\vec{k}) = \frac{1}{2\omega} \quad \text{i.e. } \beta + \beta^* = 1$$

Hence

$$\sigma(\vec{k}) = \frac{1 + ia}{4\omega} \qquad \rho(\vec{k}) = -\frac{1 - ia}{4\omega} \quad \text{for } \omega^2 < 0$$

and

$$\sigma(\vec{k}) = \frac{1}{2\omega} \qquad \rho(\vec{k}) = 0 \qquad \text{for } \omega^2 > 0$$

with arbitrary real a give rise to a one parameter family of Poincaré invariant quantizations. (Evidently we are dealing with pseudounitary representations of the Poincaré group.)

5. Summary

Tachyon quantizations are obtained by constructing representations of the tachyon commutation relations. These are causal: field operators at spacelike separation will commute with each other. The construction proceeds

from an ansatz for the two-point function; all higher truncated Wightman functions are set equal to zero.

Insistence on a translation invariant ground state necessitates an indefinite metric for the representation space in the imaginary frequency regime.

An explicit realization of the algebra in terms of creation and annihilation operators shows that e.g. rotation invariance of the ground state is sufficient to make the canonical algebra reducible. In contrast to reducible representations encountered elsewhere, the reducibility is not lifted by the action of the Hamiltonian: the space time algebra is not larger than the canonical one at a fixed time, and the field obeys the appropriate equation of motion:

$$(\Box + m^2)\ A(x) = 0.$$

While the explicit construction of the field algebra is anything but patently covariant it encompasses a one-parameter subfamily of Poincaré invariant representations. Thus consistent causal Poincaré invariant quantizations of the free tachyon field are obtained in a systematic way.

REFERENCES

1. Feinberg, G., Phys. Rev. 159, 1089 (1967).

2. Dhar, J. and E.C.G. Sudarshan, Phys. Rev. 174, 1808 (1968).

3. Tanaka, S., Progr. Theor. Phys. 24, 171 (1960).

4. Schroer, B., Phys. Rev. D3, 1764 (1971).

5. Schroer, B., and J. A. Swieca, Phys. Rev. D2, 2938 (1970).

6. Eachus, W.J.,"A Solvable Model in Quantum Field Theory" Unpublished thesis, Syracuse 1970.

7. Weinberg, J., private communication

8. Greenberg, O.W. and D. Bhaumik, "Parton Model with Variable Intermediate State Parton Mass" Univ. of Maryland preprint, May 1971.

LOCALITY AND COVARIANCE IN QED AND GRAVITATION

GENERAL PROOF OF GUPTA-BLEULER TYPE FORMULATIONS

F. Strocchi
Princeton University
Princeton, New Jersey

The aim of the present paper is to discuss the quantization of the electromagnetic and gravitational field in the framework of Wightman's theory[1]. The motivation for this is to clarify the difficulties arising from the zero mass in connection with the gauge group. All the known ways of quantizing the e.m. and the gravitational potential have in fact some unpleasant features. For example, the local and covariant Gupta-Bleuler formulation[2] of QED requires an indefinite metric Hilbert space whereas the Coulomb or radiation gauge quantization[3] uses non-local and non-covariant fields. The natural question is whether these difficulties arise because one insists on some unnecessary assumptions or there is some general property which makes them unavoidable. The impression one gets from the literature is that the difficulties connected with the quantization of the electromagnetic potential have a rather accidental origin. It seems in fact that all the troubles arise because one tries to impose the Lorentz condition $\partial_\mu A^\mu = 0$ and to work in a positive metric Hilbert space. None of the above conditions are really necessary. Even classically, there is no need for imposing the Lorentz condition in the Maxwell's equations

$$\Box A_\mu - \partial_\mu \partial^\nu A_\nu = 0$$

Similar considerations hold for the quantization of the gravitational potential with the Hilbert-Lorentz condition $\partial_\mu h^{\mu\nu} + \alpha \partial^\nu h^\mu{}_\mu = 0$ playing essentially the same role as the Lorentz condition in QED[4]. The literature on the subject is very rich, but all the formulations are based on

some definite choices or assumptions and one might hope that different choices may provide better solutions and avoid the unpleasant features of the known formulations. One should perhaps add that even in the free field case, there is a large freedom in the representation of the field operators because of the zero mass and one may consider representations which are not equivalent to the Fock representation.

The purpose of this paper is to provide general statements about the quantization of the e.m. and gravitational field without assuming the spectral condition, the temperedness of the fields, the uniqueness of the vacuum state, the Fock representation, the positive definiteness of the metric in the Hilbert space and consequently the unitarity of the Poincaré representation. Each of these assumptions could in fact turn out to be unnecessary. We will show under very general assumptions that the characteristic features of Gupta-Bleuler formulation like indefinite metric and unphysical states are in fact unescapable features of any local and/or covariant quantization of the electromagnetic potential. Conversely, any quantization avoiding indefinite metric and unphysical states like the Coulomb or radiation gauge formulation must be based on non-local and non-covariant fields. Similar statements are proved for the quantization of the gravitational potential.

