LECTURES IN THEORETICAL PHYSICS

VOLUME X-A

QUANTUM THEORY AND STATISTICAL PHYSICS

LECTURES IN THEORETICAL PHYSICS

Volume X-A

Quantum Theory and Statistical Physics

Volume X-B

High Energy Physics and Fundamental Particles

LECTURES IN THEORETICAL PHYSICS . 10 . Boulder . 1967

Volume X-A

QUANTUM THEORY AND STATISTICAL PHYSICS

Edited by

ASIM O. BARUT and WESLEY E. BRITTIN

Department of Physics and Astrophysics
University of Colorado

GORDON AND BREACH, SCIENCE PUBLISHERS
New York • London • Paris

Library of Congress Catalog Card Number: 59-13034

Published by GORDON AND BREACH, Science Publishers, Inc.
150 Fifth Avenue, New York, N. Y. 10011

Editorial office for Great Britain:

Gordon and Breach, Science Publishers, Ltd.
8 Bloomsbury Way
London W.C.1

Editorial office for France:

Gordon & Breach
7-9 rue Emile Dubois
Paris 14e

Distributed in Canada by:

The Ryerson Press
299 Queen Street West
Toronto 2B, Ontario

Printed in the United States of America

THE LECTURERS

D. G. Currie, University of Maryland

W. Hunziker, Eidgenössischen Technischen Hochschule, Zürich

A. Isihara, State University of New York

T. F. Jordan, University of Pittsburgh

Bernhard Klaiber, Princeton University

Jean-Paul Marchand, University of Denver

P. T. Matthews, Imperial College, London

Elliott W. Montroll, University of Rochester

George J. Papadopoulos, University of Leeds, England

G. L. Sewell, Queen Mary College, London

Y. Takahashi, Dublin Institute for Advanced Studies, Dublin

Jan Tarski, Université de Paris

John G. Taylor, Mathematical Institute, Oxford

PREFACE

The Proceedings of the Tenth Boulder Summer Institute for Theoretical Physics come in two volumes. Volume XA is devoted to lectures in a wide variety of areas of theoretical physics, from quantum field theory to statistical mechanics, and from group theory to nonlinear differential equations, which were presented during the first seven weeks of the Institute. These lectures review some of the recent advances made in these fields, and we hope that they will appeal to a wide audience of physicists.

The second volume, XB, contains lectures delivered during the Fourth Boulder Conference on Particle Physics. The Conference was held during the last three weeks of the Institute. Traditionally this Conference brings together both experimentalists and theorists to discuss the latest developments in a leisurely and detailed manner.

The Institute was supported in part by a grant from the National Science Foundation and in part by the University of Colorado.

We thank the lecturers and the participants for their collaboration in the final realization of these Volumes, and Mrs. Ann Cofer for her conscientious and expert typing of the manuscripts.

A. O. Barut
Wesley E. Brittin
Editors,

CONTENTS

MATHEMATICAL THEORY OF MULTI-PARTICLE QUANTUM SYSTEMS[†]

W. Hunziker
Seminar für Theoretische Physik der
Eidgenössischen Technischen Hochschule
Zürich

I. Construction of Hamiltonians

Formally, a system of N particles interacting by two-body forces is described by a Hamiltonian

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{\ell} V_\ell(x_\ell) = H_o + V,$$

acting on the Hilbert space $\mathcal{H} = L^2(R^{3N})$ (ℓ labelling the pairs of particles). However, H generates a one-parameter unitary group

$$U(t) = e^{-iHt},$$

which describes the dynamics of the system only if it is self-adjoint. To define a model, therefore, means to construct a self-adjoint Hamiltonian. Once this is achieved, the basic existence theorem is established: the initial value problem connected with the Schrödinger equation has a unique solution.

In many cases, the following theorem is sufficient for the definition of a self-adjoint Hamiltonian:

Theorem 1 (Kato[1]): On a Hilbert space $\mathcal{H}$, let A be self-adjoint and B symmetric and relatively bounded with respect to A, i.e., $D(A) \subset D(B)$[‡] and

$$||Bu|| \leq \alpha||Au|| + \beta||u|| \qquad (I.1)$$

for some $\alpha < 1$, $\beta < \infty$ and all $u \in D(A)$. Then A+B (defined on D(A)) is self-adjoint. Also, if A is bounded below, so is A+B.

[†] Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

[‡] D(A) = domain of the operator A; R(A) = range of A.

*

The proof follows from the relations

$$z - (A+B) = [\,1 - B(z-A)^{-1}](z-A), \tag{I.2}$$

$$\|B(z-A)^{-1}\| \leq \alpha \|A(z-A)^{-1}\| + \beta \|(z-A)^{-1}\|$$

$$= \alpha \sup_{x\in\sigma(A)} \frac{x}{|z-x|} + \beta \sup_{x\in\sigma(A)} \frac{1}{|z-x|},$$

which hold for all (complex) $z \notin \sigma(A)$ = spectrum of A. First, let $z = i\kappa$, κ real. Then

$$\|B(z-A)^{-1}\| \leq \alpha + \frac{\beta}{|\kappa|} < 1$$

for $|\kappa|$ sufficiently large, so that $[1 - B(z-A)^{-1}]$ has range $\mathcal{H}$ and a bounded inverse. This implies

$$R(i\kappa - (A+B)) = R(i\kappa - A) = \mathcal{H},$$

hence, A+B is self-adjoint by standard theorems.[2)]

To prove the second part of the theorem, suppose that A is bounded below by 0. Then, for Re $z < 0$, we have

$$\|B(z-A)^{-1}\| \leq \alpha + \frac{\beta}{|\mathrm{Re}\ z|} < 1 \tag{I.3}$$

for Re z sufficiently large negative. Then, by (I.2), the resolvent $(z - (A+B))^{-1}$ exists, i.e., $z \notin \sigma(A+B)$, which shows that A+B is bounded below.

*

Applying this theorem, we set $A = H_0$, $B = V$. H_0 and V are both defined as multiplication operators: H_0 in momentum space, V in x-space, on their natural domains.

Definition:[3)] V is a Kato-potential if it is infinitely small with respect to H_0, i.e., if for any $\alpha > 0$ there exists $\beta(\alpha) < \infty$ such that (I.1) holds.

Thus defined, the Kato-potentials form a linear set: If V is any finite linear combination of Kato-potentials, then $H_0 + V$ is self-adjoint.

Theorem 2 (Nelson[3]): Let $V(\vec{x}_1 \ldots \vec{x}_N) = U(y_1 \ldots y_m)$, where $y_1 \ldots y_m$ are m linearly independent combinations of the particle-coordinates $\vec{x}_1 \ldots \vec{x}_N$ $(m \leq 3N)$. Then V is a Kato-potential if

$$U(\cdot) \in L^p(R^m) + L^\infty(R^m) \quad \dagger)$$

with

$$2 \leq p \quad \text{and} \quad \frac{m}{2} < p.$$

*

In the case of a two-body potential we have

$$V(\vec{x}_1 \ldots \vec{x}_N) = V_{ik}(\vec{x}_i - \vec{x}_k),$$

hence $m = 3$ and the condition is

$$V_{ik}(\cdot) \in L^2(R^3) + L^\infty(R^3).$$

On the other hand, for a k-body potential with $k > 2$, $m = 3(k-1)$ and the restriction on p is $3/2(k-1) < p$.

*

Singular potentials.

For a two-body potential $V(r) \sim r^{-n}$ $(r \to 0)$, the condition $V \in L^2$ requires $n < 3/2$. This excludes potentials with strong repulsive singularities of the kind employed in statistical mechanics. In such cases, the potential is no longer bounded with respect to H_o. However, a self-adjoint Hamiltonian can still be constructed by the method of Friedrichs-extension.[2] This Hamiltonian then cannot be written as $H_o + V$, since H_o and V are not defined, in general, on all of D(H). For later use, we shall show, however, that $H_o^{\frac{1}{2}}$ (or the momentum operators) is still defined on D(H) and can be estimated in terms of H.

Let $V = V_S + V_R$, where V_R is a Kato-potential and where the singular part V_S satisfies the conditions

$$D(V_S) \cap D(H_o) = D \quad \text{is dense}$$

$$a \leq V_S(\vec{x}_1 \ldots \vec{x}_N) \quad \text{for some real } a.$$

$\dagger)$ This means that $U(y_1 \ldots y_m) = U^p(y_1 \ldots y_m) + U^\infty(y_1 \ldots y_m)$ a.e., with $U^p(\cdot) \in L^p(R^m)$ and $U^\infty(\cdot) \in L^\infty(R^m)$.

Since V_R is a Kato-potential, $\frac{1}{2} H_o + V_R$ is bounded below, which implies

$$0 \leq \tfrac{1}{2} (u, H_o u) + (u, u) \leq \left(u, (H_o + V_R + b)u\right)$$

for some real b and all $u \in D$. Adding on the right the non-negative form $(u, V_S u) - a(u,u)$, we obtain

$$0 \leq \tfrac{1}{2} (u, H_o u) + (u,u) \leq (u, (H' + c)u) \tag{I.4}$$

for all $u \in D$, where $c = b - a$ and $H' = H_o + V$ (on D). Thus we can define:

$$H + c = \text{Friedrichs-extension of } H' + c.$$

The inequality (I.4) is then preserved in the following sense:

$$D\left((H+c)^{\frac{1}{2}}\right) \subset D\left(H_o^{\frac{1}{2}}\right)$$

and

$$\tfrac{1}{2} \| H_o^{\frac{1}{2}} u \|^2 + \| u \|^2 \leq \| (H + c)^{\frac{1}{2}} u \|^2 \tag{I.5}$$

for all $u \in D(H+c)^{\frac{1}{2}})$.

*

The question of how to treat attractive singular potentials may not be relevant to physics; it is, however, quite interesting from the point of view of the correspondence principle.

If $H_o + V$ is still densely defined, it is symmetric and real, and has therefore always a (generally not unique) self-adjoint extension. But Nelson[3] has shown that the "correct" generator of the time-translations U(t) need not be among these extensions. He treats the case $N = 2$ with

$$V(r) = -\frac{g}{r^2} \qquad (g > 0),$$

which classically leads to capture for sufficiently small angular momentum (collapse of the pair after a finite time). He constructs the propagator $\exp(-iHt) = U(t)$ by using Feynman-integrals, and shows that in an (invariant) subspace of sufficiently small angular momentum U(t) is no longer a unitary group but a contraction-semigroup. This means that probability is dissipated—corresponding to the

classical capture process. Therefore, the Hamiltonian which gives the correct classical limit is not self-adjoint.

II. Cluster-Properties

Maybe the most characteristic feature of multiparticle systems is that they can break up into parts which move independently, if they are far separated from each other. It is clear, therefore, that the dynamics of an N-particle system contains, as a certain limit, the dynamics of any subsystem of less than N particles.

To formulate this, let $D=(C_1 \ldots C_n)$ be a partition of the set $(1 \ldots N)$ into n "clusters" $C_1 \ldots C_n$. The operator $U_D(a_1 \ldots a_n)$ which translates each cluster C_k by $a_k=(a_k^o, \vec{a}_k)$ in space-time is

$$U_D(a_1 \ldots a_n) = \prod_{k=1}^{n} e^{-iH_k a_k^o + i\vec{P}_k \vec{a}_k} , \qquad (II.1)$$

$H_k, \vec{P}_k$ being the energy and momentum of the subsystem C_k. Let $U(a)$ be the corresponding operator for the system as a whole (trivial partition into one cluster). Then the cluster property is expressed by the following theorem:

Theorem 3:[4] If all the pair-potentials V_ℓ are of the form

$$V_\ell(\cdot) \in L^2(R^3) + L^p(R^3) \qquad (II.2)$$

with $2 \leq p < 3$, then

$$U_D(-a_1 \ldots -a_n) U(a) U_D(a_1 \ldots a_n)\psi \longrightarrow U_D(a \ldots a)\psi$$

as

$$\min_{i \neq k} |a_i - a_k| \to \infty$$

($|a|$ = Euclidean distance in R^4), for all $\psi \in \mathcal{H}$ and uniformly in a.

*

Remarks.

First of all, since $U(o,\vec{a}) = U_D(o,\vec{a};\ldots o,\vec{a})$ commutes with $U_D(a_1 \ldots a_n)$, we can restrict ourselves to pure time-translations where the theorem reads

$$e^{-iHt} U_D(a_1 \ldots a_n)\psi \longrightarrow e^{-iH_D t} U_D(a_1 \ldots a_n)\psi ,$$

with

$$H_D = \sum_{k=1}^{n} H_k = H \text{ minus all interactions linking different clusters.}$$

In other words, if, in any initial state ψ, we separate the clusters (by applying $U_D(a_1 \ldots a_n)$) sufficiently far in space-time, then the subsequent motion will be arbitrarily close to the motion of the system of non-interacting clusters (with the same initial state) uniformly in t.

This uniformity in t could not be expected from classical mechanics: consider two particles A,B, where A is at rest and B has an initial position $\vec{x}$ and an initial velocity $\vec{v}$ which will make it collide with A. If we separate A and B by shifting B from $\vec{x}$ to $\vec{x} - \vec{v}a$ $(a > 0)$, the two particles will still collide after a finite time, regardless of how large we choose a. In quantum mechanics, however, the probability of a collision will decrease for increasing a, due to the spreading of the wave-packets.

Proof of Theorem 3: It suffices to prove the theorem on the dense set spanned by the states of the form

$$\psi(\vec{x}_1 \ldots \vec{x}_N) = \prod_{k=1}^{n} e^{-\frac{1}{2}(\vec{y}_k - \vec{b}_k)^2} \varphi_k(z_k),$$

where $\vec{b}_k$ varies over R^3 and φ_k over $D(h_k)$, h_k being the internal energy of C_k, and where $\vec{y}_k$ and z_k denote the coordinates of the center-of-mass of C_k and internal coordinates of C_k, respectively. Now we have

$$\left(e^{-iHt} - e^{-iH_D t}\right) U_D(a_1 \ldots a_n)\psi$$

$$= -ie^{-iHt} \int_0^t d\tau \, e^{iH\tau} I_D e^{-iH_D\tau} U_D(a_1 \ldots a_n)\psi,$$

where $I_D = H - H_D$ is the sum of all interactions linking different clusters. (Since $\psi \in D(H_o)$, $\exp(-iH_D t)\, U_D(a_1 \ldots a_n)\psi \in D(H_o)$, so that, by Theorem 1, the integrand is well-defined and continuous in t. The integral is therefore defined as a Riemann-integral.) Therefore,

$$\left\| \left(e^{-iHt} - e^{-iH_D t} \right) U_D(a_1 \ldots a_n)\psi \right\| \leq \int_0^t d\tau \; N(a_1 \ldots a_n, \tau), \qquad \text{(II.3)}$$

where

$$N(a_1 \ldots a_n, \tau) = \| I_D e^{-iH_D \tau} U_D(a_1 \ldots a_n)\psi \| .$$

An elementary calculation of the propagation of free wave-packets yields

$$\left| \left(e^{-iH_D \tau} U_D(a_1 \ldots a_n)\psi \right) (\vec{x}_1 \ldots \vec{x}_N) \right|^2$$

$$= \prod_{k=1}^{n} \mu_k^{3/2} e^{-\mu_k (\vec{y}_k + \vec{a}_k - \vec{b}_k)^2} \left| \varphi_k(\tau + a_k^o, z_k) \right|^2 ,$$

where

$$\mu_k = \left(1 + (\tau_k + a_k^o)^2 / M_k^2 \right)^{-1} ,$$

M_k = total mass of C_k, and $\varphi_k(t, \cdot) = \exp(-ih_k t)\varphi_k$. For notational convenience we now set $\vec{b}_k = 0$ and we only estimate the contribution to $N(a_1 \ldots a_n, \tau)$ of a pair-potential V linking the clusters C_1 and C_2. Then we obtain

$$N^2(a_1 \ldots a_n, \tau) = \text{const.} \int dz_1 dz_2 \, dx \, dy \left| \varphi_1(\tau + a_1^o, z_1)\varphi_2(\tau + a_2^o, z_2) \right|^2$$

$$\times \left| V(x) \right|^2 (\mu_1 \mu_2)^{3/2} e^{-\mu_1 (\vec{y}_1 - \vec{a}_1)^2 - \mu_2 (\vec{y}_2 - \vec{a}_2)^2} ,$$

where x is the relative coordinate of the pair linked by V and y is the coordinate of the center-of-mass of the subsystem (C_1, C_2), so that

$$\vec{y}_1 = \vec{y} + \alpha \vec{x} + \overrightarrow{\beta_1 z_1}$$

$$\vec{y}_2 = \vec{y} + (\alpha - 1)\vec{x} + \overrightarrow{\beta_2 z_2} ,$$

α, β_1, β_2 denoting constants depending on the masses. Carrying out the y-integration, we are left with

$$N^2(a_1 \ldots a_n, \tau) = \text{const.} \int dz_1 dz_2 dx |\varphi_1(\tau + a_1^\circ, z_1)|^2 |\varphi_2(\tau + a_2^\circ, z_2)|^2$$

$$\times |V(x)|^2 \mu^{3/2} e^{-\mu(\vec{x} - (\vec{a}_1 - \vec{a}_2) + \overrightarrow{\beta_1 z_1} + \overrightarrow{\beta_2 z_2})^2}, \qquad \text{(II.4)}$$

with

$$\mu^{-1} = \mu_1^{-1} + \mu_2^{-1} = \left(2 + (\tau + a_1^\circ)^2 / M_1^2 + (\tau + a_2^\circ)^2 / M_2^2\right).$$

First, let us assume now that $V(\cdot) \in L^2(R^3)$. Estimating the exponential by 1, we then obtain

$$N(a_1 \ldots a_n, \tau) \leq \mu^{3/4} \cdot \text{const.} \, \|\varphi_1\| \, \|\varphi_2\| \, \|V(\cdot)\|_2 . \qquad \text{(II.5)}$$

Hence,

$$\int_{-\infty}^{+\infty} N(a_1 \ldots a_n, \tau) d\tau \leq \text{const.} \int_{-\infty}^{+\infty} \left(2 + \left(\frac{\tau + a_1^\circ}{M_1}\right)^2 + \left(\frac{\tau + a_2^\circ}{M_2}\right)^2\right)^{-3/4}$$

$$\to 0 \quad \text{for} \quad |a_1^\circ - a_2^\circ| \to \infty .$$

This, together with (II.3), proves the cluster-property for purely time-like separations. On the other hand, for purely space-like separation, it suffices to prove $N(a_1 \ldots a_n, \tau) \to 0$ for $|\vec{a}_1 - \vec{a}_2| \to \infty$ ($a_1^\circ, a_2^\circ, \tau$ fixed), by (II.5) and by the dominated convergence theorem.[2] But the integrand in (II.4) is bounded by

$$|\varphi_1(\tau + a_1^\circ, z_1)|^2 |\varphi_2(\tau + a_2^\circ, z_2)|^2 |V(x)|^2$$

uniformly in $\vec{a}_1, \vec{a}_2$, and vanishes (pointwise) as $|\vec{a}_1 - \vec{a}_2| \to \infty$. Hence the integral (II.4) vanishes in this limit, again by the dominated convergence theorem. A more elaborate treatment shows, in fact, that

$$\int_{-\infty}^{+\infty} N(a_1 \ldots a_n, \tau) d\tau \to 0$$

if only the Euclidean distance $|a_1 - a_2|$ in R^4 tends to ∞.

If $V(\cdot) \in L^2(R^3) + L^p(R^3)$, $2 \leq p < 3$, we can choose the L^p-part V_p of $V(\cdot)$ arbitrarily small in the L^p-norm. Let N_p be the contribution to N, in (II.4), of V_p. Then it suffices to show that, for any $\varepsilon > 0$,

$$\int_{-\infty}^{+\infty} N_p(a_1 \ldots a_n, \tau) d\tau < \varepsilon \qquad \text{(II.6)}$$

for all $a_1 \ldots a_n$, provided that $\| V_p(\cdot) \|_p$ is sufficiently small. To get this, we apply the Hoelder-inequality to the x-integral in (II.4), obtaining

$$N_p^2(a_1 \ldots a_n, \tau) \leq \mu^{3/2} \text{ const. } \|\varphi_1\|^2 \|\varphi_2\|^2$$

$$\times \| V_p(\cdot) \|_p^2 \left(\int_0^\infty dr\, r^2\, e^{-\mu q r^2} \right)^{1/q},$$

with $q^{-1} + 2p^{-1} = 1$ $(2 \leq p)$, or, since the last integral is proportional to $\mu^{-3/2q}$,

$$N_p^2(a_1 \ldots a_n, \tau) \leq \text{ const. } \mu^{3/p} \| V_p(\cdot) \|_p^2 ,$$

which, for $p < 3$, implies (II.6).

*

Cluster-properties in momentum space.

The operator which gives each cluster C_k an additional momentum $\vec{a}_k$ is defined by

$$\left(U_D(\vec{a}_1 \ldots \vec{a}_n)\psi \right)(\vec{x}_1 \ldots \vec{x}_N) = \prod_{k=1}^{n} e^{i\vec{a}_k \vec{y}_k} \psi(\vec{x}_1 \ldots \vec{x}_N),$$

$\vec{y}_k$ being again the coordinate of the center-of-mass of C_k. Separating the clusters in momentum space means that they are given large relative velocities, and, again, the clusters become dynamically independent as the separation goes to ∞:

Theorem 4: Under the hypothesis of Theorem 3,

$$U_D(-\vec{a}_1 \ldots -\vec{a}_n) e^{-iHt} U_D(\vec{a}_1 \ldots \vec{a}_n)\psi \to e^{-iH_D t}\psi$$

as

$$\min_{i \neq k} |\vec{a}_i - \vec{a}_k| \to \infty ,$$

for any $\psi \in \mathcal{H}$ and uniformly in $-\infty < t < +\infty$.

Proof: Proceeding as in the proof of Theorem 3, we obtain

$$N^2(\vec{a}_1 \ldots \vec{a}_n, \tau) = \text{const.} \int dz_1 dz_2 dx \, |\varphi_1(\tau, z_1)|^2 \, |\varphi_2(\tau, z_2)|^2$$

$$\times \, |V(x)|^2 \mu^{3/2} e^{-\mu(\vec{x} - \tau((\vec{a}_1/M_1) - (\vec{a}_2/M_2)) + \overrightarrow{\beta_1 z_1} + \overrightarrow{\beta_2 z_2})^2}$$

where μ is obtained from the corresponding μ in (II.5) by setting $a_1^o = a_2^o = 0$. As before, we conclude that

$$\int_{-\infty}^{+\infty} N(\vec{a}_1 \ldots \vec{a}_n, \tau) d\tau \to 0$$

as $|\vec{a}_i - \vec{a}_k| \to \infty$, by twice applying the dominated convergence theorem.

*

The rate of convergence in the space-like cluster-properties.

The cluster-properties established so far were based on the decay of wave-packets, and, as pointed out before, this is why we obtained results which were stronger than what could be expected from classical mechanics. However, classical mechanics will show up in the rate of convergence to the various cluster-limits. This rate will depend on the directions in which we separate the clusters, and the directions of fast convergence will depend on the initial momenta of the clusters. Also, fast convergence cannot be expected for any initial state ψ, but only for states which describe well-localized clusters.

Localized states.[5]

The problem is to find a set of states, describing well-localized particles, which is invariant under time-translations. Since the Schrödinger equation is parabolic, it is useless to define localization in terms of the support of wave-functions. What we may use, instead, are the expectation values of arbitrary monomials in the coordinates of the particles (i.e., moments of the probability-distribution in x-space). In order to see what can be expected, we first look at the case of a free particle in one dimension. There we find

$$\| x^n e^{-ip^2 t/2m} \psi \| = \| e^{-ip^2 t/2m} \left(x + \frac{pt}{m} \right)^2 \psi \|$$

$$= \| \left(x + \frac{pt}{m} \right)^n \psi \| .$$

Therefore, a suitable set of localized states for this case is

$$D_n = \bigcap_{k+\ell \le n} D(x^k p^\ell). \tag{II.7}$$

D_n is invariant under $\exp(-i(p^2t/2m))$, and the expectation value of x^k for $0 \le k \le 2n$ in a state $\psi \in D_n$ exists and does not increase faster than const.$|t|^k$ as $|t| \to \infty$.

For the case of interacting particles, it turns out that a possible generalization of (II.7) is

$$D_n = \bigcap_{|k|+\ell \le n} D(x^k H^\ell), \tag{II.8}$$

where

$$x^k = x_1^{k_1} \cdot x_2^{k_2} \cdots x_N^{k_N}$$

and $|k| = k_1 + k_2 + \ldots + k_N$, k_i integer ≥ 0. On D_n, we introduce the norm

$$\|\psi\|_n = \sum_{|k|+\ell \le n} \|x^k H^\ell \psi\|. \tag{II.9}$$

Since x^k, H^ℓ are closed, D_n, normed by $\| \ \|_n$, is complete. The justification for calling the states $\psi \in D_n$ localized is the following theorem:

Theorem 5:[5] Let

$$H = \sum_{k=1}^{N} \frac{p_k^2}{2m_k} + V(\vec{x}_1 \ldots \vec{x}_N),$$

where V is a Kato-potential. Then

a) D_n is invariant under the unitary group e^{-iHt}.

b) For any $\psi \in D_n$, $e^{-iHt}\psi$ is continuous in t in the of the norm $\| \ \|_n$, and there exists a constant c_n such that

$$\|e^{-iHt}\psi\|_n \le c_n(1 + |t|)^n \|\psi\|_n.$$

*

An asymptotic expansion:

The second tool which is needed to estimate the rate of convergence to the cluster-limits is an asymptotic expansion for integrals of the type

$$g(t) = \int d^3p\, e^{-ip^2t/2} f(\vec{p})$$

in inverse powers of t. Such an expansion can be found by successive partial integrations:

$$g(t) = -4\psi i \sum_{k=1}^{n} t^{-(k+\frac{1}{2})} F_{2k-1}(0)\,(\Delta^{k-1}f)(0)$$
$$- it^{-(n+\frac{1}{2})} \int d^3p\, \frac{F_{2n-1}(p\sqrt{t})}{p}\,(\Delta^n f)(\vec{p}), \qquad \text{(II.10)}$$

where $F_n(r)$ is the multiple Fresnel-integral

$$F_n(r) = \int_r^\infty dr_1 \int_{r_1}^\infty dr_2 \cdots \int_{r_{n-1}}^\infty dr_n\, e^{-ir_n^2/2}$$

which has the properties

$$F_{2k-1}(0) = \frac{(-i)^{k-\frac{1}{2}}(k-3/2)!}{\sqrt{2}\,(k-1)!}$$

$$|F_n(r)| \leq c_n(1+r)^{-n} \qquad (r \geq 0)$$

for some constant c_n. This expansion holds for $t \geq 0$; the case $t<0$ is easily obtained by complex conjugation. The first term is the usual asymptotic expression obtained by the method of stationary phase and exhibits the familiar $t^{-3/2}$-decay.

*

With Theorem 5 and with the expansion (II.10) as the essential tools, one can now derive the following estimates for the rate of convergence in the spatial cluster-properties:

Theorem 6: Assumptions:

(1) For any pair ℓ and any non-negative integer n, let

$$\int d^3x |\vec{x}|^n |V_\ell(\vec{x})|^2 < \infty .$$

(2) Let $D = (C_1 \ldots C_n)$ be a partition of $(1 \ldots N)$ into n clusters, and choose ψ of the form

$$\psi(\vec{x}_1 \ldots \vec{x}_N) = \prod_{k=1}^{n} \phi_k(\vec{r}_k)\varphi_k(z_k)$$

with

$$\tilde{\phi}_k \in C_o^\infty(R^3) \quad (\tilde{\phi}_k = \text{Fourier transform of } \phi_k)$$

and

$$\varphi_k \in D_\infty(C_k) \equiv \bigcap_{n=0}^{\infty} D_n(C_k),$$

where $D_n(C_k)$ is the set of localized states (II.8) for the internal degrees of freedom of C_k.

(3) Let $\vec{a}_1 \ldots \vec{a}_n$ be such that if the supports of the functions $\tilde{\phi}_k$ (considered as free rigid bodies) start to move with velocities $\vec{a}_k/M_k$ (M_k = mass of C_k), they will never collide in the future. (Note that this implies that the $\tilde{\phi}_k$ have non-overlapping supports.)

Then

$$U_D(-\lambda\vec{a}_1, \ldots -\lambda\vec{a}_n)e^{-iHt}U_D(\lambda\vec{a}_1 \ldots \lambda\vec{a}_n)\psi \to e^{-iH_D t}\psi$$

as $\lambda \to +\infty$, faster than any inverse power of λ and uniformly in $0 \le t < +\infty$.

*

III. Time-Dependent Scattering Theory

For any decomposition $D = (C_1 \ldots C_n)$,

$$e^{-iH_D t}\psi \tag{III.1}$$

describes a motion of non-interacting clusters $C_1 \ldots C_n$. The basis of scattering theory is the fact that any such motion uniquely characterizes a motion of the fully interacting N-particle system, in the sense that the two motions become asymptotically the same as $t \to +\infty$ (or $t \to -\infty$). Symbolically: to every "asymptote" (meaning a motion of non-interacting clusters) there exists a unique "orbit" (meaning a motion of the full N-particle system) having this asymptote for $t \to +\infty$ (or for $t \to -\infty$). This is the content of the following theorem:

Theorem 7 (Hack[6]): Under the hypothesis of Theorem 3, the strong limits

$$\Omega_D^{\pm} = \lim_{t \to \pm\infty} e^{iHt} e^{-iH_D t} \tag{III.2}$$

exist on all of $\mathcal{H}$.

Corollary: The operators $\Omega_D^{\pm}$ are isometric and satisfy

$$e^{-iHt} \Omega_D^{\pm} = \Omega_D^{\pm} e^{-iH_D t} \quad (-\infty < t < +\infty).$$

Therefore, the ranges $R_D^{\pm}$ of $\Omega_D^{\pm}$ reduce the unitary group $\exp(-iHt)$ and the parts of H in $R_D^{\pm}$ are unitarily equivalent to H_D.

*

Proof: It suffices to prove convergence on the dense set of states employed in the proof of Theorem 3. Then we have

$$\Omega_D^{\pm}\psi = \psi + i \int_0^{\pm\infty} d\tau \, e^{iH\tau} I_D e^{-iH_D\tau} \psi,$$

provided that the integral converges. But this follows from the estimates of Section II, by which

$$\| I_D \exp(-iH_D\tau)\psi \| = N(0,\ldots,0,\tau)$$

is integrable over $-\infty < \tau < +\infty$. The corollary is obtained by going to the limit $s \to \pm\infty$ in

$$e^{-iHt} e^{iHs} e^{-iH_D s} \psi = e^{iH(s-t)} e^{-iH_D(s-t)} e^{-iH_D t} \psi.$$

*

Usually, in scattering theory, the asymptotic behavior of the orbits is characterized more precisely than by (III.1): One is interested only in the case where the subsystems $(C_1 \ldots C_n)$ are "fragments," i.e., either single particles or composite subsystems possessing at least one bound state. A channel, α, then specifies a decomposition into fragments $F_1 \ldots F_n$ and assigns a definite bound state φ_k to each composite fragment F_k:

$$\alpha = \begin{pmatrix} F_1 \dots F_n \\ \varphi_1 \dots \varphi_n \end{pmatrix},$$

($\varphi_k = 1$ if F_k = single particle). By selecting a maximal set of mutually orthogonal bound states for each possible fragment, we obtain a countable set of channels. The states in channel α are defined as

$$\psi_\alpha(\vec{x}_1 \dots \vec{x}_N) = \phi_\alpha(\vec{y}_1 \dots \vec{y}_n) \prod_{k=1}^{k} \varphi_k(z_k), \tag{III.3}$$

where $\vec{y}_k$, z_k are the coordinates of the center-of-mass of F_k and internal coordinates of F_k, φ_k = bound state of F_k specified by α, and $\phi_\alpha(\cdot) \in L^2(R^{3n})$ is arbitrary. It is convenient to normalize the bound states such that

$$\|\psi_\alpha\|^2 = \int dy_1 \dots dy_n |\phi_\alpha(\vec{y}_1 \dots \vec{y}_n)|^2, \tag{III.4}$$

so that the space D_α of channel states (III.4) is essentially $L^2(R^{3n})$. On D_α, H_D reduces to

$$H_\alpha = \sum_{k=1}^{n} \left(\frac{P_k^2}{2M_k} + \epsilon_k \right) \tag{III.5}$$

where $\vec{P}_k$, M_k, ϵ_k are the total momentum, the total mass and the bound-state energy of F_k. The system is now said to be in channel α for $t \to \pm\infty$ if its orbit has an asymptote

$$e^{-iH_\alpha t} \psi_\alpha, \qquad \psi_\alpha \in D_\alpha \tag{III.6}$$

for $t \to \pm\infty$. Equation (III.6) simply describes a free motion of n particles $F_1 \dots F_n$ with masses $M_1 \dots M_n$ and internal energies $\epsilon_1 \dots \epsilon_n$, and Theorem 7 states that any such asymptote uniquely defines an orbit having this asymptote for $t \to +\infty$ or $t \to -\infty$, namely

$$e^{-iHt} \Omega_\alpha^{\pm} \psi_\alpha \to e^{-iH_\alpha t} \psi_\alpha \qquad (t \to \pm\infty),$$

where $\Omega_\alpha^{\pm}$ is the restriction to D_α of $\Omega_D^{\pm}$, $D = (F_1 \dots F_n)$.

The remaining part of the scattering formalism is based on the fact that orbits which are in different channels as $t \to +\infty$ (or $t \to -\infty$) are orthogonal:

$$R_\alpha^\pm \perp R_\beta^\pm \quad \text{if} \quad \alpha \neq \beta, \tag{III.7}$$

where $R_\alpha^\pm$ is the range of $\Omega_\alpha^\pm$ or, equivalently,

$$\lim_{|t|\to\infty} \left(e^{-iH_\alpha t}\psi_\alpha , e^{-iH_\beta t}\psi_\beta \right) = 0 \tag{III.8}$$

for $\alpha \neq \beta$ and all $\psi_\alpha \in D_\alpha$, $\psi_\beta \in D_\beta$.

Proof: First, consider the case where the fragments in the channels α, β are the same, so that α differs from β only in the assignment of bound states to these fragments. By the orthogonality of bound states, we then have $D_\alpha \perp D_\beta$ and the scalar product in (III.8) vanishes identically in t. If the fragments in the channels α, β are not the same, it is convenient to define $H_{\alpha,\beta}$ by (III.5) on all of $\mathcal{H}$. Up to a constant, $H_\alpha - H_\beta$ is then a quadratic form of the particle-momenta $\vec{p}_1 \ldots \vec{p}_N$ which does not vanish identically. By a linear transformation $(\vec{p}_1 \ldots \vec{p}_N) \to (\vec{\pi}_1 \ldots \vec{\pi}_N)$, we can diagonalite this form:

$$(H_\alpha - H_\beta)(\vec{p}_1 \ldots \vec{p}_N) = \sum_{i=1}^{N} \lambda_i \pi_i^2 + \text{const.},$$

with $\lambda_1 \neq 0$. It is obviously sufficient to establish (III.8) for ψ_α, ψ_β in a dense set of $\mathcal{H}$. Such a set is spanned by the states

$$\psi(\vec{\pi}_1 \ldots \vec{\pi}_N) = \prod_{i=1}^{N} \psi^{(k)}(\vec{\pi}_k), \qquad \psi^{(k)} \in \mathcal{S}(R^3),$$

and for ψ_α, ψ_β of this form we obtain

$$\left| \left(e^{-iH_\alpha t}\psi_\alpha , e^{-iH_\beta t}\psi_\beta \right) \right|$$

$$\leq \text{const.} \left| \int d\pi_1 e^{-i\lambda_1 \pi_1^2 t} \overline{\psi_\alpha^{(1)}(\vec{\pi}_1)} \psi_\beta^{(1)}(\vec{\pi}_1) \right|,$$

which vanishes like $|t|^{-3/2}$ as $|t| \to \infty$, as can be seen from the asymptotic expansion given in Section II.

*

To define the S-operator, we now introduce the Hilbert space

$$\mathcal{H}' = \bigoplus_{\alpha} D_{\alpha} . \tag{III.9}$$

Note that, in general, $\mathcal{H}'$ cannot be viewed as a subspace of $\mathcal{H}$, since the D_{α} are not mutually orthogonal subspaces of $\mathcal{H}$. We define total wave-operators $\Omega^{\pm}$ mapping $\mathcal{H}'$ into $\mathcal{H}$ by

$$\Omega^{\pm}\psi = \sum_{\alpha} \Omega_{\alpha}^{\pm}\psi_{\alpha} , \tag{III.10}$$

where $\psi \in \mathcal{H}'$ and ψ_{α} is the component of ψ in D_{α}. Since the $\Omega_{\alpha}^{\pm}$ are isometric from D_{α} into $\mathcal{H}$, with mutually orthogonal ranges, $\Omega^{\pm}$ is isometric from $\mathcal{H}'$ into $\mathcal{H}$. Its adjoint, defined by

$$(\varphi, \Omega^{\pm}\psi)_{\mathcal{H}} = (\Omega^{\pm *}\varphi, \psi)_{\mathcal{H}'} ,$$

is therefore an operator mapping $\mathcal{H}$ onto $\mathcal{H}'$, characterized by

$$\Omega^{\pm *}\varphi = (\Omega^{\pm})^{-1}\varphi \quad \text{if} \quad \varphi \in R^{\pm} ,$$

$$\Omega^{\pm *}\varphi = 0 \quad \text{if} \quad \varphi \perp R^{\pm} ,$$

where

$$R^{\pm} = \bigoplus_{\alpha} R_{\alpha}^{\pm}$$

is the range (in $\mathcal{H}$) of $\Omega^{\pm}$. Two S-operators can now be defined by

$S = \Omega^{+}\Omega^{-*}$ = operator mapping $\mathcal{H}$ into $\mathcal{H}$, unitary if and only if $R^{+} = R^{-} = \mathcal{H}$,

$S' = \Omega^{+*}\Omega^{-}$ = operator mapping $\mathcal{H}'$ into $\mathcal{H}'$, unitary if and only if $R^{+} = R^{-}$.

S is the S-operator introduced by Jauch,[7] while $\mathcal{H}'$ and S' have first been introduced by Berezin, Faddeev and Minlos.[8] S' has a simple interpretation: each element $\psi = \{\psi_{\alpha}\} \in \mathcal{H}'$ defines an asymptote

$$\sum_{\alpha} e^{-iH_{\alpha}t}\psi_{\alpha} ,$$

i.e., a superposition of freely moving fragments. By Theorem 7 there

exists a unique orbit having this asymptote for $t \to -\infty$, namely,

$$e^{-iHt}\Omega^{-}\psi \to \sum_{\alpha} e^{-iH_\alpha t}\psi_\alpha \qquad (t \to -\infty).$$

If $\Omega^{-}\psi \in R^{+}$ (which is the case if $R^{+}=R^{-}$), this orbit also has an asymptote as $t \to +\infty$:

$$e^{-iHt}\Omega^{-}\psi \to \sum_{\alpha} e^{-iH_\alpha t}\psi'_\alpha \qquad (t \to +\infty),$$

where ψ'_α is the component in D_α of

$$\psi' = S'\psi .$$

Therefore, S' gives the asymptotic behaviour in the future in terms of the asymptotic behaviour of the past—and this is precisely what we observe in scattering experiments. Indeed, from the point of view of pure S-matrix theory, $\mathcal{H}'$ is the Hilbert space of the system and S' the operator characterizing the system, while $\mathcal{H}$, H and the Schrödinger equation are considered merely as tools for the construction of S' which one tries to replace by something else.

The conditions $R^{+}=R^{-}$ and $R^{+}=R^{-}=\mathcal{H}$, which are equivalent to the unitarity of S' and S, respectively, are the conditions of asymptotic completeness. In our symbolic language, $R^{+}=R^{-}$ means that every orbit which has an asymptote for $t \to -\infty$ also has one for $t \to +\infty$, and vice versa, while $R^{+}=R^{-}=\mathcal{H}$ means that every orbit has two asymptotes: one for $t \to -\infty$ and one for $t \to +\infty$. It is clear, therefore, that the unitarity of the S-matrix expresses nontrivial dynamical properties of the system and does not follow, as is claimed sometimes, simply from the conservation of probability. In fact, there exists so far no general proof of asymptotic completeness for nonrelativistic multiparticle systems.

*

Roughly speaking, Theorem 7 requires that any pair-potential V(r) is less singular than $r^{-3/2}$ as $r \to 0$ and decreases faster than r^{-1} at infinity. It can be generalized, however, to cover also the case of potentials with strong repulsive singularities and the case of Coulomb-potentials.

Singular potentials.†

Kupsch and Sandhas[9] have shown how to treat singular potentials of arbitrary sign in the case $N=2$. In order to extend their

†An extended version of this section will be published in Helvetica Physica Acta.

method to $N > 2$, we have to restrict ourselves to the case where the strong singularities of the potentials are repulsive.

Theorem 8: For any pair ℓ, let $V_\ell = V_{\ell,R} + V_{\ell,S}$, where the "regular" part $V_{\ell,R}$ satisfies the assumptions of Theorem 3 while the "singular" part $V_{\ell,S}$ satisfies the conditions imposed on singular potentials in Section I and, in addition, has compact support (in R^3). Then the assertion of Theorem 7 remains valid.

Proof: Let R be such that $V_{\ell,S}(\vec{x}) = 0$ for $|\vec{x}| > R$ and all ℓ. Introduce a cut-off $F(r)$ with the properties

$$F(r) \in C^\infty[0,\infty)$$

$$0 \le F(r) \le 1 \qquad (0 \le r < \infty)$$

$$F(r) = 0 \quad \text{for} \quad r \le R, \quad F(r) = 1 \quad \text{for} \quad r \ge R+1.$$

Denoting with F_ℓ the multiplication operator $F(|\vec{x}_\ell|)$, we then have

$$e^{iHt} e^{-iH_D t}\psi = e^{iHt} \prod_\ell F_\ell e^{-iH_D t}\psi + e^{iHt}\Big(1 - \prod_\ell F_\ell\Big) e^{-iH_D t}\psi, \tag{III.11}$$

where ℓ runs over all pairs linking different clusters of D. $0 \le F_\ell \le 1$ implies

$$0 \le 1 - \prod_\ell F_\ell \le \sum_\ell (1 - F_\ell),$$

and from this we obtain the following estimate for the second term on the right-hand side of (III.11):

$$\Big\| e^{iHt}\Big(1 - \prod_\ell F_\ell\Big) e^{-iH_D t}\psi \Big\| \le \sum_\ell \| (1 - F_\ell) e^{-iH_D t}\psi \| \to 0$$

for $|t| \to \infty$, since $1 - F_\ell \in C_0^\infty(R^3)$ and therefore acts like a decent two-body potential in the non-singular case. On the other hand, we can write the first term in (III.11) as

$$e^{iHt}\prod_{\ell} F_\ell\, e^{-iH_D t}\psi$$

$$= \prod_{\ell} F_\ell\psi + i\int_0^t d\tau\, e^{iH\tau}\Big(H\prod_{\ell} F_\ell - \prod_{\ell} F_\ell H_D\Big) e^{-iH_D\tau}\psi\,, \tag{III.12}$$

and

$$\Big(H\prod_{\ell} F_\ell - \prod_{\ell} F_\ell H_D\Big) = \Big[H_o, \prod_{\ell} F_\ell\Big] + \sum_{\ell'} V_{\ell'}\prod_{\ell} F_\ell\,, \tag{III.13}$$

where ℓ, ℓ' both run over all pairs linking different clusters. In order to show that the integral in (III.12) converges as $|t|\to\infty$, we estimate separately the contributions from the various terms in (III.13): Since

$$|V_{\ell'}\prod_{\ell} F_\ell| \le |V_{\ell'}F_{\ell'}| \le |V_{\ell'}|_R\,,$$

the second term in (III.13) gives a convergent contribution to the integral, as in the non-singular case. On the other hand,

$$\Big[H_o, \prod_{\ell} F_\ell\Big]$$

is a sum of terms of the form

$$\prod_{\ell'} G_{\ell'}(\vec{x}_{\ell'})A$$

where $A=1$ or $A=$ total momentum of one of the clusters or $A=$ internal momentum in one of the clusters, and where at least one $G_{\ell'}$, say G_ℓ, has compact support (being a derivative of F_ℓ). In the first two cases, A commutes with H_D, hence

$$\Big\|\prod_{\ell'} G_{\ell'}A\, e^{-iH_D\tau}\psi\Big\| \le \text{const.}\,\big\|G_\ell\, e^{-iH_D\tau}A\psi\big\|,$$

which is integrable over $-\infty<\tau<+\infty$, since again G_ℓ acts like a non-singular two-body potential. Finally, we have to deal with the case where A is an internal momentum of one of the clusters, say of C_1. Instead of (II.5), we then obtain the estimate

$$N(0\ldots0,\tau) \le \mu^{3/4}\,\text{const.}\,\big\|A\,e^{-ih_1\tau}\varphi_1\big\|\;\|\varphi_2\|\;\|G_\ell(\cdot)\|_2\,,$$

where $\mu(\tau)^{3/4}$ is integrable over $-\infty < \tau < +\infty$. It suffices to show, therefore, that $\| A \exp(-ih_1\tau)\varphi_1 \|$ is uniformly bounded in τ. But this follows by applying (I.5) to the internal Hamiltonian h_1 (which is also a Friedrichs-extension):

$$\| A\, e^{-ih_1\tau} \varphi_1 \| \leq \text{const.} \, \| h_1^{o\frac{1}{2}} e^{-ih_1\tau} \varphi_1 \|$$

$$\leq \text{const.} \, \| (h_1 + c)^{\frac{1}{2}} e^{-ih_1\tau} \varphi_1 \|$$

$$= \text{const.} \, \| (h_1 + c)^{\frac{1}{2}} \varphi_1 \| ,$$

where h_1^o is the internal kinetic energy of C_1.

*

Coulomb-potentials.

We give a brief account of the work by Dollard[10] on the Coulomb potential: Let

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m_i} + \sum_{\ell} V_\ell + \sum_{i<k} \frac{e_i e_k}{|\vec{x}_{ik}|} ,$$

where the two-body potentials V_ℓ satisfy the conditions of Theorem 3. Let $\alpha = (F_1 \ldots F_n; \varphi_1 \ldots \varphi_n)$ be an arbitrary channel of the system and let $q_1 \ldots q_n$ be the total charges of the fragments $F_1 \ldots F_n$. It then turns out that the asymptotically "free" motion in channel α is disturbed by the long-range Coulomb-interactions between these charged fragments. More precisely, the free propagator $\exp(-iH_\alpha t)$ is to be replaced by

$$e^{-iH_{\alpha,c}(t)} \equiv e^{-i(H_\alpha t + \epsilon(t) \sum_{r<s} (q_r q_s / v_{rs}) \log|t|} ,$$

where r,s label the fragments and v_{rs} is the relative velocity of F_r with respect to F_s:

$$v_{rs} = \frac{|M_s \vec{P}_r - M_r \vec{P}_s|}{M_r M_s} ,$$

M_r, $\vec{P}_r$ being the total mass and the total momentum of F_r, and where $\epsilon(t) = \text{sgn } t$. Clearly, the propagator $\exp[-iH_{\alpha,c}(t)]$ is well-defined

as a multiplication operator in momentum-space. It is still unitary, but does not define a one-parameter unitary group.

Theorem 9 (Dollard[10]): If, for some $\epsilon > 0$, the bound states $\varphi_r(z_r)$ are in the domain of $|z_r|^{\epsilon}$, where $|z_r|$ is the Euclidean distance in the space of internal coordinates of F_r, then

$$\lim_{t \to \pm\infty} e^{iHt} e^{-iH_{\alpha,c}(t)} \psi_\alpha \equiv \Omega_\alpha^{\pm} \psi_\alpha$$

exists for all $\psi_\alpha \in D_\alpha$.

*

The complete proof of this is contained in Dollard's thesis.[11] We only give a heuristic argument to make plausible the modification $H_\alpha t \to H_{\alpha,c}(t)$. As in the case of short-range forces, one tries to prove that

$$\left\| \left(H - \frac{d}{dt} H_{\alpha,c}(t) \right) e^{-iH_{\alpha,c}(t)} \psi_\alpha \right\| \tag{III.14}$$

$$= \left\| \left(\sum V_{ik} + \sum \frac{e_i e_k}{|\vec{x}_{ik}|} - \sum_{r<s} \frac{q_r q_s}{v_{rs}|t|} \right) e^{-iH_{\alpha,c}(t)} \psi_\alpha \right\|$$

vanishes like $|t|^{-s}$, $s > 1$, as $|t| \to \infty$, where (i,k) runs over all pairs of particles linking different fragments. If propagated by $\exp[-iH_{\alpha,c}(t)]$, the fragments will be far separated for $|t| \to \infty$, so that their interaction is essentially the Coulomb-interaction of point-charges $q_1 \ldots q_n$:

$$\sum \frac{e_i e_k}{|\vec{x}_{ik}|} = \sum_{r<s} \frac{q_r q_s}{\rho_{rs}} + \text{short-range forces},$$

where ρ_{rs} is the distance between the centers-of-mass of F_r and F_s. In classical terms, $\rho_{rs} = v_{rs} \cdot |t|$ as $|t| \to \infty$, which makes it plausible that the second and third sums on the right-hand side of (III.14) cancel as $|t| \to \infty$, up to forces of short range.

*

Asymptotic completeness for weak forces:

In the limit of weak forces, the question of asymptotic completeness can be settled by showing the convergence of the Dyson

expansion. For $N=2$, this result is due to Prosser.[12] Prosser's argument also suffices for $N=3$, while in the case $N \geq 4$ more detailed estimates of the terms in the perturbation expansion are needed.[13] We shall restrict ourselves to the case $N=3$. Let

$$H = H_o + \lambda \sum_{\ell} V_{\ell} = H_o + \lambda V.$$

If the coupling-constant λ is sufficiently small, we expect the following situation: Neither the N-particle-system nor any of its subsystems possess bound states. Therefore, in scattering theory there is only one channel $\alpha = 0$ in which all particles are asymptotically free as $|t| \to \infty$. Then $D_o = \mathcal{K}' = \mathcal{K}$, and the system is asymptotically complete if $R^{\pm} = \mathcal{K}$, i.e., if $\Omega^{\pm}$ is not only isometric but unitary, or, equivalently, if $\Omega^{\pm *}$ is isometric too. From Theorem 7, we can only conclude that

$$\Omega^{\pm *} = \lim_{t \to \pm\infty} e^{iH_o t} e^{-iHt} \tag{III.15}$$

<u>weakly</u> on $\mathcal{K}$, which does not imply that the limit is isometric. This follows, however, if we can show that the right hand side of (III.15) converges strongly (on a dense set), and to establish this we use the Dyson expansion for $\Omega^*(t) \equiv \exp(iH_o t)\exp(-iHt)$. Starting from the integral equation

$$\Omega^*(t) = 1 - i\lambda \int_o^t e^{iH_o t_1} V e^{-iHt_1} dt_1$$

$$= 1 - i\lambda \int_o^t e^{iH_o t_1} V e^{-iH_o t_1} \Omega^*(t_1) dt_1,$$

we obtain by iteration the formal series

$$\Omega^*(t) = \sum_{u=0}^{\infty} \Omega^{*(n)}(t) \lambda^n \tag{III.16}$$

$$\Omega^{*(n)}(t) = (-i)^n \int_o^t dt_1 \int_o^{t_1} dt_2 \cdots \int_o^{t_{n-1}} dt_n \, e^{iH_o t_1} V e^{-iH(t_1 - t_2)} V \cdots$$

$$\times e^{-iH_o(t_n - t_{n-1})} V e^{-iH_o t_n}.$$

Theorem 10 (Prosser[12]): For each ℓ, let $V_\ell(\cdot) \in L^1(R^3) \cap L^\infty(R^3)$, and let $N \leq 3$. Then, for sufficiently small $|\lambda|$, the perturbation series (III.16) converges strongly and absolutely on a dense set D, uniformly in $-\infty < t < +\infty$. Moreover, the strong limit of $\Omega^*(t)$ as $|t| \to \infty$ exists and can be taken term by term.

Proof: Since $V_\ell(\cdot) \in L^\infty(R^3)$, the operators V_ℓ are bounded and $\Omega^{*(n)}(t)$ therefore well-defined. By splitting each V_ℓ into two factors

$$V_\ell = A_\ell B_\ell$$

$$A_\ell = |V_\ell|^{\frac{1}{2}}, \qquad B_\ell = (\operatorname{sgn} V_\ell)|V_\ell|^{\frac{1}{2}},$$

one easily finds the estimate

$$\|\Omega^{*(n)}(t)\psi\| \leq \sum_{\ell_1 \ldots \ell_n} \|A_{\ell_1}\| \prod_{k=1}^{n-1} \int_0^\infty dt\, K_{\ell_k \ell_{k+1}}(t) \int_0^\infty N_{\ell_n}(t)dt,$$

where

$$K_{\ell m}(t) = \|B_\ell e^{-iH_o t} A_m\|$$

$$N_\ell(t) = \|B_\ell e^{-iH_o t}\psi\|.$$

To prove the theorem, it is obviously sufficient to show that, for ψ in a suitable dense set, $K_{\ell m}(t)$ and $N_\ell(t)$ are integrable over $0 \leq t < \infty$. $V_\ell(\cdot) \in L^1(R^3)$ implies $B_\ell(\cdot) \in L^2(R^3)$; therefore, $N_\ell(t)$ is integrable over $0 \leq t < \infty$ by the estimates of Section III. Also, since the operators B_ℓ and A_ℓ are bounded, the functions $K_{\ell m}(t)$ are uniformly bounded in t. Therefore, it remains to show that $K_{\ell m}(t)$ decreases faster than t^{-1} as $t \to \infty$:

1. $\ell = m$:

$$\|B_\ell e^{-iH_o t} A_\ell\| = \|B_\ell e^{-iH_{o,\ell} t} A_\ell\|,$$

where $H_{o,\ell}$ is the internal kinetic energy of the pair ℓ, whence

$$\langle \vec{x}| e^{-iH_{o,\ell} t} |\vec{x}'\rangle = -i\left(\frac{m}{2\pi i t}\right)^{3/2} e^{(im(\vec{x}-\vec{x}')^2)/2t},$$

$\vec{x}$, m being the relative coordinate and the reduced mass of the pair ℓ, respectively. Let $\vec{y}$ be the coordinate of the center-of-mass of

the pair ℓ and $\vec{z}$ the coordinate of the third particle. Then, for any $\psi \in \mathcal{K}$,

$$\left|\left(B_\ell e^{-iH_{o,\ell}t} A_\ell \psi\right)(\vec{x},\vec{y},\vec{z})\right|$$

$$\leq |V_\ell(\vec{x})|^{\frac{1}{2}} \left|\frac{m}{2\pi t}\right|^{3/2} \int d^3x' |V_\ell(\vec{x}')|^{\frac{1}{2}} |\psi(\vec{x}',\vec{\lambda},\vec{z})|$$

$$\leq |V_\ell(\vec{x})|^{\frac{1}{2}} \left|\frac{m}{2\pi t}\right|^{3/2} \|V_\ell(\cdot)\|_1^{\frac{1}{2}} \|\psi(\cdot,\vec{y},\vec{z})\|_2,$$

and therefore

$$K_{\ell\ell}(t) \leq \left|\frac{m}{2\pi t}\right|^{3/2} \|V_\ell(\cdot)\|_1.$$

2. $\ell \neq m$: Let $\ell = (12)$, $m = (23)$. Then

$$\|B_\ell e^{-iH_o t} A_m\| = \|B_\ell e^{-iH_{o,2}t} A_m\|,$$

when $H_{o,2}$ is the kinetic energy of the second particle, since the propagator of particle 1 commutes with A_m and the propagator of particle 3 commutes with B_ℓ. In the same way as before, one finds

$$K_{\ell m}(t) \leq \left|\frac{m}{2\pi t}\right|^{3/2} \|V_\ell(\cdot)\|_1^{\frac{1}{2}} \|V_m(\cdot)\|_1^{\frac{1}{2}},$$

where m is now the mass of particle 2.

*

Remarks.

1. It is easy to see why the method fails for $N>3$. Then one encounters terms like

$$B_{12} e^{-iH_o t} A_{34} = e^{-iH_{o,34}t} B_{12} A_{34} e^{-iH_{o,12}t},$$

where $H_{o,12}$ and $H_{o,34}$ are the kinetic energies of the subsystems (12) and (34). This implies

$$K_{12,34}(t) = \|B_{12} A_{34}\|$$

which is independent of t and not integrable, therefore, over $0 \leq t < \infty$. A more elaborate grouping of the factors in $\Omega^{*(n)}(t)$ is then necessary before taking the norms.[13]

2. The condition $V_\ell(\cdot) \in L^\infty$ is pure luxury in order to make $\Omega^{*(n)}(t)$ obviously well-defined. Making use of the smoothing-properties of the free propagator in x-space, one can include unbounded potentials such as, for example, Yukawa-potentials.

IV. Cluster-Properties of the S-Matrix, Cross-Sections

Let $D = (C_1 \ldots C_n)$ be a decomposition of $(1 \ldots N)$. For the translation operator which serves to separate the clusters in space-time we choose

$$U_D(a_1 \ldots a_n) = \prod_{k=1}^{n} \exp i\Big(\vec{P}_k \vec{a}_k - \frac{P_k^2}{2M_k} a_k^o\Big), \qquad \text{(IV.1)}$$

where $a_k = (a_k^o, \vec{a}_k) \in R^4$. This operator translates each cluster C_k by a_k without changing its internal state, i.e., it differs from the translation-operator (II.1) insofar as we have replaced the internal propagators for the subsystems C_k by 1. Theorem 3 still holds for this modified translation-operator, since only the unitarity of the internal propagators was used in its proof.

Now let $\alpha = (F_1 \ldots F_m; \varphi_1 \ldots \varphi_m)$ be a channel of the N-particle system. Definition: we write $(C_1 \ldots C_n) \subset (F_1 \ldots F_m)$ or simply $D \subset \alpha$ if $(F_1 \ldots F_m)$ is obtained by further partitioning $(C_1 \ldots C_n)$ or, equivalently, if each fragment F_k belongs to a definite cluster. If this is the case, then α reduces to a definite subchannel α_k for each subsystem C_k. For example, if $C_k = (F_1, F_2, F_3)$, then

$$\alpha_k = (F_1, F_2, F_3; \varphi_1, \varphi_2, \varphi_3).$$

Obviously, we then have

$$\mathcal{H} = \bigotimes_{k=1}^{n} \mathcal{H}_k \qquad D_\alpha = \bigotimes_{k=1}^{n} D_{\alpha_k}$$

where $\mathcal{H}_k$ and $D_{\alpha_k} \subset \mathcal{H}_k$ are the Hilbert space and the space of channel-states of C_k, respectively. If no interactions between the clusters are present—i.e., if H_D is the Hamiltonian of the system, then the wave operators $\Omega_\alpha^\pm$ can be simply expressed in terms of the wave operators of the independent subsystems C_k:

$$\Omega_\alpha^\pm = \bigotimes_{k=1}^{n} \Omega_{\alpha_k}^\pm .$$

Therefore, we expect this to be the cluster-limit of the wave operator for large separations of the clusters:

Theorem 11[4]: Under the hypothesis of Theorem 3, we have

$$\Omega_\alpha^\pm U_D(a_1 \ldots a_n) \to \bigotimes_{k=1}^{n} \Omega_{\alpha_k}^\pm U_D(a_1 \ldots a_n)$$

as

$$\min_{i \neq k} |a_i - a_k| \to \infty ,$$

strongly on D_α, for all α and all $D \subset \alpha$.

Proof: Let $\Omega_\alpha(t) = \exp(iHt)\exp(-iH_\alpha t)$ and similarly for $\Omega_{\alpha_k}(t)$. Then

$$\left\| \left[\Omega_\alpha(t) - \bigotimes_{k=1}^{n} \Omega_{\alpha_k}(t) \right] U_D(a_1 \ldots a_n) \psi_\alpha \right\|$$

$$= \left\| \left(e^{iHt} - e^{iH_D t} \right) U_D(a_1 \ldots a_n) e^{-iH_\alpha t} \psi_\alpha \right\| \to 0$$

as $\min|a_i - a_k| \to \infty$, uniformly in $-\infty < t < +\infty$. This would follow directly from Theorem 3 if, instead of $\exp(-iH_\alpha t)\psi_\alpha$, we had a time-independent state φ. However, going through the proof of Theorem 3, it is easily seen that this time-dependence does no harm. The uniformity in t then implies the desired result for the limits of $\Omega_\alpha(t)$ and of

$$\bigotimes_{k=1}^{n} \Omega_{\alpha_k}(t) \quad \text{as} \quad t \to \pm\infty .$$

*

It is now easy to derive cluster-properties for the S-operators. For the operator S this is found in Reference 4, while for S' we find the following:

Theorem 12: Let $S'_{\beta\alpha}$ be the part of S' mapping D_α into D_β, and let $D \subset \alpha$. Then, under the hypothesis of Theorem 3,

$$S'_{\beta\alpha} U_D(a_1 \ldots a_n) \to \begin{cases} \bigotimes_{k=1}^{n} S'_{\beta_k \alpha_k} U_D(a_1 \ldots a_n) & \text{if } D \subset \beta \\ 0 & \text{if } D \not\subset \beta \end{cases}$$

as

$$\min_{i\neq k} |a_i - a_k| \to \infty,$$

strongly on D_α.

*

In the case $D \subset \beta$, this means that for large separations of the clusters, the scattering process $\alpha \to \beta$ factors into independent processes $\alpha_k \to \beta_k$ for each isolated subsystem C_k. On the other hand, $D \not\subset \beta$ implies that one of the fragments in the final channel β contains particles from different clusters. Therefore, the probability of production for such a fragment vanishes in the limit of large separation of the clusters.

Theorem 12 suggests an expression of $S'_{\beta\alpha}$ in terms of "connected parts" $R_{\beta\alpha}$ by the usual truncation procedure:

$$S'_{\beta\alpha} = \sum_{\substack{D \\ D\subset\alpha,\, D\subset\beta}} R_{\beta_1\alpha_1} R_{\beta_2\alpha_2} \cdots R_{\beta_n\alpha_n}. \qquad \text{(IV.2)}$$

Since the trivial decomposition always satisfies $D\subset\alpha$, $D\subset\beta$, this defines the connected parts $R_{\beta\alpha}$ recursively in terms of the $S'_{\beta\alpha}$:

$$R_{\beta\alpha} = S'_{\beta\alpha} - \sum_{\substack{D\subset\alpha,\, D\subset\beta \\ n\geq 2}} R_{\beta_1\alpha_1} \cdots R_{\beta_n\alpha_n}.$$

(This reduces to the usual $S'_{\beta\alpha} = \delta_{\beta\alpha} + R_{\beta\alpha}$ in the case where the ingoing channel α contains only two fragments, where $\delta_{\beta\alpha} = 0$ for $\beta \neq \alpha$ and $\delta_{\alpha\alpha}$ = identity map on D_α.) In terms of the connected parts $R_{\beta\alpha}$, Theorem 12 then simply reads

$$R_{\beta\alpha} U_D(a_1 \cdots a_n) \longrightarrow 0 \qquad \text{(IV.3)}$$

as

$$\max_{i\neq k} |a_i - a_k| \longrightarrow \infty,$$

if $D\subset\alpha$, strongly on D_α.

*

$R_{\beta\alpha}$ maps D_α into D_β and still conserves energy and momentum. Formally, one would therefore introduce scattering amplitudes by

$$\langle \vec{p}'_1 \ldots \vec{p}'_m | R_{\beta\alpha} | \vec{p}_1 \ldots \vec{p}_n \rangle = \delta(\vec{P}' - \vec{P})\,\delta(E' - E)\langle \vec{p}'_1 \ldots \vec{p}'_m | T_{\beta\alpha} | \vec{p}_1 \ldots \vec{p}_n \rangle ,$$

(IV.4)

where $\vec{p}_1 \ldots \vec{p}_n$ and $\vec{p}'_1 \ldots \vec{p}'_m$ are the momenta of the fragments in channel α and β, respectively, and $\vec{P}, \vec{P}', E, E'$ the total momenta and the total energies in the two channels. Equations (IV.2) and (IV.4) then express S' in terms of scattering amplitudes. In this way, the amplitudes of subsystems C_k appear explicitly in S', describing the possible disconnected processes.

*

Cross-sections.

We want to define cross-sections, for collisions with an arbitrary number of fragments in the incident channel, in a way which clearly exhibits that the existence of cross-sections is a consequence of the spatial cluster-properties of the S-matrix.

First, we define an impact parameter in classical terms, which parametrizes (up to translations) all possible collisions of or ingoing fragments $F_1 \ldots F_n$ with fixed momenta $\vec{p}_1 \ldots \vec{p}_n$:

Let C be the 3n-dimensional configuration-space of the particles $F_1 \ldots F_n$ with coordinates $y = (\vec{y}_1 \ldots \vec{y}_n)$. On C, we define the metric

$$dy^2 = \frac{1}{2} \sum_{i=1}^{n} m_i d\vec{y}_i^{\,2} . \qquad \text{(IV.5)}$$

The transition to the center-of-mass coordinates

$$\vec{y}'_i = \vec{y}_i - \frac{1}{M} \sum_{k=1}^{n} m_k \vec{y}_k \qquad \text{(IV.6)}$$

(m_i = mass of F_i, $M = m_1 + \ldots + m_n$) is then the orthogonal projection onto the $3(n-1)$-dimensional subspace

$$C' = \Big\{ y \Big| \sum_{i=1}^{n} m_i \vec{y}_i = 0 \Big\} .$$

The free motion of the incoming fragments $F_1 \ldots F_n$ is described classically by a straight trajectory

$$y(t) = y(0) + vt \qquad \text{(IV.7)}$$

in C, where $v = (v_1 \ldots v_n)$, v = velocity of F_i, or, projected onto C'

by

$$y'(t) = y'(0) + wt, \tag{IV.8}$$

where w_i is the velocity of F_i in the center-of-mass frame. This projected trajectory characterizes the collision up to translation. Since its direction is given by the fixed momenta $(\vec{p}_1 \ldots \vec{p}_n)$, we can use as an impact parameter a the point of intersection of this trajectory with any $(3n-4)$-dimensional plane $A \subset C'$ not parallel to w. We choose

$$A = w^{\perp} \quad (\text{in } C').$$

Then, in C, A is also orthogonal to v (in the sense of (IV 5), and therefore characterized by

$$A = \Big\{ y \Big| \sum_{i=1}^{n} m_i \vec{y}_i = 0, \quad \sum_{i=1}^{n} \vec{p}_i \vec{y}_i = 0 \Big\}. \tag{IV.9}$$

The impact parameter a is thus simply the (orthogonal) projection of $y'(0)$ (or $y'(t)$) onto A, and its length is given by

$$(a,a) = \min_t \sum_{i=1}^{n} \frac{m_i}{2} \vec{y}_i^{\,2}(t).$$

Therefore, $a = 0$ only if $y'(t) = 0$ for some t, i.e., if all the particles $F_1 \ldots F_n$ collide simultaneously. In general, (a,a) is a measure of the closeness the particles could reach in free motion. True n-particle collisions therefore happen for small $|a|$ only, while collisions between less than n particles are possible for arbitrarily large $|a|$.

In terms of this impact parameter a, a classical cross-section may be defined as follows: Consider a statistical ensemble of collisions with fixed momenta $\vec{p}_1 \ldots \vec{p}_n$ in the ingoing channel α, such that the impact parameters a are distributed over the plane A with uniform density n (in the sense of the metric (IV.5)). Let $N(\Omega)$ be the total number of events in which the system is finally in channel β with momenta $(\vec{p}'_1 \ldots p'_m) \in \Omega$. Then the cross-section is defined as

$$\sigma_{\beta\alpha}(\Omega, \vec{p}_1 \ldots \vec{p}_n) = \frac{N(\Omega)}{n}.$$

In order that this is finite, the region Ω cannot be arbitrary; it must be chosen such that $N(\Omega)$ receives no contribution from processes which are not true n-particle collisions and which may happen for

arbitrarily large $|a|$. There are two types of such processes:

a) Disconnected processes:

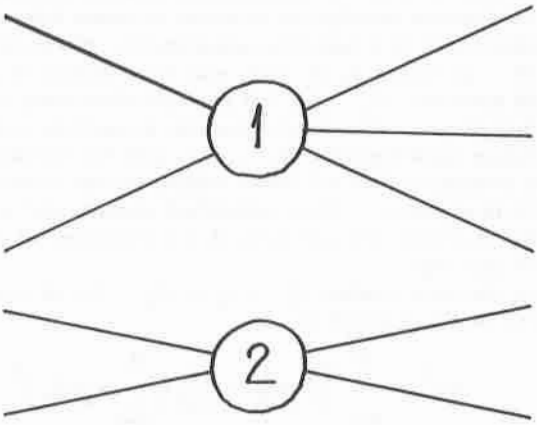


Figure 1

b) Rescattering processes:

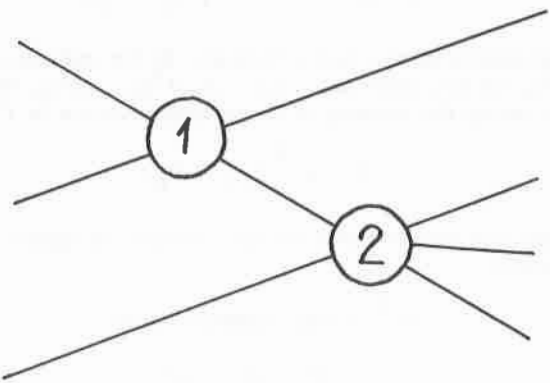


Figure 2

While it is easy to enumerate the possible disconnected processes (see (IV.2)), we know of no general rule giving the rescattering processes which have to be avoided. In any case, however, the processes of both types are restricted by the fact that energy and momentum are conserved for each single process $1,2\ldots$, and this restricts the final momenta to a set of measure zero on the energy-momentum shell. It is possible, therefore, to choose Ω such

that it does not intersect this set. For the familiar case $n=2$, this simply means that Ω must not contain the forward direction.

Let us discuss briefly the relation between this cross-section and the counting rate in a specific experiment. We consider a target consisting of ν particles F_1 at rest, and $n-1$ beams of particles $F_2 \ldots F_n$ with momenta $\vec{p}_1 \ldots \vec{p}_n$ and particle-densities $\rho_1 \ldots \rho_n$ converging on this target. The candidates for n-particle collisions are then all possible combinations $(F_1 \ldots F_n)$, and the counting rate is evaluated by assuming that all these combinations behave like isolated N-particle-systems. This simplified picture will predict the large majority of collisions correctly if the densities of beams and target are not too high.

Using the coordinates $\vec{y}_2 \ldots \vec{y}_n$ of $F_2 \ldots F_n$ as coordinates in C', the metric in C' is given by

$$ds^2 = \sum_{k=2}^{n} \frac{m_k}{2} d\vec{y}_k^{\,2} - \frac{1}{2M} \Big(\sum_{k=2}^{n} m_k d\vec{y}_k \Big)^2$$

(CM-kinetic energy!), which yields

$$g = \Big(2^{(1-n)} M^{-1} m_1 m_2 \ldots m_n \Big)^3$$

for the determinant of the metric tensor. In the volume-element $d^3y_2 \ldots d^3y_n$ of C', there are $\nu \rho_2 \ldots \rho_n \, d^3y_2 \ldots d^3y_n$ combinations $(F_1 \ldots F_n)$; hence the density of these combinations in C' is

$$\rho = g^{-\frac{1}{2}} \nu \rho_2 \ldots \rho_n .$$

$\rho |w|$ is then the flux of these points through the impact-parameter plane A, where

$$|w|^2 = \text{CM-kinetic energy}$$

$$= \sum_{k=2}^{n} \frac{\vec{p}_k^{\,2}}{2m_k} - \frac{\vec{P}^2}{2M} ,$$

with $\vec{P} = \vec{p}_2 + \ldots + \vec{p}_n$. The counting rate for n-particle collisions producing the fragments of channel β with momenta $(\vec{p}'_1 \ldots \vec{p}'_m) \in \Omega$ is then

$$\mu(\Omega) = \sigma_{\beta\alpha}(\Omega, 0, \vec{p}_2 \ldots \vec{p}_n) \rho |w| .$$

*

Finally, we define the cross-section in quantum mechanics and relate it to the S-matrix. Instead of fixed initial momenta $\vec{p}_1 \ldots \vec{p}_n$, we consider a state

$$\psi_\alpha = \phi_\alpha(\vec{p}_1 \ldots \vec{p}_n) \prod_{k=1}^{n} \varphi_k(z_k) \in D_\alpha ,$$

normalized to $\|\psi_\alpha\| = 1$, and use the expectation values $\langle \vec{p}_i \rangle$ of the ingoing momenta to define the impact-parameter plane A as in the classical case:

$$A = \Big\{ y \Big| \sum_{i=1}^{n} m_i \vec{y}_i = 0, \qquad \sum_{i=1}^{n} \langle \vec{p}_i \rangle \vec{y}_i = 0 \Big\} .$$

Of course, we cannot attribute a definite impact parameter to the state ψ_α. However, we can define a variation of the impact-parameter by a, $a \in A$, as the transformation

$$\phi_\alpha \to T(a)\phi_\alpha , \qquad T(a) = e^{i \sum_{i=1}^{n} \vec{p}_i \vec{a}_i} .$$

The cross-section $\sigma_{\beta\alpha}(\Omega, \phi_\alpha)$ is now defined as follows: For given ϕ_α, consider an ensemble of collisions with initial states $T(a_k)\phi_\alpha$, in which the impact-parameters a_k are distributed uniformly over A with density n. Let Ω and $N(\Omega)$ be as in the classical case. Then

$$\sigma_{\beta\alpha}(\Omega, \phi_\alpha) = \frac{N(\Omega)}{n} .$$

This is easily expressed in terms of the S-matrix:

$$\begin{aligned} \sigma_{\beta\alpha}(\Omega, \phi_\alpha) &= \lim_{n \to \infty} \frac{1}{n} \sum_k \| P_\Omega S'_{\beta\alpha} T(a_k)\phi_\alpha \|^2 \\ &= \int_A da \, \| P_\Omega S'_{\beta\alpha} T(a)\phi_\alpha \|^2 , \end{aligned} \tag{IV.10}$$

where P_Ω is the projection operator on D_β corresponding to the final region Ω, and da the volume-element on A defined by the metric (IV.5). In (IV.10) we could replace $S'_{\beta\alpha}$ by its connected part $R_{\beta\alpha}$, since Ω has to be chosen so as to avoid contributions from disconnected processes. By (IV.3), $\| R_{\beta\alpha} T(a)\phi_\alpha \|^2 \to 0$ for $|a| \to \infty$, but not fast enough, in general, to make this integrable over A. The reason

is that $R_{\beta\alpha}$ still contains the rescattering processes, which die out too slowly as $|a|$ increases. However, if Ω also avoids the contributions from these processes, we can expect that $\|P_\Omega R_{\beta\alpha} T(a)\phi_\alpha\|^2$ is integrable over A. A general proof of this, however, is not available.

Proceeding formally from (IV.10) and (IV.4), we can express the cross-section in terms of the scattering amplitude:

$$\sigma_{\beta\alpha}(\Omega,\phi_\alpha) = \int_A da \int_\Omega dp' \int dp\, dp''\, \delta(P-P')\delta(P'-P'')$$

$$\times \delta(E-E')\delta(E'-E'')\langle p'|T_{\beta\alpha}|p\rangle\overline{\langle p'|T_{\beta\alpha}|p''\rangle} \qquad \text{(IV.11)}$$

$$\times \exp\Big[i\sum_{k=1}^{n}(\vec{p}_k-\vec{p}_k'')\vec{a}_k\Big]\phi_\alpha(p)\overline{\phi_\alpha(p'')},$$

with obvious abbreviations. Using the CM-velocities w,w" instead of p,p", we can express the exponent in terms of the scalar product corresponding to (IV.5):

$$\sum_{k=1}^{n}(\vec{p}_k-\vec{p}_k'')\vec{a}_k = 2(w-w'',a),$$

so that

$$\int_A da\, e^{2i(w-w'',a)} = \pi^{(3n-4)}\delta(w_A-w_A''), \qquad \text{(IV.12)}$$

where w_A, w_A'' are the projections onto A of w,w" and where the δ-function is normalized with respect to integration over A. The δ-functions in (IV.11) and (IV.12) now imply

$$P'' = P, \qquad w_A'' = w_A, \qquad \text{(IV.13)}$$

$$E'' = \frac{P''^2}{2M} + (w'',w'') + \epsilon_\alpha = E = \frac{P^2}{2M} + (w,w) + \epsilon_\alpha, \qquad \text{(IV.14)}$$

ϵ_α being the sum of the bound state energies in channel α. In view of (IV.13), (IV.14) reduces to

$$w_N = \pm w_N'', \qquad \text{(IV.15)}$$

where w_N, w_N'' are the components of w,w" orthogonal to A. In order that the integral (IV.11) only receives a contribution from the

points where $w_N = w_N''$, we now postulate that

$$\phi_\alpha(p) = 0 \quad \text{for} \quad (w, \langle w\rangle) \le 0. \tag{IV.16}$$

This means that the support of the w-distribution is entirely on one side of the plane A—the side determined by $\langle w\rangle$, or, equivalently, that the spread in the w-distribution is smaller than $|\langle w\rangle|$. The δ-functions then simply imply that $p = p''$, so that only the absolute squares of $\langle p'|T_{\beta\alpha}|p\rangle$ and of $\phi_\alpha(p)$ enter in the cross-section. The remaining phase-space integral over p" is easily found to be

$$\int dp''\,\delta(P-P'')\,\delta\Big[(w_N,w_N) - (w_N'',w_N'')\Big]\,\delta(w_A - w_A'') = \frac{\prod_{i=1}^{n}(2m_i)^{3/2}}{(2M)^{3/2}}\;\frac{1}{2|w_N|},$$

where $|w_N|$ is the length—in the sense of (IV.5)—of w_N:

$$|w_N|^2 = \frac{(w,\langle w\rangle)^2}{(\langle w\rangle,\langle w\rangle)^2} = \frac{\Big(\sum_{i=1}^{n}\frac{\vec p_i\langle\vec p_i\rangle}{2m_i} - \frac{\vec P\langle\vec P\rangle}{2M}\Big)^2}{\sum_{i=1}^{n}\frac{\langle\vec p_i\rangle^2}{2m_i} - \frac{\langle\vec P\rangle^2}{2M}}.$$

The result is, therefore,

$$\sigma_{\beta\alpha}(\Omega,\phi_\alpha) = \frac{1}{2}\pi^{3n-4}\;\frac{\prod_{i=1}^{n}(2m_i)^{3/2}}{(2M)^{3/2}}\int_\Omega dp'\int dp$$

$$\times\,\delta(E-E')\,\delta(\vec P-\vec P')\;\frac{\Big(\sum_{i=1}^{n}\frac{\langle\vec p_i\rangle^2}{2m_i} - \frac{\langle\vec P\rangle^2}{2M}\Big)^{\frac{1}{2}}}{\sum_{i=1}^{n}\frac{\vec p_i\langle\vec p_i\rangle}{2m_i}\;\frac{\vec P\langle\vec P\rangle}{2M}}$$

$$\times\,|\langle p'|T_{\beta\alpha}|p\rangle|^2\,|\phi_\alpha(p)|^2.$$

Finally, if we specialize the momentum-distribution $|\phi_\alpha(p)|^2$ to a δ-function (which is consistent, in view of (III.4), with the normalization $\|\psi_\alpha\| = 1$), we obtain the differential cross-section

$$d\sigma_{\beta\alpha}(\vec{p}_1' \ldots \vec{p}_m'; \vec{p}_1 \ldots \vec{p}_n)$$

$$= \frac{1}{2}\pi^{3n-4} \frac{\prod_{i=1}^{n} (2m_i)^{3/2}}{(2M)^{3/2}} \delta(\vec{P} - \vec{P}') \delta(E - E')$$

$$\times \Big(\sum_{k=1}^{n} \frac{\vec{p}_k^{\,2}}{2m_k} - \frac{\vec{P}^2}{2M} \Big)^{-\frac{1}{2}} |\langle p' | T_{\beta\alpha} | p \rangle|^2 \, d^3p_1' \ldots d^3p_m' .$$

V. Time-Independent Scattering Theory

As far as scattering theory is concerned, this section will be an exposition of problems rather than of results. To get familiar with some of these problems, we briefly discuss the case $N = 2$.