1. Basic assumptions

The basic assumptions which serve as a definition of the problem are the following. (We consider the electromagnetic case first.)

i) The fields $A_\mu(x)$, $\mu = 0,1,2,3$, may be defined as operator valued distributions (not necessarily tempered), for which the Fourier transform may be defined. They are supposed to act in a Hilbert space $\mathcal{H}$ equipped with a nondegenerate[5] sesquilinear hertean form $\eta = \eta^\dagger$ ($\dagger$ denotes the Hilbert space adjoint.)

ii) There exists a representation $\{a\} \to U(a)$ of the space time translation group in $\mathcal{H}$ such that

$$U(a)A_\mu(f)U(a)^{-1} = A_\mu(f_a), \quad f_a(x) \equiv f(x-a)$$

and the operators U are unitary with respect to η

$$U^{\dagger}\,\eta U = \eta$$

iii) There is a vector ψ_o, called the vacuum state, in the domain of the fields $A_\mu(f)$, invariant under U(a).

$$U(a)\psi_o = \psi_o$$

A few comments on the above assumptions may be useful. Assumption (i) is nothing but the statement that $A_\mu(x)$ can be defined as quantum fields. No assumption is made about the type of distribution except for the requirement that the Fourier transform exists. For example $A_\mu(x)$ could be a strictly local field in the sense of Jaffe [6]. The introduction of the sesquilinear form η is done in order to cover the most general use. It might be necessary, in the quantization of the electromagnetic potential, to define all the physically meaningful quantities such as transition probabilities, vacuum expectation values, etc., in terms of a "product" (,) defined by

$$(\psi_1,\psi_2) = \langle \eta\ \psi_1,\psi_2\rangle$$

where <,> is the natural scalar product in $\mathcal{H}$. Assumptions (ii) and (iii) look rather mild and it seems difficult to think of a quantum field theory in which they do not hold.

2. Quantization of Maxwell's equation and Lorentz covariance

The assumptions made up to now apply to a general vector field (for example a massive vector field). The connection with free quantum electrodynamics is given by the Maxwell's equations

$$\partial_\mu F^{\mu\nu} = 0 \qquad\qquad \epsilon_{\mu\nu\rho\sigma}\partial^{\nu}F^{\rho\sigma} = 0 \qquad\qquad (2.1)$$

They have the advantage of not involving unphysical states or subsidiary conditions. It is therefore reasonable to try to impose eqs. 2.1, and see which of the basic assumption of quantum field theory conflict with them.

The first property we want to discuss is covariance under the Lorentz group. Up to now no assumption was made about the transformation properties of $A_\mu(x)$ under the Lorentz group. ($A_\mu(x)$ could be a non-covariant field like in the Coulomb or radiation gauge.) We will therefore try to assume that

iv) The two point function $(\psi_o, A_\mu(x)A_\nu(y)\psi_o) \equiv W_{\mu\nu}(x-y)$ transforms covariantly under the Lorentz group

$$W_{\mu\nu}(x) = \Lambda^{-1}{}_{\mu}^{\alpha}\ \Lambda^{-1}{}_{\nu}^{\beta}\ W_{\alpha\beta}(\Lambda x)$$

One may then prove the following:

Statement 1.[7] In any quantum field theory satisfying conditions (i), (ii), (iii), (iv) (weak local commutativity is not required).

(a) The Maxwell's equations cannot even hold as weak equations on the vacuum state

$$\partial_\mu F^{\mu\nu}\psi_o = 0, \qquad \epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma}\psi_o = 0 \tag{2.2}$$

Otherwise the two point function $(\psi_o, F_{\mu\nu}(x)F_{\rho\sigma}(y)\psi_o)$ vanishes.

(b) If the Maxwell's equations are required to hold as mean values in $D_o \equiv \{$ set of vectors obtained by applying polynomials in the smeared fields $F_{\mu\nu}(f)$ to $\psi_o \}$

$$(\Phi, \partial_\mu F^{\mu\nu}(f)\psi) = 0, \qquad \forall \Phi, \psi \in D_o \ .$$

then the metric η cannot be positive definite in D_o, and there is a subspace $\mathcal{H}'' \subset D_o$ of vectors of vanishing η-norm

$$(\Phi, \Phi) = \langle \eta\Phi, \Phi\rangle = 0 \qquad \forall \Phi \in \mathcal{H}''$$

and

$$\partial_\mu F^{\mu\nu}(f) D_o \subset \mathcal{H}''$$

(c) the metric η cannot be semidefinite ($\eta \geqslant 0$) in $\mathcal{H}$; i.e., there must be vectors in $\mathcal{H}$ with negative η-norm: $(\psi, \psi) = \langle \eta\psi, \psi\rangle < 0$

Proof.

(a) One considers the two-point function

$$(\psi_o , A_\mu(x)F_{\rho\sigma}(y)\psi_o) \equiv F_{\mu\rho\sigma}(x-y) \qquad (2.3)$$

Eqs. (2.2) imply

$$(\Box F_{\mu\nu})(f)\psi_o = [\partial_\mu(\partial F)_\nu - \partial_\nu(\partial F)_\mu]\psi_o + \partial^\tau \epsilon_{\tau\mu\nu\sigma}(\partial F^d)^\sigma \psi_o = 0$$

$((\partial F)_\nu \equiv \partial^\rho F_{\rho\nu}, (\partial F^d)^\sigma \equiv \partial_\lambda \epsilon^{\lambda\sigma\alpha\beta} F_{\alpha\beta})$. Therefore one has

$$\Box F_{\mu\rho\sigma}(x) = 0 , \qquad p^2 \tilde{F}_{\mu\rho\sigma}(p) = 0 ,$$

where $\tilde{F}$ denotes the Fourier transform of F. Hence

$$\text{supp } \tilde{F}_{\mu\rho\sigma} \subset \bar{V}^+ \cup \bar{V}^-$$

Now the forward and backward cone are regularly separated[8] with respect to their intersection (the origin), which is a compact set. Then by Malgrange's theorem[9] one may write

$$\tilde{F}_{\mu\rho\sigma}(p) = \tilde{F}^+_{\mu\rho\sigma}(p) + \tilde{F}^-_{\mu\rho\sigma}(p) \qquad (2.4)$$

where

$$\text{supp } \tilde{F}^+_{\mu\rho\sigma} \subset \bar{V}^+, \qquad \text{supp } \tilde{F}^-_{\mu\rho\sigma} \subset \bar{V}^-$$

The Fourier transform of eq. (2.4) gives

$$F_{\mu\rho\sigma}(x) = F^+_{\mu\rho\sigma}(x) + F^-_{\mu\rho\sigma}(x)$$

and both $F^\pm_{\mu\rho\sigma}$ transform covariantly under the Lorentz group, since by assumption iv) so does $F_{\mu\rho\sigma}$.

Because of the support properties $F^{\pm}_{\mu\rho\sigma}(x)$ may be regarded as boundary values[6] of two analytic functions $F^{\pm}_{\mu\rho\sigma}(z)$, F^{+} being analytic in the forward tube $\mathfrak{I}^{+}$ and F^{-} being analytic in the backward tube $\mathfrak{I}^{-}$

$$\mathfrak{I}^{\pm}\equiv\{z, \quad -\infty < \mathrm{Re}z < \infty, \quad \mathrm{Im}z\in V^{\pm}\}$$

According to assumption iv), $F^{\pm}_{\mu\rho\sigma}(z)$ yield two representations of the complex Lorentz group L_4 (C) and by the Araki-Hepp[10] theorem, they may be written in the following form

$$F^{\pm}_{\mu\rho\sigma}(z) = g_{\mu\rho}z_{\sigma}F_1{}^{\pm}(z) + g_{\mu\sigma}z_{\rho}F_2{}^{\pm}(z) + g_{\rho\sigma}z_{\mu}F_3{}^{\pm}(z)$$

$$+ \epsilon_{\mu\rho\sigma\lambda}z^{\lambda}F_4{}^{\pm}(z), \qquad \text{for } z\in\mathfrak{I}^{\pm}$$

where $g_{\mu\nu}$ is the metric tensor and $F_i{}^{\pm}(z)$ are invariant functions

$$F_i{}^{\pm}(z) = F_i{}^{\pm}(\Lambda z), \qquad \Lambda\in L_4(C)$$

By using essentially the same argument discussed elsewhere[11] and the antisymmetry of $F_{\rho\sigma}(x)$ one may write $F^{\pm}_{\mu\rho\sigma}(z)$ also in the following form

$$F^{\pm}_{\mu\rho\sigma}(z) = g_{\mu\sigma}\frac{\partial}{\partial z_{\rho}} - g_{\mu\rho}\frac{\partial}{\partial z_{\sigma}}F^{\pm}(z) + \epsilon_{\mu\rho\sigma\lambda}\frac{\partial}{\partial z^{\lambda}}G^{\pm}(z)$$

and by going to the boundary value one has

$$F_{\mu\rho\sigma}(x) = (g_{\mu\sigma}\partial_{\rho} - g_{\mu\rho}\partial_{\sigma})F + \epsilon_{\mu\rho\sigma\lambda}\partial^{\lambda}G \qquad (2.5)$$

where $F(x) = F^{+}(x) + F^{-}(x)$ and $G = G^{+}(x) + G^{-}(x)$ are Lorentz invariant distributions.