The starting point is the resolvent equation linking $G(z) = (z - H)^{-1}$ and $G_o(z) = (z - H_o)^{-1}$:

$$G(z) = G_o(z) + G_o(z) V G(z), \tag{V.1}$$

or the Lippmann-Schwinger equation for the transition-operator $T(z)$:

$$T(z) = V + V G_o(z) T(z), \tag{V.2}$$

which are connected by

$$T(z) = (z - H_o)\Big[G(z) - G(z_o) \Big] (z - H_o) \tag{V.3}$$

$$G(z) = G_o(z) + G_o(z) T(z) G_o(z). \tag{V.4}$$

In momentum space, the kernel of the Lippmann-Schwinger equation is

$$\langle \vec{p} | V G_o(z) | \vec{q} \rangle = \tilde{V}(\vec{p} - \vec{q}) \Big(z - \frac{\vec{q}^{\,2}}{2m} \Big)^{-1},$$

where $\tilde{V}$ is the Fourier-transform of V. For $V(\cdot) \in L^2(R^3)$ and $z \notin \sigma(H_o)$, this is a Hilbert-Schmidt kernel (HS-kernel):

$$\|VG_o(z)\|_{HS}^2 \equiv \int d^3p\, d^3q\, |\langle \vec{p}|\, VG_o(z)|\, \vec{q}\rangle|^2$$

$$= \|V(\cdot)\|_2^2\, 4\pi \int_o^\infty dq\, q^2 |z - q^2/2m|^{-2} < \infty . \qquad (V.5)$$

Also, $VG_o(z)$ is holomorphic in $z \notin \sigma(H_o)$ (i.e., in the cut plane avg $z \neq 0$), and vanishes in norm as $\mathrm{Re}\, z \to -\infty$ (by (I.3) with $B=V$, $A=H_o$), so that $[1 - VG_o(z)]^{-1}$ exists for sufficiently large negative $\mathrm{Re}\, z$. Now we use the following lemma:

Lemma 1: Let $A(z)$ be a holomorphic function of the complex variable z in an open, connected region G, whose values are compact operators on a Hilbert-space, and suppose that, for some $z_o \in G$, $[1 - A(z_o)]^{-1}$ exists. Then $[1 - A(z)]^{-1}$ is meromorphic in $z \in G$ (i.e., it admits a Laurent-expansion with a non-vanishing radius of convergence (in the sense of the norm) for every $z \in G$).

*

A proof can be found in Reference 14. Applying this to the Lippmann-Schwinger kernel, we conclude that $[1 - VG_o(z)]^{-1}$ is meromorphic for $z \notin \sigma(H_o)$. Furthermore, if $z_o \notin \sigma(H_o)$ is a pole of $[1 - VG_o(z)]^{-1}$, the homogeneous equation

$$VG_o(z_o)\varphi = \varphi \qquad (V.6)$$

has nontrivial solutions which form a subspace of finite dimension, since $VG_o(z_o)$ is compact. Then $\psi = G_o(z_o)\varphi$ is a nontrivial solution of $\psi = G_o(z_o)V\psi$, or of

$$(z_o - H_o)\psi = V\psi , \qquad (V.7)$$

and vice-versa: a nontrivial solution ψ of (V.7) leads to a nontrivial solution $\varphi = (z_o - H_o)\psi$ of (V.6). Therefore, the poles of $[1 - VG_o(z)]^{-1}$ and the eigenvalues of H (in $z \notin \sigma(H_o)$) are in one-to-one correspondence, and so are the bound states and the solutions of (V.6). It follows that the poles of $[1 - VG_o(z)]^{-1}$ are real and that the eigenvalues of H are of finite multiplicity and cannot accumulate except at $z=0$. Furthermore, since

$$G(z) = G_o(z) + G_o(z)\Big[1 - VG_o(z)\Big]^{-1} VG_o(z),$$

these poles are the only singularities of $G(z)$ for $z \notin \sigma(H_o)$, i.e., the part of $\sigma(H)$ not contained in $\sigma(H_o)$ consists of these eigenvalues only. (Note that, by the corollary of Theorem 7, $\sigma(H_o) \subset \sigma(H)$).

Now let us look at the Lippmann-Schwinger equation in the following way:

$$\langle \cdot | T(z) | \vec{p}' \rangle = \tilde{V}(\cdot - \vec{p}') + V G_o(z) \langle \cdot | T(z) | \vec{p}' \rangle , \tag{V.8}$$

i.e., we look at $\langle \cdot | T(z) | \vec{p}' \rangle$ as an $L^2(R^3)$-valued function of z and $\vec{p}'$. Since $\tilde{V}(\cdot) \in L^2(R^3)$, $\tilde{V}(\cdot - \vec{p}')$ is such a function, depending continuously on $\vec{p}'$, and with L^2-norm independent of $\vec{p}'$. Therefore, $\langle \cdot | T(z) | \vec{p}' \rangle$ is meromorphic in $z \notin \sigma(H_o)$, with poles at the eigenvalues of H, bounded in norm by some constant $C(z) < \infty$ for $z \notin \sigma(H)$ uniformly in $\vec{p}'$, and continuous in $\vec{p}' \in R^3$. Inserting this estimate on the right side of (V.8) and using the Schwarz inequality, we see that $\langle \vec{p} | T(z) | \vec{p}' \rangle - \tilde{V}(\vec{p} - \vec{p}')$ is meromorphic in $z \notin \sigma(H_o)$, continuous in $\vec{p}, \vec{p}'$ and bounded by $C(z) < \infty$ for $z \notin \sigma(H)$, uniformly in $\vec{p}$ and $\vec{p}'$.

With this information on the T-matrix elements, we can already justify, to some extent, the usual formulae of time-independent scattering theory, for example:

$$(\Omega^{\pm}\psi)(\vec{p}) = \psi(\vec{p}) + \lim_{\varepsilon \downarrow o} \int d^3q \, \frac{\langle p | T(q^2/2m \mp i\varepsilon) | \vec{q} \rangle}{(2m)^{-1}(q^2 - p^2) \mp i\varepsilon} , \tag{V.9}$$

where, by our estimates, the right hand side is well-defined for $\varepsilon > 0$ if, for example, $\psi \in C_o^{\infty}(R^3)$, and where the (strong) convergence for $\varepsilon \downarrow 0$ follows from time-dependent scattering theory. Similarly, with $R = S' - 1$:

$$(\psi, R\varphi) = \lim_{\varepsilon \downarrow o} -2i \int d^3p \, d^3q \, \frac{\varepsilon}{(E - E')^2 + \varepsilon^2} \times \overline{\psi(\vec{p})} \langle \vec{p} | T\left(\tfrac{1}{2}(E + E') + \frac{i\varepsilon}{2}\right) | \vec{p}' \rangle \varphi(\vec{p}') , \tag{V.10}$$

where again the right hand side is well-defined for $\varepsilon > 0$ and $\psi, \varphi \in C_o^{\infty}(R^3)$.

As an example, we give a formal derivation of (V.9), omitting the justification for each intermediate step. By Theorem 7

$$\Omega^{\pm}\psi - \psi = \lim_{t \to \pm\infty} i \int_o^t e^{iH\tau} V e^{-iH_o\tau} \psi \, d\tau$$

$$= \lim_{\varepsilon \to \pm o} i \int_o^{\pm\infty} e^{-\varepsilon\tau} e^{iH\tau} V e^{-iH_o\tau} \psi \, d\tau$$

in the sense of the norm. To show that the last integral is identical

with the integral in (V.9), we use contour-integration:

$$\varphi \equiv i\int_0^{\pm\infty} d\tau \ldots = i(2\pi i)^{-2} \oint_T dz_1 \oint_T dz_2 \int_0^{\pm\infty} d\tau\, e^{-(\varepsilon - iz_4 + iz_2)\tau}$$

$$\times\, G(z_1)VG_0(z_2)\psi\,,$$

where T is a contour circling $\sigma(H)$ in the positive sense, inside the strip $|\mathrm{Im}\, z| < \varepsilon/2$. Carrying out the τ-integration first, we obtain a term $(\varepsilon - iz_1 + iz_2)^{-1}$. As a function of z_1, this has a pole at $z_1 = z_2 - i\varepsilon$, outside T. Deforming the z_1-contour T, we pick up only the residue from this pole:

$$\varphi = (2\pi i)^{-1} \oint_T dz_2\, G(z_2 - i\varepsilon)VG_0(z_2)\psi\,.$$

Now we use the relation $G(z)V = G_0(z)T(z)$ and obtain

$$\varphi(\vec{p}) = -(2\pi i)^{-1} \oint_T dz_2 \int d^3q\, \frac{\langle \vec{p}|\, T(z_2 - i\varepsilon)\,|\vec{q}\rangle \psi(\vec{q})}{(z - i\varepsilon - p^2/2m)(z - q^2/2m)}\,.$$

Since the pole at $z_2 = q^2/2m$ is the only singularity of the integrand inside T, (V.9) results by taking the residue of this pole. To justify this formal derivation, one can consider matrix-elements between suitable states. The formula (V.10) is obtained similarly by starting from

$$(\psi, S'\varphi) = \lim_{t \to +\infty} \left(\psi, e^{iH_0 t} e^{-2iHt} e^{iH_0 t} \varphi\right)$$

$$= \lim_{\varepsilon \downarrow 0} \varepsilon \int_0^{\infty} e^{-\varepsilon t} \left(\psi, e^{iH_0 t} e^{-2iHt} e^{iH_0 t} \varphi\right).$$

*

The main step—and the most difficult one to prove—is to go beyond (V.10) and to express the S-matrix element in terms of boundary-values of T-matrix elements as the energy approaches the continuous spectrum of H. Formally,

$$\lim_{\varepsilon \downarrow 0} \frac{\varepsilon}{(E-E')^2 + \varepsilon^2} = \pi\, \delta(E - E')\,,$$

so that one obtains the well-known relation:

$$(\psi, R\varphi) = -2\pi i \int d^3p\, d^3p'\, \delta\Big(\frac{p^2}{2m} - \frac{p'^2}{2m}\Big)\overline{\psi(\vec{p})}$$

$$\times \langle \vec{p} | T\Big(\frac{p^2}{2m} + io\Big) | \vec{p}' \rangle \varphi(\vec{p}'). \qquad (V.11)$$

This actually follows from (V.10) if

$$\langle \vec{p} | T\Big(\frac{1}{2}\Big(\frac{p^2}{2m} + \frac{p'^2}{2m}\Big) + \frac{i\varepsilon}{2}\Big) | \vec{p}' \rangle$$

is continuous in $\vec{p}, \varepsilon, \vec{p}'$ in the region $\operatorname{supp}\psi \times [0,1] \times \operatorname{supp}\varphi$, and attains its boundary-value for $\varepsilon \downarrow 0$ uniformly in $\vec{p}$ and $\vec{p}'$. The real problem of time-independent scattering theory is to obtain this kind of information on the boundary-values of T-matrix elements.

For the case $N=2$, this is comparatively easy. We use a trick similar to the one used in Section IV in the case of weak potentials, and write

$$T(z) = A\,U(z)B$$

$$A = |V|^{\frac{1}{2}} \qquad B = (\operatorname{sgn} V)|V|^{\frac{1}{2}}$$

so that $AB=V$. The Lippmann Schwinger equation for $U(z)$ is then

$$U(z) = 1 + BG_o(z)A\,U(z), \qquad (V.12)$$

and in x-space the kernel of the operator $BG_o(z)A$ is

$$-(4\pi)^{-1} B(\vec{x}) \frac{e^{i\sqrt{z}\,|\vec{x}-\vec{x}'|}}{|\vec{x}-\vec{x}'|} A(\vec{x}'). \qquad (\operatorname{Im}\sqrt{z} > 0).$$

The important point now is that this is a Hilbert-Schmidt kernel for all z in the closed cut plane $\operatorname{Im}\sqrt{z} \geq 0$, provided that

$$\int d^3x\, d^3x' \frac{|V(\vec{x})|\,|V(\vec{x}')|}{|\vec{x}-\vec{x}|^2} < \infty,$$

which we now assume. Furthermore, it is easily seen that $BG_o(z)A$ is holomorphic in the open cut plane $\operatorname{Im}\sqrt{z} > 0$ and continuous (even in the HS-norm) in the closed cut plane $\operatorname{Im}\sqrt{z} \geq 0$.

The crucial point is now to obtain control over the solutions of the homogeneous equation

$$\psi = BG_0(z)A\psi \qquad (V.13)$$

for $\mathrm{Im}\sqrt{z} \geq 0$. If $\mathrm{Im}\sqrt{z} > 0$, one can show again that these solutions are in one-to-one correspondence to the bound states of H. On the two boundaries of the cut $\arg z = 0$, however, the situation is more involved. There it can be shown[15] that a nontrivial solution of (V.13) gives rise to a purely (ingoing or outgoing) radiative solution of the time-independent Schrödinger equation with energy z. Conservation of probability then requires that the radiative term of this solution vanishes, i.e., that the solution decreases faster than $|\vec{x}|^{-1}$ as $|\vec{x}| \to \infty$. As for the free wave-equation, one can then show (under some additional assumptions on V) that this forces the solution to vanish identically, provided that $z > 0$. (For $z = 0$, there may be nontrivial solutions of (V.13) not corresponding to bound states, but to a "resonance at zero energy.") The upshot of all this is that

$$U(z) = 1 + W(z)$$

where W(z) is a Hilbert-Schmidt operator, holomorphic for $z \notin \sigma(H)$ and continuous in the closed cut plane $\mathrm{Im}\, z \geq 0$ except for the (negative) eigenvalues of H and possibly for $z = 0$. Then,

$$\langle \vec{p}\,|\,T(z)\,|\,\vec{p}\,'\rangle - \tilde{V}(\vec{p} - \vec{p}\,') = \int d^3q\, d^3r\, \tilde{A}(\vec{p} - \vec{q})\, \langle \vec{q}\,|\,W(z)\,|\,\vec{r}\rangle\, \tilde{B}(\vec{r} - \vec{p}\,'),$$

and if $\tilde{A}, \tilde{B} \in L^2(R^3)$ (i.e., $V(\cdot) \in L^1(R^3)$), this is bounded by $\|V\|_1 \, \|W(z)\|_{HS}$ and continuous in z in the region described above, uniformly in $\vec{p}, \vec{p}\,'$. This immediately gives the justification of (V.11).

In this account of the case $N = 2$, our aim was to point out the typical problems of time-independent scattering theory in their simplest form. We did not try, therefore, to manage with the least restrictive assumptions on $V(\vec{x})$. (For spherically symmetric potentials, everything could be based on the conditions

$$\int_0^\infty r^n |V(r)|\, dr < \infty \quad \text{for} \quad n = 1, 2,$$

familiar from partial-wave analysis.)

*

A consistent time-independent scattering theory for general N does not yet exist, but the situation has been very much improved by Faddeev's[16] solution for $N = 3$, which represents the simplest case

with a nontrivial multiparticle structure, and is therefore a big step towards a solution of the general case. We shall not present here Faddeev's analysis, but we want to discuss some attempts to generalize the off-energy shell-part of the preceding argument for $N=2$ (i.e., the part where $z \notin \sigma(H_o)$) to general N. These results are due mainly to Weinberg[17] and van Winter.[18]

The N-particle Green's function.

In this section, we work exclusively in the center-of-mass frame of the N-particle system: $\mathcal{K}=L^2(R^{3N-3})$ denotes the Hilbert-space corresponding to the internal degrees of freedom of the system. H, H_D, etc., are then considered as operators on $\mathcal{K}$. Theorem 7 still holds, with obvious modifications, and its corollary says that $\sigma(H_D) \subset \sigma(H)$ for any decomposition $D=(C_1 \ldots C_n)$ of the system $(1 \ldots N)$. If we want to indicate the number of clusters in D, we write D_n instead of D. For $n \geq 2$, $\sigma(H_{D_n})$ is continuous, extending from some real number to $+\infty$, since the relative kinetic energy of the non-interacting clusters can take any positive values. Therefore, $\sigma(H)$ contains the continuum

$$\sigma_c = \bigcup_{D_n, n \geq 2} \sigma(H_{D_n}) = [\varepsilon, +\infty)$$

for some $\varepsilon \leq 0$. In the following, we always tacitly assume that $z \notin \sigma_c$; that corresponds to the condition $z \notin \sigma(H_o)$ in the case $N=2$.

In order to apply Lemma 1, we want to derive an integral equation for the N-particle Green's function $G(z)=(z-H)^{-1}$ with a compact kernel. We start from the resolvent equation

$$G(z) = G_o(z) + G_o(z) V G(z),$$

and solve this by iteration:

$$G(z) = \sum_{n=o}^{\infty} \sum_{\ell_1 \ldots \ell_n} G_o V_{\ell_1} G_o \ldots V_{\ell_n} G_o. \qquad (V.14)$$

This is legitimate for sufficiently large negative $\mathrm{Re}\, z$, since, by (I.3) for $A=G_o$, $B=V_\ell$, there exists an $M<0$ such that

$$\| V_\ell G_o(z) \| < \binom{N}{2}^{-1} \quad \text{for } \mathrm{Re}\, z < M \text{ and all } \ell,$$

because V_ℓ is a Kato-potential. For these z, the series (V.14) is absolutely convergent and can therefore be rearranged in any way we please. We now analyze the series in terms of graphs: the term

$G_oV_{23}G_oV_{12}G_oV_{3N}G_o$, for example, is represented by the graph:

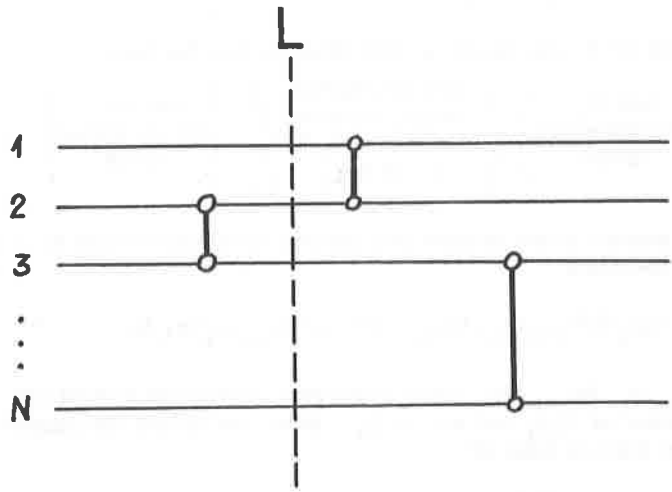


Figure 3

Classification of graphs:

1. Each graph g defines a decomposition D(g): two particles belong to the same cluster if their lines in g belong to the same connected part of g.

2. g is called D-disconnected if $D \subset D(g)$, i.e., if D(g) is obtained by further partitioning D. For $\mathrm{Re}\, z < M$, we obviously have

$$\sum (\text{all D-disconnected graphs}) = (z - H_D)^{-1} \equiv G_D(z).$$

3. Let us cut a graph g by a vertical line L (see Figure 3). Then D(L) denotes the decomposition defined by the subgraph to the left of L. First, let L be to the left of all interactions; then D(L) is the finest possible decomposition: $D_N = (1)(2)\ldots(N)$. Shifting now L gradually from left to right, until it is to the right of all interactions, D(L) takes a sequence of values

$$S(g) = \left(D_N, D_{N-1}, \ldots D_k \right), \qquad (V.15)$$

where $D_{i+1} \supset D_i$, and $D_k = D(g)$. So every graph g uniquely defines a sequence S(g) of type (V.15). On the other hand, a given sequence S of this type is generally obtained from an infinite number of graphs. For any such S, we define (for $\mathrm{Re}\, z < M$)

$$G_S(z) = \sum (\text{all } g \text{ with } S(g) = S).$$

By definition of S, any graph g with $S(g)=S$ has the form

$$\left[\prod_{i=N}^{k+1} \left(\begin{array}{c}\text{any } D_i\text{-}\\ \text{disconnected}\\ \text{graph}\end{array}\right)\left(\begin{array}{c}\text{any interaction}\\ \text{linking different}\\ \text{clusters of } D_i \text{ but}\\ \text{not of } D_{i-1}\end{array}\right)\right]\left(\begin{array}{c}\text{any } D_k\text{-}\\ \text{disconnected}\\ \text{graph}\end{array}\right),$$

where the factors in the product are ordered from left to right as i decreases. Therefore,

$$G_S(z) = G_{D_N}(z) V_{D_N D_{N-1}} G_{D_{N-1}}(z) \ldots V_{D_{k+1} D_k} G_{D_k}(z), \qquad (V.16)$$

where $V_{D_{i+1}D_i} = I_{D_{i+1}} - I_{D_i}$ = sum of all pair-potentials linking different clusters of D_{i+1} but not of D_i. Next, we define the disconnected part $D(z)$ of $G(z)$ by

$$D(z) = \sum_{\substack{S=(D_N \ldots D_k)\\ k\geq 2}} G_S(z) \qquad (V.17)$$

(= sum of all disconnected graphs), and the connected part $C(z)$ by

$$C(z) = \sum_{S=(D_N \ldots D_1)} G_S(z). \qquad (V.18)$$

Since $G(z)$ is the sum over all $G_S(z)$, we thus obtain $G(z)=D(z)+C(z)$, or

$$G(z) = D(z) + I(z)\, G(z), \qquad (V.19)$$

where

$$I(z) = \sum_{S=(D_N \ldots D_1)} G_{D_N}(z) V_{D_{N-1}} G_{D_{N-1}}(z) \ldots V_{D_2 D_1}. \qquad (V.20)$$

(This is obtained from (V.16) by noting that $G_{D_1}(z)=G(z)$, since D_1 is the trivial decomposition into one cluster.)

Now we note that $D(z)$ and $I(z)$ are holomorphic in $z \notin \sigma_c$, if defined by (V.16), (V.17) and (V.20). Therefore, the equation (V.19), which was established only for $\mathrm{Re}\, z < M$, extends by analyticity to all $z \notin \sigma(H)$.

The next step is to show that $I(z)$ is a Hilbert-Schmidt operator for $z \notin \sigma_c$, provided that $V_\ell(\cdot) \in L^2(R^3)$ for all ℓ.

Proof:[19] We consider a single term in (V.20) and write it for short as

$$G_N V_{N,N-1} G_{N-1} \cdots G_{k+1} V_{k+1,k} G_k \cdots V_{21}.$$

Let A_k be the product of the first factors up to, and including, G_k, and let $\mathcal{K}_k$ be the Hilbert space of the decomposed system with the Hamiltonian H_{D_k}, $D_k = (C_1 \ldots C_k)$, without the translational degrees of freedom of the independent subsystems C_i—i.e., $\mathcal{K}_k = L^2(R^{3N-3k})$. We now prove by induction that

$$\begin{aligned} &A_k \text{ is a Hilbert-Schmidt operator on } \mathcal{K}_k \\ &\text{for } z \notin \sigma(H_{D_k}), \text{ and} \\ &\|A_k(z)\|_{HS} \leq \text{const.} \, |\operatorname{Re} z|^{-1} \text{ as } \operatorname{Re} z \to -\infty. \end{aligned} \tag{V.21}$$

For $k = N$, R^{3N-3k} contains only the vector 0, so $\mathcal{K}_N$ = Hilbert space of complex numbers and $A_N(z) = z^{-1}$, which satisfies (V.21). Now, let $N \geq k > 1$ and suppose that (V.21) holds. Let V be a pair-potential linking the two clusters of D_k—say C_1 and C_2—which are united to a single cluster in D_{k-1}. By elementary computation one finds

$$\|A_k(z)V\|_{HS}^2 = \|V(\cdot)\|_2^2 \, 4\pi \int_0^\infty dp \, p^2 \|A_k\left(z - \frac{p^2}{2\mu}\right)\|_{HS}^2, \tag{V.22}$$

where the HS-norm on the left refers to $\mathcal{K}_{k-1}$, the one on the right to $\mathcal{K}_k$, and where μ is the reduced mass of the two clusters C_1, C_2. (This is simply the generalization of (V.5).) By (V.21), the integral converges and vanishes as $\operatorname{Re} z \to -\infty$. Using

$$\begin{aligned} \|A_{k-1}(z)\|_{HS} &= \|A_k(z) V_{k,k-1} G_{k-1}(z)\|_{HS} \\ &\leq \|A_k(z) V_{k,k-1}\|_{HS} \, \|G_{k-1}(z)\|, \end{aligned}$$

we see that $A_{k-1}(z)$ also satisfies (V.21), since

$$\|G_{k-1}(z)\| \leq \text{const.} \, |\operatorname{Re} z|^{-1}$$

as $\operatorname{Re} z \to -\infty$. For $k = 1$, (V.22) then shows that $I(z)$ is a Hilbert-Schmidt operator if $z \notin \sigma_c$.

*

Since $I(z)$ is holomorphic for $z \notin \sigma_c$ and vanishes (even in HS-norm) as $\mathrm{Re}\, z \to -\infty$, we can apply Lemma 1 to the functional equation (V.19) and obtain:

Theorem 13:[14] The part of $\sigma(H)$ in the complement of σ_c consists of eigenvalues only, which are of finite multiplicity and can accumulate at most at the lower end of σ_c.

Proof: The first part is immediate since, by (V.19) and Lemma 1, $G(z)$ is meromorphic in $z \notin \sigma_c$ ($D(z)$ is holomorphic!). To show the finite multiplicity, let $z_o \notin \sigma_c$ be a pole of $G(z)$. Then

$$\lim_{z \to z_o} G(z)(z - z_o) = P = \text{projection onto the subspace of eigenstates with eigenvalue } z_o.$$

Passing to the limit $z \to z_1$ in the equation (obtained from (V.19))

$$(z - z_o)G(z) = (z - z_o)D(z) + I(z)(z - z_o)G(z),$$

we obtain

$$P = I(z_o)P. \qquad (V.23)$$

Hence P is a compact projection and therefore of finite rank.

*

Remarks.

1. In Reference 14, Theorem 13 is proved if the pair-potentials are only locally square-integrable and vanish arbitrarily slowly at infinity. The fact that $\sigma_c \subset \sigma(H)$ then cannot be inferred from time-dependent scattering theory, but follows from the spatial cluster properties of the system.

2. Theorem 13 has some applications to the bound-state problem. First, it provides a basis for the customary perturbation formalism for bound states, which has been justified, notably by Kato,[20] for isolated eigenvalues of finite multiplicity. Secondly, one can derive from Theorem 13 some information about the behaviour of bound-state wave functions at infinity (in configuration space); they can be shown to belong to D_n (see (II.8)) for any n.[14]

*

Faddeev's equations.

Equation (V.23) shows that $(z - H)\psi = 0$, $z \notin \sigma_c$, implies

$$\psi = I(z)\psi.$$

The reverse, however, is not true for $N \geq 3$, i.e., the equivalence we had between (V.6) and (V.7) for $N = 2$ is lost. This is quite a serious drawback, since in the case $N = 2$ we actually used the equivalence with the (time-independent) Schrödinger equation to obtain control over the nontrivial solutions of the homogeneous integral equation. The merit of Faddeev's formulation is precisely that this equivalence is preserved.[16] He splits the three-particle Green's function into components:

$$G(z) = G_0(z) + \sum_\ell R_\ell(z),$$

where

$$R_\ell(z) = \sum (\text{all graphs with } V_\ell \text{ as the leftmost interaction})$$

$$= G_0 V_\ell G.$$

The components R_ℓ then satisfy the linked set of equations

$$R_\ell = G_0 V_\ell G_\ell + G_\ell V_\ell \sum_{n \neq \ell} R_m,$$

where $G_\ell = (z - H_0 - V_\ell)^{-1}$. This set of equations now has the two desired properties:

A) The second iterated kernel is connected, and thus defines a HS-operator for $z \notin \sigma_c$ (on the direct sum of three identical copies of $\mathcal{H}$).

B) The homogeneous equation

$$\psi_\ell = G_\ell(z) V_\ell \sum_{m \neq \ell} \psi_m$$

is equivalent to the Schrödinger equation

$$(z - H)\psi = 0,$$

where ψ and ψ_ℓ are connected by

$$\psi_\ell = G_0 V_\ell \psi$$

$$\psi = \sum_\ell \psi_\ell .$$

This allows the application of Lemma 1, which still holds if only some power of A(z) is compact.

*

Faddeev and Yakubowski[21] have found a generalization of Faddeev's equations to arbitrary N, which has the property that some power of the kernel is compact and for which the homogeneous equation is equivalent to the Schrödinger equation. Hopefully, a time-independent scattering theory for non-relativistic N-particle systems can be developed with this tool. Some progress in this direction has already been made by Hepp.[13]

References

1. T. Kato, Trans. A.M.S. 70, 195 (1951).
2. F. Riesz, B. Sz. Nagy, Functional Analysis, Frederick Unger, New York (1955).
3. E. Nelson, J. Math. Phys. 5, 332 (1964).
4. W. Hunziker, J. Math. Phys. 6, 6 (1965).
5. W. Hunziker, J. Math. Phys. 7, 300 (1966).
6. M. N. Hack, Nuovo Cim. 13, 231 (1959).
7. J. M. Jauch, Helv. Phys. Acta. 31, 661 (1958).
8. F. A. Berezin, L. D. Faddeev, R. A. Minlos, Proceedings of the Fourth All-Union Mathematical Conference held in Leningrad 1961.
9. J. Kupsch, W. Sandhas, Comm. Math. Phys. 2, 147 (1966).
10. J. D. Dollard, J. Math. Phys. 5, 729 (1964).
11. J. D. Dollard, thesis, Princeton University (1963).
12. R. T. Prosser, J. Math. Phys. 5, 708 (1964).
13. K. Hepp, private communication.
14. W. Hunziker, Helv. Phys. Acta. 39, 451 (1966).
15. T. Ikebe, Arch. Ratl. Mech. Anal. 5, 1 (1960).
16. L. D. Faddeev, Mathematical Aspects of the Three-Body Problem in the Quantum Theory of Scattering, Israel program of scientific translations, Jerusalem (1965).
17. S. Weinberg, Phys. Rev. 133, B232 (1964).
18. C. van Winter, Mat. Fys. Skr. Dan. Vid. Selsk. 2, No. 8 (1964 (1964); Mat. Fys. Skr. Dan. Vid. Selsk. 2, No. 10 (1965).
19. W. Hunziker, Phys. Rev. 135, B800 (1964).
20. T. Kato, Perturbation theory for linear operators, Springer, New York (1966).
21. Yakubowski, to appear in Soviet Nuclear Physics.

RIGOROUS RESULTS IN SCATTERING THEORY†

Jean-Paul Marchand‡
Institute of Theoretical Physics
University of Geneva
Switzerland

Abstract

We discuss the energy-dependence of scattering in Jauch's axiomatic frame of simple scattering theory by means of the simplest soluble mathematical models.

I. Introduction

Among the phenomena which can be described by (nonrelativistic) Quantum Mechanics the scattering processes are sufficiently distinguished as to justify a separate axiomatic treatment in accordance, of course, with the general principles of Quantum Mechanics.

A "time-dependent" picture of scattering is that of a wave-packet which moves freely in the remote past, then undergoes an interaction with some scattering center, and finally again moves freely. The so-called scattering operator would then relate the past and the future asymptotically free states, and the main problem of scattering theory would be to express this operator through the given interaction (or vice versa).

In the frame of an abstract Hilbert space formulation of Quantum Mechanics this idea has been expressed as follows:[1] Let H_o and V be two self-adjoint operators in the Hilbert space $\mathcal{H}$ corresponding to the kinetic energy of the free wave packets and the scattering interaction respectively. Then the system $\{H_o, H=H_o+V\}$ describes a scattering process (without bound states) if to any $\psi \in \mathcal{H}$ there exist two elements $\varphi^{in}, \varphi^{out} \in \mathcal{H}$ such that

$$\lim_{t \mapsto \{\mp\}} || e^{-iHt}\psi - e^{-iH_o t}\varphi^{\{in\ out\}} || = 0. \qquad (I.1)$$

---

†Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

‡Present address: Department of Mathematics and Physics, University of Denver, Denver, Colorado.

This asymptotic condition postulates that in norm the physical state $\psi_t = e^{-iHt}\psi$ tends to either of the asymptotic free states

$$\varphi_t^{\{\substack{\text{in}\\\text{out}}\}} = e^{-iH_0t}\varphi^{\{\substack{\text{in}\\\text{out}}\}} .$$

We may then define the two wave operators

$$\Omega_\pm : \varphi^{\{\substack{\text{in}\\\text{out}}\}} \to \psi ,$$

and the fundamental problem of expressing the scattering operator

$$S : \varphi^{\text{in}} \to \varphi^{\text{out}}$$

in terms of the interaction V is solved if we know how the wave operators depend on V.

A scattering system is, of course, not entirely described by the two energy-operators $\{H_0, H\}$ alone. But if we assume that there exists a set of other operators A_i such that both $\{H_0, A_i\}$ and $\{H, A_i\}$ form a complete set of commuting observables (for instance the angular momenta L and L_3 in the case of a spinless particle in a spherical potential) and that the spectra of A_i are discrete (as is the case for the quantum numbers ℓ, m), then both H_0 and H are reduced by the simultaneous eigenspaces of the operators A_i and none of the propagators e^{-iH_0t} and e^{-iHt} can lead the system outside these subspaces.

So the scattering problem may, in this case, be treated separately in each eigenspace of the other variables, the reductions of H_0 and H to these spaces having simple spectra. This holds, of course, only as long as we are merely interested in the energy-dependence of scattering.

The actual derivation of the scattering theory is most easily carried out, not in the abstract space $\mathcal{H}$, but in that space of square integrable (vector-) functions in which the free Hamiltonian H_0 appears as a multiplication operator, i.e., in the so-called direct integral of $\mathcal{H}$ with respect to H_0

$$\int \mathcal{H}(\lambda) d\lambda .$$

In this "spectral representation with respect to H_0" the S-operator which commutes with H_0 reduces to a direct integral of operators $\{S(\lambda)\}$ where each $S(\lambda)$ acts on the "energy shell" $\mathcal{H}(\lambda)$.

If we now assume that the spectrum Λ_0 of H_0 is simple (placing ourselves in a subspace with fixed other quantum numbers), the S-operator becomes a function of H_0 alone and acts in $\mathcal{L}^2$ as multiplication by $S(\lambda)$.

It is this scattering function depending on the energy which we shall derive in terms of the interaction and this is reached essentially in the following two steps: (1) The so-called scattering amplitude $R(\lambda)=S(\lambda)-1$ can be related to the diagonal elements of the "integral kernel" of the operator $V\Omega_+$. (2) This kernel can in turn be obtained from V alone by means of a singular integral equation for Ω.

What this exactly means will be explained in the first part of this series of lectures (Chapters III and IV), where the form of $S(\lambda)$ and the integral equation for Ω_+ will be rigorously derived. In the second part (Chapters V and VI) we solve the scattering problem explicitly for the simplest models which have the property that the kernel of $V\Omega_+$ is separable and that the integral equations for the wave operators become trivially (i.e., algebraically) soluble. This happens whenever the interaction operator V has a finite-dimensional range (so-called finite-rank potentials).

I find these easy explicit solutions of the scattering problem amusing and I think that, despite the point to be not "realistic," they provide a valuable information about the mathematical structure of the theory as well as about the fundamental ideas underlying its physical interpretation. In particular, the scattering amplitude and its analytic construction, the cross-section and the phase-shift may be discussed explicitly and the concept of resonance be given a precise meaning. Furthermore, the effect of bound states on these quantities can be clearly exhibited. The case of bound states in the free rather than the total Hamiltonian which "decay" under the influence of V will be given a special attention. We shall see how energy and life-time of these decays are related to the poles in the second sheet of the scattering amplitude and to the parameters of the resonances (energy and line width).

There is still another mathematical aspect of these explicit solutions of the scattering problem which may be interesting: they link the spectral representations with respect to H_o and H together. We may, for instance, define a particular spectral representation of a state $\psi \in \mathcal{H}$ with respect to H by putting it equal to the H_o -representation of the state φ^{in} which corresponds to it asymptotically in the remote past. Since then $\varphi^{in}=\Omega_+^*\psi$ and Ω_+^* is explicitly known in the H_o-representation in terms of the interaction V, the H-representation of ψ can be explicitly expressed by its H^o-representation and V.

This is, of course, a special case of the much more general mathematical problem of relating the spectral representations with respect to two arbitrary linear operators with equal spectra. But it is remarkable that in our particular models the transformation formulae appear in the form of Stieltjes-type integrals which are frequently encountered, for instance, in the theory of renormalization.

II. Mathematics

A. Spectral Representation

Consider a self-adjoint operator A in a Hilbert space $\mathcal{H}$ which has a simple spectrum Λ. Let $E(\Delta)$ be its spectral measure and $g \in \mathcal{H}$ a cyclic element with respect to A, i.e., an element satisfying $E(\Delta)g \neq 0$ for all Δ's. Then the function $\sigma(\Delta) \equiv \| E(\Delta)g \|^2$ is a numerically valued measure with maximal spectral type (with respect to A), and the following theorem holds:

There exists an isometric mapping h between the abstract Hilbert space $\mathcal{H}$ and the space $\mathcal{L}_\sigma^2(\Lambda)$ of σ-square integrable functions on the spectrum of A

$$\psi \in \mathcal{H} \quad \overset{h}{\longleftrightarrow} \quad \langle \lambda | \psi) \in \mathcal{L}_\sigma^2(\Lambda)$$

such that A becomes a multiplication operator in $\mathcal{L}_\sigma^2$:

$$A\psi \in \mathcal{H} \quad \overset{h}{\longleftrightarrow} \quad \lambda \langle \lambda | \psi) \in \mathcal{L}_\sigma^2(\Lambda) \, .$$

A function $u(A)$ of A may then be defined by

$$u(A)\psi \in \mathcal{H} \quad \longleftrightarrow \quad u(\lambda)\langle \lambda | \psi) \in \mathcal{L}_\sigma^2(\Lambda) \, ,$$

and the square integrable function $\langle \lambda | \psi)$ is called a spectral representation of the element $\psi \in \mathcal{H}$ with respect to the operator A. (For the proof cf. Reference 2, p. 197.)

The spectral representation with respect to A depends, of course, on the choice of the cyclic element g. Let g_1, g_2 be two cyclic elements with respect to A. Then the measures $\sigma_1(\Delta)$, $\sigma_2(\Delta)$ are absolutely continuous with respect to each other, i.e., they admit the same zero-subsets of Λ, and there exists between the two corresponding spectral representations a canonical isometry

$$ {}_1\langle \lambda | \psi) \in \mathcal{L}_{\sigma_1}^2(\Lambda) \longleftrightarrow \; {}_2\langle \lambda | \psi) = \sqrt{\rho(\lambda)} \, {}_1\langle \lambda | \psi) \in \mathcal{L}_{\sigma_2}^2(\Lambda) \, ,$$

where $\rho(\lambda)$ is the Radon-Nikodym derivative of $\sigma_1(\Delta)$ with respect to $\sigma_2(\Delta)$.

If the spectrum Λ of A is absolutely continuous, then all so-called distribution functions $\sigma(\lambda) \equiv \sigma((-\infty, \lambda))$ are absolutely continuous functions of λ and the derivative $(d\sigma(\lambda))/d\lambda$ exists. In this case we thus obtain a spectral representation in the space $\mathcal{L}^2(\Lambda)$ of Lebesgue square integrable functions.