Now if one imposes the Maxwell equations, in the weak form

$$\partial_{\mu}F^{\mu\nu}(f)\psi_o = 0, \qquad \epsilon_{\mu\nu\rho\sigma}\partial^{\lambda}F^{\rho\sigma}(f)\psi_o = 0$$

one obtains the following restrictions for F and G

$$0 = \partial^{\rho} F_{\mu\rho\sigma}(x) = (\Box g_{\mu\sigma} - \partial_{\mu}\partial_{\sigma})\, F(x) \tag{2.6}$$

$$0 = \epsilon^{\lambda\nu\rho\sigma}\partial_{\nu} F_{\rho\sigma}(x) = (\Box g_{\lambda\mu} - \partial_{\lambda}\partial_{\mu})G(x) \tag{2.7}$$

The above differential equations for the Lorentz invariant distributions F and G have no solution apart from the trivial one[12]

$$F(x) = \text{const} \qquad G(x) = \text{const} \tag{2.8}$$

Eqs. (2.8) imply

$$(\psi_o\,,\ F_{\mu\nu}(x)\ F_{\rho\sigma}(y)\ \psi_o) = 0 \tag{2.9}$$

(b) If Maxwell's equations are required to hold only as mean values on D_o one gets

$$(\Phi,\ \partial_{\mu} F^{\mu\nu}(f)\ \psi) = 0 \qquad \forall \Phi,\ \psi \in D_o$$

Then by choosing $\Phi = \partial_{\mu} F^{\rho\sigma}(f)\psi$. one has

$$(\Phi, \Phi) = 0 \tag{2.10}$$

Since for $\psi = \psi_o$ one cannot have

$$\Phi = \partial_{\mu} F^{\mu\nu}(f)\psi_o = 0 \qquad \forall \text{ test function } f$$

as discussed at point a), eq. (2.10) proves that there are vectors in D_o with zero η-norm and that

$$\partial_{\mu} F^{\mu\nu}(f) D_o \subset \mathcal{K}''$$

(c) This follows from the following remark. A non-degenerate hermitean sesquilinear form η in a Hilbert space $\mathcal{K}$ cannot have vanishing expectation value on a non-zero vector, unless η is indefinite. One has in fact

$$(\psi+\lambda\Phi,\psi+\lambda\Phi)=(\psi,\psi)+|\lambda|^2(\Phi,\Phi)+2\mathrm{Re}[\lambda(\psi,\Phi)]$$

$$=(\psi,\psi)+2\mathrm{Re}[\lambda(\psi,\Phi)] \qquad (2.11)$$

for any vector ψ, if $\Phi \in \mathcal{K}''$. The left hand side is clearly indefinite unless $(\psi,\Phi) = 0$. Now, one cannot have

$$(\psi,\Phi) = 0 \qquad \forall\psi \in \mathcal{K}$$

because η is non-degenerate. Thus, eq. (2.11) implies that there are vectors in $\mathcal{K}$ with negative η-norm.

The content of Statement 1 is slightly strengthened by the following

Corollary If the metric η is semidefinite ($\eta \geq 0$) in D_0 and $F_{\mu\nu}(f)$ is hermitean with respect to η ($F^{\dagger}_{\mu\nu}(f)\eta=\eta F_{\mu\nu}(f)$, $\dagger$ denoting the Hilbert space adjoint), then eq. (2.2) implies that all the Wightman functions of $F_{\mu\nu}$ vanish.

Proof. Since, by Statement 1, the two-point function of $F_{\mu\nu}$ vanishes

$$F_{\mu\nu}(f)\psi_0 \in \mathcal{K}''$$

Now, if $\Phi \in D_0$ and $(\Phi,\Phi) = 0$, then $(\psi,\Phi) = 0 \quad \Psi \in D_0$, provided that η is non negative in D_0. Therefore

$$(\psi_0, F_{\mu\nu}(f_1)\ldots F_{\rho\sigma}(f_n)\psi_0) =$$

$$=(F_{\lambda\tau}(f_{n-1})\ldots F_{\mu\nu}(f_1)\psi_0,\ F_{\rho\sigma}(f_n)\psi_0) = 0$$

Remark. Statement 1 proves that any quantization of Maxwell's equations satisfying conditions i), ii), iii), in which the potential $A_\mu(x)$ transforms as a four vector, must share all the essential features of Gupta-Bleuler formulation. These results make use of very general properties of quantum field theory and may be regarded as a proof[13] that the Gupta-Bleuler formulation is inavoidable if one wants a covariant theory.