We should note that the functions of $\mathfrak{L}^2_\sigma(\Lambda)$ are defined on Λ only up to subsets $\Lambda_o \subset \Lambda$ of σ-measure zero. This raises the following question: Let $\psi_n \in \mathcal{H}$ be a sequence of elements in $\mathcal{H}$ which converge strongly to an element $\psi \in \mathcal{H}$. What can be said about the convergence of the corresponding spectral representations $\langle \lambda | \psi_n \rangle$ and $\langle \lambda | \psi \rangle$? The relevant theorem[3] which is a consequence of Fatou's Lemma is the following:

Let $\psi(\lambda)$ be a particular representative of the class of square integrable functions corresponding to the element $\langle \lambda | \psi \rangle$ of $\mathfrak{L}^2$. Then it is possible to choose representatives $\psi_n(\lambda)$ of the elements $\langle \lambda | \psi_n \rangle$ such that a subsequence $\psi_{n_k}(\lambda)$ of $\psi_n(\lambda)$ converges almost everywhere pointwise to $\psi(\lambda)$. This theorem constitutes the adequate mathematical paraphrase of the operation of exchanging limits and "bras" and gives a precise meaning to the equation

$$\langle \lambda | \operatorname*{s-lim}_{n\to\infty} \psi_n \rangle = \lim_{n\to\infty} \langle \lambda | \psi_n \rangle .$$

We may now ask the question, "How does a linear operator T in $\mathcal{H}$ look in a spectral representation with respect to A?" To that purpose we have to investigate the functional relation between $\langle \lambda | \psi \rangle$ and $\langle \lambda | T\psi \rangle$. It may so happen that T acts as an integral operator with a kernel $K_T(\lambda, \lambda')$ which is an $\mathfrak{L}^2_\sigma$-function in both variables (i.e., in λ for almost all λ' fixed and in λ' for almost all λ fixed):

$$\langle \lambda | T\psi \rangle = \int_\Lambda K_T(\lambda, \lambda') \langle \lambda' | \psi \rangle d\sigma(\lambda') .$$

Such an operator is called a Carleman integral operator and we may use the notation $K_T(\lambda, \lambda') = \langle \lambda | T | \lambda' \rangle$ for its kernel and call $\langle \lambda | T | \lambda' \rangle$ the spectral representation of T with respect to A. In more general cases we may still use Dirac's notation $\langle \lambda | T | \lambda' \rangle$ but it then becomes a symbolical short-hand as seen, for instance, when writing down the "kernel" of the multiplication operator A itself: $\langle \lambda | A | \lambda' \rangle = \lambda \delta(\lambda - \lambda')$.

B. Change of Spectral Representation

One of the fundamental problems in scattering theory is that of connecting the spectral representations with respect to two different operators A and B which possess the same simple spectrum. A and B are then unitarily equivalent, i.e., there exist unitary operators U in $\mathcal{H}$ such that $B = UAU^{-1} = UAU^*$. If U is explicitly known in the spectral representation with respect to A, we may define a spectral representation with respect to B in the following natural way:

$$ {}_B\langle\lambda\,|\,\psi\rangle \equiv {}_A\langle\lambda\,|\,U^*\psi\rangle . $$

It is sufficient to verify that B appears as a multiplication operator in this new representation:

$$ {}_B\langle\lambda\,|\,B\psi\rangle = {}_A\langle\lambda\,|\,U^*B\psi\rangle = {}_A\langle\lambda\,|\,AU^*\psi\rangle = \lambda\,{}_A\langle\lambda\,|\,U^*\psi\rangle = \lambda\,{}_B\langle\lambda\,|\,\psi\rangle . $$

An analogous relation holds between the spectral representations of an operator T with respect to A and B. The exact statement is: If T is a Carleman integral operator with respect to B on a domain D_T, then U^*TU is a Carleman integral operator with respect to A on a domain U^*D_T and the relation

$$ {}_B\langle\lambda\,|\,T|\lambda'\rangle_B = {}_A\langle\lambda\,|\,U^*TU|\lambda'\rangle_A $$

holds. In fact, we have

$$ {}_A\langle\lambda\,|\,U^*TU\cdot U^*\psi\rangle = {}_A\langle\lambda\,|\,U^*T\psi\rangle = {}_B\langle\lambda\,|\,T\psi\rangle = \int_\Lambda {}_B\langle\lambda\,|\,T|\lambda'\rangle_B\,{}_B\langle\lambda'|\psi\rangle d\sigma(\lambda') $$

$$ = \int_\Lambda {}_B\langle\lambda\,|\,T|\lambda'\rangle_B\,{}_A\langle\lambda'\,|\,U^*\psi\rangle\,d\sigma(\lambda') $$

for all elements $U^*\psi$ in U^*D_T.

C. Abstract Integration in $\mathcal{H}$ [4]

At any given time t the state of a physical system may be described in $\mathcal{H}$ by a density operator W_t or, in simple cases, by a state vector $\psi_t \in \mathcal{H}$. The evolution of a state leads thus in a natural manner to the concept of vector-valued and operator-valued functions in $\mathcal{H}$. Many results in scattering theory can be formulated and derived in the abstract space and thus have the advantage of manifest representation independence. But to that purpose we need the concept of (abstract) integration of vector-valued functions.

Consider the time axis R as a measure space. The σ-ring Σ of measurable subsets of R are the Borel sets on the real line generated by countable union and intersection from the intervals, and the measure is the Lebesque measure generated from the natural lengths of the intervals. Using the standard concept of measurable numerical valued functions on R we may now define the integration of vector- and operator-valued functions in $\mathcal{H}$ as follows:

Definition II.1: A vector-valued function ψ_t is integrable with respect to t iff

(1) (φ, ψ_t) is a measurable numerical valued function of t for all $\varphi \in \mathcal{H}$.

(2) $\int \|\psi_t\| dt$ exists and is finite.

Definition II.2: An operator-valued function A_t whose range contains only bounded linear operators in $\mathcal{H}$ is integrable with respect to t iff $A_t\psi$ is t-integrable for all $\psi \in \mathcal{H}$.

Lemma II.1: If ψ_t is t-integrable then there exists $\psi_0 \in \mathcal{H}$ such that $\int_R (\varphi, \psi_t) dt = (\varphi, \psi_0)$ for all $\varphi \in \mathcal{H}$. ψ_0 is called the integral of ψ_t and we write $\psi_0 = \int_R \psi_t dt$.

Proof: The functional $F(\varphi) \equiv \int (\varphi, \psi_t) dt$ is bounded and linear on $\mathcal{H}$. In fact,

$$\int (\varphi, \psi_t) dt \leq \int |(\varphi, \psi_t)| dt \leq \int \|\varphi\| \cdot \|\psi_t\| dt = \|\varphi\| \int \|\psi_t\| dt < \infty;$$

the linearity is obvious. According to Riesz' theorem, there exists a $\psi_0 \in \mathcal{H}$ such that $F(\varphi) = (\varphi, \psi_0)$.

Lemma II.2: If A_t is t-integrable then there exists A_0 such that $\int_R A_t\psi \, dt = A_0\psi$ for all $\psi \in \mathcal{H}$. A_0 is called the integral of A_t and we write $A_0 = \int_R A_t dt$.

Proof: The form $F(\varphi, \psi) \equiv \int (\varphi, A_t\psi) dt$ is a bilinear functional on $\mathcal{H}$. Thus there exists a closed operator A_0 such that $F(\varphi, \psi) = (\varphi, A_0\psi)$. If the functional is bounded, i.e., if $|F(\varphi, \psi)| \leq M\|\varphi\| \cdot \|\psi\|$, then the closed linear operator A_0 is defined everywhere and thus bounded.

Lemma II.3: If ψ_t is t-integrable and T a bounded linear operator, then $T\int \psi_t dt = \int T\psi_t dt$.

Proof: First we show that $T\psi_t$ is integrable. From the integrability of ψ_t it follows that $(T^*\varphi, \psi_t)$ is a t-measurable function for all $\varphi \in \mathcal{H}$. So $(\varphi, T\psi_t)$ is measurable. Furthermore

$$\int \|T\psi_t\| dt \leq \int \|T\| \cdot \|\psi_t\| dt = \|T\| \int \|\psi_t\| dt < \infty,$$

and according to Definition II.1 the statement is proved. Now we have, according to the definition of the integral of $T\psi_t$:

$$\int (\varphi, T\psi_t) dt = (\varphi, \int T\psi_t dt).$$

On the other hand, using the definition of the t-integral of ψ_t

$$\int (\varphi, T\psi_t)dt = \int (T^*\varphi, \psi_t)dt = (T^*\varphi, \int \psi_t dt) = (\varphi, T\int \psi_t dt).$$

So $(\varphi, T\int \psi_t dt) = (\varphi, \int T\psi_t dt)$ for all $\varphi \in \mathcal{H}$ whence the result.

D. The Spectral Representation of Abstract Integrals

Combining the concepts of spectral representation and abstract integration, the following question arises: In what precise sense is it true that the spectral representation of the abstract integral of a vector-valued function equals the integral of the spectral representation? The answer is given by the following lemma which we cite without proof (cf. Reference 4).

Lemma II.4: Let ψ_t be a t-integrable vector-valued function in $\mathcal{H}$ and and $\langle\lambda|\psi_t\rangle$ its spectral representation in a function space $\mathfrak{L}^2_\sigma(\Lambda)$. Then Then there exists a numerical valued function $\psi(t,\lambda)$ on the direct product $R \times \Lambda$ such that for any fixed $t \in R$ $\psi(t,\lambda)$ is a representative of $\langle\lambda|\psi_t\rangle$ and $\int\psi(t,\lambda)dt$ is a representative of $\langle\lambda|\int\psi_t dt)$.

This theorem gives the precise meaning to the operation

$$\langle\lambda|\int\psi_t dt) = \int\langle\lambda|\psi_t\rangle dt$$

of exchanging bras and vector-integrals.

III. The Simple Scattering System

A. Definition

A simple scattering system is according to Jauch[1] described by a free Hamiltonian H_0 (kinetic energy of the non-interacting particles) and a total Hamiltonian $H = H_0 + V$ which satisfy the conditions:

(1) There exist two wave operators (=asymptotic condition)

$$\Omega_\pm \equiv \underset{t\to\mp\infty}{\text{s-lim}} \; V_t^* U_t$$

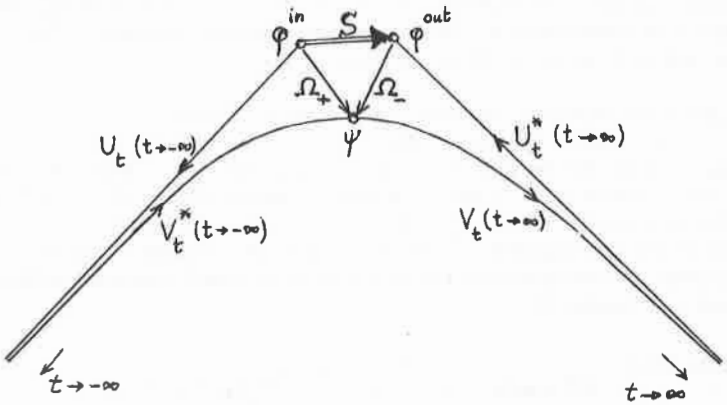
where $U_t \equiv e^{-iH_0t}$ and $V_t \equiv e^{-iHt}$ are the free and the physical propagator, respectively.

(2) There exists the scattering operator

$$S \equiv \Omega_-^*\Omega_+$$

with the same domain as Ω_+.

Condition (1) exactly expresses the fact mentioned already in the Introduction that (in the absence of bound states) to any physical state ψ there exist asymptotically free states φ^{in}, φ^{out} satisfying Eq.(I.1), and conversely. If we symbolize the free trajectories U_t by straight and the physical trajectories V_t by curved lines, this may be pictured as follows:



The scattering operator as defined in (2) takes then the required physical significance $S:\varphi^{in}\to\varphi^{out}$ mentioned in the Introduction.

The following properties of the wave and scattering operators are easily verified: Let P_0 and P be two projectors in Hilbert space which project on the subspaces $P_0\mathcal{H}$ and $P\mathcal{H}$ corresponding to the absolutely continuous parts of the spectra of H_0 and H, i.e., the maximal projectors reducing the operators H_0 and H in such a way that their respective parts P_0H_0 and PH have absolutely continuous spectra. Then

(1) $\Omega_\pm$ are partial isometries with domain $P_0\mathcal{H}$ and range $P\mathcal{H}$, and we have $\Omega_\pm^*\Omega_\pm = P_0$ and $\Omega_\pm\Omega_\pm^* = P$.

(2) $PH = \Omega_\pm(P_0H_0)\Omega_\pm^*$.

(3) The scattering operator exists on $P_0\mathcal{H}$, commutes with P_0H_0 and is unitary in $P_0\mathcal{H}$.

As we pointed out in the Introduction, the energy dependence of scattering can in many cases be treated in those subspaces of $\mathcal{H}$ which are obtained by fixing the value of all other quantum numbers (e.g., ℓ and m). In these subspaces, the reduction of the Hamiltonians H_0 and H become operators with simple spectra, and this will permit us to use the concept of spectral representation in the simple form introduced in Chapter II.

An immediate consequence is that any operator commuting with H_0 is a function of H_0 alone, and this holds in particular for the scattering operator, as already noted in the Introduction.

For simplicity we shall derive the scattering theory first under the additional assumption that no bound states exist in either H_0 or H. This means we pose $P_0 = P = I$. The wave operators now become unitary operators and establish the unitary equivalence of H_0 and H, which now have identical absolutely continuous spectra. This restriction will be removed at the end of Chapter IV.

B. Abstract Integral Formulae for the Wave Operators

We shall first derive some preparatory relations of scattering theory which can be expressed in a representation independent way. The first lemma relates the resolvent operator $R(z) = (H - z \cdot I)^{-1}$ of H to the propagator V_t and the second expresses the wave operators through the propagators. Their proof may be considered as good exercises in manipulating the concept of abstract integrals (introduced in Chapter II).

Lemma III.1:

$$R(\lambda + i\mu)\psi = i \int_0^\infty e^{i(\lambda + i\mu)t} V_t \psi \, dt \quad [\mu > 0].$$

Proof: In order to give a precise meaning to the right hand side, it is first of all necessary to show that the vector-valued function $\psi_t \equiv e^{i(\lambda + i\mu)t} V_t \psi$ satisfies the two conditions for integrability (cf. Definition II.1). This is easy since, (1) (φ, ψ_t) is a continuous function of t and thus measurable (for all $\varphi \in \mathcal{H}$), and (2)

$$\int_0^\infty \|\psi_t\| \, dt = \int_0^\infty e^{-\mu t} \|\psi\| \, dt \leq \frac{1}{\mu} \|\psi\| .$$

Now let E_λ be the spectral family of H and let us remind the representation

$$(\varphi, u(H)\psi) = \int u(\lambda) d(\varphi, E_\lambda \psi)$$

for the function $u(H)$. The left hand side in the lemma now becomes

$$\left(\varphi, R(\lambda + i\mu)\psi\right) = \int_\Lambda \frac{1}{\lambda' - \lambda - i\mu} \, d(\varphi, E_\lambda \psi) = \int_\Lambda \int_0^\infty e^{-i(\lambda' - \lambda - i\mu)t} d(\varphi, E_{\lambda'} \psi) dt$$

and the right hand side

$$\left(\varphi,\ i\int_o^\infty \psi_t dt\right) = i\int_o^\infty (\varphi,\psi_t)dt = i\int_o^\infty e^{i(\lambda+i\mu)t}(\varphi,V_t\psi)dt$$

$$= i\int_o^\infty dt\, e^{i(\lambda+i\mu)t}\int_\Lambda e^{-i\lambda' t} d(\varphi,E_{\lambda'}\psi).$$

Both sides coincide if the integrals can be exchanged in the last expression. But this follows from the fact that the second integral is bounded by the t-independent number $\|\varphi\|\cdot\|\psi\|$ and thus converges uniformly with respect to t.

A similar expression may be found for $R(\lambda - i\mu)$

Lemma III.2:

$$\Omega_-^*\psi = \underset{\varepsilon\to o}{\text{s-lim}}\ \varepsilon\int_o^\infty e^{-\varepsilon t}\, U_t^* V_t\psi\ dt.$$

Proof: The vector-valued function $\psi_t \equiv e^{-\varepsilon t} U_t^* V_t\psi$ is integrable. In fact, (φ,ψ_t) is measurable for all $\varphi\in\mathcal{H}$ and

$$\int_o^\infty \|\psi_t\| dt \le \|\psi\| \int_o^\infty e^{-\varepsilon t}\, dt = \frac{1}{\varepsilon}\|\psi\|.$$

Now

$$\left(\varphi,\ \underset{\varepsilon\to o}{\text{s-lim}}\ \varepsilon\int_o^\infty \psi_t dt\right) = \lim_{\varepsilon\to o}\ \varepsilon\int_o^\infty e^{-\varepsilon t}(\varphi,U_t^*V_t\psi)dt$$

for the right hand side and

$$(\varphi,\Omega_-^*\psi) = \lim_{t\to\infty}(\varphi,U_t^*V_t\psi)$$

for the left hand side. The equality follows from the observation that (Abel's theorem):

$$\lim_{t\to\infty} f(t) = \lim_{\varepsilon\to o}\ \varepsilon\int_o^\infty e^{-\varepsilon t} f(t)\, dt.$$

Similar representations may be found for $\Omega_\pm$ and Ω_+^*.[1)]

IV. Spectral Representation of the Simple Scattering System

A. General Remarks

We turn now to the solution of the main problem in scattering theory, namely that of calculating the scattering operator S from the

interaction V. (The converse problem consisting of establishing the so-called phase-equivalent potentials out of a given S-operator is, though mathematically very interesting, of minor physical importance.) As we already mentioned in the Introduction, this problem is essentially solved whenever the wave operators are known in terms of the interaction, S being easily related to the operator Ω_+. Another way of solving the scattering problem consisting of relating S to V and the resolvents (which themselves are functions of H_o and V) will also be sketched briefly.

These relations will be expressed in the spectral representation with respect to H_o, and there is no need to search for more abstract formulations since S is a function of H_o alone (for fixed other quantum numbers). The solution of the scattering problem, i.e., the knowledge of the wave operators in H_o-representation provides us, however, immediately with a natural way of passing from an H_o-representation to an H-representation, and conversely. In fact, let these representations be denoted by ${}_o\langle\lambda|\psi)$ and $\langle\lambda|\psi)$, respectively. Since the wave operators $\Omega_\pm$ establish the unitary equivalence of H_o and H by $H=\Omega_\pm^* H_o \Omega_\pm$, they also establish a change of spectral representation either from H_o to H:

$$ {}_\pm\langle\lambda|\psi) \equiv {}_o\langle\lambda|\Omega_\pm^*\psi), $$

${}_\pm\langle\lambda|\psi)$ being, according to Chapter II, two particular spectral representations with respect to H, or from H to H_o:

$$ {}^\pm_o\langle\lambda|\psi) \equiv \langle\lambda|\Omega_\pm\psi), $$

${}^\pm_o\langle\lambda|\psi)$ being two particular spectral representatives with respect to H_o. The scattering operator clearly relates the two representations ${}^\pm_o\langle\lambda|$ by

$$ {}^-_o\langle\lambda|S\psi) = {}^-_o\langle\lambda|\Omega_-^*\Omega_+\psi) = \langle\lambda|\Omega_+\psi) = {}^+_o\langle\lambda|\psi) $$

which may also be written

$$ {}^+_o\langle\lambda|\varphi^{in}) = {}^-_o\langle\lambda|\varphi^{out}). $$

So the asymptotic in-states look in (+)-representation like the out-states look in (-)-representation. Furthermore, if T is an integral operator in H_o-representation, $\Omega_\pm T \Omega_\pm^*$ are integral operators in the two H-representations related by

$$ {}_o\langle\lambda|T|\lambda'\rangle_o = {}_\pm\langle\lambda|\Omega_\pm T\Omega_\pm^*|\lambda'\rangle_\pm , $$

and conversely.

We shall first derive a basic formula (IV.1) which may be considered the key to the whole problem.

B. The Key Formula

Let us start from the

Lemma IV.1: $${}_o\langle\lambda|\Omega_-^*\psi) = -i \lim_{\varepsilon\to o} \varepsilon\ {}_o\langle\lambda|R(\lambda+i\varepsilon)\psi).$$

Proof: We use the interchange of strong limits and t-integrals in $\mathcal{H}$ with the bras of the spectral representation in the precise sense established in Chapter II. According to Lemma III.1 the right hand side becomes

$$-i \lim_{\varepsilon\to o} \varepsilon\ {}_o\langle\lambda|R(\lambda+i\varepsilon)\psi) = -i \lim_{\varepsilon\to o} \varepsilon\ {}_o\langle\lambda|i\int_o^\infty e^{i(\lambda+i\varepsilon)t}V_t\psi\,dt)$$

$$= \lim_{\varepsilon\to o} \varepsilon \int_o^\infty e^{i(\lambda+i\varepsilon)t}\ {}_o\langle\lambda|V_t\psi)dt$$

and the left hand side, according to Lemma III.2,

$${}_o\langle\lambda|\Omega_-^*\psi) = {}_o\langle\lambda|s\text{-}\lim_{\varepsilon\to o} \varepsilon\int_o^\infty d^{-\varepsilon t}U_t^*V_t\psi\,dt) = \lim_{\varepsilon\to o} \varepsilon\int_o^\infty e^{-\varepsilon t}\,{}_o\langle\lambda|U_t^*V_t\psi)dt$$

$$= \lim_{\varepsilon\to o} \varepsilon\int_o^\infty d^{i(\lambda+i\varepsilon)t}\,{}_o\langle\lambda|V_t\psi)dt.$$

This establishes the relation. A similar one can be found for Ω_+^*.

Consider now the two Hamiltonians H_o and $H = H_o + V$ and their corresponding resolvents $R_o(z) = (H_o - z\cdot I)^{-1}$ and $R(z) = (H - z\cdot I)^{-1}$ which are related through the so-called second resolvent relation $R(z) = R_o(z)(I - VR(z)) = (I + R(z)V)R_o(z)$. Using the fact that $R_o(z)$ is diagonal in H_o-representation, we may now recast the relation of Lemma IV.1 and its analogue for Ω_+^* into the following appealing forms:

$${}_o\langle\lambda|\Omega_\pm^*\psi) = {}_o\langle\lambda|\psi) - \lim_{\varepsilon\to o}\ {}_o\langle\lambda|VR(\lambda\mp i\varepsilon)\psi) \qquad \text{(IV.1)}$$

There exists a dual equation to (IV.1). In fact, if (H_o, H) is a scattering system, (H, H_o) is one, too, and all relations remain unchanged provided the substitutions $R_o \longleftrightarrow R$, $V \longleftrightarrow -V$, $\Omega_\pm \longleftrightarrow \Omega_\pm^*$ are simultaneously carried out and that the H_o-representation ${}_o\langle\lambda|$ is replaced by the H-representations ${}_\pm\langle\lambda|$. With these substitutions,

(IV.1) transforms into

$$ {}_{\pm}\langle\lambda|\Omega_{\pm}\psi) = {}_{\pm}\langle\lambda|\psi) + \lim_{\epsilon\to o} {}_{\pm}\langle\lambda|VR_o(\lambda \mp i\epsilon)\psi). \qquad \text{(IV.1')} $$

Relations (IV.1) and (IV.1') may be called the key formulae of scattering theory since they contribute the natural starting point to the derivation of all the essential results such as the Lippmann-Schwinger integral equations for the wave operators and the scattering amplitudes. It should, however, be stressed once more that the form of these relations is essentially representation dependent and that they cannot be converted into pure operator relations between resolvents and wave operators, since the parameters λ appearing in the resolvents are explicitly linked to the spectral variable λ.

C. Integral Equations for the Wave Operators

Up to this point all results were derived under the sole assumption that H_o and H form a scattering system, i.e., that the wave operators $\Omega_{\pm}$ exist. Let us now make the additional mathematical assumption that the operators $V\Omega_{\pm}$ and $\Omega_{\pm}^{*}V$ are Carleman integral operators in the H_o-representation. This, of course, implicitly restricts the possible forms of the interaction V and has to be verified separately in any explicit model.

An immediate consequence is, according to Section (A), that $\Omega_{\pm}V$ and $V\Omega_{\pm}^{*}$ are integral operators in the H-representation and that these different representations are linked by

$$ {}_{\pm}\langle\lambda|\Omega_{\pm}V|\lambda'\rangle_{\pm} = {}_{o}\langle\lambda|V\Omega_{\pm}|\lambda'\rangle_{o}; \quad {}_{\pm}\langle\lambda|V\Omega_{\pm}^{*}|\lambda'\rangle_{\pm} = {}_{o}\langle\lambda|\Omega_{\pm}^{*}V|\lambda'\rangle_{o}. \qquad \text{(IV.2)} $$

From the key formulas (IV.1) and (IV.1') we immediately derive the following integral relations for the wave operators:

(1) Substituting $\Omega_{\pm}\psi$ into ψ in (IV.1)

$$ {}_{o}\langle\lambda|\Omega_{\pm}\psi) = {}_{o}\langle\lambda|\psi) + \lim_{\epsilon\to o} {}_{o}\langle\lambda|VR(\lambda \mp i\epsilon)\Omega_{\pm}\psi) $$

$$ = {}_{o}\langle\lambda|\psi) + \lim_{\epsilon\to o} {}_{o}\langle\lambda|V\Omega_{\pm}R_o(\lambda \mp i\epsilon)\psi) $$

$$ = {}_{o}\langle\lambda|\psi) + \lim_{\epsilon\to o} \int \frac{{}_{o}\langle\lambda|V\Omega_{\pm}|\lambda'\rangle_{o}}{\lambda'-\lambda\pm i\epsilon} \, {}_{o}\langle\lambda'|\psi)d\lambda'. $$

(2) Translating (IV.1) into the H-representations ${}_{\pm}\langle\lambda|$ according to Section (A):

$$
\begin{aligned}
{}_{\pm}\langle\lambda|\Omega_{\pm}\psi\rangle &= {}_{\pm}\langle\lambda|\psi\rangle + \lim_{\varepsilon\to 0}\ {}_{\pm}\langle\lambda|\Omega_{\pm}VR(\lambda \mp i\varepsilon)\psi\rangle \\
&= {}_{\pm}\langle\lambda|\psi\rangle + \lim_{\varepsilon\to 0}\int \frac{{}_{\pm}\langle\lambda|\Omega_{\pm}V|\lambda'\rangle_{\pm}}{\lambda' - \lambda \pm i\varepsilon}\ {}_{\pm}\langle\lambda'|\psi\rangle d\lambda' = {}_{0}\langle\lambda|\psi\rangle .
\end{aligned}
\tag{IV.3}
$$

(3) Substituting $\Omega_{\pm}^{*}\psi$ into ψ in (IV.1'):

$$
\begin{aligned}
{}_{\pm}\langle\lambda|\Omega_{\pm}^{*}\psi\rangle &= {}_{\pm}\langle\lambda|\psi\rangle - \lim_{\varepsilon\to 0}\ {}_{\pm}\langle\lambda|V\Omega_{\pm}^{*}R(\lambda \mp i\varepsilon)\psi\rangle \\
&= {}_{\pm}\langle\lambda|\psi\rangle - \lim_{\varepsilon\to 0}\ {}_{\pm}\langle\lambda|V\Omega_{\pm}^{*}R(\lambda \mp i\varepsilon)\psi\rangle \\
&= {}_{\pm}\langle\lambda|\psi\rangle - \lim_{\varepsilon\to 0}\int \frac{{}_{\pm}\langle\lambda|V\Omega_{\pm}^{*}|\lambda'\rangle_{\pm}}{\lambda' - \lambda \pm i\varepsilon}\ {}_{\pm}\langle\lambda'|\psi\rangle d\lambda' .
\end{aligned}
$$

(4) Translating (IV.1') into the H_0-representation ${}_{0}\langle\lambda|$ according to Section (A):

$$
\begin{aligned}
{}_{0}\langle\lambda|\Omega_{\pm}^{*}\psi\rangle &= {}_{0}\langle\lambda|\psi\rangle - \lim_{\varepsilon\to 0}\ {}_{0}\langle\lambda|\Omega_{\pm}^{*}VR_{0}(\lambda \mp i\varepsilon)\psi\rangle \\
&= {}_{0}\langle\lambda|\psi\rangle - \lim_{\varepsilon\to 0}\int \frac{{}_{0}\langle\lambda|\Omega_{\pm}^{*}V|\lambda'\rangle_{0}}{\lambda' - \lambda \pm i\varepsilon}\ {}_{0}\langle\lambda'|\psi\rangle = {}_{\pm}\langle\lambda|\psi\rangle .
\end{aligned}
\tag{IV.3'}
$$

Relations (IV.3) and (IV.3') are transformation formulae between the spectral representations ${}_{0}\langle\lambda|\psi\rangle$ and ${}_{\pm}\langle\lambda|\psi\rangle$, which are, however, still implicit in the sense that they are not expressed by the interaction V alone but rather by the kernels $\Omega_{\pm}V$ and $\Omega_{\pm}^{*}V$ where $\Omega_{\pm}$ and $\Omega_{\pm}^{*}$ themselves depend on V.

D. The Scattering Operator in Terms of the Resolvents[5)]

Consider the operator

$$
R \equiv S - I = \Omega_{-}^{*}\Omega_{+} - I = \operatorname*{s-lim}_{t\to -\infty} \Omega_{-}^{*}V_{t}^{*}U_{t} - I = \operatorname*{s-lim}_{t\to -\infty} U_{t}^{*}(\Omega_{-}^{*} - I)U_{t} .
$$

Using Lemma IV.1 and noting that

$$i \lim_{\varepsilon \to o} {}_o\langle \lambda \,|\, R_o(\lambda + i\varepsilon)\psi \rangle = {}_o\langle \lambda \,|\, \psi \rangle$$

we obtain for the H_o-representation of R:

$$\begin{aligned}
{}_o\langle \lambda \,|\, R\psi \rangle &= \lim_{t \to -\infty} e^{i\lambda t}\, {}_o\langle \lambda \,|\, (\Omega_-^* - I) U_t \psi \rangle \\
&= - \lim_{t \to -\infty} e^{i\lambda t} \lim_{\varepsilon \to o} i\varepsilon\, {}_o\langle \lambda \,|\, \Big(R(\lambda + i\varepsilon) - R_o(\lambda + i\varepsilon) \Big) U_t \psi \rangle \\
&= - \lim_{t \to \infty} e^{i\lambda t} \lim_{\varepsilon \to o} i\varepsilon\, {}_o\langle \lambda \,|\, R_o(\lambda + i\varepsilon) V R(\lambda + i\varepsilon) U_t \psi \rangle \\
&= \lim_{t \to -\infty} e^{i\lambda t} \lim_{\varepsilon \to o} {}_o\langle \lambda \,|\, V R(\lambda + i\varepsilon)\, U_t \psi \rangle .
\end{aligned}$$

We note that this relation is a rigorous consequence of the axioms and does not make any use of the additional assumption about integral kernels. But if we now use once more the second resolvent equation and assume that V and $VR(\lambda + i\varepsilon)V$ are expressible as integral operators in H_o-representation, we can develop the expression further:

$$\begin{aligned}
{}_o\langle \lambda \,|\, R\psi \rangle &= \lim_{t \to -\infty} e^{i\lambda t} \lim_{\varepsilon \to o} {}_o\langle \lambda \,|\, V\Big(I - R(\lambda + i\varepsilon) V \Big) R_o(\lambda + i\varepsilon) U_t \psi \rangle \\
&= \lim_{t \to -\infty} \lim_{\varepsilon \to o} \int {}_o\langle \lambda \,|\, V - VR(\lambda + i\varepsilon)V \,|\, \lambda' \rangle \, \frac{e^{-i(\lambda' - \lambda)t}}{\lambda' - \lambda - i\varepsilon}\, {}_o\langle \lambda' \,|\, \psi \rangle d\lambda'
\end{aligned}$$

and using

$$\lim_{t \to -\infty} \lim_{\varepsilon \to o} \int \frac{e^{-i(\lambda' - \lambda)t}}{\lambda' - \lambda - i\varepsilon} \varphi(\lambda') d\lambda' = -2\pi i\, \varphi(\lambda)$$

for some class of test-functions φ, we obtain the known formula

$${}_o\langle \lambda \,|\, R\psi \rangle = -2\pi i\, {}_o\langle \lambda \,|\, V - VR(\lambda + io)V \,|\, \lambda \rangle_o\, {}_o\langle \lambda \,|\, \psi \rangle .$$

The validity-question for these manipulations needs, of course, more care.[6]

E. The Scattering Operator in Terms of $V\Omega_+$

If we substitute $\Omega_+\psi$ for ψ in the key formula (IV.1), we obtain another representation for the scattering operator:

$$
\begin{aligned}
{}_o\langle\lambda|S\psi) &= {}_o\langle\lambda|\Omega_-^*\Omega_+\psi) = {}_o\langle\lambda|\Omega_+\psi) + \lim_{\varepsilon\to o}\ {}_o\langle\lambda|VR(\lambda+i\varepsilon)\Omega_+\psi) \\
&= {}_o\langle\lambda|\psi) - \lim_{\varepsilon\to o}\ {}_o\langle\lambda|V\Omega_+R_o(\lambda-i\varepsilon)\psi) + \lim_{\varepsilon\to o}\ {}_o\langle\lambda|V\Omega_+R_o(\lambda+i\varepsilon)\psi),
\end{aligned}
$$

and if we again assume that $V\Omega_+$ is an integral operator in H_o-representation

$$
\begin{aligned}
{}_o\langle\lambda|S\psi) &= {}_o\langle\lambda|\psi) + \lim_{\varepsilon\to o}\int {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o\left(\frac{1}{\lambda'-\lambda-i\varepsilon} - \frac{1}{\lambda'-\lambda-i\varepsilon}\right) {}_o\langle\lambda'|\psi)d\lambda' \\
&= {}_o\langle\lambda|\psi) + \lim_{\varepsilon\to o}\int {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o\ \frac{-2i\varepsilon}{(\lambda'-\lambda)^2+\varepsilon^2}\ {}_o\langle\lambda'|\psi)d\lambda' \\
&= {}_o\langle\lambda|\psi) - 2\pi i\ {}_o\langle\lambda|V\Omega_+|\lambda\rangle_o\ {}_o\langle\lambda|\psi).
\end{aligned}
$$

We see that S is diagonal in H_o as was to be expected from the fact that these two operators commute. The above relation may thus symbolically be written

$$
{}_o\langle\lambda|S|\lambda'\rangle_o = \delta(\lambda-\lambda')S(\lambda); \quad S(\lambda) = 1 - 2\pi i\,{}_o\langle\lambda|V\Omega_+|\lambda\rangle_o \qquad \text{(IV.4)}
$$

where $S(\lambda)$ is the S-matrix at energy λ. (In our simplification in which all other variables have been dropped these matrices reduce, of course, to ordinary numbers.)

F. Scattering Amplitude, Cross-Section and Phase-Shift

We now define those concepts which relate the scattering "matrix" $S(\lambda)$ to experiments—in principle. They are, like S, functions of the energy alone. Furthermore, since we are mainly interested in the mathematical structure and merely the general aspects of their physical interpretation we shall introduce them as dimensionless quantities. Note, for instance, that the dimension of the total cross-section depends on the dimension of the space in which the scattering process takes place while our treatment of scattering should not depend on that dimension. So we define:

(1) Scattering Amplitude:
(not to be confused with the resolvent R(z))

$$R(\lambda) \equiv S(\lambda) - 1 = -2\pi i \, {}_o\langle \lambda | V\Omega_+ | \lambda \rangle_o . \qquad \text{(IV.5a)}$$

(2) (Total) Cross Section:

$$\sigma(\lambda) = \tfrac{1}{4} |R(\lambda)|^2 = \pi^2 \left| {}_o\langle \lambda | V\Omega_+ | \lambda \rangle_o \right|^2 . \qquad \text{(IV.5b)}$$

(3) Phase-Shift:

$$\delta(\lambda) \equiv \frac{1}{2i} \ln S(\lambda) . \qquad \text{(IV.5c)}$$

The scattering amplitude and the cross-section are then, in terms of the phase shift:

$$R(\lambda) = S(\lambda) - 1 = e^{2i\delta(\lambda)} - 1 = 2i e^{i\delta(\lambda)} \sin\delta(\lambda); \quad \sigma(\lambda) = \sin^2 \delta(\lambda) . \qquad \text{(IV.6)}$$

G. Modifications in Presence of Bound States

In the definition of the simple scattering system we postulated the existence of the two wave operators

$$\Omega_\pm = \underset{t \to \mp\infty}{\text{s-lim}} \, V_t^* U_t .$$

They have the property of being partial isometrics connecting the absolutely continuous parts $P_o H_o$ and PH of the two Hamiltonians. We then supposed for commodity that H_o and H contain no bound states, i.e., that $P_o = P = I$. We shall now drop this restriction and ask the question how the scattering amplitude and the cross-section are modified in the presence of bound states.

In reinterpreting the derivations of Chapter III the following remark suffices: The domains $P_o\mathfrak{H}$ and $P\mathfrak{H}$ of the wave operators $\Omega_\pm$ and $\Omega_\pm^*$ can be extended to the entire space $\mathfrak{H}$ by postulating that they should project into zero those components which are orthogonal to P_o and P. Being pedantic, this could be systematically achieved by replacing Ω and Ω^* by $P\Omega P_o$ and $P_o\Omega^* P$. It is then easily seen that all abstract formulae of Chapter III remain valid.

The situation becomes less trivial for the relations in this chapter which have been expressed in the spectral representations with respect to H_o and H.

Any element $\psi \in \mathfrak{H}$ can be decomposed in two ways according to

$$\psi = P_o\psi + (I - P_o)\psi = P\psi + (I - P)\psi .$$

The continuous parts $P_o\psi$ and $P\psi$ are still spectrally represented by square integrable functions ${}_o\langle\lambda|\psi\rangle$ and $\langle\lambda|\psi\rangle$ such that H_o and H act as multiplication operators, and the interrelation of these representations is still

$$ {}_\pm\langle\lambda|\psi\rangle = {}_o\langle\lambda|\Omega_\pm^*\psi\rangle; \quad {}_o\langle\lambda|\psi\rangle = {}_\pm\langle\lambda|\Omega_\pm\psi\rangle; $$

but no information about the discrete parts of the element ψ is contained in these relations. In particular an operator T can be an integral operator with respect to H_o and H only in the continuous subspaces $P_o\mathfrak{H}$ and $P\mathfrak{H}$. In these subspaces we still have the property that if T is an integral operator in the H_o-representation in the space $P_o\mathfrak{H}$ then $\Omega_\pm T\Omega_\pm^*$ are integral operators in the H-representations in the space $P\mathfrak{H}$, and their kernels are related by

$$ {}_o\langle\lambda|T|\lambda'\rangle_o = {}_\pm\langle\lambda|\Omega_\pm T\Omega_\pm^*|\lambda'\rangle_\pm . $$

It follows from this that the integral relations for the wave operators derived in Section (C) are to be regarded as relations which hold only between elements of the spaces $P_o\mathfrak{H}$ or $P\mathfrak{H}$ in which the integral kernels act. This yields an essential change in formulae (IV.3) and (IV.3′) which now read

$$ {}_\pm\langle\lambda|\Omega_\pm P\psi) = {}_\pm\langle\lambda|\psi) + \int \frac{{}_\pm\langle\lambda|\Omega_\pm V|\lambda'\rangle_\pm}{\lambda' - \lambda \pm i0} \; {}_\pm\langle\lambda'|\psi)d\lambda' = {}_o\langle\lambda|P\psi) \qquad \text{(IV.7)} $$

$$ {}_o\langle\lambda|\Omega_\pm^* P_o\psi) = {}_o\langle\lambda|\psi) - \int \frac{{}_o\langle\lambda|\Omega_\pm^* V|\lambda'\rangle_o}{\lambda' - \lambda \pm i0} \; {}_o\langle\lambda'|\psi)d\lambda' = {}_\pm\langle\lambda|P_o\psi). \qquad \text{(IV.7′)} $$

Equation (IV.4) for the S-matrix remains formally unchanged, but the "kernel" ${}_o\langle\lambda|S|\lambda'\rangle_o$ should be considered as acting only between elements in $P_o\mathfrak{H}$ in accordance to the fact that $S = \Omega_-^*\Omega_+$ vanishes on the discrete part of H_o and is unitary only in $P_o\mathfrak{H}$.

The natural question which arises here is that of the effect produced by the bound states in the scattering amplitude. Before treating particular models, let us remember Levinson's Theorem which can be proved in the general context of this theory.[7] It states that, for a continuous spectrum of H_o extending from 0 to ∞

$$ \delta(\infty) - \delta(0) = \pi(n_o - n) \qquad \text{(IV.8)} $$

where n_o and n are the dimensions of the spaces $(I - P_o)\mathfrak{H}$ and $(I - P)\mathfrak{H}$, i.e., the number of bound states of H_o and H. An analysis

of the scattering amplitude thus indeed furnishes information about the bound states.