3. Quantization of Maxell's equations and weak local commutativity

The next property we want to investigate is Weak Local Commutativity (WLC), which was not assumed up to now. We will discuss the possibility of writing the Maxwell's equations in the weak form (2,2), in terms of a weakly local field operator $A_\mu(x)$. We will not assume that $A_\mu(x)$ transforms covariantly under the Lorentz group but only that

v) The two-point function $F_{\mu\nu,\rho\sigma}(x-y) \equiv (\psi_o, F_{\mu\nu}(x)F_{\rho\sigma}(y)\psi_o)$ transforms covariantly under the Lorentz group

$$F_{\mu\nu,\rho\sigma}(x-y) = \Lambda_\mu^{-1\alpha}\,\Lambda_\nu^{-1\beta}\,\Lambda_\rho^{-1\gamma}\,\Lambda_\sigma^{-1\delta}\,F_{\alpha\beta,\gamma\delta}(\Lambda(x-y))$$

Moreover we will assume that

vi) $$(\psi_o, [A_\mu(f), A_\nu(g)]\psi_o) = 0$$

whenever the supports of the test functions f and g are spacelike to one another.

The above equation may be empty unless one makes some assumption on the class of test functions for which $A_\mu(f)$ is defined. We will therefore assume that $A_\mu(x)$ are operator valued distributions for which WLC may be defined. For example the strictly local fields introduced by Jaffe[6] satisfy this condition.

One may then prove the following

Statement 2.[14] In any quantum field theory satisfying conditions i), ii), iii), v), and vi),

(a) The Maxwell's equations cannot even hold as weak equations on the vacuum state

$$\partial_\mu F^{\mu\nu}\psi_o = 0\,, \qquad \epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma}\psi_o = 0 \quad (2.2)$$

Otherwise the vacuum expectation value

$$(\psi_o\,,\ [F_{\mu\nu}(x),\ F_{\rho\sigma}(y)]\psi_o)$$

should vanish
b) and c) of Statement 1 also hold.

Proof.

(a) One considers the following commutator

$$(\psi_o\,,\ [A_\mu(x),\ F_{\rho\sigma}(y)]\psi_o) = F_{\mu\rho\sigma}(x-y)$$

The first step is to show that $F_{\mu\rho\sigma}(x-y)$ transforms covariantly under the Lorentz group, i.e.

$$F_{\mu\rho\sigma}(x-y) = \Lambda_\mu^{-1\alpha}\ \Lambda_\rho^{-1\beta}\ \Lambda_\sigma^{-1\gamma}\ F_{\alpha\beta\gamma}(\Lambda(x-y)) \qquad (2.12)$$

To this purpose it is convenient to introduce the following distribution

$$\mathfrak{F}_{\mu\rho\sigma}(x,\Lambda)\equiv\Lambda_\mu^{\ \alpha}\ \Lambda_\rho^{\ \beta}\ \Lambda_\sigma^{\ \gamma}\ F_{\alpha\beta\gamma}(\Lambda^{-1}x) - F_{\mu\rho\sigma}(x)$$

so that eq. (2.12) is proved by showing that $\mathfrak{F}_{\mu\rho\sigma}$ vanishes. An immediate consequence of assumption iv is that the four dimensional curl of $\mathfrak{F}_{\mu\rho\sigma}$, with respect to the index μ, vanishes. This implies that $\mathfrak{F}_{\mu\rho\sigma}$ may be written in the following form

$$\mathfrak{F}_{\mu\rho\sigma}(x,\Lambda) = \partial_\mu \mathfrak{F}_{\rho\sigma}(x,\Lambda)$$

Eqs. (2.2) imply that the supp $\overline{\mathfrak{F}}_{\rho\sigma}$ is contained $\overline{V^+ \cup V^-}$ and therefore by Malgrange's theorem one may split $\mathfrak{F}_{\rho\sigma}$ in the following way

$$\mathfrak{F}_{\rho\sigma}(x) = \mathfrak{F}^+_{\rho\sigma} - \mathfrak{F}^-_{\rho\sigma}$$

where

$$\text{supp}\ \overline{\mathfrak{F}^\pm_{\rho\sigma}} \subset \overline{V^\pm}$$

$\mathfrak{F}_{\rho\sigma}(x)$ are therefore boundary values of analytic functions $\mathfrak{F}^\pm_{\rho\sigma}(z)$, analytic in $\mathfrak{I}^\pm$ respectively. ($\mathfrak{I}^\pm$ denote the forward and backward tube).

Moreover by condition vi)

$$\mathfrak{F}^{+}_{\rho\sigma}(x) = \mathfrak{F}^{-}_{\rho\sigma}(x) \qquad \text{for } x^2 < 0$$

Then, by a known theorem[15], there is an analytic function $G_{\rho\sigma}(z)$ such that

$$G_{\rho\sigma}(z) = \mathfrak{F}^{+}_{\rho\sigma}(z) \qquad \text{for } z \in \mathfrak{I}^{+}$$

$$= \mathfrak{F}^{-}_{\rho\sigma}(z) \qquad \text{for } z \in \mathfrak{I}^{-}$$

Furthermore $G_{\rho\sigma}(z)$ is analytic in a neighborhood of Jost's points and by Streater's theorem[17] $G_{\rho\sigma}(z)$ can be analytically continued to $\mathfrak{I}'$, the extended tube.

We may now show that $G_{\rho\sigma}(z) = 0$. To this purpose we note that

$$\sum_{\rho,\sigma=0}^{3} z^{\rho} z^{\sigma} G_{\rho\sigma}(z) = 0 \qquad (2.13)$$

since $G_{\rho\sigma}$ is antisymmetric in ρ,σ. Therefore, putting

$$F_{\rho} = \sum_{\sigma=0}^{3} z^{\sigma} G_{\rho\sigma}(z)$$

we have

$$\sum_{\sigma=0}^{3} z^{\rho} F_{\rho}(z) = 0 \qquad \text{in } \mathfrak{I}' \qquad (2.14)$$

Now $\mathfrak{I}'$ contains the intervals of the form $\{z^0=0,\ z^i\neq 0,\ z^j=0,\ j\neq i\}$ and on those intervals eq. (2.14) becomes

$$z^i F_i(z) = 0 \qquad \text{(no sum over i!)} \qquad (2.15)$$

i.e.

$$F_i(z) = 0 \qquad (2.16)$$

By analyticity, eqs. (2.16) hold throughout $\mathfrak{I}'$ and by eq. (2.14) one also gets $F_o(z)=0$ throughout $\mathfrak{I}'$. In a similar way one also proves that $G_{\rho\sigma}(z)=0$.

Having proved that $F_{\mu\rho\sigma}(x)$ transform covariantly under the Lorentz group one may repeat the argument given in Statement 1 and conclude that eqs. (2.2) imply

$$F_{\mu\rho\sigma}(x) = 0$$

and

$$F_{\mu\nu,\rho\sigma}(x-y) \equiv (\psi_o, [F_{\mu\nu}(x), F_{\rho\sigma}(y)]\psi_o) = 0 \qquad (2.17)$$

This concludes the proof of a); b) and c) follow in the same way discussed in Statement 1. It is clear from eq. (2.17) that if eqs. (2.2) hold one would get a trivial theory and therefore one is forced to abandon eqs. (2.2). The conclusion can be strengthened if one assumes that the Fourier transform of the two-point function $(\psi_o, F_{\mu\nu}(x) F_{\rho\sigma}(y)\psi_o)$ has support in $\overline{V^+}$. Eq. (2.17) implies in fact $\tilde{F}_{\mu\nu,\rho\sigma}(p)=0$. Therefore, if one writes

$$\tilde{F}_{\mu\nu,\rho\sigma}(p) = \tilde{F}^+_{\mu\nu,\rho\sigma}(p) + \tilde{F}^-_{\mu\nu,\rho\sigma}(p) \qquad (2.18)$$

with

$$\text{supp } \tilde{F}^{\pm}_{\mu o} \subset \overline{V}^{\pm}$$

one has

$$\tilde{F}^{\pm}_{\mu\nu,\rho\sigma}(p) = 0 \qquad (2.19)$$

Now the two-point function

$$W_{\mu\nu,\rho\sigma}(x-y) \equiv (\psi_o, F_{\mu\nu}(x) F_{\rho\sigma}(y)\psi_o)$$

satisfies the spectral condition

$$\text{supp } W_{\mu\nu,\rho\sigma}(p) \subset \overline{V}^+$$

and therefore $\tilde{F}^{\pm}_{\mu\nu,\rho\sigma}$ in eq. (2.18) are the Fourier transforms $W_{\mu\nu,\rho\sigma}(x-y)$ and $-W_{\rho\sigma,\mu\nu}(y-x)$, respectively. Then, by eq. (2.19)

$$(\psi_o, F_{\mu\nu}(x)F_{\rho\sigma}(y)\psi_o) = 0$$

In conclusion, by combining Statement 1 and 2, any quantization of the Maxwell's equations using a weakly local and/or covariant potential $A_\mu(x)$ has all the essential features of the Gupta-Bleuler formulation. In this respect, Statement 1 and 2 are a general proof of Gupta-Bleuler formulation. Conversely any theory in which Maxwell's equations hold as operator equations must necessarily use a non-local and non-covariant potential. Thus a result which was known for the Coulomb or radiation gauge is shown to hold in general, whenever eqs. (2.2) hold.