The two main bound-state situations arising in physics are those where either H_0 or H have bound states. Only the second one corresponds to what usually is called a bound state system, while the first one may be, in case the eigenvalues λ_0 are imbedded in the continuum of H_0, considered as a system describing the decay of the H_0-bound states χ corresponding to the energies λ_0 through the interaction V. This decay takes place according to a decay law

$$P(t) = \left(\chi, e^{-iHr}\chi\right) \qquad \text{(IV.9)}$$

whose absolute square expresses, for a fixed time t, the probability that the system it found in the undecayed state χ at time t, if it was created in that state at time zero.

V. A Soluble Model without Boundstate in H_0

A. Definition

We shall now treat the simplest non-trivial model of a scattering system which can be calculated entirely, without any approximation techniques. The general expression (IV.4) for the scattering matrix is still implicit insofar as it contains the wave operator Ω_+ which itself depends on the interaction V. The crux of such a soluble model is to permit a separation of the integral kernel of $V\Omega_+$ into the two kernels of V and Ω_+ such that the integral relation for Ω_+ established in Chapter IV can be solved.

Let H_0 be a free energy Hamiltonian whose spectrum is supposed absolutely continuous and connected (the real axis, for instance). The interaction V will simply be a rank one projection operator P_φ which projects the Hilbert space $\mathcal{H}$ onto the one-dimensional subspace generated by the element $\varphi \in \mathcal{H}$ which we suppose to be normalized. According to a theorem of T. Kato (Reference 8, p. 540), the system $(H_0, H = H_0 + P_\varphi)$ is a simple scattering system, i.e., there exist wave operators $\Omega_\pm$ which are partial isometries connecting the absolutely continuous parts of the operators H_0 and H by

$$PH = \Omega_\pm H_0 \Omega_\pm^*$$

where the operator P projects away the possible discrete eigenspaces of H. We are thus placed in the case $P_0 = I$ and P arbitrary.

B. Solution of the Integral Relations for the Wave Operators

First we prove that our system satisfies the additional mathematical postulate introduced in Chapter IV.

Lemma V.1: The operators $V\Omega_+$ and Ω_+^*V are Carleman integral operators in the H_o-representation, and their kernels are

$$ {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o = {}_o\langle\lambda|\varphi\rangle\,\overline{{}_+\langle\lambda'|\varphi\rangle}\,; \quad {}_o\langle\lambda|\Omega_+^*V|\lambda'\rangle_o = {}_+\langle\lambda|\varphi\rangle\,\overline{{}_o\langle\lambda'|\varphi\rangle}\,. \tag{V.1} $$

Proof: The H_o-representation of $V\Omega_+\psi$ may be directly calculated:

$$ {}_o\langle\lambda|V\Omega_+\psi\rangle = {}_o\langle\lambda|(\varphi,\Omega_+\psi)\varphi\rangle = (\varphi,\Omega_+\psi)\,{}_o\langle\lambda|\varphi\rangle = (\Omega_+^*\varphi,\psi)\,{}_o\langle\lambda|\varphi\rangle $$

$$ = \int {}_o\langle\lambda|\varphi\rangle\,\overline{{}_o\langle\lambda'|\Omega_+^*\varphi\rangle}\,{}_o\langle\lambda'|\psi\rangle d\lambda' = \int {}_o\langle\lambda|\varphi\rangle\,{}_+\langle\lambda'|\varphi\rangle\,{}_o\langle\lambda'|\psi\rangle\,. $$

Hence,

$$ {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o = {}_o\langle\lambda|\varphi\rangle\,\overline{{}_+\langle\lambda'|\varphi\rangle}\,. $$

The Carleman property of this kernel follows from

$$ \int |{}_o\langle\lambda|V\Omega_\pm|\lambda'\rangle_o|^2 d\lambda = |{}_+\langle\lambda'|\varphi\rangle|^2 \int |{}_o\langle\lambda|\varphi\rangle|^2 d\lambda = |{}_+\langle\lambda'|\varphi\rangle|^2 \cdot \|\varphi\|^2 $$

which is finite for almost all λ', and

$$ \int |{}_o\langle\lambda|V\Omega_\pm|\lambda'\rangle_o|^2 d\lambda' = |{}_o\langle\lambda|\varphi\rangle|^2 \int |{}_+\langle\lambda'|\varphi\rangle|^2 d\lambda' = |{}_o\langle\lambda|\varphi\rangle|^2 \cdot \|P\varphi\|^2 $$

which is finite for almost all λ.

In order to verify the second part, we may first calculate $V\Omega_+^*P\psi$ in the H-representation:

$$ {}_+\langle\lambda|V\Omega_+^*P\psi\rangle = (\varphi,\Omega_+^*P\psi)\,{}_+\langle\lambda|\varphi\rangle = (\Omega_+\varphi,P\psi)\,{}_+\langle\lambda|\varphi\rangle $$

$$ = \int {}_+\langle\lambda|\varphi\rangle\,\overline{{}_+\langle\lambda'|\Omega_+\varphi\rangle}\,{}_+\langle\lambda'|\psi\rangle d\lambda' = \int {}_+\langle\lambda|\varphi\rangle\,\overline{{}_o\langle\lambda'|\varphi\rangle}\,{}_+\langle\lambda'|\psi\rangle d\lambda' \,, $$

hence the kernel

$$ {}_+\langle\lambda|V\Omega_+^*|\lambda'\rangle = {}_+\langle\lambda|\varphi\rangle\,\overline{{}_o\langle\lambda'|\varphi\rangle} $$

which again satisfies the Carleman properties. But according to Eq. (IV.2),

$$ {}_+\langle\lambda|V\Omega_+^*|\lambda'\rangle_+ = {}_o\langle\lambda|\Omega_+^*V|\lambda'\rangle_o\,. $$

The explicit form (V.1) for the integral kernels still has the disadvantage to contain the element φ in two different spectral representations. But Eq. (IV.7') (or (IV.3') since $P_o = I$ in our model) which links the two representations reduces now, for the particular element φ, to an algebraic equation. In fact, substituting φ for ψ in (IV.7') and inserting the kernels (V.1) we can solve for ${}_+\langle\lambda|\varphi\rangle$:

$$ {}_+\langle\lambda|\varphi\rangle = \frac{{}_o\langle\lambda|\varphi\rangle}{1 + \int \frac{|{}_o\langle\lambda'|\varphi\rangle|^2}{\lambda' - \lambda + io}\, d\lambda'} . \qquad (V.2) $$

Inserting (V.2) into (V.1), the integral kernels read now, in terms of ${}_o\langle\lambda|\varphi\rangle$ alone:

$$ {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o = \frac{{}_o\langle\lambda|\varphi\rangle\overline{{}_o\langle\lambda'|\varphi\rangle}}{1 + \int \frac{|{}_o\langle\lambda''|\varphi\rangle|^2\, d\lambda''}{\lambda'' - \lambda' - io}}; \quad {}_o\langle\lambda|\Omega_+^* V|\lambda'\rangle_o = \frac{{}_o\langle\lambda|\varphi\rangle\overline{{}_o\langle\lambda'|\varphi\rangle}}{1 + \int \frac{|\langle\lambda'|\varphi\rangle|^2\, d\lambda'}{\lambda' - \lambda + io}} . $$

(V.3)

With this all spectral representations of the wave operators can now be expressed in terms of the spectral function ${}_o\langle\lambda|\varphi\rangle$ which is explicitly known for the given interaction $V = P_\varphi$.

Instead of expressing the H-representation of φ by its H_o-representation, we could as well have done the other way around, using (IV.7) instead of (IV.7'). An analogous calculation would then lead us, by (IV.2), (IV.3) and (IV.9), to the inverse of the formula (V.2)

$$ {}_o\langle\lambda|P\varphi\rangle = \frac{{}_+\langle\lambda|\varphi\rangle}{1 - \int \frac{|{}_+\langle\lambda'|\varphi\rangle|^2\, d\lambda'}{\lambda' - \lambda + io}} , \qquad (V.2') $$

and substituting one into the other would lead to functional identities which may be of some intrinsic mathematical interest in the theory of Hilbert transforms.

C. Scattering Amplitude, Cross-Section and Phase-Shift

The scattering quantities (IV.5) are now immediately obtained when inserting (V.3) into (IV.5):

$$R(\lambda) = -\frac{2\pi i X(\lambda)}{1+\int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda-i0}}\,;\quad \sin^2\delta(\lambda) = \sigma(\lambda) = \frac{\pi^2 X^2(\lambda)}{\left(1+\mathcal{P}\int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda}\right)^2 + \pi^2 X^2(\lambda)} \tag{V.4}$$

where we introduced the notation $X(\lambda)$ for the positive function $|{}_o\langle\lambda|\varphi\rangle|^2$.

It would be interesting to get an inversion formula for (V.4) which would yield a procedure of reconstructing a rank-one potential P_φ out of given scattering amplitudes. This particular mathematical problem is unsolved, and it is rather improbable that a solution always exists because this would mean that such a simple potential as P_φ could produce a scattering amplitude of arbitrary complexity.

Before going into a discussion of (V.4) let us inquire about the spectrum of H.

D. Possible Existence of H-Bound States[9)]

Although it will turn out that the above model is not very enlightening in the context of the bound state problem, we shall rapidly review here the question of possible eigenstates χ of the operator H.

Lemma V.2: H has a bound state χ with energy $\tilde{\lambda}$ iff

$$\text{(a)}\quad \int \frac{X(\lambda)d\lambda}{\lambda-\tilde{\lambda}} = -1\,; \qquad \text{(b)}\quad \int \frac{X(\lambda)d\lambda}{(\lambda-\tilde{\lambda})^2} < \infty\,.$$

The (unnormalized) eigenstate χ with energy $\tilde{\lambda}$ is then given by

$${}_o\langle\lambda|\chi\rangle = \frac{{}_o\langle\lambda|\varphi\rangle}{\lambda-\tilde{\lambda}}\,. \tag{V.5}$$

Proof: The element χ has to satisfy the following eigenstate-equation:

$${}_o\langle\lambda|H\chi\rangle = \lambda\,{}_o\langle\lambda|\chi\rangle + (\varphi,\chi)\,{}_o\langle\lambda|\varphi\rangle = \tilde{\lambda}\,{}_o\langle\lambda|\chi\rangle.$$

This equation makes sense for all λ only if $(\varphi,\chi)\neq 0$ and we may solve for ${}_o\langle\lambda|\chi\rangle$:

$${}_o\langle\lambda|\chi\rangle = -\frac{(\varphi,\chi)\,{}_o\langle\lambda|\varphi\rangle}{\lambda-\tilde{\lambda}}\,.$$

Multiplying both sides by $\overline{{}_o\langle\lambda|\varphi)}$ and integrating, we obtain

$$(\varphi,\chi) = -(\varphi,\chi)\int \frac{X(\lambda)d\lambda}{\lambda - \tilde{\lambda}} .$$

The arbitrary phase-factor (φ,χ) which fixes the length of ψ drops out and we are left with the necessary condition (a) on φ. Condition (b) follows from the postulate that χ is square-integrable:

$$\int |{}_o\langle\lambda|\chi)|^2 d\lambda = |(\varphi,\chi)|^2 \int \frac{X(\lambda)d\lambda}{(\lambda - \tilde{\lambda})^2} < \infty .$$

These conditions are sufficient for the existence of a bound state in H. In fact,

$${}_o\langle\lambda|\chi) \equiv \frac{{}_o\langle\lambda|\varphi)}{\lambda - \tilde{\lambda}}$$

is square integrable according to (b) and it satisfies the eigenvalue equation according to (a).

Lemma V.3: If the spectrum Λ_o of H_o is the positive half-axis, any possible discrete eigenvalue $\tilde{\lambda}$ of H is imbedded in Λ_o.

Proof: The contrary: $\tilde{\lambda} \leq 0$ implies

$$\int \frac{X(\lambda)d\lambda}{\lambda - \tilde{\lambda}} \geq 0$$

which contradicts condition (a).

Of course, we should verify that a $\varphi \in \mathcal{H}$ actually exists such that $H = H_o + P_\varphi$ admits bound states. According to Lemma V.2 the absolute square of its H_o-representation $X(\lambda) = |{}_o\langle\lambda|\varphi)|^2$ must satisfy conditions (a) and (b) (and $\int X(\lambda)d\lambda = \|\varphi\|^2 = 1$). A necessary condition for $X(\lambda)$ is clearly that it should vanish for some $\tilde{\lambda} \in \Lambda_o$ as follows from (b). This is, however, not sufficient. But consider the following example: Let $X'(\lambda)$ be a positive non-vanishing integrable function on Λ_o and define

$$X(\lambda) = \begin{cases} 0 & \text{on an open interval } (a,b) \subset \Lambda_o \\ \alpha X'(\lambda) & \text{on the rest of } \Lambda_o \quad [\alpha > 0] \end{cases} .$$

Clearly, α can be so chosen that $\int X(\lambda)d\lambda = 1$. Next, (b) is certainly satisfied for any $\tilde{\lambda} \in (a,b)$. In order to show that (a) can also be satisfied for some $\tilde{\lambda}$ in (a,b), consider the left-hand side of (a) as a

function of $\tilde{\lambda} \in (a,b)$. It splits into the two parts

$$Y_1(\tilde{\lambda}) \equiv \int_0^a \frac{X(\lambda)d\lambda}{\lambda - \tilde{\lambda}} ; \qquad Y_2(\tilde{\lambda}) \equiv \int_b^\infty \frac{X(\lambda)d\lambda}{\lambda - \tilde{\lambda}}$$

with the following properties: Y_1 is negative, continuous and monotonously decreasing to $-\infty$ if $\tilde{\lambda}$ tends to a, while Y_2 is positive, continuous and monotonously increasing if $\tilde{\lambda}$ tends to b. So they add up to a continuous function increasing monotonously from $-\infty$ to ∞, as $\tilde{\lambda}$ moves from a to b. So there exists at least one value $\tilde{\lambda}$ in (a,b) for which condition (a) is satisfied and which is thus a discrete eigenvalue of H.

So we may recapitulate the essential bound state features of the model as follows: for the existence of a discrete energy eigenvalue $\tilde{\lambda}$ in H it is necessary that the H_0-spectral representation of φ vanishes at $\tilde{\lambda}$. If such a value exists it is always imbedded in the continuous spectrum Λ_0 of H. There is, in fact, no hope of producing an isolated negative H-bound state by a rank-one potential.

We did not bother about the question as to whether or not in the spectrum of H there may appear a part which is neither discrete nor absolutely continuous. It is known that its impossibility hinges on a certain Hölder condition imposed on $X(\lambda)$ which we shall describe more thoroughly in an analogous situation met in the model of Chapter VI.

The presence of bound states in H has to be considered as a rather accidental fact in this model since it supposes that the function $X(\lambda)$ vanishes in an interval, and its discussion has not much physical relevance as may also be concluded from the discussion of the cross-section (V.4) given in the next section.

E. Discussion of the Scattering Quantities (V.4)

In order to interpret formulas (V.4), let us introduce the analytic function

$$h(z) \equiv 1 + \int_0^\infty \frac{X(\lambda')d\lambda'}{\lambda' - z} .$$

Furthermore, we define

$$h(\lambda) \equiv \tfrac{1}{2}\Big(h(\lambda + i0) - h(\lambda - i0)\Big) = 1 + \mathrm{P}\int_0^\infty \frac{X(\lambda')d\lambda'}{\lambda' - \lambda} .$$

In terms of these functions the scattering quantities (V.4) read

$$R(\lambda) = -\frac{2\pi i\, X(\lambda)}{h(\lambda+io)}\ ; \quad \sin^2\delta(\lambda) = \sigma(\lambda) = \frac{\pi^2 X^2(\lambda)}{(h(\lambda))^2 + \pi^2 X^2(\lambda)}\ .$$

If we assume $X(\lambda) \neq 0$ on the spectrum Λ_o, excluding thus bound states in H, the function $\sin^2\delta(\lambda)$ has the following properties:

(a) It does not vanish on Λ_o except at 0 and ∞.

(b) It is positive and never exceeds 1, and it reaches this limit exactly if $h(\lambda) = 0$.

Consider now the phase-shift

$$\delta(\lambda) = \frac{1}{2i}\,\ell n(R(\lambda) + 1)$$

and its derivative

$$\delta'(\lambda) = \frac{1}{2i}\,\frac{R'(\lambda)}{R(\lambda)+1} = \frac{\pi(X(\lambda)h'(\lambda) - X'(\lambda)h(\lambda))(h(\lambda)+i\pi X(\lambda))}{(h^2(\lambda) + 2\pi i\, X(\lambda)h(\lambda) - \pi^2 X^2(\lambda))(h(\lambda) - i\pi X(\lambda))}\ . \tag{V.6}$$

$\delta(\lambda)$ is a multivalued function of $\sigma(\lambda)$ and, according to (a), we may fix its determination by $\delta(0) = 0$. Since $\sin^2\delta(\lambda)$ never vanishes on Λ_o, the values of $\delta(\lambda)$ remain in either the interval $(0,\pi)$ or $(0,-\pi)$ and reach, for $\lambda \to \infty$, the points 0 or $\pm\pi$. Furthermore, it admits the values $\pm\pi/2$ exactly if $h(\lambda)$ vanishes.

In order to establish which alternative actually holds in our example, we may determine the sign of $\delta'(\lambda)$ at that point λ where $\delta(\lambda)$ crosses $|\pi/2|$ for the first time, i.e., at the first zero of $h(\lambda)$. Equation (V.6) reads for these zeros:

$$\delta'(\lambda) = \frac{h'(\lambda)}{\pi X(\lambda)}\ . \tag{V.7}$$

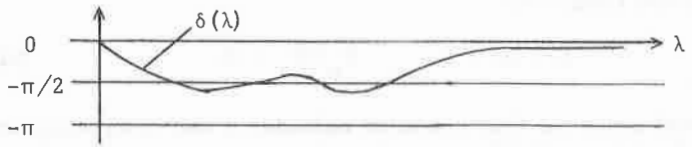
Now it follows from the fact that $h(\lambda)$ is positive at 0 and ∞ that $h(\lambda)$ admits a pair number of zeros, that $h'(\lambda)$ is negative for the first zero of $h(\lambda)$, and that the sign of $h'(\lambda)$ alternates when progressing from one zero to another.

This establishes now uniquely the behaviour of the phase-shift.

(a) $\delta(0) = \delta(\infty) = 0$

(b) $0 > \delta(\lambda) > -\pi$ for $0 < \lambda < \infty$

(c) $\delta(\lambda)$ crosses the line $-\pi/2$, from above and below alternatively, exactly at those points where $\sigma(\lambda)$ reached its upper limit.



Note that (a) is in accordance with Levinson's theorem (IV.8).

In the presence of H-bound states some of the above conclusions fail. It is still true that

$$h(\lambda) = 1 + P\int_0^\infty \frac{X(\lambda')d\lambda'}{\lambda' - \lambda} = 0 \qquad (V.8)$$

is a necessary condition on $\sigma(\lambda)$ to reach its maximum, but it is no longer sufficient. In fact, if $\tilde{\lambda}$ is an H-bound state, we have, simultaneously with (V.8), which is only the first condition of Lemma V.2

$$\int_0^\infty \frac{X(\lambda')d\lambda'}{(\lambda' - \tilde{\lambda})^2} < \infty .$$

Hence (using Lemma V.3), $X(\tilde{\lambda}) = 0$. So $\sin^2\delta(\lambda)$ is not 1 at these points although $h(\lambda)$ goes through zero (in fact, it can be shown to be 0 rather than 1!), and so is it not true that it never crosses the line 0 nor that $\delta(\lambda)$ crosses the line $-\pi/2$ an equal number of times from above and below. This can also be inferred from Levinson's Theorem, which in the case of one H-bound state yields

$$\delta(\infty) - \delta(0) = -\pi .$$

We close the discussion of this model here, its aim being merely a first acquaintance with the subject. Many points treated here will reappear in a similar form in the model treated in the next chapter which will, however, be more significant from a physical view and which will be particularly well suited for a systematic introduction of the concept of resonance.

VI. Soluble Decay Models

A. Definition of a Model with One Decaying State[10)]

Let H_0 be a self-adjoint operator with an absolutely continuous spectrum Λ_0 extending from 0 to ∞ and a discrete eigenvalue $\lambda_0 > 0$ embedded in Λ_0. Let χ be the normalized eigenvector of H_0 with energy λ_0, and P_0 the projector on its orthogonal complement. Finally, let the interaction V be a self-adjoint operator which

satisfies the conditions

$$(\chi, V\chi) = 0; \qquad P_o V P_o = 0. \tag{VI.1}$$

In physical terms the first condition excludes a self-interaction of the bound state χ, and the second excludes an interaction of the continuous (scattered) states among themselves.

Lemma VI.1: The range of V is two-dimensional (i.e., V is a rank-two potential) and the system $(H_o, H = H_o + V)$ is a simple scattering system.

Proof: Any element $\psi \in \mathcal{H}$ can be decomposed into the parts $P_o\psi$ and $(I - P_o)\psi = (\chi, \psi)\chi$. According to (VI.1), $VP_o\psi$ is orthogonal to $P_o\mathcal{H}$ and thus lies in the one-dimensional space spanned by χ, while $V(I - P_o)\psi = (\chi, \psi)V\chi$ lies in the one-dimensional space spanned by the element $V\chi$. This proves the first part of the lemma. The second follows again from the theorem of Kato[8] according to which a system of two operators H_o and H differing by a potential V which is defined everywhere and has finite-dimensional range is a simple scattering system. In particular, the absolutely continuous part of the spectrum of H coincides with Λ_o.

B. Scattering Quantities and Decay Law

Lemma VI.2: The operators $V\Omega_+$ and $\Omega_+^* V$ are Carleman integral operators in the H_o-representation of the space $P_o\mathcal{H}$, and their kernels are

$$ {}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o = {}_o\langle\lambda|V\chi)\overline{{}_+\langle\lambda'|\chi)}; \quad {}_o\langle\lambda|\Omega_+^* V|\lambda'\rangle_o = {}_+\langle\lambda|\chi)\,\overline{{}_o\langle\lambda'|V\chi)} \tag{VI.2}$$

where we now denote by ${}_o\langle\lambda|\psi)$ and ${}_+\langle\lambda|\psi)$ the spectral representations of $P_o\psi$ and $P\psi$ with respect to H_o and H respectively.

Proof: From (VI.1) follows

$$ {}_o\langle\lambda|V\Omega_+P_o\psi) = (\chi, \Omega_+P_o\psi)\,{}_o\langle\lambda|V\chi) = (\Omega_+^*\chi, P_o\psi)\,{}_o\langle\lambda|V\chi)$$

$$= \int {}_o\langle\lambda|V\chi)\,\overline{{}_o\langle\lambda'|\Omega_+^*\chi)}\,{}_o\langle\lambda'|P_o\psi)d\lambda'$$

$$= \int {}_o\langle\lambda|V\chi)\,\overline{{}_+\langle\lambda'|\chi)}\,{}_o\langle\lambda'|P_o\psi)d\lambda'.$$

Hence ${}_o\langle\lambda|V\Omega_+|\lambda'\rangle_o = {}_o\langle\lambda|V\chi\rangle\overline{{}_+\langle\lambda'|\chi\rangle}$. Similarly,

$$ {}_o\langle\lambda|\Omega_+^*VP_o\psi\rangle = {}_+\langle\lambda|VP_o\psi\rangle = (\chi, VP_o\psi)\,{}_+\langle\lambda|\chi\rangle $$

$$ = \int {}_+\langle\lambda|\chi\rangle\,\overline{{}_o\langle\lambda'|V\chi\rangle}\,{}_o\langle\lambda'|P_o\psi\rangle d\lambda'. $$

Hence ${}_o\langle\lambda|\Omega_+^*V|\lambda'\rangle_o = {}_o\langle\lambda|\chi\rangle\overline{{}_o\langle\lambda'|V\chi\rangle}$. The Carleman properties follow as in Lemma V.1 from the finiteness of the norms $\|V\chi\|$ and $\|\chi\|$.

In order to express the scattering amplitude entirely in terms of the interaction function ${}_o\langle\lambda|V\chi\rangle$, we have first to transform ${}_+\langle\lambda|\chi\rangle$, which appears in the kernel of $V\Omega_+$, into H_o-representation. Equation (IV.7') which relates the two representations is, however, not directly applicable to the bound state χ since it is defined only on the continuous states $P_o\psi$. But here we may now make explicit use of the first condition (VI.1) according to which $V\chi \in P_o\mathcal{H}$. After having inserted the expression of Lemma VI.2 for the kernel of Ω_+^*V, Eq. (IV.7') reads now, for $\psi = V\chi$:

$${}_+\langle\lambda|V\chi\rangle = {}_o\langle\lambda|V\chi\rangle - {}_+\langle\lambda|\chi\rangle\int\frac{X(\lambda')d\lambda'}{\lambda'-\lambda-io}$$

where we have introduced the positive function $X(\lambda) \equiv |{}_o\langle\lambda|V\chi\rangle|^2$. On the other hand, we may write

$${}_+\langle\lambda|V\chi\rangle = {}_+\langle\lambda|H\chi\rangle - {}_+\langle\lambda|H_o\chi\rangle = (\lambda-\lambda_o)\,{}_+\langle\lambda|\chi\rangle$$

and, solving these two equations for ${}_+\langle\lambda|\psi\rangle$,

$${}_+\langle\lambda|\chi\rangle = \frac{{}_o\langle\lambda|V\chi\rangle}{\lambda-\lambda_o+\int\frac{X(\lambda')d\lambda'}{\lambda'-\lambda+io}}\,. \qquad \text{(VI.3)}$$

With this and Lemma VI.2, the scattering quantities (IV.5) become

$$R(\lambda) = -\frac{2\pi i X(\lambda)}{\lambda-\lambda_o+\int\frac{X(\lambda')d\lambda'}{\lambda'-\lambda-io}}\;;$$

$$\sin^2\delta(\lambda) = \sigma(\lambda) = \frac{\pi^2X^2(\lambda)}{\left(\lambda-\lambda_o+\mathcal{P}\int\frac{X(\lambda')d\lambda'}{\lambda'-\lambda}\right)^2+\pi^2X^2(\lambda)} \qquad \text{(VI.4)}$$

and the decay law (IV.9) for the H_0-bound state χ:

$$P(t) = \left(\chi, e^{-iHt}\chi\right) = \int \left|{}_+\langle\lambda|\chi)\right|^2 e^{-i\lambda t} d\lambda$$

$$= \int \frac{X(\lambda)e^{-i\lambda t}d\lambda}{\left(\lambda - \lambda_0 + \beta\int \frac{X(\lambda')d\lambda'}{\lambda' - \lambda}\right)^2 + \pi^2\chi^2(\lambda)} . \qquad \text{(VI.5)}$$

We should immediately remark that these explicit results are a rigorous consequence of conditions (VI.1) alone; in particular, no a priori assumption has been made on the existence of bound states in H and their validity is independent of the problem whether such bound states exist or not.

If we insert (VI.3) into the kernels (VI.2) and (VI.2) into (IV.7'), we remark that we obtain an explicit transformation formula from the H_0- to the H-representation of an arbitrary element $\psi \in \mathcal{H}$ which depends only on the interaction function ${}_0\langle\lambda|V\chi)$. In fact, (IV.7') provides the passage $H_0 \to H$ for the part $P_0\psi$ of ψ and (VI.3) for the part $(I - P_0)\psi = (\chi, \psi)\chi$. With some effort, the inverse transformation could also be established using such relations as (IV.2) and (IV.7) where (IV.7) does not take account of the possible H-bound states which have to be separately considered.

C. Possible Existence of H-Bound States

Lemma VI.3:

(a) If $X(\lambda) \neq 0$ on the spectrum Λ_0 then there does not exist any discrete eigenvalue $\tilde{\lambda}$ imbedded in the continuous spectrum of H.

(b) If

$$\lambda_0 \geq \int \frac{X(\lambda)}{\lambda} d\lambda$$

then there does not exist any discrete eigenvalue $\tilde{\lambda}$ outside the continuous spectrum Λ of H (i.e., on the negative real axis).

(c) If $X(\lambda) = 0$ for $\lambda = 0$ (as follows from (b)) and if $X(\lambda)$ satisfies the Hölder condition

$$H(\mu): \left|X(\lambda') - X(\lambda)\right| \leq \rho|\lambda' - \lambda|^{\mu}; \quad [0 < \mu < 1],$$

then there does not exist any singular (i.e., continuous but not absolutely continuous) part in the spectrum of H.

Proof: First we write down the action of the potential V in H_o-representation. For the continuous part we obtain, according to (VI.1),

$$ {}_o\langle\lambda|V\psi) = {}_o\langle\lambda|V(I-P_o)\psi) = (\chi,\psi)\,{}_o\langle\lambda|V\chi) $$

and for the discrete part

$$ (\chi,V\psi) = (\chi, VP_o\psi) = (V\chi, P_o\psi) = \int \overline{{}_o\langle\lambda|V\chi)}\,{}_o\langle\lambda|\psi)d\lambda . $$

The eigenvalue equations for H read in this representation

$$ \left.\begin{aligned} {}_o\langle\lambda|H\psi) &= \lambda\,{}_o\langle\lambda|\psi) + (\chi,\psi)\,{}_o\langle\lambda|V\chi) = \tilde{\lambda}\,{}_o\langle\lambda|\psi) \\ (\chi,H\psi) &= \lambda_o(\chi,\psi) + \int \overline{{}_o\langle\lambda|V\chi)}\,{}_o\langle\lambda|\psi)d\lambda = \tilde{\lambda}(\chi,\psi) \end{aligned}\right\} \qquad (VI.6) $$

for the continuous and the discrete part respectively. Their solution is

$$ {}_o\langle\lambda|\psi) = -\frac{(\chi,\psi)\,{}_o\langle\lambda|V\chi)}{\lambda-\tilde{\lambda}} \;;\; (\chi,\psi) = -\frac{\int {}_o\langle\lambda|V\chi)\,{}_o\langle\lambda|\psi)d\lambda}{\lambda_o-\tilde{\lambda}} . \qquad (VI.7) $$

The normalizability condition for ψ implies

$$ \|\psi\|^2 = \int |{}_o\langle\lambda|\psi)|^2 d\lambda + |(\chi,\psi)|^2 = |(\chi,\psi)|^2\left(1+\int\frac{X(\lambda)d\lambda}{(\lambda-\tilde{\lambda})^2}\right) < \infty . $$

If $\tilde{\lambda}$ belongs to Λ_o the integral diverges and the normalizability can not be satisfied unless $(\chi,\psi)=0$ which, however, contradicts the first equation (VI.6). This proves part (a). Inserting the first equation (VI.7) into the second we obtain

$$ 1 = \frac{\int\frac{X(\lambda)}{\lambda-\tilde{\lambda}}d\lambda}{\lambda_o-\tilde{\lambda}} , $$

and in the case $\tilde{\lambda} \le 0$ this is majorized by

$$ \frac{1}{\lambda_o}\int\frac{X(\lambda)}{\lambda}\,d\lambda . $$

This proves (b). The proof of (c), which serves to exclude a physically pathological situation, is more subtle and will be omitted here.[8)]

D. The Concept of Resonance

We introduce the analytic function

$$h(z) \equiv z - \lambda_0 + \int_{\Lambda_0} \frac{X(\lambda)d\lambda}{\lambda - z}$$

which has a cut along the spectrum Λ_0 of H_0 with the discontinuity

$$h(\lambda+i0) - h(\lambda-i0) = \left(\lambda - \lambda_0 + \mathcal{P}\int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda} + i\pi X(\lambda)\right)$$

$$-\left(\lambda - \lambda_0 + \mathcal{P}\int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda} - i\pi X(\lambda)\right) = 2\pi i X(\lambda).$$

Furthermore, we again define the real function

$$h(\lambda) \equiv \tfrac{1}{2}\left(h(\lambda+i0)+h(\lambda-i0)\right) = \lambda - \lambda_0 + \mathcal{P}\int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda}.$$

According to (VI.4) the scattering cross-section now reads

$$\sigma(\lambda) = \sin^2\delta(\lambda) = \frac{\pi^2 X^2(\lambda)}{h^2(\lambda)+\pi^2 X^2(\lambda)}.$$

This positive function is limited above by the value 1 and admits that value only for the zeros of $h(\lambda)$.

We first assume that the three conditions of the Lemma VI.3 hold and that the spectrum of H is therefore absolutely continuous. These three conditions imply respectively that

(a) the vanishing of $h(\lambda)$ is also sufficient for $\sigma(\lambda) = \sin^2\delta(\lambda)$ to reach its maximum 1;

(b) there exists at least one zero λ_r of $h(\lambda)$ with $h'(\lambda_r) > 0$;

(c) if the interaction is small enough then this zero is unique, while for arbitrary interactions there exists always an odd number of zeros λ_r and the signs of $h'(\lambda_r)$ alternate.

The first statement follows clearly from the fact that $X(\lambda)$ cannot vanish on the spectrum. Furthermore, $h(\lambda)$ is a continuous function tending to ∞ for $\lambda \to \infty$. Since on the other hand $h(0) \leq 0$ according to condition (b) of Lemma VI.3, the statement (b) follows immediately. In order to prove the first part of (c), we remark that it follows from the Hölder condition $|X(\lambda') - X(\lambda)| \leq \rho|\lambda' - \lambda|^\mu$ with $\mu < 1$ according to a known lemma[11] that its Stieltjes transform

$$Y(\lambda) \equiv \mathrm{P}\int \frac{X(\lambda')d\lambda'}{\lambda' - \lambda}$$

also satisfies a Hölder condition $|Y(\lambda') - Y(\lambda)| \leq \rho|\lambda' - \lambda|^{\mu}$ with the same exponent μ. A second zero of $h(\lambda)$ can, however, arise only if the absolute derivative of $Y(\lambda)$ becomes larger than that of $(\lambda - \lambda_0)$, i.e., if $\rho' > 1$. But introducing the parameter g in the interaction (coupling constant): $X(\lambda) \to gX(\lambda)$; $Y(\lambda) \to gY(\lambda)$, we see that ρ (and thus ρ') can be made arbitrarily small. So there exists a limit coupling for which $\rho' = 1$ and, if g is below, no other zero exists for $h(\lambda)$. The second part of (c) follows again from $h(0) \leq 0$ and $h(\infty) = \infty$.

In order to interpret all of this in terms of the phase-shift, we calculate $\delta'(\lambda)$ as in Chapter V. Formula (V.6) obtained there remains formally true and we obtain

$$\delta'(0) = -\frac{\pi X'(0)}{h(0)} \geq 0; \quad \operatorname{sign} \delta'(\lambda) = \operatorname{sign} \frac{h'(\lambda)}{\pi X(\lambda)} = \operatorname{sign} h'(\lambda).$$

If we choose again the determination $\delta(0) = 0$ for the phase-shift, which is possible since $\sin^2\delta(0) = 0$, we obtain the following properties for $\delta(\lambda)$:

(a) it increases at the origin and stays in the interval $(0, +\pi)$ since $\sin^2\delta(\lambda)$ never vanishes except at 0 and ∞;

(b) it crosses the line $\pi/2$ from below and above, alternatively, every time $h(\lambda)$ vanishes, and, since this happens in an odd number of points,

(c) it reaches π for $\lambda \to \infty$.

We remark that this corresponds to Levinson's theorem (IV.8) when applied to the case of one H_0-bound state and no bound state in H. If H contains a bound state, condition (b) of Lemma VI.3 fails, and we now conclude from $h(0) \geq 0$ and $h(\infty) = \infty$ that $h(\lambda)$ vanishes in a pair number of points, hence we would infer $\delta(0) = \delta(\infty) = 0$ again in correspondence to Levinson's theorem.

One might question the physical significance of the signs of $\delta'(\lambda)$ appearing in the maxima of $\sigma(\lambda)$. In this regard we may invoke the relation

$$\delta'(\lambda) = \tfrac{1}{2} Q(\lambda) \qquad \text{(VI.8)}$$

which links the derivative of the phase-shift to the so-called delay time of the physical wave with respect to the asymptotic free wave at a given energy. (For a proof in terms of the concepts developed here see Reference 12.) Since the sign of this quantity depends on the question of whether or not the potential is attractive or repulsive, we may finally recapitulate what we obtained in this section in the following natural terminology:

$\sigma(\lambda)$ has a (true) resonance whenever it reaches its absolute maximum 1. The corresponding energies (= the zeros of $h(\lambda)$) are the resonance energies λ_r. According to relation (VI.8) we call the resonances attractive or repulsive whenever the sign of $\delta'(\lambda_r)$ is positive or negative. In the special case $\delta'(\lambda_r) = 0$ we say to be in the presence of a double (attractive and repulsive) resonance.

The previous discussion may then be resumed in the proposition: for weak enough interactions there exists exactly one attractive resonance; for stronger interactions new resonances appear in pairs of attractive and repulsive ones. The difference in the number of attractive and repulsive resonances equals the difference in the number of H_o- and H-bound states.

The puzzling fact that additional resonances in the one particle decay cannot be excluded in a model satisfying all the axioms of a simple scattering system indicates that the latter are incomplete from a physical sight. The missing point seems to be related to the question of locality. This can be seen as follows: In the case of local potentials it is known[13] that the delay time $Q(\lambda)$ is limited below. According to (V.6) and (VI.8), it seems therefore that some locality requirement could provide the necessary lower bound for $h'(\lambda)$ needed in order to exclude additional resonances.

E. Discussion of the Decay Law

In terms of the analytic function $h(z)$ introduced in the last section, the decay law (VI.1) can be written

$$P(t) = \frac{1}{2\pi i} \int_o^\infty \frac{h(\lambda+io) - h(\lambda-io)}{h(\lambda+io)h(\lambda-io)} e^{-i\lambda t} d\lambda$$

$$= \frac{1}{2\pi i} \int_o^\infty \left(\frac{1}{h(\lambda-io)} - \frac{1}{h(\lambda+io)} \right) e^{-i\lambda t} d\lambda . \qquad \text{(VI.9)}$$

Instead of discussing this expression directly, we shall use the analytic properties of the integrand in order to deform the integration path in an appropriate manner. An easy property of $h(z)$ is given by:

Lemma VI.4: If H contains no bound state, then the function $h(z)$ is regular analytic in the entire z-plane except on the spectrum Λ_o of H_o (and H) where it has a cut, and it does not admit any zeros in its first sheet.

Proof: The first part follows from very general features of the Cauchy integrals and Hilbert transforms.[11] Next we see immediately that

the function h(z) does not vanish for complex arguments $z=\lambda+i\mu$, since its imaginary part

$$\mu\left(1+\int \frac{X(\lambda')d\lambda'}{(\lambda'-\lambda)^2+\mu^2}\right)$$

vanishes only for $\mu=0$. Finally, suppose $z=\lambda$ is on the negative real axis. Now it is the real part of h(z) which cannot vanish. In fact, suppose it does:

$$h(\lambda) = \lambda - \lambda_0 + \int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda} = 0.$$

Inserting into this condition (b) of Lemma VI.3:

$$\lambda_0 \geq \int \frac{X(\lambda')d\lambda'}{\lambda'} \geq \int \frac{X(\lambda')d\lambda'}{\lambda'-\lambda}$$

leads to the inequality $\lambda \geq \lambda_0$ which contradicts the fact that λ is negative. This argument shows clearly that the absence of zeros in h(z) on the negative axis depends essentially on the absence of H-bound states with negative energy.

Now let us continue h(z) analytically through the cut from above. We define thus the second sheet h^{II} of h by

$$h^{II}(\lambda - i0) \equiv h^{I}(\lambda + i0). \qquad \text{(VI.10)}$$

In order to write $h^{II}(z)$ down explicitly, we make at this point the important additional postulate that the interaction function $X(\lambda) \equiv |{}_0\langle \underline{\lambda} | V\chi\rangle|^2$ permits an analytic continuation X(z) which is regular in the entire open z-plane. With this we verify easily that

$$h^{II}(z) = z - \lambda_0 + \int \frac{X(\lambda)d\lambda}{\lambda - z} + 2\pi i\, X(z)$$

is the unique analytic function being connected to $h^{I}(z)$ by (VI.10).

The general belief is that $h^{II}(z)$ admits one zero in the negative half-plane which approaches $z=\lambda_0$ for decreasing interaction strength. This has been established rigorously for particular choices of the interaction (e.g., $X(\lambda)=g\sqrt{\lambda}\, e^{-\alpha\lambda}$ as encountered in the Lee model). In these particular models the zero of h^{II} moves with increasing coupling constant g along quite characteristic trajectories, all of which show the common property that the zero reappears on the negative axis of the first sheet if the interaction becomes strong enough to produce there an H-bound state. The uniqueness problem of this zero has, however, not been solved in entire generality.

and it is doubtful whether uniqueness could be deduced, for instance, on the basis of the conditions of Lemma VI.3 alone. So we shall leave this question open here and turn to the physical significance of these zeros, if they exist, in connection with the decay law.

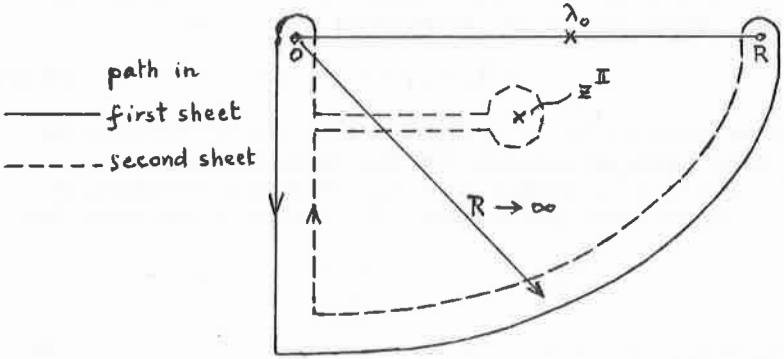
The reason we discussed the extension of the function h into the complex plane is that it permits us to rewrite the expression for P(t) in a new form which is better suited for an interpretation of the decay in terms of its essential physical parameters—at least in the limit of weak interaction. This procedure has been known for a long time[14] and we shall merely sketch it here without going into details.

The integral (VI.9) may be written

$$\frac{1}{2\pi i} \int_{\Lambda_o} \frac{e^{-i\lambda t}}{h(z)}\, dz$$

and the integration path

may be further transformed as follows (here we assume that h(z) admits a unique zero z^{II} in the second sheet):



This new path clearly consists of three distinct parts: the residuum at z^{II}, the negative imaginary axis part, and the great circles. Calculating the first we obtain

$$\frac{1}{2\pi i} \oint_{\Lambda} \frac{e^{-izt}}{h(z)}\, dz = \frac{1}{2\pi i} \lim_{r\to\infty} \int_0^{2\pi} \frac{e^{-i(z^{II}+re^{i\theta})t}}{re^{i\theta} h'(z^{II})}\, ire^{i\theta} d\theta = \frac{e^{-iz^{II}t}}{h'(z^{II})} \,. \tag{VI.11}$$

The other two parts depend linearly on $X(\lambda)$ and thus on the coupling constant g; that means they are of the order of magnitude of the square norm $\|V\chi\|^2$. So they may be considered as corrections (which become small for weak interactions) of the main, exponential term $e^{-iz^{II}t}$, and it becomes clear that we may call

$$\lambda_d \equiv \text{Re } z^{II} \text{ the decay energy and}$$
$$1/\tau \equiv -\text{Im } z^{II} \text{ the reciprocal life-time} \qquad \text{(VI.12)}$$

of the state χ.

Before interpreting this more thoroughly in the limit of weak interaction, let us have a glance at the question of the "poles in the S-matrix."

F. Analytic Continuation of the Scattering Amplitude

If we make the additional postulate about the analyticity of $X(z)$ as stated in the last section, the analytic function

$$R(z) \equiv -\frac{2\pi i X(z)}{h(z)}$$

is well-defined and may be considered as the analytic continuation of the scattering amplitude which should now be written (cp. (VI.4)) $R(\lambda + i0)$ rather than $R(\lambda)$.

$R(z)$ is a multi-valued function and, since $X(z)$ is supposed to be regular in the entire z-plane, the previous discussions of the zeros of $h(z)$ amounts now to a discussion of the poles of $R(z)$, and thus of the poles of the analytic continuation $S(z) = 1 + R(z)$ of the scattering matrix. So we obtain the proposition that the analytic continuation of the S-matrix $S(z)$ is regular in the entire z-plane of its first sheet cut along the spectrum Λ_o of H_o and that its possible poles z^{II} in the second sheet have the physical significance (VI.12).

G. The Limit of Weak Interaction

If we replace X by gX (or V by $\sqrt{g}\cdot V$) and suppose that the coupling constant $g > 0$ is small ($g \ll 1$), we may discuss the scattering quantities in various degrees of approximation with respect to g.

Let us introduce the "line width Δ" of a resonance by $\sigma(\lambda_r + \Delta) = \frac{1}{2}\sigma(\lambda_r)$; then the two resonance parameters (λ_r, Δ) are linked to the two decay parameters $(\lambda_d, 1/\tau)$ by

$$\lambda_r \overset{g}{\sim} \lambda_d \overset{g}{\sim} \lambda_o ; \qquad \Delta \overset{g^2}{\sim} 1/\tau \overset{g^2}{\sim} \pi g X(\lambda_o)$$

where $a \overset{g^\mu}{\sim} b$ means a differing from b by a term of order g^μ (note that Δ and $1/\tau$ are themselves already of order g). Furthermore, the scattering cross-section (VI.4) can be approximated by a so-called Breit-Wigner resonance

$$\sigma(\lambda) \overset{g^2}{\sim} \frac{\Delta^2}{(\lambda - \lambda_r)^2 + \Delta^2}$$

in a "near resonance" domain $|\lambda - \lambda_r|$ of order g, and the decay law (VI.5) by the exponential

$$P(t) \overset{g}{\sim} e^{-i(\lambda_d - i/\tau)t}$$

where we assumed that for small interaction the S-matrix pole $z^{II} = \lambda_d - i/\tau$ is unique.

For the proof see, for example, Reference 15. In writing down the degrees of approximation we supposed that $\mu = 1$ in the Hölder condition for $X(\lambda)$ and its Stieltjes transform $Y(\lambda)$. For arbitrary $0 < \mu < 1$ we would have to replace

$$g^2 \quad \text{by} \quad g^{1+\mu}.$$

H. The Model with n Bound States in H_o

We generalize now the above model in order to describe the decay of n H_o-bound states. The interest of this consists in showing how the single resonances and decay laws "interact" among each other. The procedure used above remains essentially the same: Let H_o still denote a self-adjoint operator with a continuous spectrum Λ_o extending from 0 to ∞, and now let χ_ν be the n bound states of H_o with their energies λ_ν embedded in Λ_o. With this the first condition (VI.1) on the potential reads

$$(\chi_\mu, V\chi_\nu) = 0 \qquad \text{for all } \mu, \nu,$$

while the second remains unchanged. This system is still a scattering system, the interaction being now of rank 2n.

If we introduce the generalized notation

$$X_{\mu\nu}(\lambda) \equiv \overline{{}_o\langle\lambda|V\chi_\mu\rangle}\ {}_o\langle\lambda|V\chi_\nu\rangle; \quad h_{\mu\nu}(z) \equiv (z - \lambda_\mu)\delta_{\mu\nu} + \int \frac{X_{\mu\nu}(\lambda)d\lambda}{\lambda - z},$$

the calculation of the scattering amplitude, cross-section and the n decay laws $P_\nu(t) = (\chi_\nu, e^{-iHt}\chi_\nu)$ can be carried out along the same lines stated earlier (for details, cp. Reference 15), and we obtain

$$R(\lambda) = 2\pi i \sum_{\mu,\nu} (-1)^{\mu+\nu} X_{\mu\nu}(\lambda) \frac{\mathrm{Min}_{\mu\nu} h(\lambda - io)}{\mathrm{Det}\, h(\lambda - io)} \tag{VI.13}$$

$$\sigma(\lambda) = \frac{1}{2}\left(n - \mathrm{Re}\frac{\sum_{\mu,\nu}(-1)^{\mu+\nu} h_{\mu\nu}(\lambda + io)\mathrm{Min}_{\mu\nu} h(\lambda - io)}{\mathrm{Det}\, h(\lambda - io)}\right) \tag{VI.14}$$

$$P_\nu(t) = \frac{1}{2\pi i}\int_\Lambda \left(\frac{\mathrm{Min}_{\nu\nu} h(\lambda - io)}{\mathrm{Det}\, h(\lambda - io)} - \frac{\mathrm{Min}_{\nu\nu} h(\lambda + io)}{\mathrm{Det}\, h(\lambda + io)}\right) e^{-i\lambda t} d\lambda . \tag{VI.15}$$

Furthermore, if we require the additional analyticity postulate for the functions $X_{\mu\nu}(z)$ and if we suppose that to every energy eigenvalue λ_ν there exists exactly one pole z_ν^{II} in the second sheet of the scattering amplitude, the "residual" parts of $P_\nu(t)$ (cp. (VI.11) in the case of one decaying state) become

$$P_\nu^{res}(t) = \sum_\mu \frac{e^{iz_\mu^{II} t}}{H_\nu'(z_\mu^{II})}; \qquad H_\nu(z) \equiv \frac{\mathrm{Det}\, h(z)}{\mathrm{Min}_{\nu\nu} h(z)} . \tag{VI.16}$$

It is easily verified by putting $n=1$ that our previous expressions (VI.4) and (VI.9) for $\sigma(\lambda)$ and $P(t)$ are special cases of (VI.14) and (VI.15). For $P(t)$ this is immediate; let us write it down for $\sigma(\lambda)$:

$$\sigma(\lambda) = \tfrac{1}{2} - \tfrac{1}{2}\,\mathrm{Re}\,\frac{h(\lambda + io)}{h(\lambda - io)} = \tfrac{1}{2} - \tfrac{1}{2}\,\frac{h^2(\lambda) - \pi^2 X^2(\lambda)}{h^2(\lambda) + \pi^2 X^2(\lambda)} \tag{VI.17}$$

$$= \frac{\pi^2 X^2(\lambda)}{h^2(\lambda) + \pi^2 X^2(\lambda)} . \tag{VI.17'}$$

For not too strong interactions it can be shown that $\sigma(\lambda)$ has exactly n true resonances and the scattering amplitude n simple poles in the second sheet. The interesting point is, however, that the line form (VI.13) is not merely a superposition of n one-particle line forms (VI.4), nor are the residual parts (VI.15) of the decay laws just simple exponentials with the characteristics of the single poles. Let us look at this in the example $n=2$.

I. Coupling of Two Resonances and Decay Laws

With the simplified notation

$$h^{\pm}_{\mu\nu}(\lambda) \equiv h_{\mu\nu}(\lambda \pm i0)$$

the expressions (VI.14) and (VI.16) read for $n=2$

$$\sigma(\lambda) = \frac{1}{2}\Big(2 - \mathrm{Re}\,\frac{h^{+}_{11}h^{-}_{22} - h^{+}_{12}h^{-}_{21} - h^{+}_{21}h^{-}_{12} + h^{+}_{22}h^{-}_{11}}{h^{-}_{11}h^{-}_{22} - h^{-}_{12}h^{-}_{21}}(\lambda)\Big)$$

$$P_1^{\mathrm{res}}(t) = \frac{e^{-iz_1^{II}t}}{\frac{d}{dz}\Big(\frac{h_{11}h_{22}-h_{12}h_{21}}{h_{22}}\Big)(z_1^{II})} + \frac{e^{-iz_2^{II}t}}{\frac{d}{dz}\Big(\frac{h_{11}h_{22}-h_{12}h_{21}}{h_{22}}\Big)(z_2^{II})};$$

$$P_2^{\mathrm{res}}(t) = \frac{e^{-iz_1^{II}t}}{\frac{d}{dz}\Big(\frac{h_{11}h_{22}-h_{12}h_{21}}{h_{11}}\Big)(z_1^{II})} + \frac{e^{-iz_2^{II}t}}{\frac{d}{dz}\Big(\frac{h_{11}h_{22}-h_{12}h_{21}}{h_{11}}\Big)(z_2^{II})}.$$

(VI.18)

The expression for $\sigma(\lambda)$ may be discussed in various orders of the coupling constant g, starting from the observation that the cross-terms $h_{12}(z)$ and $h_{21}(z)$ are small in order g compared to the diagonal terms $h_{11}(z)$ and $h_{22}(z)$ since, for instance,

$$h_{11}(z) = z - \lambda_1 + g\int \frac{\chi_{11}(\lambda)d\lambda}{\lambda - z} \underset{\sim}{1} z - \lambda_1;\quad h_{12}(z) = g\int \frac{\chi_{12}(\lambda)d\lambda}{\lambda - z} \underset{\sim}{g} 0.$$

So we would obtain for $\sigma(\lambda)$ in the lowest approximation

$$\sigma(\lambda) \sim \frac{1}{2}\Big(2 - \mathrm{Re}\Big(\frac{h^{+}_{11}}{h^{-}_{11}} + \frac{h^{+}_{22}}{h^{-}_{22}}\Big)\Big) = \frac{1}{2}\Big(1 - \mathrm{Re}\,\frac{h^{+}_{11}}{h^{-}_{11}}\Big) + \frac{1}{2}\Big(1 - \mathrm{Re}\,\frac{h^{+}_{22}}{h^{-}_{22}}\Big)$$

which is just the superposition of two uncoupled one-particle resonances (VI.17), while in the next higher approximation

$$\sigma(\lambda) \sim \frac{\pi^2 g^2 (X_1 h_{22} + X_2 h_{11})^2}{\pi^2 g^2 (X_1 h_{22} + X_2 h_{11})^2 + (h_{11} h_{22})^2}$$

is not the sum of two terms (VI.17').

In the case of the decay laws a similar discussion of higher order approximations can be done if one takes into account their non-residual terms which we neglected in (VI.18). We shall not do this here, but we shall show for weak interaction how the decay laws $P_1^{res}(t)$ and $P_2^{res}(t)$ tend to "their own" exponential laws.

For $P_1^{res}(t)$ this consists, for instance, in showing that the denominator in the second term of (VI.18)

$$\frac{d}{dz}\left(\frac{h_{11}h_{22} - h_{12}h_{21}}{h_{22}}\right)(z_2^{II}) \qquad \text{(VI.19)}$$

tends to infinity if $g \to 0$. According to (VI.13) the zeros of the nominator in (VI.19) are equal to the poles z_1^{II}, z_2^{II} of the scattering amplitude $R(\lambda)$, and they tend to λ_1, λ_2 for $g \to 0$. Furthermore, the unique zero of h_{22} tends also to λ_2. So we find, for small coupling, a pole of

$$\left(\frac{h_{11}h_{22} - h_{12}h_{21}}{h_{22}}\right)$$

in proximity of the zero z_2^{II} which shows clearly that (VI.19) becomes large at z_2^{II} (but not at z_1^{II}!).

References

1. J. M. Jauch, Helv. Phys. Acta 31, 136 (1958).
2. N. I. Achieser and I. M. Glasmann, "Theorie der linearen Operatoren im Hilbertraum," Akademie-Verlag, Berlin.
3. B. Misra, Lecture Notes at the University of Geneva (unpublished).
4. These results on abstract integration (Bochner integrability) are found, e.g., in Hille and Phillips, "Functional Analysis and Semi-Groups," Amer. Math. Soc. Colloquium Publications, Vol. XXXI. I learned them in an informal seminar given by B. Misra in 1964.
5. The content of this section emerged from a discussion with B. Misra.
6. Clasine van Winter, Mat.-Fys. Dansk. Vid. Selsk. 35 (1966).
7. J. M. Jauch, Helv. Phys. Acta 30, 143 (1957).
8. T. Kato, "Perturbation Theory of Linear Operators," Springer (1966).

9. W. Amrein kindly procured the use of his notes on this subject which I reproduce here freely.
10. This model has been introduced by K. O. Friedrichs, Comm. Pure Appl. Math. 1, 378 (1948).
11. N. I. Muskhelishvili, "Singular Integral Equations," P. Noordhoff, Groningen (1953).
12. J. M. Jauch and J. P. Marchand, Helv. Phys. Acta 40, 217 (1967).
13. E. P. Wigner, Phys. Rev. 98, 145 (1955).
14. G. Hoehler, Zeits. für Physik 152, 546 (1958).
15. J. P. Marchand, Helv. Phys. Acta 37, 475 (1964).

INTERACTIONS IN RELATIVISTIC CLASSICAL PARTICLE MECHANICS[†]

D. G. Currie
University of Maryland, College Park, Maryland and Northeastern University, Boston, Massachusetts

T. F. Jordan
University of Pittsburgh
Pittsburgh, Pennsylvania

Contents

[†]Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

I. Introduction

We are interested in interactions for relativistic particles. Are they possible? What do they look like? We will consider these questions for a classical-mechanical system of a fixed number of particles, usually two. We will see that even there, these questions are not entirely trivial and the answers are not exactly what our prejudices would lead us to expect.

We want to consider interactions which can be described by a Hamiltonian. For relativistic invariance, we will have the kinematic assumption of Lorentz transformations for the particle positions, and the dynamic assumption that the equations of motion are Lorentz invariant.

In view of the difficulty of handling interactions in relativistic quantum theories of fields or particles, we think it might be worthwhile to learn what interactions look like in relativistic classical particle mechanics. For example, relativistic quantum theories of a fixed number of particles have been developed for scattering[1] and for bound states (as in relativistic quark models[2]). These theories say nothing about Lorentz transformations of the particle positions. We believe that this cannot be justified by invoking the uncertainty principle. This suggests investigating the analogous problem in classical mechanics. We will see that in a classical system, Lorentz transformations of particle positions are incompatible with canonical representations of the Poincaré group, and we will see how to correctly represent the Poincaré group.

Later, when we are in a position to be more precise and go into more detail, we will outline the history of this subject and say more about its relation to quantum mechanics.

References.

1. T. F. Jordan, A. J. Macfarlane, and E. C. G. Sudarshan, Phys. Rev. 133, B487 (1964); R. Fong and J. Sucher, J. Math. Phys. 5, 456 (1964); T. F. Jordan, J. Math. Phys. 5, 1345 (1964); F. Coester, Helv. Phys. Acta 38, 7 (1965).
2. N. D. Son and J. Sucher, Phys. Rev. 153, 1496 (1967).

II. Equations of Motion

World-Line Conditions.

In the following sections, we will use the relations which guarantee that the points which compose the world-line transform as events—the world-line conditions.[1),2)] These are

$$\mathcal{P}_i x_j = -\delta_{ij} \qquad \mathcal{J}_i x_j = \epsilon_{ijk} x_k$$

$$\mathcal{H} x_i = v_i \qquad \mathcal{K}_i x_j = x_i v_j$$

where P_i, J_i, H and K_i generate infinitesimal spatial translations, spatial rotations, time translations, and Lorentz transformations, respectively. These operators act upon the position of the particle, evaluated at a given time. The first three of these relations are the infinitesimal representation of the familiar effects of translation, rotation and time translation on the particle position. The fourth relation, governing the Lorentz rotation, is perhaps less familiar although it is derived in the same manner as the first three. In order to derive this relation let us consider the transformation properties of the position under an infinitesimal Lorentz transformation. Thus we have the function $x_i(t)$ in the first frame and we wish to obtain the new function $x_i'(t')$ in the new frame. If we consider the particle position at $t' = 0$ in the transformed frame, we are considering an event with coordinates

$$X_i' = x_i'(t' = 0) \qquad t' = 0.$$

In the original frame, this event will have coordinates given by

$$X_i = X_i' - \alpha_j t' = X_i'$$

$$t = t' - \alpha_j X_j' = -\alpha_j X_j'.$$

This point is on the world-line in the old frame, so it must be on the world-line in the new frame. Thus the spatial coordinates of this point are given by

$$X_i = x_i(t = -\alpha_j X_j').$$

Given that the position function has sufficient analyticity for a power series expansion, we have

$$x_i(t = -\alpha_j X_j') = x_i(t=0) - \alpha_j X_j' \left[\frac{d}{dt} x_i(t)\right]_{t=0}$$

$$= x_i(t=0) - \alpha_j X_j v_i(t=0) = x_i - \alpha_j x_j v_i$$

where, by comparison with the above expressions, we have

$$x_i'(t'=0) = X_i' = X_i = x_i(t=0) - \alpha_j x_j(t=0) v_i(t=0)$$

or

$$K_j x_i = -\frac{x_i' - x_i}{\alpha_j} = +x_j v_i .$$

Having obtained the transformation of the particle position under the Lorentz transformation, we now ask how does the particle velocity transform under an infinitesimal Lorentz transformation. One may determine this from either of two different approaches. The most direct, but algebraically complicated, is to make a Lorentz transformation on an entire world-line, express the world-line in terms of the new time, and then take the time derivative with respect to the new time to obtain the new velocity at $t'=0$ in terms of the variables in the old frame at $t=0$. The second, more formal but simpler, method is to use the commutation relations of the generators of the infinitesimal Lorentz transformations.[3] Thus we have, for $t=0$,

$$K_i v_j = K_i \mathcal{K} x_j = \mathcal{K} K_i x_j + [K_i \mathcal{K}] x_j = \mathcal{K}(x_i v_j) + P_i x_j$$

$$= v_i v_j + x_i a_j - \delta_{ij} .$$

In the same manner, one may obtain the transformation of the acceleration, and so forth,

$$K_i a_j = x_i \dot{a}_j + 2 v_i a_j + a_i v_j$$

$$K_i \dot{a}_j = x_i \ddot{a}_j + 3 v_i \dot{a}_j + 3 a_i a_j + \dot{a}_i v_j .$$

Note that these are all for the coordinates of a single particle.

Equations of Motion and Poincaré Invariance.

The dynamics of the particle system will be described by equations of motion for the position of the particle. The acceleration of each particle (at time t) is a function of the position and velocity of all the particles (at the same time t). For two particles, the equations of motion have the form

$$\ddot{\underline{x}}^1(t) = \underline{f}^1\left(\underline{x}^1(t), \dot{\underline{x}}^1(t), \underline{x}^2(t), \dot{\underline{x}}^2(t)\right)$$

$$\ddot{\underline{x}}^2(t) = \underline{f}^2\left(x^2(t), \dot{x}^2(t), x^1(t), \dot{x}^1(t)\right).$$

In a single frame, these equations may be solved as differential equations in t to obtain the position as a function of time. This is the coordinate time, not the proper time, and a dot denotes differentiation with respect to time. The latin indices and vector symbols refer to three vectors.

We now consider the criteria in terms of which the equations of motion are called invariant. First, we transform into the new frame each of the points which compose the world-lines. More precisely, since we employ the passive view, we obtain a new set of coordinates for each point of the world-line. These new points make up a world-line, from which we may obtain the new transformed position as a function of the new time. If the new position functions satisfy the same set of equations of motion (that is, if the acceleration function has the same functional form in terms of the new position components), then the equations of motion or the acceleration functions will be called invariant.

We will find that there are many invariant acceleration functions, that is, this condition for Lorentz invariance is weaker than one might suppose. This condition eliminates only as many possible solutions (equations of motion) as the requirement of Galilean invariance eliminates when one considers the non-relativistic equations of motion.

Einstein Causality.

As is obvious from the previous comments, the acceleration at a particular time is expressed as a function of the position and velocity of another distant particle at the same time. That is, the theory has the form of an action-at-a-distance or a single time formalism. In this frame the system is causal in the traditional sense—that a given set of initial conditions specify the entire solution (world-lines). Let us now consider the property of Einstein causality—that is, the requirement that no information be propagated at a velocity greater than that of light. A very clear and complete discussion of background ideas of causality has been given by Havas.[4] Our equations of motion do not satisfy manifest Einstein causality—that is, from the form of the equations, the requirement of Einstein causality is not automatically satisfied. Our present approach will be to investigate first the implications of Poincaré invariance, leaving until later the consideration of the Einstein causality. The single time formalism is particularly interesting in that it permits one to obtain a Hamiltonian, which in turn may permit a more or less conventional quantization of the theory in a single frame.

At this point, let us consider how the question of Einstein causality would be investigated. Since we are watching the interaction of the particles from outside the system, we are not free to arbitrarily alter the motion of one of the particles. Thus in order to test

the assumption of Einstein causality, we permit a testing particle to interact and alter the world-line of the original two particles. This interaction may take place through the same interaction which is being considered between the original particles, or it may be another admissible interaction. We then ask at what time the modification of the world-line of the particle affected by the testing particle starts to affect the world-line of the other original particle. It is clear that the question of Einstein causality requires careful consideration and definitions, and also requires a knowledge of the form of solutions which are invariant. Thus the study of Einstein causality is either the consideration of three-body interactions, or the consideration of the class of interactions which are compatible with each other.

Of course, there is one class of theories which does possess manifest Einstein causality, and this class is the local field theories.

The particle theory which would seem to satisfy manifest Einstein causality is the Feynman-Wheeler classical electrodynamics,[5] that is, the electrodynamics with the field eliminated, so it is a direct particle interaction theory. This has the same structure as classical electrodynamics, in which one has manifest Einstein causality due to the fact that the interaction is transmitted by the field which travels at the velocity of light, and the field produces no forces upon a particle until it arrives, so the theory would seem to guarantee manifest Einstein causality. This was assumed until Dirac was able to obtain exact solutions for the motion of a particle in a field[6] with the proper inclusion of the radiation reaction terms. The resulting equations of motion for the particle are third order in the time, and for most of the resultant world-lines the particle, after interaction with the field, accelerates to the velocity of light. The only solutions which do not have this "runaway" property have a "preacceleration," that is, they start to accelerate (to react to the field) before the field applies a force to the particle. Thus this theory which appeared on inspection to satisfy Einstein manifest causality, when exact solutions are obtained, is the worst violator. Quantum electrodynamics is the other theory which should have manifest Einstein causality. It is supposed to limit as $\hbar$ vanishes to classical electrodynamics, but the solutions are not yet explicit enough to comment on this question.

Application of Invariance Conditions to a Specific Example.

We will now consider an example of an interaction and determine whether or not it is invariant.[7),8)] In particular, we will solve for the world-lines, then transform each point which composes the world-line to the new frame, and obtain a new function (new position as a function of the new time). If this new function is a solution to the same equations of motion, then the equations of motion will be called invariant.

As an example, let us consider the equation of motion[7)-12)] given by

$$\ddot{x}^1 = -(v^1 - v^2)^2/(x^1 - x^2) \qquad \ddot{x}^2 = +(v^1 - v^2)^2/(x^1 - x^2)$$

in one space dimension. To simplify somewhat, let us express these equations of motion in "collective coordinates" defined by

$$x = (x^1 - x^2)/2 \qquad X = (x^1 + x^2)/2,$$

and then the collective acceleration functions have the form

$$a = -vv/x \qquad A = 0.$$

These are rather simple differential equations in time which may be solved by sight, or at least without too much staring, by noting that

$$0 = ax + vv = \frac{1}{2}\frac{d}{dt}\frac{d}{dt}x^2.$$

So we have for the solutions of the particle motion

$$x = \sqrt{a + bt} \qquad X = A + Bt$$

in our given frame where $t^1 = t^2 = t$. For the individual particles the coordinates are given by

$$x^1 = A + Bt^1 + \sqrt{a + bt^1} \qquad x^2 = A + Bt^2 - \sqrt{a + bt^2}$$

where in our frame $t^1 = t^2 = t$. To determine the positions as a function of time in the new frame, we must transform the coordinates of each point in the world-line so we have

$$x^n = \bar{x}^n \cosh \alpha + \bar{t}^n \sinh \alpha$$

$$t^n = \bar{t}^n \cosh \alpha + \bar{x}^n \sinh \alpha$$

where the bar refers to the new frame. Using the first expression, we get

$$\bar{x}^n = (x^n - \bar{t}^n \sinh \alpha)/\cosh \alpha,$$

and now substituting into this the general solution to the equations of motion, $x^n(t)$, we obtain

$$\bar{x}^1 = \left(A + Bt^1 + \sqrt{a+bt^1} - \bar{t}^1 \sinh\alpha\right)\Big/\cosh\alpha$$

$$\bar{x}^2 = \left(A + Bt^2 - \sqrt{a+bt^2} - \bar{t}^2 \sinh\alpha\right)\Big/\cosh\alpha,$$

and, finally, using the transformation equations to eliminate t^n in favor of E^n and $\bar{x}^n$ yields

$$\bar{x}^n = \Big[A+B\{\bar{t}^n\cosh\alpha + \bar{x}^n\sinh\alpha\} - (-1)^n\sqrt{a+b\{\bar{t}^n\cosh\alpha+\bar{x}^n\sinh\alpha\}} - \bar{t}^n\sinh\alpha\Big]\Big/\cosh\alpha.$$

This may be solved for $\bar{x}^n$ to obtain an expression of the form

$$\bar{x}^n = \tilde{A} + \tilde{B}\bar{t}^n - (-1)^n\sqrt{\tilde{a}+\tilde{b}\bar{t}^n}$$

where the $\tilde{A}$ for $n=1$ and the A for $n=2$ are the same, and so on. If we now take $\bar{t}^1 = \bar{t}^2 = \bar{t}$, these expressions for the new position in terms of the new time have the same form as the world-lines in terms of the old time so they obey the same acceleration functions, i.e., an expression of the form

$$\frac{d}{d\bar{t}}\frac{d}{d\bar{t}}\bar{x}^n(\bar{t}) = -\bar{v}(\bar{t})\bar{v}(\bar{t})\Big/\bar{x}(\bar{t}).$$

The appearance of the tilded quantities indicates that we have a different set of initial conditions in the new frame. In particular

$$\tilde{A} = \left(A\cosh\alpha - AB\sinh\alpha + b\sinh\alpha\right)/\varkappa^2$$

$$\tilde{B} = \left\{B(\cosh^2\alpha + \sinh^2\alpha) - \sinh\alpha\cosh\alpha(1+B^2)\right\}/\varkappa^2$$

$$\tilde{a} = \left\{a\cosh\alpha^2 + b^2\sinh^2\alpha + Ab\sinh\alpha\cosh\alpha - AB\sinh^2\alpha - 2ab\sinh\alpha\cosh\alpha\right\}/\varkappa$$

$$\tilde{b} = \left\{b\cosh\alpha - bB\sinh\alpha\right\}/\varkappa^4$$

where $\varkappa = \cosh\alpha - B\sinh\alpha$.

Thus we conclude that this particular equation of motion is invariant. Note that it is not manifestly invariant—that is, it is not composed of scalar products of four-vectors.

We see that invariant equations of motion exist. Further, we see that the simple requirement of invariance does not guarantee other properties we might associate with invariance. In particular, this force law causes the particle velocity to exceed the velocity of light. To see this, consider the velocity of the individual particles

$$v' = B - \frac{b/2}{\sqrt{a + bt}} \qquad v' = B + \frac{b/2}{\sqrt{a + bt}} .$$

At the time $t = -a/b$ the velocity of each of the particles is infinite. On the other hand, the asymptotic velocity of the particles is not infinite and, in fact, the asymptotic relative velocity vanishes. Thus the Lorentz invariance of our equations of motion does not prevent particle motion at velocities exceeding that of light. The asymptotic relative velocity vanishes for every set of initial conditions, another pathology of this model, resulting in the lack of a "complete set of asymptotic states."

Differential Invariance Conditions.

As a general method for investigation of the invariance of a given acceleration function f, the procedure just considered is not very practical. This is because one must solve the differential equations for the world-line in order to determine if a given acceleration function is invariant. We shall now state the conditions for Lorentz invariance as differential equations which the acceleration functions must satisfy.

Let us note at this point that we in general also require that the acceleration functions be invariant under space translation. By means of a rather obvious calculation, we may show that this implies that the acceleration functions be independent of the mean position $\underline{X} = (\underline{x}_1 + \underline{x}_2)/2$. Further, we in general require that the acceleration function be invariant under rotation so that the acceleration function depends only upon vector dot products of the form

$$\underline{x}\cdot\underline{x},\ \underline{x}\cdot\underline{v},\ \underline{v}\cdot\underline{v},\ \underline{x}\cdot\underline{V},\ \underline{v}\cdot\underline{V},\ \underline{V}\cdot\underline{V}.$$

Finally, the invariance under time translation requires that the acceleration functions are not explicitly functions of the time. However, for certain of the following calculations, we will express the acceleration in terms of the individual coordinates, leaving until a later time the imposition of the requirement of space translation invariance.

We have already determined how the acceleration of the n^{th} particle transforms under an infinitesimal Lorentz transformation, that is,

$$\mathcal{K}_i a_j^n = x_i^n \dot{a}_j^n + 2v_i^n a_j^n + a_i^n v_j^n .$$

Since we seek a function which depends only upon x_i^n and v_i^n, we must re-express the terms a_j^n and $\dot{a}_j^n$ in this form. The former may be written as $f_j^n(\underline{x}^1, \underline{x}^2, \underline{v}^1, \underline{v}^2)$ and for the latter we have

$$\dot{a}_j^n = \frac{d}{dt} f_j^n = \frac{\partial f_j^n}{\partial x_k^1} v_k^1 + \frac{\partial f_j^n}{\partial x_k^1} v_k^2 + \frac{\partial f_j^n}{\partial v_k^1} f_k^1 + \frac{\partial f_j^n}{\partial v_k^2} f_k^2 .$$

Thus we obtain for the transformation of the acceleration

$$\mathcal{K}_i a_j^n = x_i^n \left\{ \frac{\partial f_j^n}{\partial x_k^1} v_k^1 + \frac{\partial f_j^n}{\partial x_k^2} v_k^2 + \frac{\partial f_j^n}{\partial v_k^1} f_k^1 + \frac{\partial f_j^1}{\partial v_k^2} f_k^2 \right\} + 2v_i^n f_j^n + f_i^n a_j^n .$$

Now let us consider what happens to the right hand side of the expression $a_j^n = f_j^n$. The transformation of the acceleration function is given by

$$\mathcal{K}_i f_j^n = \mathcal{K}_i f_j^n(x^1, x^2, v^1, v^2)$$

$$= \frac{\partial f_j^n}{\partial x_k^1} \mathcal{K}_i x_k^1 + \frac{\partial f_j^n}{\partial x_k^2} \mathcal{K}_i x_k^2 + \frac{\partial f_j^n}{\partial v_k^1} \mathcal{K}_i v_j^1 + \frac{\partial f_j^n}{\partial v_k^2} \mathcal{K}_i v_k^2 ,$$

but we have seen that the transformation of the position and velocity is given by

$$\mathcal{K}_i x_k^n = x_i^n v_k^n \qquad \mathcal{K}_i v_k^n = x_i^n a_k^n + v_i^n v_k^n - \delta_{ij} = x_j^n f_k^n + v_i^n v_k^n - \delta_{ij} ,$$

so we have

$$\mathcal{K}_i f_j^n = \frac{\partial f_j^n}{\partial x_k^1} x_i v_k + \frac{\partial f_j^n}{\partial x_k^2} x_i^2 v_k^2 + \frac{\partial f_j^n}{\partial v_k^1} \left\{ x_i f_k + v_i v_k - \delta_{ik} \right\}$$

$$+ \frac{\partial f_j^n}{\partial v_k^2} \left\{ x_i^2 f_k^2 + v_i^2 v_k^2 - \delta_{ik} \right\} .$$

Setting $\mathcal{K}_i a_j^n$ equal to $\mathcal{K}_i f_j^n$, we obtain a differential statement, the invariance conditions, expressing the condition under

which the acceleration function is invariant under infinitesimal Lorentz rotations.

Since any acceleration function which satisfies these equations will be invariant, any solution to these equations will be invariant and to obtain a set including all invariant acceleration functions, we need only solve this non-linear partial differential equation. The difficulty in obtaining solutions to such equations is the reason one does not find a wealth of examples in the usual textbook discussions of relativistic invariance.

Since we shall later want to use the collective coordinates, let us write down these invariance conditions in that form:[3),13),11)]

$$x_i\dot{F}_j + 2V_if_j + 2v_iF_j + F_iv_j + f_iV_j - \frac{\partial f_j}{\partial x_k}V_kx_i - \frac{\partial f_j}{\partial v_k}\left(F_kx_i + v_kV_i + V_kv_i\right)$$

$$- \frac{\partial f_j}{\partial V_k}\left(V_kV_i - \delta_{ki}\right) - \frac{\partial f_j}{\partial V_k}\left(f_kx_i + v_kv_i\right) = 0,$$

$$2V_iF_j + F_iV_j - \frac{\partial F_j}{\partial x_k}V_kx_i - \frac{\partial F_j}{\partial v_k}\left(F_kx_i + v_kV_i + V_kv_i\right) - \frac{\partial F_j}{\partial V_k}\left(V_kV_i - \delta_{ik}\right)$$

$$+ \left\{x_i\dot{f}_j + 2v_if_j + f_iv_j - \frac{\partial F_j}{\partial V_k}\left(f_kx_i + v_kv_i\right)\right\} = 0.$$

We can now return to the example considered in the previous section, namely,

$$a = f = -vv/x \quad A = F = 0$$

and check that it is a solution to these Lorentz invariance conditions. We must first "drop the indices" on the invariance conditions and obtain the one-dimensional form:

$$xVf_x + (xF + 2vV)f_v + (xf + VV + vv - 1)f_V = xvF_x + xfF_v + xFF_V + 3Vf + 3vF$$

$$xVF_x + (xF + 2vV)F_v + (xf + VV + vv - 1)F_V = xvf_x + xff_v + xFf_V + 3vf + 3VF.$$

Since F vanishes, all of its derivatives do, thus simplifying equations. Now simply taking the derivatives of f, and substituting them in the first relation demonstrates that the first is obviously satisfied, and the second is satisfied, if we remember that $f = -v^2/x$ (the second expression in this case contains the real nonlinearity).

We have again considered this explicit example of an invariant set of acceleration functions. This example is not particularly physical, due, in part, to the fact that it was obtained by requiring that the equations have a very simple type of solution. We now consider a few more interesting examples. Again, since they were obtained by putting rather severe restrictions on the invariance condition in order to solve them, we cannot be particularly unhappy that these equations of motion are not the most general possible interaction.

Our first class of more interesting examples is the case of electromagnetic interactions when one of the particles is infinitely heavy, that is, it has vanishing acceleration. We first consider an interaction which is essentially the Coulomb interaction. When we calculate the electric field at the "test" charge, that is, our finite mass charge in terms of the present position of the infinite mass particle, rather than the retarded position, we obtain

$$\underline{E} = \frac{e_2 \underline{x}}{\sqrt{1-v_2^2}}\left\{x\cdot x + \frac{(\underline{x}\cdot\underline{v}^2)}{\sqrt{1-v^2 v^2}}\right\}^{-3/2}$$

where the "2" refers to the infinite mass particle. The magnetic field has the usual form

$$\underline{H} = \underline{v}^2 \times \underline{E}.$$

The Lorentz force now has the form, where we have used $\dot{v} = f$,

$$e_1\underline{E} + e\cdot\underline{v}^1 \times \underline{H} = \frac{dp^1}{dt} = \frac{d}{dt}\left(\underline{v}^1 \Big/ \sqrt{1-\underline{v}^1\cdot\underline{v}^1}\right)$$

$$= \underline{f}^1 \Big/ \sqrt{1-\underline{v}^1\cdot\underline{v}^1} + v^1(\underline{v}^1\cdot\underline{f}^1)\Big/\sqrt{1-\underline{v}^1\cdot\underline{v}^1}^{\,3}.$$

This may be solved for f to obtain

$$\underline{f}^1 = \underline{x}(1-\underline{v}^1\cdot\underline{v}^2) - (\underline{v}^1-\underline{v}^2)(\underline{x}\cdot\underline{v}^1)\,\frac{\sqrt{1-v^1\cdot v^1}}{\sqrt{1-v^2\cdot v^2}}\,\frac{e_1 e_2}{m}\left\{\underline{x}\cdot\underline{x} + \frac{(\underline{x}\cdot\underline{v}^2)^2}{1-\underline{v}^2\cdot\underline{v}^2}\right\}^{-3/2}.$$

If our minus sign finite mass test particle is a magnetic monopole instead of a charge, we may go through the same procedure and obtain

$$\frac{dp'}{dt} = e_1\underline{H} - e_1\underline{v}^1 \times \underline{E} = -e_1(\underline{v}^1-\underline{v}^2)\times\underline{E}$$

so

$$\underline{f}^1 = \left\{\underline{x} \times (\underline{v}^1 - \underline{v}^2) - \underline{v}^1(\underline{x}\cdot\underline{v}^1 \times \underline{v}^2)\right\} \frac{\sqrt{1-\underline{v}_1\cdot\underline{v}_1}}{\sqrt{1-\underline{v}^2\cdot\underline{v}^2}} \frac{e_1e_2}{m}$$

$$\left\{\underline{x}\cdot\underline{x} + \frac{(\underline{x}\cdot\underline{v}^2)^2}{(1-\underline{v}^2\cdot\underline{v}^2)}\right\}^{-3/2}.$$

In both cases, we have a term

$$\underline{x}\cdot\underline{x} + \frac{(\underline{x}\cdot\underline{v}^2)^2}{1-\underline{v}^2\cdot\underline{v}^2}$$

which, for any choice of the vectors $\underline{v}_1$ and $\underline{v}_2$, behaves like $x\cdot x$ for large enough separations. In our present case, this is sufficient to demonstrate that these forces may be chosen to vanish for sufficiently large separation regardless of their velocities.