4. Quantization of Einstein's equation: locality and covariance

To simplify the discussion we will consider the Einstein's equations in the weak field approximation in vacuo. This does not seem to be a limitation of the argument since one expects that the general theory will admit the weak field approximation as a limit

$$R_{\mu\nu}(f)\psi_o = 0 \quad \text{when } T_{\mu\nu}(f)\psi_o = 0$$

The assumptions which define the problem in this case are the same as those discussed in Section 2 with $A_\mu(x)$ replaced by the gravitation potential $h_{\mu\nu}(x)$ and $F_{\mu\nu}(x)$ replaced by the Einstein's tensor $R_{\mu\nu\rho\sigma}(x)$. One may then prove the following.

Statement 3. In any quantum field theory satisfying condition i), ii), iii), with the gravitational potential defined

either as a covariant operator (assumption iv))
or a weakly local operator (assumptions v), vi))

(a) The Einstein's equation cannot even hold in the weak form

$$R_{\mu\nu}(f)\psi_o = 0$$

Otherwise the two-point function $(\psi_o, R_{\mu\nu\rho\sigma}(f) R_{\alpha\beta\gamma\delta}(g)\psi_o)$ vanishes.

(b) If the Einstein's equations are required to hold as mean values in D_o = {set of vectors which are obtained by applying polynomials in the smeared fields, $R_{\mu\nu\rho\sigma}(f)$ to the vacuum}

$$(\Phi, R_{\mu\nu}(f)\psi) = 0 \qquad \forall \Phi, \psi \in D_o$$

then the metric η cannot be positive definite in D_o, there is a subspace $\mathcal{K}'' \subset D_o$ of vectors of zero η-norm and

$$R_{\mu\nu}(f)\ D_o \subset \mathcal{K}''$$

(c) The metric η cannot be semidefinite ($\eta \geq 0$) in $\mathcal{K}$, i.e. there must be vectors ψ with negative η-norm.

Proof.

The proof is similar to that of Statement 1 and 2. For details we refer the reader to reference (18).

5. Remarks

A possible objection to the difficulties encountered in a local and/or covariant quantization of the free Maxwell's and Einstein's equations may be that: a) one may quantize the Maxwell's equations in a local and covariant way by using only the fields $F_{\mu\nu}(x)$, without introducing the potential $A_\mu(x)$ and the indefinite metric is not necessary. b) even if $A_\mu(x)$ is introduced, as suggested by the second Maxwell's equation ($\epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma} = 0$), one may think that the interacting case is totally different from the free field theory and the above difficulties may disappear.

There are indications that the above remarks do not get into the core of the problem for the following reasons: a') One does not know how to formulate a local interacting theory (even in the Lagrangian approach) in terms of $F_{\mu\nu}$ only. Thus, even if the free field case can be formulated in terms of $F_{\mu\nu}$ only, there is little chance that this will be possible in the interacting case. Actually, there are

strong indications that S-matrix elements must involve A_μ in order to account for bremstrahlung processes. b') If, as one reasonably hopes, the interacting theory will allow the definition of asymptotic fields, $A_\mu^{in/out}(x)$ should be free fields and for them Statements 1 and 2 apply. The arguments a') and b') can be supplied by a rigorous statement To this purpose, let $j^\nu(x)$ denote the electromagnetic current, associated with the local charge operator so that

$$[\int j^\circ(\vec{x},f) f_R(\vec{x}) f_d(x^\circ) d^3x\, dx_\circ\,,\ \varphi(g)] = q\varphi(g) \qquad (2.20)$$

for R sufficiently large, where $\int f_d(x^\circ) dx^\circ = 1$,

$f_R(\vec{x})=1$ for $|\vec{x}| < R$, $f_R(\vec{x})=0$ for $|\vec{x}| > R+\epsilon$, $f_d(x^\circ)=0$

for $|x_\circ| > d$, and $\varphi(x)$ is local field carrying charge q. Then one may prove the following [19].

Statement 4. In any quantum field theory in which a charged field φ is defined as a local field in a Hilbert space $\mathcal{H}$ equipped with a non-degenerate metric η, the Maxwell's equations

$$\partial_\mu F^{\mu\nu}(f) = j^\nu(f), \qquad \epsilon_{\mu\nu\rho\sigma}\partial^\nu F^{\rho\sigma}(f) = 0$$

cannot hold as operator equation in $\mathcal{H}$. Moreover, if $\mathcal{H}'$ is a linear manifold $\subset \mathcal{H}$, stable under $a^\nu(f) \equiv \partial_\mu F^{\mu\nu}(f) - j^\nu(f)$

$$a^\nu(f)\mathcal{H}' \subset \mathcal{H}'$$

and such that $\eta \geq 0$ in $\mathcal{H}'$ and

$$(\Phi, a^\nu(f)\psi) = 0 \qquad \forall \Phi, \psi \in \mathcal{H}',\ \psi \in D_{\varphi(f)}$$

then η cannot be positive definite in $\mathcal{H}'$ and/or semi-definite in $\mathcal{H}$ unless

$$(\Phi, \varphi(f)\psi) = 0 \qquad \forall \Phi, \psi \in \mathcal{H}'$$

(this means that $\mathcal{H}'$ has zero charge).