If we return to the Lorentz invariance conditions, a bit of playing with the Lorentz force indicates that we may write a generalized Lorentz force of the form

$$\underline{f}^1 = \left\{\underline{x}\,(1-\underline{v}^1\cdot\underline{v}^2) - \underline{v}(\underline{x}\cdot\underline{v}^1)\right\}\frac{\sqrt{1-\underline{v}^1\cdot\underline{v}^1}}{\sqrt{1-\underline{v}^2\cdot\underline{v}^2}}\, g\left[(\underline{x}\cdot\underline{x})(1-\underline{v}^2\cdot\underline{v}^2) + (\underline{x}\cdot\underline{v}^2)^2\right],$$

where we have an arbitrary function g replacing the -3/2 power that appears in the Lorentz force. To explore some of the properties of this force, we may turn to the frame in which the infinite mass particle is at rest. In this frame, the force is then given by

$$f^1 = \left\{\underline{x} - (\underline{x}\cdot\underline{v}^1)\underline{v}^1\right\}\frac{e_1e_2\sqrt{1-\underline{v}^1\cdot\underline{v}^1}}{m_1}\, g(\underline{x}\cdot\underline{x}),$$

which may be written

$$\frac{d}{dt}\left(\underline{v}^1 \Big/ \sqrt{1-\underline{v}^1\cdot\underline{v}^1}\right) = \underline{x}\,\frac{e_1e_2}{m_1}\, g(x\cdot x),$$

i.e., an arbitrary central potential.

These examples indicate that the usual external fields may be properly described within the present discussion of particle interaction. However, we would be more interested in examples where both

particles have finite mass. One class of examples may be obtained by making the additional requirement that the multiplicative coupling constant of the force may take on an arbitrary value. This requirement simplifies the Lorentz invariance conditions so solutions of this simplified condition contain forces of the form

$$\underline{f}^1 = (\underline{v}^1 - \underline{v}^2)\,\frac{\sqrt{1-\underline{v}^1\cdot\underline{v}^1}^{\,3}}{\sqrt{(1-\underline{v}^1\cdot\underline{v}^1)(1-\underline{v}^2\cdot\underline{v}^2)-(1-\underline{v}^1\cdot\underline{v}^2)^2}}\,g(\beta),$$

where the variable β has the form

$$\beta = \underline{x}\cdot\underline{x} + \frac{(x\cdot v^2)^2}{1-\underline{v}^2\cdot\underline{v}^2} -$$

$$- \frac{(\underline{x}\cdot\underline{v}_1)^2(1-\underline{v}^2\cdot\underline{v}^2) - 2(\underline{x}\cdot\underline{v}^1)(\underline{x}\cdot\underline{v}^2)(1-\underline{v}^1\cdot\underline{v}^2) + \dfrac{(\underline{x}\cdot\underline{v}^2)^2(1-\underline{v}^1\cdot\underline{v}^2)^2}{(1-\underline{v}^2\cdot\underline{v}^2)}}{\dfrac{(1-\underline{v}^1\cdot\underline{v}^1)^2 - (1-\underline{v}^1\cdot\underline{v}^2)^2}{(1-\underline{v}^1\cdot\underline{v}^1)(1-\underline{v}^2\cdot\underline{v}^2)}}\,.$$

For large separations, this variable has the form

$$\{\text{positive coefficient depending on } v\}(x\cdot x),$$

so this force may be defined so that it vanishes as the separation goes to infinity in all frames, as did the Lorentz force.

In order to obtain some feeling for the general solution to the Lorentz invariance conditions, let us consider a "power series" solution[3] to the Lorentz invariance condition in one space dimension. We will simplify some of the algebra in these considerations by restricting the interaction to be invariant under parity, and the particles to be identical. Under the parity transformation, we have

$$x \to -x \qquad X \to -X$$
$$v \to -v \qquad V \to -V$$
$$a \to -a \qquad A \to -A.$$

This requirement of parity invariance is the one-dimensional remnant of the rotation invariance in three dimensions. The particle identity, or symmetry under particle interchange induces the transformation

$$x^1 \to x^2 \quad x^2 \to x^1 \quad v^1 \to v^2 \quad v^2 \to v^1$$

or

$$\begin{aligned} x &\to -x & X &\to -X \\ v &\to -v & V &\to -V \\ a &\to -a & A &\to -A. \end{aligned}$$

Combining these, we have

$$\begin{aligned} x &\to x & X &\to -X \\ v &\to v & V &\to -V \\ a &\to a & A &\to -A. \end{aligned}$$

Now let us expand the acceleration functions in power series in V. Thus we have

$$f(x,v,V) = f_0(x,v) + f_1(x,v)V + \tfrac{1}{2} f_2(x,v)V^2 + \ldots$$

$$F(x,v,V) = F_0(x,v) + F_1(x,v)V + \tfrac{1}{2} F_2(x,v)V^2 + \ldots$$

where the translation invariance has removed the X dependence. The above symmetry conditions imply

$$f_1 = f_3 = f_{odd} = 0$$

$$F_0 = F_2 = F_{even} = 0.$$

Now let us substitute into the invariance conditions and use the fact that the invariance conditions are true to each order in V. Thus we have for zeroth order in the second invariance condition

$$(xf_0 + v^2 - 1)F_1 = xvf_{0,x} + xf_0 f_{0,v} + 3vf_0,$$

which yields for F_1

$$F_1(x,v) = \frac{xvf_{0,x} + xf_0 f_{0,v} + 3vf}{xf_0 + v^2 - 1}.$$

For zeroth order expression in the other equation we find an identity which is the result of our assumption that the particles are identical.

If the particles were not identical, we would have a restrictive condition on the non-vanishing f_{odd} and F_{even}.

Now, going to first order in V in the same manner, we obtain

$$f_2 = \Big[xF_1F_1 + 3f_0 + 3vF_1 + xvF_{1,x} + xf_0F_{1,v} - xf_{0,x} - (xF_1 + 2v)f_{0,v}\Big] \Big/ \Big(xf_0 + v^2 - 1\Big)$$

and an identity from the first order part of the other equation.

To second order, we obtain

$$F_3 = 2\Big[x\Big(F_{1,x} + F_1F_{1,v} + \tfrac{1}{2}f_2F_1 - \tfrac{1}{2}vf_{2,x} - \tfrac{1}{2}f_0f_{2,v} - \tfrac{1}{2}f_2f_{0,v} - f_2F_1\Big) + 2vF_{1,v} - 2F_1 - \tfrac{3}{2}vf_2\Big] \Big/ \Big(xf_0 + v^2 - 1\Big).$$

From the general structure of these relations, we now see that all the higher order coefficients may be determined as functionals of $f_0(x,v)$. Thus, f_0 is the only free "parameter" entering the theory. It is an arbitrary function of two variables and it completely determines the equations of motion in every frame.

Let us now recall the class of acceleration functions which are invariant under the Galilean transformation. In this case we have, for identical particles,

$$a_G = g(x,v) \qquad A_G = G(x,v) = 0.$$

The accelerations are thus specified by an arbitrary function of two variables. Thus we see that the "number of possible acceleration functions" which are invariant under the Lorentz transformation is the same as the number which is invariant under the Galilean transformation.

The statements of the preceding two paragraphs are conditional upon the existence or convergence of the expansions which we have been considering. They might accidentally not converge, except of course for those classes of examples which we have displayed. However, the facts turn out to be otherwise.[14] There exists, by the Cauchy-Kowalski theorem, solutions of these equations of the same dimensionality as have been displayed in the power series analysis at every point in x,v,V space. Further, if the acceleration function has the form

$$f^n = (1 - v^n v^n)^{1+\epsilon}\gamma^n(x, v^1, v^2);$$

where $\gamma^1(x,1,v^2)<\infty$, $\gamma^2(x,v^1,1)<\infty$ and $\epsilon>0$, then we can show that there exists a global solution—that is, the solution at one point (x,v,V) may be continued for all values of x, v and V. In terms of the power series, it says that any reasonable $f_0(x,v)$ produces a convergent power series in V. Since the class that we have here does not include the example $a=-vv/x$, there are obviously somewhat more solutions than elaborated above, i.e., the condition can be weakened in some manner. The primary result is, however, to demonstrate the existence of many invariant equations of motion.

Since we shall shortly go on to a discussion of similar matters in terms of the Hamiltonian and phase space, let us review the situation as expressed in terms of the acceleration functions.

First, the equations of motion are called invariant if the resultant world-lines, when transformed point by point, satisfy an equation of motion of the same form in the new frame.

This requirement may be restated in terms of a differential equation which acts upon the x, v, and V of the acceleration function (equation of motion). These are the Lorentz invariance conditions.

A somewhat limited class of closed form examples has been found of invariant force functions. A power series expression which is guaranteed to exist has been found for a very large class of solutions.

Let us now consider these invariant equations of motion. There are several additional properties we will require of them. If we are to deal with scattering solutions, we might well demand "weak separability," that is, stated for two particles, if the particles are far apart, then a change in the motion of one should scarcely affect the motion of the other.[3] This property will find reflection in the Hamiltonian problem when we might ask for strong separability: that is, we will require that the Hamiltonian for the two particle system, for very large separation, has the form of the sum of individual particle Hamiltonians.

Finally, we would wish to delve into the question of Einstein causality, and determine the relation between the invariance property and the Einstein causality.

References.

1. D. G. Currie, J. Math. Phys. 4, 1470 (1963).
2. D. G. Currie, T. F. Jordan, E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963).
3. D. G. Currie, Phys. Rev. 142, 817 (1966).
4. P. Havas, Boston Studies in the Philosophy of Science, Vol. 4 (1965).
5. A. D. Fokker, Z. F. Physik 58, 386 (1929); J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945) and 21, 425 (1949).
6. P. A. M. Dirac, Proc. Roy. Soc. (London) A167, 148 (1939).

7. D. G. Currie and E. J. Saletan (1966), unpublished.
8. D. G. Currie, Bull. Am. Phys. Soc. 11, 356 (1966).
9. D. G. Currie, Ann. Phys. 36, 104 (1966).
10. E. H. Kerner, Phys. Rev. Letters 16, 667 (1966).
11. D. G. Currie and T. F. Jordan, Phys. Rev. Letters 16, 1210 (1966).
12. R. N. Hill, preprint.
13. R. N. Hill, J. Math. Phys. 8, 201 (1967).
14. D. G. Currie, to be published.

III. Hamiltonians and Conservation Laws

Conservation of Momentum and Angular Momentum

We will continue to study a classical-mechanical system of two particles. We have seen how interactions can be described by relativistically invariant equations of motion which specify the accelerations as functions of the positions and velocities. Now we will explore some of their properties. Most of this investigation will involve putting the equations of motion in Hamiltonian form and seeing what kind of structure we get. Before we do that, we can learn something about conservation of momentum and angular momentum by working directly with the Newtonian equations of motion and their invariance conditions.

We will show that if the accelerations are not zero, the constants of the motion include neither the conventional total momentum nor the conventional total angular momentum as defined for free particles or for particles in a field—in other words, the kinematic particle momentum and angular momentum. These quantities could have the same values before and after a collision by being asymptotic limits of constants of the motion. The constants of the motion would depend on the interaction. They could be the momentum and angular momentum which correspond to the generators of space translations and rotations (about which more will be said later).

Let $\underline{x}^1$ and $\underline{x}^2$ be the positions of the particles, $\underline{v}^1$ and $\underline{v}^2$ their velocities, and let

$$\underline{u}^n = m_n \underline{v}^n \left(1 - \underline{v}^{n2}\right)^{-\frac{1}{2}}$$

for $n = 1,2$, with m_1 and m_2 positive numbers. Let $\underline{x} = \underline{x}^1 - \underline{x}^2$ and $\underline{v} = \underline{v}^1 - \underline{v}^2$. Consider equations of motion

$$\frac{d\underline{u}^n}{dt} = \underline{\phi}^n(\underline{u}^1, \underline{u}^2, \underline{x})$$

for $n = 1,2$. Translation invariance implies that $\underline{\phi}^1$ and $\underline{\phi}^2$ depend on the positions of the particles only as functions of the relative position $\underline{x}$. Rotation invariance implies that $\underline{\phi}^1$ and $\underline{\phi}^2$ rotate as vectors when $\underline{u}^1, \underline{u}^2$, and $\underline{x}$ are rotated.

The conditions for Lorentz invariance are

$$(-1)^n x_j \phi^{n'}_{\ell} \partial \phi^n_k / \partial u^{n'}_{\ell} + \delta_{jk}\left(\underline{u}^n \cdot \underline{\phi}^n\right)\left(\underline{u}^{n2} + m_n^2\right)^{-\frac{1}{2}}$$

$$- u^n_j \left(\underline{u}^{n2} + m_n^2\right)^{-\frac{1}{2}} \phi^n_k + x_j \left(\underline{u}^{n'2} + m_{n'}^2\right)^{-\frac{1}{2}} u^{n'}_{\ell} \partial \phi^n_k / \partial x_{\ell}$$

$$- \left(\underline{u}^{n2} + m_n^2\right)^{\frac{1}{2}} \partial \phi^n_k / \partial u^n_j - \left(\underline{u}^{n'2} + m_{n'}^2\right)^{\frac{1}{2}} \partial \phi^n_k / \partial u^{n'}_j = 0$$

for $j,k,\ell = 1,2,3$ and $n,n' = 1,2$ with n' different from n (if $n=1$ then $n'=2$ and if $n=2$ then $n'=1$); the repeated index ℓ implies a sum. These conditions are derived in the same way as for the accelerations, from the usual Lorentz transformation of space-time coordinates, and from the requirement that for an infinitesimal Lorentz transformation the change of $d\underline{u}^n/dt$ is the same as the change of $\underline{\phi}^n$ as a function of $\underline{u}^1$, $\underline{u}^2$ and $\underline{x}$. We use these conditions to obtain the following.

Theorem 1: The conventional total momentum $\underline{u}^1 + \underline{u}^2$ is not a constant of the motion unless both $\underline{\phi}^1$ and $\underline{\phi}^2$ are zero.
Proof: Suppose $\underline{u}^1 + \underline{u}^2$ is a constant of the motion. Then $\underline{\phi}^1 + \underline{\phi}^2 = 0$. Let $\underline{e}$ be any three-vector orthogonal to $\underline{x}$. Multiplying the conditions for Lorentz invariance by e_j, summing for $j=1,2,3$, and adding the result for $n=1,2$ yields

$$\left(\underline{u}^{1^2} + m_1^2\right)^{-\frac{1}{2}} \left\{ (\underline{u}^1 \cdot \underline{\phi}^1)\underline{e} - (\underline{u}^1 \cdot \underline{e})\underline{\phi}^1 \right\}$$

$$+ \left(\underline{u}^{2^2} + m_2^2\right)^{-\frac{1}{2}} \left\{ (\underline{u}^2 \cdot \underline{\phi}^2)\underline{e} - (\underline{u}^2 \cdot \underline{e})\underline{\phi}^2 \right\} = 0$$

or

$$\left\{ \left(\underline{u}^{1^2} + m_1^2\right)^{-\frac{1}{2}} \underline{u}^1 - \left(\underline{u}^{2^2} + m_2^2\right)^{-\frac{1}{2}} \underline{u}^2 \right\} \times (\underline{e} \times \underline{\phi}^1) = 0$$

which means that $\underline{e} \times \underline{\phi}^1$ is collinear with the relative velocity

$$\underline{v} = \underline{v}^1 - \underline{v}^2 = \left(\underline{u}^{1\,2} + m_1^2\right)^{-\frac{1}{2}} \underline{u}^1 - \left(\underline{u}^{2\,2} + m_2^2\right)^{-\frac{1}{2}} \underline{u}^2 .$$

If $\underline{\phi}^1$ is not zero, then $\underline{e}$ is orthogonal to $\underline{v}$, which means that $\underline{x}$ and $\underline{v}$ are collinear, and $\underline{\phi}^1$ is orthogonal to $\underline{v}$.

We assume that $\underline{\phi}^1$ and $\underline{\phi}^2$ are functions which are differentiable enough for the Lorentz-invariance conditions to be meaningful. We do not consider singular $\underline{\phi}^1$ and $\underline{\phi}^2$ which are zero for almost all values of $\underline{u}^1$, $\underline{u}^2$ and $\underline{x}$. Thus if $\underline{\phi}^1$ and $\underline{\phi}^2$ are nonzero only when the relative position $\underline{x}$ is collinear with the relative velocity $\underline{v}$, we say that $\underline{\phi}^1$ and $\underline{\phi}^2$ are zero.

Theorem 2: The conventional total angular momentum $\underline{x}^1 \times \underline{u}^1 + \underline{x}^2 \times \underline{u}^2$ is not a constant of the motion unless both $\underline{\phi}^1$ and $\underline{\phi}^2$ are zero.

Proof: Suppose $\underline{x}^1 \times \underline{u}^1 + \underline{x}^2 \times \underline{u}^2$ is a constant of the motion. Then $\underline{x}^1 \times \underline{\phi}^1 + \underline{x}^2 \times \underline{\phi}^2 = 0$. Let $\underline{e}$ be any three-vector orthogonal to $\underline{x}$. Multiplying the conditions for Lorentz invariance by $e_j \epsilon_{k\ell m} x^n_\ell$, summing for $j,k,\ell = 1,2,3$, using

$$\underline{v}^n = \underline{u}^n \left(\underline{u}^{n\,2} + m_n^2\right)^{-\frac{1}{2}},$$

and adding the results for $n = 1,2$ yields the m^{th} component of

$$(\underline{v}^1 \cdot \underline{\phi}^1)\underline{e} \times \underline{x}^1 + (\underline{v}^2 \cdot \underline{\phi}^2)\underline{e} \times \underline{x}^2 - (\underline{e} \cdot \underline{v}^1)\underline{\phi}^1 \times \underline{x}^1$$

$$- (\underline{e} \cdot \underline{v}^2)\underline{\phi}^2 \times \underline{x}^2 = 0.$$

Taking the scalar product of this with $\underline{e}$, we get

$$(\underline{e} \cdot \underline{v})\underline{x}^1 \times \underline{\phi}^1 \cdot \underline{e} = 0.$$

If $\underline{\phi}^1$ is not zero, then $\underline{e} \cdot \underline{v}$ is zero, which means that $\underline{x}$ and $\underline{v}$ are collinear.

These techniques can be used to show also that $m_1 \underline{v}^1 + m_2 \underline{v}^2$ is not a constant of the motion unless the accelerations are zero.[1] This is an example of a statement which is true for three-dimensional space but not for one-dimensional space. We have seen a one-dimensional example of accelerations with the property that $v^1 + v^2$ is a constant of the motion. In three dimensions these appear as singular accelerations which are nonzero only when the relative position $\underline{x}$ is collinear with the relative velocity $\underline{v}$.

Another statement which is true for three-dimensional space but not for one-dimensional space is the theorem that the conventional total momentum $\underline{u}^1 + \underline{u}^2$ is not a constant of the motion unless both $\underline{\phi}^1$ and $\underline{\phi}^2$ are zero. For a one-dimensional counter-example, consider constant functions

$$du^1/dt = \phi^1 = b$$

and

$$du^2/dt = \phi^2 = -b$$

with b a real number. It is easy to see that they satisfy the Lorentz-invariance conditions for one-dimensional space. This example violates the parity requirement that ϕ^1 and ϕ^2 change sign when v^1, v^2 and x change sign. We might expect to get this property from rotation invariance when a three-dimensional system is cut down to one dimension.

The lack of conservation of $\underline{u}^1 + \underline{u}^2$ and $\underline{x}^1 \times \underline{u}^1 + \underline{x}^2 \times \underline{u}^2$ is one reason why we do not emphasize the conventional momentum variables $\underline{u}^1$ and $\underline{u}^2$. When we put the equations of motion in Hamiltonian form, we will want canonical momenta $\underline{p}^1$ and $\underline{p}^2$ such that $\underline{p}^1 + \underline{p}^2$ and $\underline{x}^1 \times \underline{p}^1 + \underline{x}^2 \times \underline{p}^2$ are constants of the motion. Evidently $\underline{p}^1$ and $\underline{p}^2$ will not be the same as $\underline{u}^1$ and $\underline{u}^2$. They could be the same asymptotically when the particles are widely separated and not interacting. This would imply that $\underline{p}^1$ and $\underline{p}^2$ are functions of the relative position of the particles as well as functions of their velocities.

Generators of Noncanonical Transformations and Their Lie Brackets

Now we want to put the equations of motion in Hamiltonian form. This involves choosing canonical momenta as functions of the positions and velocities. Then the Poincaré group is represented by the transformations of positions and canonical momenta which follow from the transformations of positions and velocities. We expect that the canonical momenta can be chosen to be invariant under space translations and vectors under space rotations, so for this part of the Poincaré group we will have the usual transformations, which are canonical. But there is no reason to expect the Lorentz transformations to be canonical. In fact, we will see that there can be no interaction if the whole Poincaré group is represented by canonical transformations. Therefore, we must learn how to work with infinitesimal generators of noncanonical transformations.

Consider a classical-mechanical system of N particles described by positions $\underline{q}^n$ and canonical momenta $\underline{p}^n$ for $n = 1, 2, \ldots N$. Suppose we have a one-parameter group of transformations of this phase space. This means that there are real functions $\underline{q}^n(u, \underline{q}, \underline{p})$

and $\underline{p}^n(u, \underline{q}, \underline{p})$ of the phase-space variables $\underline{q}$ and $\underline{p}$, which depend also on a real parameter u, such that

$$\underline{q}^n(0, \underline{q}, \underline{p}) = \underline{q}^n$$

$$\underline{p}^n(0, \underline{q}, \underline{p}) = \underline{p}^n$$

$$\underline{q}^n(u+t, \underline{q}, \underline{p}) = \underline{q}^n\Big(u, \underline{q}(t, \underline{q}, \underline{p}), \underline{p}(t, \underline{q}, \underline{p})\Big)$$

$$\underline{p}^n(u+t, \underline{q}, \underline{p}) = \underline{p}^n\Big(u, \underline{q}(t, \underline{q}, \underline{p}), \underline{p}(t, \underline{q}, \underline{p})\Big).$$

For example, these could be transformations in time, or they could be Lorentz transformations.

Each one-parameter group of transformations has an infinitesimal generator. Let

$$[\underline{q}^n, \hat{K}] = [\partial \underline{q}^n(u)/\partial u]_{u=0}$$

$$[\underline{p}^n, \hat{K}] = [\partial \underline{p}^n(u)/\partial u]_{u=0}.$$

These are real functions of the phase-space variables $\underline{q}$ and $\underline{p}$. Let $\hat{K}$ denote this set of functions. This is the generator. From the group property we get

$$\partial \underline{q}^n(u, \underline{q}, \underline{p})/\partial u = \lim_{t\to 0} (1/t)\Big\{\underline{q}^n(u+t, \underline{q}, \underline{p}) - \underline{q}^n(u, \underline{q}, \underline{p})\Big\}$$

$$= \lim_{t\to 0} (1/t)\Big\{\underline{q}^n(u, \underline{q}(t), \underline{p}(t)) - \underline{q}^n(u, \underline{q}, \underline{p})\Big\}$$

$$= \sum_{m=1}^{N} \sum_{j=1}^{3} \Big\{[q_j^m, \hat{K}]\, \partial \underline{q}^n(u, \underline{q}, \underline{p})/\partial q_j^m +$$

$$+ [p_j^m, \hat{K}] \partial \underline{q}^n(u, \underline{q}, \underline{p})/\partial p_j^m\Big\}$$

and similarly

$$\partial \underline{p}^n(u)/\partial u = \sum_{m=1}^{N} \sum_{j=1}^{3} \Big\{[q_j^m, \hat{K}]\, \partial \underline{p}^n(u)/\partial q_j^m + [p_j^m, \hat{K}]\, \partial \underline{p}^n(u)/\partial p_j^m\Big\}.$$

The functions $\underline{q}^n(u, \underline{q}, \underline{p})$ and $\underline{p}^n(u, \underline{q}, \underline{p})$ are the solutions of these differential equations which equal $\underline{q}^n$ and $\underline{p}^n$ at $u=0$. Thus the

one-parameter group of transformations is characterized by its generator.

For a function $F(\underline{q},\underline{p})$ of the phase-space variables $\underline{q}$ and $\underline{p}$, let

$$[F,\hat{K}] = \sum_{m=1}^{N}\sum_{j=1}^{3}\left\{[q_j^m,\hat{K}]\,\partial F/\partial q_j^m + [p_j^m,\hat{K}]\,\partial F/\partial p_j^m\right\}.$$

Then

$$\partial \underline{q}^n(u)/\partial u = [\underline{q}^n(u),\hat{K}]$$

$$\partial \underline{p}^n(u)/\partial u = [\underline{p}^n(u),\hat{K}]$$

$$\partial F\Big(\underline{q}(u),\underline{p}(u)\Big)/\partial u = \sum_{n=1}^{N}\sum_{k=1}^{3}\left\{\left[\partial F\Big(\underline{q}(u),\underline{p}(u)\Big)/\partial q_k^n(u)\right]\partial q_k^n(u)/\partial u\right.$$

$$\left. + \left[\partial F\Big(\underline{q}(u),\underline{p}(u)\Big)/\partial p_k^n(u)\right]\partial p_k^n(u)/\partial u\right\}$$

$$= \sum_{m,n=1}^{N}\sum_{j,k=1}^{3}\left\{\left[\partial F/\partial q_k^n(u)\right]\left(\left[q_j^m,\hat{K}\right]\partial q_k^n(u)/\partial q_j^m\right.\right.$$

$$\left. + \left[p_j^m,\hat{K}\right]\partial q_k^n(u)/\partial p_j^m\right) + \left[\partial F/\partial p_k^n(u)\right]$$

$$\left.\cdot\left(\left[q_j^m,\hat{K}\right]\partial p_k^n(u)/\partial q_j^m + \left[p_j^m,\hat{K}\right]\partial p_k^n(u)/\partial p_j^m\right)\right\}$$

$$= \left[F\Big(\underline{q}(u),\underline{p}(u)\Big),\ \hat{K}\right].$$

The generator $\hat{K}$ is determined by the set of functions $[F,\hat{K}]$ for any complete set of independent functions $F(\underline{q},\underline{p})$ of the phase-space variables.

When we say that a transformation of the phase-space variables $\underline{q}$ and $\underline{p}$ is canonical, we mean that it preserves Poisson brackets. The transformations generated by $\hat{K}$ are canonical if and only if there is a function $K(\underline{q},\underline{p})$ of the phase-space variables such that

$$\left[q_k^n,\hat{K}\right] = \partial K/\partial p_k^n$$

and

$$\left[p_k^n, \hat{K}\right] = -\partial K/\partial q_k^n.$$

Suppose there is such a function K. Then $[F,\hat{K}]$ is the Poisson bracket $[F,K]$ for any function $F(\underline{q},\underline{p})$ of the phase-space variables. Using Poisson brackets and the Jacobi identity, we get

$$(\partial/\partial u)\left[q_j^m(u), p_k^n(u)\right] = \left[\left[q_j^m(u), K\right], p_k^n(u)\right] + \left[q_j^m(u), \left[p_k^n(u), K\right]\right]$$

$$= \left[\left[q_j^m(u), p_k^n(u)\right], K\right].$$

This implies that as a function of u the Poisson bracket $[q_j^m(u), p_k^n(u)]$ does not change from the value $\delta_{mn}\delta_{jk}$ which it has at $u=0$. It follows similarly that

$$\left[q_j^m(u), q_k^n(u)\right] = 0 = \left[p_j^m(u), p_k^n(u)\right].$$

Thus the transformations are canonical. Conversely, suppose the transformations are canonical. Taking the derivative of

$$\left[q_j^m(u), p_k^n(u)\right] = \delta_{mn}\delta_{jk}$$

with respect to u at $u=0$, we get

$$(\partial/\partial q_k^n)\left[q_j^m, \hat{K}\right] + (\partial/\partial p_j^m)\left[p_k^n, \hat{K}\right] = 0.$$

Similarly, from $[q_j^m(u), q_k^n(u)] = 0$ we get

$$(\partial/\partial p_j^m)\left[q_k^n, \hat{K}\right] - (\partial/\partial p_k^n)\left[q_j^m, \hat{K}\right] = 0$$

and from $[p_j^m(u), p_k^n(u)] = 0$ we get

$$(\partial/\partial q_k^n)\left[p_j^m, \hat{K}\right] - (\partial/\partial q_j^m)\left[p_k^n, \hat{K}\right] = 0.$$

These are the integrability conditions for existence of a function K.

Let $\hat{K}_1$ and $\hat{K}_2$ be generators for two one-parameter groups of transformations. Their Lie bracket $[\hat{K}_1, \hat{K}_2]$ is defined by

$$\left[F,[\hat{K}_1,\hat{K}_2]\right] = \left[[F,\hat{K}_1],\hat{K}_2\right] - \left[[F,\hat{K}_2],\hat{K}_1\right]$$

for functions $F(\underline{q},\underline{p})$ of the phase-space variables. If $\hat{K}_1$ and $\hat{K}_2$ are two of the generators for a many-parameter group of transformations, for example, the ten-parameter Poincaré group, then $[\hat{K}_1,\hat{K}_2]$ is one of the brackets which characterizes the infinitesimal structure of the group. This bracket is related to the commutator of the two one-parameter groups of transformations in the following standard way.[2] Suppose that to a function $F(\underline{q},\underline{p})$ we apply first a transformation generated by $\hat{K}_1$, second a transformation generated by $\hat{K}_2$, using the same infinitesimal value u for the parameter of both, and then apply the inverse of the first transformation followed by the inverse of the second transformation, keeping terms to second order in u. The result is

$$F + u^2\left\{\left[[F,\hat{K}_1],\hat{K}_2\right] - \left[[F,\hat{K}_2],\hat{K}_1\right]\right\}.$$

(Terms involving only one of the generators cancel because they just contribute to a transformation followed by its inverse. Also two terms $u^2[[F,\hat{K}_1],\hat{K}_2]$ cancel.)

The definition of $[\hat{K}_1,\hat{K}_2]$ looks like a Jacobi identity. Suppose $\hat{K}_1$ and $\hat{K}_2$ generate canonical transformations. Let K_1 and K_2 be functions such that $[F,\hat{K}_1]$ and $[F,\hat{K}_2]$ are the Poisson brackets $[F,K_1]$ and $[F,K_2]$ for all functions F. Then, because Poisson brackets satisfy the Jacobi identity, we see from the definition that $[F,[\hat{K}_1,\hat{K}_2]]$ is the Poisson bracket $[F,[K_1,K_2]]$ for all functions F, with the Poisson bracket $[K_1,K_2]$ replacing $[\hat{K}_1,\hat{K}_2]$. Thus all brackets reduce to Poisson brackets in the case of canonical transformations.

In the general case of noncanonical transformations there are two kinds of brackets: the derivative $[F,\hat{K}]$ of a function $F(\underline{q},\underline{p})$ for transformations generated by $\hat{K}$, and the Lie bracket $[\hat{K}_1,\hat{K}_2]$ of two generators. It is convenient to use bracket notation for both. The definition of the Lie bracket provides a Jacobi identity involving both kinds of brackets. Both kinds of brackets can be viewed as commutators by using operators on the Hilbert space of square-integrable functions of the phase-space variables. This is outlined in the next section.

Operator Formulation

Both kinds of brackets described in the preceding section can be viewed as commutators in an operator formulation of classical mechanics.[3] Consider the Hilbert space of square-integrable functions of the phase-space variables $\underline{q}$ and $\underline{p}$. Each physical quantity is a

function $F(\underline{q},\underline{p})$ of the phase-space variables. It is represented by the operator which multiplies every function in the Hilbert space by $F(\underline{q},\underline{p})$. All of these operators commute with each other.

A generator of a one-parameter group of transformations is a Hermitian operator

$$\hat{K} = -i \sum_{m=1}^{N} \sum_{j=1}^{3} \left\{ \left[q_j^m, \hat{K} \right] \frac{\partial}{\partial q_j^m} + \left[p_j^m, \hat{K} \right] \frac{\partial}{\partial p_j^m} \right\}.$$

The bracket $[F,\hat{K}]$ for a function $F(\underline{q},\underline{p})$ is the commutator of the operators F and $\hat{K}$:

$$[F,\hat{K}] = -i(F\hat{K} - \hat{K}F).$$

In particular

$$\left[q_j^m, \hat{K} \right] = -i\left(q_j^m \hat{K} - \hat{K} q_j^m \right)$$

and

$$\left[p_j^m, \hat{K} \right] = -i\left(p_j^m \hat{K} - \hat{K} p_j^m \right).$$

The operators $\underline{q}^n$ and $\underline{p}^n$ are transformed to the operators

$$\underline{q}^n(u) = e^{iu\hat{K}} \underline{q}^n e^{-iu\hat{K}}$$

and

$$\underline{p}^n(u) = e^{iu\hat{K}} \underline{p}^n e^{-iu\hat{K}}.$$

An operator $F(\underline{q},\underline{p})$ is transformed to the operator

$$F\left(\underline{q}(u), \underline{p}(u)\right) = e^{iu\hat{K}} F(\underline{q},\underline{p}) e^{-iu\hat{K}}.$$

Then

$$\partial F\left(\underline{q}(u), \underline{p}(u)\right)/\partial u = -i\left\{ F\left(\underline{q}(u), \underline{p}(u)\right)\hat{K} - \hat{K}F\left(\underline{q}(u), \underline{p}(u)\right) \right\}$$

$$= \left[F\left(\underline{q}(u), \underline{p}(u)\right), \hat{K} \right].$$

In particular,

$$[\partial \underline{q}^n(u)/\partial u]_{u=0} = [\underline{q}^n, \hat{K}]$$

and

$$[\partial \underline{p}^n(u)/\partial u]_{u=0} = [\underline{p}^n, \hat{K}].$$

The operators $\underline{q}^n(u)$ and $\underline{p}^n(u)$ are functions of the operators $\underline{q}$ and $\underline{p}$. We have

$$\underline{q}^n(t+u) = e^{iu\hat{K}}\underline{q}^n(t)e^{-iu\hat{K}} = \underline{q}^n\Big(t, \underline{q}(u), \underline{p}(u)\Big)$$

and similarly

$$\underline{p}^n(t+u) = \underline{p}^n\Big(t, \underline{q}(u), \underline{p}(u)\Big).$$

Just as in quantum mechanics, the Lie bracket of two generators is their commutator:

$$[\hat{K}_1, \hat{K}_2] = -i(\hat{K}_1\hat{K}_2 - \hat{K}_2\hat{K}_1).$$

Generators and Bracket Relations for the Poincaré Group

Now we can use generators and Lie-bracket relations for the Poincaré group to develop a relativistic Hamiltonian formalism with noncanonical Lorentz transformations. This is what we expect to get by putting the Lorentz-invariant equations of motion in Hamiltonian form. We will establish a correspondence between invariance of the Newtonian equations of motion and the Lie-bracket relations for the Poincaré group. We will see that canonical Lorentz transformations are not required either for invariance of the equations of motion or for the relativistic Hamiltonian formalism.

Suppose the classical-mechanical description of a system of N particles admits the Poincaré group of transformations. Let $\hat{H}$, $\hat{\underline{P}}$ $\hat{\underline{J}}$, $\hat{\underline{K}}$ be the generators, of the kind described in the two preceding sections, for time translations, space translations, space rotations, and Lorentz transformations, respectively. We assume that these generators satisfy the Lie-bracket relations characteristic of the Poincaré group. These Lie-bracket relations will be the basic object of study.

We assume that the positions $\underline{q}^n$ transform as usual under space translations and rotations. This means that

$$\left[q_j^n, \hat{P}_k\right] = \delta_{jk}$$

and

$$\left[q_j^n, \hat{J}_k\right] = \epsilon_{jk\ell} q_\ell^n$$

for $n = 1, 2, \ldots N$ and $j, k, \ell = 1, 2, 3$. We assume also that the time-dependent positions transform as usual under Lorentz transformations. This means that

$$\left[q_j^n, \hat{K}_k\right] = q_k^n \left[q_j^n, \hat{H}\right]$$

for $n = 1, 2, \ldots N$ and $j, k = 1, 2, 3$.

We assume that $\hat{H}$ generates canonical transformations. This means that there is a function $H(\underline{q}, \underline{p})$ of the phase-space variables such that $[F, \hat{H}]$ is the Poisson bracket $[F, H]$ for all functions $F(\underline{q}, \underline{p})$ of the phase-space variables. We assume that space translations and rotations also are canonical. This means that there are functions $\underline{P}$ and $\underline{J}$ of the phase-space variables such that $[F, \hat{\underline{P}}]$ and $[F, \hat{\underline{J}}]$ are the Poisson brackets $[F, \underline{P}]$ and $[F, \underline{J}]$ for all functions F of the phase-space variables. Then the Lie-bracket relations involving only $\hat{\underline{P}}$ and $\hat{\underline{J}}$ are the Poisson-bracket relations

$$[P_j, P_k] = 0$$

$$[J_j, J_k] = \epsilon_{jk\ell} J_\ell$$

$$[J_j, P_k] = \epsilon_{jk\ell} P_\ell$$

and transformations of the positions $\underline{q}^n$ under space translations and rotations are characterized by the Poisson-bracket relations

$$[q_j^n, P_k] = \delta_{jk}$$

and

$$[q_j^n, J_k] = \epsilon_{jk\ell} q_\ell^n .$$

From these it follows that $\underline{P}$ and $\underline{J}$ can be put in the standard forms

$$\underline{P} = \sum_{n=1}^{N} \underline{p}^n$$

and

$$\underline{J} = \sum_{n=1}^{N} \underline{q}^n \times \underline{p}^n$$

by a canonical transformation which leaves the positions $\underline{q}^n$ unchanged.[4] Thus without loss of generality we may assume that $\underline{P}$ and $\underline{J}$ are these standard functions.

We assume that the equations

$$\underline{v}^m = [\underline{q}^m, \hat{H}] = \partial H/\partial \underline{p}^m$$

have a solution for the canonical momenta $\underline{p}^n$ as functions of the positions $\underline{q}^m$ and the velocities $\underline{v}^m$ for $m = 1,2,\ldots N$. With Hamilton's equations, the motion is determined by the same initial data as for Newton's equations: the positions and velocities at time zero.

The motivation for these assumptions is the following. We want equations of motion which specify the accelerations of the particles as functions of the positions and velocities. We want the equations of motion to be invariant under the Poincaré group of transformations. We want to put these equations of motion in Hamiltonian form. This involves choosing canonical momenta as functions of the positions and velocities. Then transformations of the canonical momenta are determined by the transformations of the positions and velocities. We expect that the canonical momenta can be chosen to be invariant under space translations and vectors under space rotations so that the standard functions $\underline{P}$ and $\underline{J}$ are suitable for generators. If there is interaction, the canonical momenta may be complicated functions of the positions and velocities, and their Lorentz transformations may be complicated. Therefore, we do not assume that Lorentz transformations are canonical.

We want constants of the motion H, $\underline{P}$, $\underline{J}$ which correspond, in the usual way, to invariance of the equations of motion under time translations, space translations, and space rotations. We do not have an equally strong motivation for associating a function of the phase-space variables with the generator $\hat{\underline{K}}$ of Lorentz transformations. That it is consistent for Lorentz transformations to be noncanonical, when time translations and space translations and rotations are canonical, is demonstrated in the following example for a single free particle.