The above statement confirms that in any quantum field theory in which charged fields are local, the Hilbert space

must be equipped with an indefinite metric η, and unphysical fields $a^{\nu}(x)$ must be introduced. (In the Gupta-Bleuler formulation $a^{\nu}=\partial^{\nu}\partial_{\mu}A^{\mu}(x)$). Thus any local theory must share all the essential features of the Gupta-Bleuler formulation. It may be worthwhile to remark that Statement 4 is proved under very general assuptions and that nowhere the existence of a field $A_{\mu}(x)$ has been used.

Acknowledgement

I would like to thank Prof. A. S. Wightman for making a Research Associateship available at Princeton University, where most of this work was done, for his encouragement and for useful suggestions.

REFERENCES

1. A. S. Wightman, Phys. Rev. 101, 860 (1956)
A. S. Wightman and L. Garding , Arkiv Fysik 28, 129 (1964)
R. Streater and A. S. Wightman, PCT, Spin and Statistics and All That (Benjamin, New York, 1964)

2. S. N. Gupta, Proc. Phys. Soc. (London) 63, 681 (1950)
K. T. Bleuler, Helv. Phys. Acta, 23, 567 (1950)

3. P. A. M. Dirac, Proc. Roy Soc. (London), A 114, 243 (1927); A 114, 710 (1927); The Principles of Quantum Mechanics, 4th ed. (Oxford U. P., New York, 1958)
J. Schwinger, Phys. Rev. 74, 1439 (1948); 127, 324 (1964)
S. Weinberg, Phys. Rev. 134, B 882 (1964); 138 B 988 (1965)

4. S. N. Gupta, Rev. Mod. Phys. 29, 334 (1957) and in Recent Developments in General Relativity (Pergamon Press, Inc., New York, 1962)

5. We recall that non-degenerate means that there are no eigenvectors of η corresponding to the eigenvalue zero. This is a much weaker statement than requiring $\eta^2=1$, which is assumed in all the physical cases using indefinite metric

6. A. M. Jaffe, Phys. Rev. Letters, 17, 661 (1966); Phys. Rev. 158, 1457 (1967)
F. Costantinescu, Lett. Nuovo Cimento, I1, 849 (1969)

7. An earlier version of this statement (Phys. Rev. 162, 1429 (1967)) was proved under stronger assumptions like the validity of the spectral condition in the whole Hilbert space $\mathcal{H}$, the invariance of the vacuum under the Lorentz group and the Unitarity of $U(\Lambda)$ with respect to η.

8. We recall the definition of regularly separated sets[9]. Let A and B be two closed sets and Λ a compact set of R^n. A and B are said to be regularly separated by Λ if either $\Lambda=\Phi$ and $\overline{A}\cap\overline{B}=\Phi$ or $\exists C>0$ and $\rho>0$ such that $\forall x\in A$, distance $(x,B)\geq C\{\text{distance } (x,\Lambda)\}^{\rho}$.

9. L. Malgrange, Séminaire Schwartz, 4e année, 1959/60, n. 21.
Malgrange's theorem is the following: let A and B be two closed sets $\subset R^n$ regularly separated with respect to the compact set $A \cap B$; then any distribution T with supp $T \subset A \cup B$ can be written as $T = T_1 + T_2$, where supp $T_1 \subset A$, supp $T_2 \subset B$.

10. K. Hepp, Helv. Phys. Acta, 36, 355 (1963)

11. F. Strocchi, Phys. Rev. 162, 1429 (1967), Lemma 1.

12. Ref. 11. Lemma 2.

13. The standard justifications of Gupta-Bleuler formulation are based on many assumptions, some of which may look questionable.

14. An earlier proof of this statement (Phys. Rev. D2, 2334 (1970) made use of the additional assumptions that the fields $A_\mu(x)$ are operator-valued tempered distributions, there is a representation $\{\Lambda\} \to U(\Lambda)$ of the Lorentz group, the operators U are unitary with respect to η, the vacuum is invariant under $U(\Lambda)$, and the spectral condition holds in $\mathcal{H}$.

15. N. N. Bogoliubov and V. S. Vladimirov, Nauchn, Dokl. Vysshei Shkoly 3, 26 (1958).

16. R. Jost, The General Theory of Quantized Fields, (Providence, R.I. 1965)

17. R. F. Streater, Journ. Math. Phys. 3, 256 (1962)

18. F. Strocchi, Phys. Rev., 166, 1302 (1968)
L. Bracci and F. Strocchi, Comm. Math. Phys., in press
L. Bracci, Thesis at Scuola Normale Superiore, Pisa, (1971)

19. For the proof see R. Ferrari, L. E. Picasso and F. Strocchi, to be published.