Free-Particle Example: Let the equation of motion $d^2\underline{q}/dt^2 = 0$ for a single free particle be described by the Hamiltonian

$$H = (1/2)\underline{p}^2.$$

Then the canonical momentum $\underline{p}$ is the velocity $\underline{v}$ because

$$\underline{v} = [\underline{q},\hat{H}] = \partial H/\partial \underline{p} = \underline{p}.$$

The transformations of $\underline{q}$ and $\underline{v}$ under space translations and rotations are the canonical transformations of $\underline{q}$ and $\underline{p}$ generated by the standard functions

$$\underline{P} = \underline{p}$$

and

$$\underline{J} = \underline{q} \times \underline{p}.$$

For Lorentz transformations we have

$$[q_j,\hat{K}_k] = q_k[q_j,\hat{H}] = q_k v_j = q_k p_j$$

and

$$[p_j,\hat{K}_k] = [v_j,\hat{K}_k] = v_j v_k - \delta_{jk} = p_j p_k - \delta_{jk}.$$

The Lorentz transformations are noncanonical. For example,

$$\partial[q_j,\hat{K}_\ell]/\partial q_k + \partial[p_k,\hat{K}_\ell]/\partial p_j = \delta_{k\ell} p_j + \delta_{j\ell} p_k + \delta_{jk} p_\ell.$$

For canonical transformations this would be zero. The generators satisfy the Lie-bracket relations for the Poincaré group, because they generate the usual transformations of $\underline{q}$ and $\underline{v}$ for a free particle; one can check this explicitly. The free-particle equation of motion is obviously invariant under the transformations of the Poincaré group. We will make more use of this example later.

For the Lie-bracket relations involving $\hat{H}$ but not $\underline{\hat{K}}$, we have the Poisson-bracket relations

$$[H,\underline{P}] = 0$$

and

$$[H,\underline{J}] = 0.$$

These just imply that H is a function of only the canonical momenta and the relative positions and is invariant under rotations.[4)]

Of those involving $\hat{K}$, the Lie-bracket relation

$$[\underline{\hat{K}}, \hat{H}] = \underline{\hat{P}}$$

is decisive. It is easy to see why. If we apply $[\hat{K}_k, \hat{H}] = \hat{P}_k$ to q^n_j, we get

$$\begin{aligned}\left[\left[q^n_j, \hat{H}\right], \hat{K}_k\right] &= \left[\left[q^n_j, \hat{K}_k\right], \hat{H}\right] - \left[q^n_j, \hat{P}_k\right] \\ &= \left[q^n_k \left[q^n_j, \hat{H}\right], \hat{H}\right] - \delta_{jk} \\ &= q^n_k \left[\left[q^n_j, \hat{H}\right], \hat{H}\right] + \left[q^n_j, \hat{H}\right]\left[q^n_k, \hat{H}\right] - \delta_{jk}.\end{aligned}$$

If we write

$$\underline{v}^n = [\underline{q}^n, \hat{H}]$$

and

$$\underline{a}^n = \left[[\underline{q}^n, \hat{H}], \hat{H}\right]$$

for velocity and acceleration, we have

$$\left[v^n_j, \hat{K}_k\right] = q^n_k a^n_j + v^n_j v^n_k - \delta_{jk}.$$

This is just the equation which characterizes the transformations of velocity derived from Lorentz transformations of the time-dependent position. Now if we apply $[\hat{K}_k, \hat{H}] = \hat{P}_k$ to $[q^n_j, \hat{H}]$, and use the equation just obtained, we get

$$\begin{aligned}\left[\left[\left[q^n_j, \hat{H}\right], \hat{H}\right], \hat{K}_k\right] &= \left[\left[\left[q^n_j, \hat{H}\right], \hat{K}_k\right], \hat{H}\right] - \left[\left[q^n_j, \hat{H}\right], \hat{P}_k\right] \\ &= q^n_k \left[\left[\left[q^n_j, \hat{H}\right], \hat{H}\right], \hat{H}\right] + 2\left[q^n_k, \hat{H}\right] \\ &\quad \cdot \left[\left[q^n_j, \hat{H}\right], \hat{H}\right] + \left[q^n_j, \hat{H}\right]\left[\left[q^n_k, \hat{H}\right], \hat{H}\right]\end{aligned}$$

because

$$\left[\left[q^n{}_j,\hat{H}\right],\hat{P}_k\right] = \left[\left[q^n{}_j,\hat{P}_k\right],\hat{H}\right] + \left[q^n{}_j,\left[\hat{H},\hat{P}_k\right]\right] = 0.$$

In terms of $\underline{v}^n$, $\underline{a}^n$, and

$$\dot{\underline{a}}^n = \left[\underline{a}^n,\hat{H}\right] = \left[\left[\left[\underline{q}^n,\hat{H}\right],\hat{H}\right],\hat{H}\right]$$

we have

$$\left[a^n{}_j,\hat{K}_k\right] = q^n{}_k\dot{a}^n{}_j + 2v^n{}_k a^n{}_j + v^n{}_j a^n{}_k.$$

This is just the equation which characterizes the transformations of acceleration derived from Lorentz transformations of the time-dependent position.

The canonical momenta are functions of the positions and velocities. With $[\hat{\underline{K}},\hat{H}] = \hat{\underline{P}}$ applied to positions we see that the Lorentz transformations of positions and canonical momenta generated by $\hat{\underline{K}}$ are just the usual Lorentz transformations of positions and velocities. The accelerations are functions of the positions and canonical momenta and therefore also functions of the positions and velocities. With $[\hat{\underline{K}},\hat{H}] = \hat{\underline{P}}$ applied to velocities we see that the Lorentz transformations of these functions are the same as Lorentz transformations of accelerations. This means that the equations of motion are invariant under Lorentz transformations.

The other Lie-bracket relations involving $\hat{\underline{K}}$ follow as a consequence of $[\hat{\underline{K}},\hat{H}] = \hat{\underline{P}}$. For example, we have

$$\left[q^n{}_i,\left[\hat{K}_j,\hat{K}_k\right]\right] = \left[\left[q^n{}_i,\hat{K}_j\right],\hat{K}_k\right] - \left[\left[q^n{}_i,\hat{K}_k\right],\hat{K}_j\right]$$

$$= -q^n{}_j\delta_{ik} + q^n{}_k\delta_{ij} = -\epsilon_{jk\ell}\epsilon_{i\ell m}q^n{}_m$$

$$= -\epsilon_{jk\ell}\left[q^n{}_i,\hat{J}_\ell\right]$$

because

$$\left[\left[q^n{}_i,\hat{K}_j\right],\hat{K}_k\right] = \left[q^n{}_j\left[q^n{}_i,\hat{H}\right],\hat{K}_k\right]$$

$$= \left[q^n{}_j,\hat{K}_k\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[\left[q^n{}_i,\hat{H}\right],\hat{K}_k\right]$$

$$= q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[\left[q^n{}_i,\hat{K}_k\right],\hat{H}\right]$$

$$- q^n{}_j\left[q^n{}_i,\left[\hat{K}_k,\hat{H}\right]\right]$$

$$= q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[q^n{}_k\left[q^n{}_i,\hat{H}\right],\hat{H}\right]$$

$$- q^n{}_j\left[q^n{}_i,\hat{P}_k\right]$$

$$= q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[q^n{}_k,\hat{H}\right]\left[q^n{}_i,\hat{H}\right]$$

$$+ q^n{}_j q^n{}_k\left[\left[q^n{}_i,\hat{H}\right],\hat{H}\right] - q^n{}_j\delta_{ik},$$

and we have

$$\left[\left[q^n{}_i,\hat{H}\right],\left[\hat{K}_j,\hat{K}_k\right]\right] = \left[\left[\left[q^n{}_i,\hat{H}\right],\hat{K}_j\right],\hat{K}_k\right] - \left[\left[\left[q^n{}_i,\hat{H}\right],\hat{K}_k\right],\hat{K}_j\right]$$

$$= -\left[q^n{}_j,\hat{H}\right]\delta_{ik} + \left[q^n{}_k,\hat{H}\right]\delta_{ij}$$

$$= -\epsilon_{jk\ell}\epsilon_{i\ell m}\left[q^n{}_m,\hat{H}\right]$$

$$= -\epsilon_{jk\ell}\left[\left[q^n{}_i,\hat{J}_\ell\right],\hat{H}\right]$$

$$= -\epsilon_{jk\ell}\left[\left[q^n{}_i,\hat{H}\right],\hat{J}_\ell\right] - \epsilon_{jk\ell}\left[q^n{}_i,\left[\hat{J}_\ell,\hat{H}\right]\right]$$

$$= -\epsilon_{jk\ell}\left[\left[q^n{}_i,\hat{H}\right],\hat{J}_\ell\right]$$

because

$$\begin{aligned}
\left[\left[\left[q^n{}_i,\hat{H}\right],\hat{K}_j\right],\hat{K}_k\right] &= \left[\left[\left[q^n{}_i,\hat{K}_j\right],\hat{H}\right],\hat{K}_k\right] - \left[\left[q^n{}_i,\left[\hat{K}_j,\hat{H}\right]\right],\hat{K}_k\right] \\
&= \left[\left[q^n{}_j\left[q^n{}_i,\hat{H}\right],\hat{H}\right],\hat{K}_k\right] - \left[\left[q^n{}_i,\hat{P}_j\right],\hat{K}_k\right] \\
&= \left[\left[q^n{}_j\left[q^n{}_i,\hat{H}\right],\hat{K}_k\right],\hat{H}\right] - \left[q^n{}_j\left[q^n{}_i,\hat{H}\right],\left[\hat{K}_k,\hat{H}\right]\right] \\
&= \left[\left[q^n{}_j,\hat{K}_k\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[\left[q^n{}_i,\hat{H}\right],\hat{K}_k\right],\hat{H}\right] \\
&\qquad - \left[q^n{}_j\left[q^n{}_i,\hat{H}\right],\hat{P}_k\right] \\
&= \left[q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[\left[q^n{}_i,\hat{K}_k\right],\hat{H}\right]\right. \\
&\qquad \left. - q^n{}_j\left[q^n{}_i,\left[\hat{K}_k,\hat{H}\right]\right],\hat{H}\right] - \delta_{jk}\left[q^n{}_i,\hat{H}\right] \\
&= \left[q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[q^n{}_k\left[q^n{}_i,\hat{H}\right],\hat{H}\right]\right. \\
&\qquad \left. - q^n{}_j\left[q^n{}_i,\hat{P}_k\right],\hat{H}\right] - \delta_{jk}\left[q^n{}_i,\hat{H}\right] \\
&= \left[q^n{}_k\left[q^n{}_j,\hat{H}\right]\left[q^n{}_i,\hat{H}\right] + q^n{}_j\left[q^n{}_k,\hat{H}\right]\left[q^n{}_i,\hat{H}\right]\right. \\
&\qquad \left. + q^n{}_j q^n{}_k\left[\left[q^n{}_i,\hat{H}\right],\hat{H}\right],\hat{H}\right] - \left[q^n{}_j,\hat{H}\right]\delta_{ik} \\
&\qquad\qquad - \delta_{jk}\left[q^n{}_i,\hat{H}\right].
\end{aligned}$$

This is the Lie-bracket relation

$$[\hat{K}_j,\hat{K}_k] = -\epsilon_{jk\ell}\hat{J}_\ell$$

applied to positions and velocities. This is sufficient to establish the Lie-bracket relation because the canonical momenta, and all

functions of the phase-space variables, are functions of the positions and velocities. The remaining Lie-bracket relations

$$[\hat{K}_j, \hat{P}_k] = \delta_{jk}\hat{H}$$

and

$$[\hat{K}_j, \hat{J}_k] = \epsilon_{jk\ell}\hat{K}_\ell$$

can be verified similarly on the positions and velocities.

Thus we see how the Lie-bracket relations correspond to invariance of the equations of motion. Lorentz invariance of the equations of motion does not require canonical Lorentz transformations with generators corresponding to functions of the phase-space variables. In particular, this is not required by the conservation law which follows, according to Noether's theorem, from Lorentz invariance of the equations of motion.[5] This can be seen explicitly in the free-particle example above where noncanonical Lorentz transformations are used for the Lorentz-invariant free-particle equations of motion.

The Hamiltonian determines the equations of motion which specify the accelerations as functions of the positions and velocities. Therefore, Lorentz invariance can be considered a property of the Hamiltonian as well as a property of the equations of motion. We have assumed that the equations

$$\underline{v}^n = \partial H/\partial \underline{p}^n$$

have solutions for the canonical momenta as functions of the positions and velocities. Therefore, given the Hamiltonian, we have only one choice for the generators $\hat{\underline{K}}$ of Lorentz transformations. They are determined by the Lorentz transformations of positions and velocities. This means that $[\hat{p}^n_j, \hat{K}_k]$ are the solution, in terms of the Hamiltonian, of the Lie-bracket relation $[\hat{\underline{K}}, \hat{H}] = \hat{\underline{P}}$ applied to the positions, which gives the infinitesimal Lorentz transformations of velocities. Then Lorentz invariance requires only that the Hamiltonian be a solution of the Lie-bracket relation $[\hat{\underline{K}}, \hat{H}] = \hat{\underline{P}}$ applied to the velocities, which gives the infinitesimal Lorentz transformations of accelerations. Thus we get differential equations to solve for the possible Hamiltonians. These equations were obtained originally from a slightly different point of view.[6]

The Hamiltonian form of the equations of motion, and the canonical form of space translations and rotations, do not depend on the choice of space-time coordinates. Because the equations of motion are invariant, they can be put in Hamiltonian form in the same

way with respect to a transformed reference frame. The same Hamiltonian function H and the same standard functions $\underline{P}$ and $\underline{J}$ can be used for the generators of time translations and space translations and rotations with respect to any reference frame gotten by a transformation in the Poincaré group. If H, $\underline{P}$, $\underline{J}$ are the total energy, momentum and angular momentum, then the same functions of the positions and canonical momenta for the transformed frame are the transformed energy, momentum and angular momentum. Thus the correspondence of generators to energy, momentum and angular momentum is independent of the choice of reference frame.

For example, consider a Lorentz transformation for an infinitisimal velocity ϵ in the z direction. The positions $\underline{q}^n$ and canonical momenta $\underline{p}^n$ are transformed to

$$\underline{q}^{n'} = \underline{q}^n + \epsilon[\underline{q}^n, \hat{K}_3]$$

and

$$\underline{p}^{n'} = \underline{p}^n + \epsilon[\underline{p}^n, \hat{K}_3]$$

and $H(\underline{q}, \underline{p})$ is transformed to $H(\underline{q}', \underline{p}')$. Time translations with respect to the transformed frame are generated by

$$\hat{H}' = \hat{H} + \epsilon[\hat{H}, \hat{K}_3].$$

Then

$$[\underline{q}^{n'}, \hat{H}'] = [\underline{q}^n, \hat{H}] + \epsilon\left[\underline{q}^n, [\hat{H}, \hat{K}_3]\right] + \epsilon\left[[\underline{q}^n, \hat{K}_3], \hat{H}\right]$$

$$= [\underline{q}^n, \hat{H}] + \epsilon\left[[\underline{q}^n, \hat{H}], \hat{K}_3\right]$$

which is the transform of

$$[\underline{q}^n, \hat{H}] = \partial H/\partial \underline{p}^n.$$

Therefore

$$[\underline{q}^{n'}, \hat{H}'] = \partial H(\underline{q}', \underline{p}')/\partial \underline{p}^{n'}$$

and similarly,

$$[\underline{p}^{n'}, \hat{H}'] = -\partial H(\underline{q}', \underline{p}')/\partial \underline{q}^{n'}.$$

Thus time translations with respect to the transformed frame are canonical with respect to the transformed positions and canonical momenta, and their generator corresponds to the same function H of the transformed positions and canonical momenta. Similarly, space translations and rotations with respect to the transformed frame are canonical with respect to the transformed positions and canonical momenta, and their generators correspond to the standard functions

$$\sum_{n=1}^{N} \underline{p}^{n'}$$

and

$$\sum_{n=1}^{N} \underline{q}^{n'} \times \underline{p}^{n'} \ .$$

The same is true when this infinitesimal Lorentz transformation is replaced by any transformation in the Poincaré group.

No-Interaction Theorem for Canonical Transformations

Now we can state exactly why we do not assume that all the transformations of the Poincaré group are canonical.

> Theorem: If the assumptions described in the preceding section are satisfied, for a classical-mechanical system of a finite number of particles, and if all the transformations of the Poincaré group are canonical, then the particle accelerations are all zero.[7),8),4),9)]

Lorentz Transformation of Energy, Momentum and Angular Momentum

One property of canonical Lorentz transformations which might be defended physically is that the way they transform the Hamiltonian function H and the generator functions $\underline{P}$ and $\underline{J}$ for space translations and rotations is the way energy, momentum and angular momentum usually transform. We will see that noncanonical Lorentz transformations need not have this property, and that if we assume it we can prove again that there is no interaction for two particles.

Suppose we think of H and $\underline{P}$ as the total energy and momentum. They are constants of the motion and are invariant under space translations. Under space rotations H is a scalar and $\underline{P}$ is a vector. In what has been assumed so far, H and $\underline{P}$ have no particular properties under Lorentz transformations. This is illustrated by the example of a single free particle with $H = (1/2)\underline{p}^2$ in which

$$[H, \hat{K}_k] = \underline{p} \cdot [\underline{p}, \hat{K}_k] = (\underline{p}^2 - 1)p_k$$

and

$$[P_j, \hat{K}_k] = p_j p_k - \delta_{jk}.$$

Should we assume that H and $\underline{P}$ transform as a four-vector under Lorentz transformations, as is usual for energy and momentum? This means that

$$[H, \hat{K}_j] = -P_j$$

and

$$[P_j, \hat{K}_k] = -\delta_{jk} H$$

for $j,k = 1,2,3$. The Lie-bracket relations

$$[\hat{H}, \hat{K}_j] = -\hat{P}_j$$

and

$$[\hat{P}_j, \hat{K}_k] = -\delta_{jk} \hat{H}$$

imply that the generators $\hat{H}$ and $\hat{\underline{P}}$ transform as a four-vector under Lorentz transformations. Now should we assume the same for the quantities H and $\underline{P}$?

Suppose we think of $\underline{J}$ as the total angular momentum. It is a constant of the motion and a vector under space rotations. Under space translations $\underline{J}$ transforms as is usual for angular momentum; we have

$$[J_j, \hat{P}_k] = \epsilon_{jk\ell} J_\ell.$$

For Lorentz transformations of $\underline{J}$ we can investigate various hypotheses:

(a) $$[J_j, \hat{K}_k] = -[J_k, \hat{K}_j]$$

(b) $$[J_j, \hat{K}_j] = 0$$

(c) $$\left[[J_j, \hat{K}_k], \hat{K}_m\right] = -\epsilon_{jk\ell}\epsilon_{\ell mn} J_n = \delta_{km} J_j - \delta_{jm} J_k.$$

None of these is implied by what has been assumed so far; this is shown by the example of the single free particle with $H = (1/2)\underline{p}^2$ in which

$$[J_j, \hat{K}_k] = \epsilon_{j\ell m}[q_\ell, \hat{K}_k]p_m + \epsilon_{j\ell m}q_\ell[p_m, \hat{K}_k]$$

$$= \epsilon_{j\ell m}q_k p_\ell p_m + \epsilon_{j\ell m}q_\ell p_m p_k - \epsilon_{j\ell m}q_\ell \delta_{mk}$$

$$= J_j p_k + \epsilon_{jk\ell}q_\ell$$

so $[J_j, \hat{K}_k]$ is not antisymmetric in the indices jk, in particular $[J_j, \hat{K}_j]$ is not zero, and furthermore

$$\left[[J_j, \hat{K}_k], \hat{K}_m\right] = J_j(p_k p_m - \delta_{km}) + (J_j p_m + \epsilon_{jm\ell}q_\ell)p_k + \epsilon_{jk\ell}q_m p_\ell.$$

All the hypotheses (a), (b) and (c) are properties of a representation of the Lorentz group by canonical transformations, or by unitary transformations in quantum mechanics, because then these brackets are the same as the Lie brackets. These properties are more or less understandable physically. For example, (b) says that the component of angular momentum in the direction of the Lorentz transformation is not changed. The combination of (a) and (c) implies that $\underline{J}$ transforms as part of an antisymmetric second-rank four-tensor. In particular, (c) is analogous to

$$\left[[H, \hat{K}_j], \hat{K}_k\right] = \delta_{jk}H$$

which says that H transforms as part of a four-vector, or to

$$\left[[P_j, \hat{K}_k], \hat{K}_m\right] = \delta_{jk}P_m.$$

The choice of hypotheses is simplified by the following.

Theorem: Hypothesis (a) is true if and only if (b) and (c) are true.

Proof: From $[\hat{K}_m, \hat{K}_k] = -\epsilon_{mk\ell}\hat{J}_\ell$ it follows that

$$\left[[J_j, \hat{K}_k], \hat{K}_m\right] - \left[[J_j, \hat{K}_m], \hat{K}_k\right] = -\epsilon_{mk\ell}\epsilon_{\ell jn}J_n = \delta_{jk}J_m - \delta_{jm}J_k$$

so if $[[J_j, \hat{K}_k], \hat{K}_m] = \epsilon_{jk\ell}A_{\ell m}$, then

$$\epsilon_{jk\ell}A_{\ell m} - \epsilon_{jm\ell}A_{\ell k} = \delta_{jk}J_m - \delta_{jm}J_k.$$

Multiplying by ϵ_{abj} and summing over j yields

$$\delta_{ak}A_{bm} - \delta_{bk}A_{am} - \delta_{am}A_{bk} + \delta_{bm}A_{ak} = \epsilon_{abk}J_m - \epsilon_{abm}J_k.$$

For $a=k$ and $b=m$ but $a\neq b$, this is $A_{bb}+A_{aa}=0$ which implies that $A_{aa}=0$ for $a=1,2,3$. For $b=m$ but $a\neq k$, $a\neq b$, and $k\neq m$, it is

$$A_{ak} = -\epsilon_{akm}J_m.$$

Therefore

$$\left[[J_j,\hat{K}_k],\hat{K}_m\right] = -\epsilon_{jk\ell}\epsilon_{\ell mn}J_n.$$

Thus (a) implies (c). Evidently (a) implies (b). From (c) it follows that the part of $[J_j,\hat{K}_k]$ which is symmetric in the indices jk is invariant under transformations generated by $\underline{\hat{K}}$. Then it is invariant under rotations because

$$[\hat{K}_j,\hat{K}_k] = -\epsilon_{jk\ell}\hat{J}_\ell.$$

Therefore the symmetric part of $[J_j,\hat{K}_k]$ is proportional to δ_{jk}. This must be zero if (b) is true. Thus (b) and (c) imply (a). This completes the proof.

We will use (a) to characterize the usual Lorentz transformations of angular momentum.

The parity transformation reflects the positions $\underline{q}^n$ and the velocities $\underline{v}^n=\partial H/\partial \underline{p}^n$. One might assume that H is invariant under the parity transformation. This implies that the parity transformation reflects the canonical momenta $\underline{p}^n$. Then H is invariant under reflection of the positions and canonical momenta.

No-Interaction Theorem Based on Transformations of Energy, Momentum and Angular Momentum

That we should not assume all the properties described in the preceding section is suggested by the following.

Theorem: For a system of two particles, there are no interactions, no accelerations, consistent with the assumption that H and $\underline{P}$ Lorentz transform as a four-vector, hypothesis (a) for Lorentz transformations of $\underline{J}$, and the previous assumptions: the Lie-bracket relations for the generators of the Poincaré group; the transformations of particle positions; canonical transformations for time dependence and for space translations and rotations. It is assumed that the Hamiltonian function H is invariant under the parity transformation and

that the canonical momenta are functions of the positions and velocities.[10]

Which assumptions are really needed for this theorem? We doubt that the parity assumption is essential. Relaxing it would probably not allow interactions.

We expect that interactions will be possible if H, $\underline{P}$ and $\underline{J}$ are not required to Lorentz transform as the conventional four-vector and tensor. This guess is based on the following considerations. We have seen how the Lie-bracket relations for the Poincaré group, plus the transformations of positions, correspond to invariance of the equations of motion which specify the accelerations as functions of the positions and velocities. It appears that these invariant equations of motion allow interactions. We expect that they can be put in Hamiltonian form, with canonical momenta which are invariant under space translations and vectors under space rotations. We expect that the equations of motion can be invariant under the parity transformation, and that the canonical momenta can be vectors under parity. This would satisfy all the assumptions of the theorem except that of Lorentz transformations of H, $\underline{P}$ and $\underline{J}$.

For two particles, invariant equations of motion allow no interactions for which either the total momentum $\underline{u}^1 + \underline{u}^2$ is a constant of the motion, with $\underline{u}^n = m_n \underline{v}^n (1 - \underline{v}^{n2})^{-1/2}$, or the total angular momentum $\underline{q}^1 \times \underline{u}^1 + \underline{q}^2 \times \underline{u}^2$ is a constant of the motion. These quantities could be conserved only in the asymptotic limits of collisions. Perhaps the physical motivations for the conventional Lorentz transformations of H, $\underline{P}$ and $\underline{J}$, based on energy, momentum and angular momentum, are meaningful only asymptotically. This might be a reasonable point of view for describing a collision, but it suggests difficulties in describing a bound system of two particles.

Interactions might be possible with H, $\underline{P}$ Lorentz transforming as a four-vector, if only $\underline{J}$ is not required to Lorentz transform as usual. For one-dimensional space, there are interactions with H and $P = p^1 + p^2$ Lorentz transforming as a "two-vector." An example is provided by the Hamiltonian

$$H = \left(p^{1\,2} + m_1^2\right)^{1/2} + \left(p^{2\,2} + m_2^2\right)^{1/2} - b(q^1 - q^2)$$

with b a real number. This yields

$$p^n = u^n = m_n v^n (1 - v^{n2})^{-1/2}$$

for $n = 1, 2$, and

$$dp^1/dt = b$$

$$dp^2/dt = -b$$

which were mentioned previously as an example showing that the conventional total momentum can be a constant of the motion for one-dimensional space. From the Lorentz transformations of q^1, q^2, v^1 and v^2, one can calculate

$$[H,\hat{K}] = -P$$

and

$$[P,\hat{K}] = -H.$$

References

1. D. G. Currie and T. F. Jordan, Phys. Rev. Letters 16, 1210 (1966).
2. L. Pontrjagin, Topological Groups, second edition, translated by A. Brown (Gordon and Breach, New York, 1966), section 52.
3. B. O. Koopman, Proc. Natl. Acad. Sci. U.S. 17, 315 (1931); J. von Neumann, Ann. of Math. 33, 587 (1932); A. Loinger, Ann. of Phys. 20, 132 (1962).
4. J. T. Cannon and T. F. Jordan, J. Math. Phys. 5, 299 (1964).
5. E. L. Hill, Rev. Mod. Phys. 23, 253 (1951).
6. D. G. Currie, Ann. of Phys. 36, 104 (1966).
7. D. G. Currie, T. F. Jordan, and E. C. G. Sudarshan, Rev. Mod. Phys. 35, 350 (1963).
8. D. G. Currie, J. Math. Phys. 4, 1470 (1963).
9. H. Leutwyler, Nuovo Cimento 37, 556 (1965).
10. T. F. Jordan, Phys. Rev., to be published.

IV. Discussion

History

Now we will trace the history of some of the ideas which have been presented in these lectures. Our primary aim is to elucidate the motivation and development of classical particle theories. However, we shall consider also some of the quantum aspects, since they provided part of the motivation, and since some of the ideas were stated originally in a quantum context. Beyond that, most of the classical work has as its aim a better understanding of the analogous quantum theory.

The interaction of two particles can be specified by giving the accelerations as functions of the relative position and relative

velocity in some particular frame. This yields equations of motion of the form

$$\underline{a}^n = \underline{f}^n(\underline{x}^1 - \underline{x}^2, \underline{v}^1 - \underline{v}^2)$$

(for $n = 1, 2$) in that frame. To obtain the equations of motion in a different frame, we can transform the world lines and then find the equations of motion which they satisfy. For Galilei transformations, we find equations of motion of the same form

$$\underline{a}'^n = \underline{f}^n(\underline{x}'^1 - \underline{x}'^2, \underline{v}'^1 - \underline{v}'^2)$$

(for $n = 1, 2$) for the variables of the new frame. The Galilei group of transformations is such that the accelerations remain the same functions of only the relative position and relative velocity. When the Galilei group is replaced by the Lorentz group, the transformation properties of the equations of motion become considerably more complicated, primarily because simultaneity is not Lorentz invariant. The accelerations are not only functions of the relative position and relative velocity, but depend on the individual particle velocities separately. Thus one could easily write down the most general Galilean-invariant equations of motion, but finding Lorentz-invariant equations of motion requires non-trivial computation.

Historically, this apparent difficulty was one of the factors which led to an emphasis on fields to describe the interaction of particles. Other factors were Einstein causality and the question of energy conservation in classical electrodynamics. In these lectures we have considered energy-conserving interactions (as, for example, Wheeler-Feynman time-symmetric classical electrodynamics), so the energy argument, which favors fields for the retarded interaction, is not relevant here. The assumption of manifest Einstein causality requires the dynamic properties of a radiation field; the field at a given time is not determined by the position and velocity of the source at the same time. However, while the use of a radiation field is a convenient and conventional approach, it does not relieve the causality difficulties, even in the retarded formulation. In 1938 Dirac[1] did the mass renormalization needed to obtain the trajectories of particles interacting with an electromagnetic field. His exact solutions for the motion of a charge revealed the difficulties associated with preacceleration. Havas[2] has shown that this failure of manifest Einstein causality is not limited to the scalar electron interacting with the electromagnetic field, as considered by Dirac; it occurs for particles with spin and fields with various transformation properties.

At about the same time that Dirac discussed the preacceleration questions, Wigner[3] showed how irreducible unitary

representations of the Poincaré group can be used to describe elementary systems, or free particles, in quantum mechanics. In 1941 Dirac[4] set up an analogous formalism for classical particles. He emphasized the representation of the Poincaré group, and did not require correct Lorentz transformations for the particle positions. (A similar formalism for fields had been considered earlier by Pryce.[5])

Dirac assumed that canonical representations of the Poincaré group should be used to describe systems of interacting particles. He introduced this assumption by requiring that Poisson-bracket relations remain invariant under the relativity transformations. He took this to mean that the transformations must be canonical, as indeed they must if the Poisson brackets are always computed with the canonical variables of a single frame. We use noncanonical Lorentz transformations, and we have a formalism in which Poisson-bracket relations are invariant. For example, the Hamiltonian and the generator functions for space translations and rotations are invariant. The point is simply that after a Lorentz transformation one must use the transformed canonical variables.

In 1953, Bakamjian and Thomas[6] developed Dirac's formalism for interacting particles by finding a general solution of the Poisson-bracket relations for the generators of the Poincaré group. They describe a huge class of interactions for which the total mass (in a representation similar to that of a free particle) is a function of suitably chosen relative variables. In 1961, Foldy[7] presented an excellent discussion of the ideas which lie behind the use of unitary representations of the Poincaré group in a formalism which says nothing about Lorentz transformations for particle positions. This kind of relativistic quantum theory of particles has been applied to scattering[8)-11)] and to bound states (as in relativistic quark models[12)]).

The spirit of this kind of theory is to abstract the Poincaré group from transformations in space-time, and then describe particles with representations of the group. The description is invariant under the group representation (for example, the equations of motion are invariant) but only the group structure identifies the representation with the physical relativity transformations. We believe that the physical relativity transformations should be identified by their transformations of a complete set of physical quantities. For two or more particles, the group representation is reducible, the generators are not a complete set of quantities, and the group structure is not sufficient. How, then, did it happen that Lorentz transformations of particle positions were generally ignored in the development of these theories?

It appears that Thomas[13)] considered manifest Einstein causality and Lorentz transformations of particle positions at the same time and, seeing the difficulty of this combination, gave up the transformation requirements on the positions. Foldy[7] states explicitly

that there is no manifest Einstein causality in his formalism and that Einstein causality is a question to consider at a later time. Foldy's reason for not considering Lorentz transformations of particle positions is that he works in quantum mechanics where the definition of the particle position is complicated by the uncertainty principle and zitterbewegung.

To make a detailed investigation of this question of the position and its transformation, we consider the analogous classical system. We have seen that with correct Lorentz transformations of particle positions in classical mechanics, there can be no interaction when the Poincaré group is represented by canonical transformations, as in Dirac's formalism. This no-interaction theorem, together with the realization that canonical transformations are not required, led us to the point of view described in these lectures. We begin with Lorentz transformations of particle positions and from them construct the representation of the Poincaré group. Our Hamiltonian formalism is developed in as close analogy to quantum mechanics as is possible, in the hope that it might be quantized in a somewhat conventional manner.

In the past few years several no-interaction theorems have been proved. Fong and Sucher[9] show that conventional Lorentz transformations of particle momenta allow no interaction in the quantum theory. Van Dam and Wigner[14] show that there is no interaction if the conventional total momentum and kinetic energy are conserved and the particles are free asymptotically. We proved a similar theorem in these lectures. We do not assume that the kinetic energy is conserved, and, since we do not use an asymptotic condition, our result holds for bound systems as well as for collisions. Van Dam and Wigner do not use equations of motion, and their proof is for 2, 3 or 4 particles. Still another kind of no-interaction theorem was proved by Ekstein[15] using manifest Einstein causality.

We have considered the instantaneous equations of motion which can be put in Hamiltonian form. There are other ways to describe interaction in relativistic classical particle mechanics. One of the earliest, developed by Fokker[16] and by Feynman and Wheeler,[17] was to eliminate the fields from classical electrodynamics. This has the problem of preacceleration discussed in connection with Einstein causality. Havas and Plebanski[18] have developed another formalism, a special case of which can be described as follows. The acceleration of particle 1 at a given point on its world line depends on the position and velocity of particle 2 at the point on its world line which has the same time as that of particle 1 in the rest frame of particle 1. Therefore, the interactions are not simultaneous. Havas and Plebanski show that the resulting series of iterations is convergent and yields solutions for the world lines. Another formalism is that of Van Dam and Wigner[19] in which the acceleration of

particle 1 at a given point on its world line depends on the positions and velocities of particle 2 at all points on its world line which are space-like with respect to that of particle 1. This can be described by manifestly Lorentz-invariant equations of motion of the form

$$m_1 \, d^2 x_{1\mu}(\tau_1)/d\tau_1^2 = \int d\tau_2 F_\mu\left(x_1(\tau_1), \dot{x}_1(\tau_1), x_2(\tau_2), \dot{x}_2(\tau_2)\right)$$

where F_μ is a function of four-vectors.

Kerner[20] and Hill[21] have proposed a Hamiltonian formalism in which the physical positions, which transform correctly, are different from the canonical coordinates. The advantage is that canonical transformations can be used for the entire Poincaré group. The big disadvantage is that the relation of the canonical variables to the physical variables will be complicated, and this will make the question of quantization that much more difficult.

Questions Still Unanswered

In the following brief review, we will list some of the problems which are not yet solved.

The spirit of these lectures has been to concentrate on the particle positions, because Lorentz transformations are defined most surely for points in space-time. Thus we first study the Newtonian equations of motion. The requirement of Lorentz invariance takes the form of nonlinear differential equations for the acceleration functions. This is a strong requirement in the sense that it eliminates the use of canonical representations of the Poincaré group for describing interactions. On the other hand, we believe that there are many Lorentz-invariant equations of motion. For one-dimensional space, we can have any acceleration in the center-of-mass frame. The proof of a similar statement for three-dimensional space is not yet complete. Beyond that, we need more examples, and physically more interesting examples. Solving the nonlinear differential equations may not be the best way to find them; perhaps another approach is needed.

It would be interesting to study in detail the connection between the instantaneous equations of motion which we use and the equations of Van Dam and Wigner.[19] This might shed some light on the questions of Einstein causality. Eventually, however, the questions of Einstein causality should be studied in a system of three or more particles.

For given invariant equations of motion, there are generally many different Hamiltonians.[22-24] Is there a physical way to choose a unique Hamiltonian? The lack of a decisive answer to this question is one reason we concentrate on particle positions rather than canonical momenta. For the same invariant Newtonian equations

of motion, there are different Hamiltonians, which means that the canonical momenta are different functions of the positions and velocities, satisfy different Hamiltonian equations of motion, and transform differently under Lorentz transformations. The situation is illustrated by the following example.[25]

For a system of two particles, consider the different Hamiltonians

$$H = \left((\underline{p}^1 + \underline{q}\lambda)^2 + m_1^2\right)^{1/2} + \left((\underline{p}^2 - \underline{q}\lambda)^2 + m_2^2\right)^{1/2}$$

for different choices of a function λ of $\underline{q}^2$ (where $\underline{q} = \underline{q}^1 - \underline{q}^2$). For any λ, the accelerations are zero, so the Hamiltonian describes a system of two free particles. For the canonical momenta we get

$$\underline{p}^1 = \underline{v}^1(1 - \underline{v}^{1^2})^{-1/2} - \underline{q}\lambda$$

and

$$\underline{p}^2 = \underline{v}^2(1 - \underline{v}^{2^2})^{-1/2} + \underline{q}\lambda.$$

Evidently the canonical momenta are constants of the motion only if λ is zero, and this is the only case in which the canonical momenta are parallel with the velocities. We can require that the canonical momenta converge asymptotically to the conventional kinematic momenta. This implies that λ decreases faster than $|\underline{q}|^{-1}$ for large $|\underline{q}|$. This leaves an infinite number of different Hamiltonians, all of which are positive. In this case, where there is no interaction, we can choose the usual Hamiltonian with $\lambda = 0$ by requiring that the energy H and the total momentum $\underline{p}^1 + \underline{p}^2$ transform as a four-vector, but this may not be possible when there is interaction.

One can require, as above, that the Hamiltonian converge asymptotically to a sum of free-particle Hamiltonians. This would seem natural for equations of motion which are weakly separable. Fong and Sucher[9] have shown that in their quantum theory this property follows from their choice of relative variables. We do not know yet whether a similar statement can be made in our Hamiltonian formalism.

If we now turn to quantum mechanics, we must first define the role of position in the quantum theory. An intuitive but incomplete approach is to require that the theory have a sensible classical limit. To illustrate this, consider at an initial time $t = 0$, a narrow gaussian wave packet whose width is proportional to $\sqrt{\hbar}$. As time unfolds, this state develops a "world packet," a small region of space about a world line. This world line is roughly defined as the mean position

of the particle as a function of time. More exactly, it is defined by letting $\hbar \to 0$, in which case this world packet shrinks to a world line with a well-defined momentum. The requirement on the quantum theory is that this limiting world line transform properly.

With this definition of the transformation of the position in a theory with unitary representations of the Poincaré group, it is possible[26] to prove that there is no interaction except generalized contact interactions. In other words, given a classical system which interacts as we have discussed (for example, an arbitrary potential $V(x)$ in the center-of-mass frame) there exists no quantum theory of the above type which limits to this classical theory as $\hbar$ goes to zero.

References

1. P. A. M. Dirac, Proc. Roy. Soc. (London) A167, 148 (1938).
2. P. Havas, Phys. Rev. 87, 305 (1952).
3. E. P. Wigner, Ann. Math. 40, 149 (1939).
4. P. A. M. Dirac, Rev. Mod. Phys. 21, 392 (1949).
5. M. H. L. Pryce, Proc. Roy. Soc. (London) A195, 62 (1948).
6. B. Bakamjian and L. H. Thomas, Phys. Rev. 92, 1300 (1953).
7. L. L. Foldy, Phys. Rev. 122, 275 (1961).
8. T. F. Jordan, A. J. Macfarlane, and E. C. G. Sudarshan, Phys. Rev. 133, B487 (1964).
9. R. Fong and J. Sucher, J. Math. Phys. 5, 456 (1964).
10. T. F. Jordan, J. Math. Phys. 5, 1345 (1964).
11. F. Coester, Helv. Phys. Acta 38, 7 (1965).
12. N. D. Son and J. Sucher, Phys. Rev. 153, 1496 (1967).
13. L. H. Thomas, Phys. Rev. 85, 868 (1952).
14. H. Van Dam and E. P. Wigner, Phys. Rev. 142, 838 (1966).
15. H. Ekstein, Comm. Math. Phys. 1, 6 (1965).
16. A. D. Fokker, Z. f. Physik 58, 386 (1929).
17. J. A. Wheeler and R. P. Feynman, Rev. Mod. Phys. 17, 157 (1945) and 21, 425 (1949).
18. P. Havas and J. Plebanski, Bull. Am. Phys. Soc. 5, 433 (1961); P. Havas, Proceedings of the 1964 International Congress of Logic, Methodology, and Philosophy of Science, edited by Y. Bar-hillel (North-Holland Publishing Company, Amsterdam, 1965).
19. H. Van Dam and E. P. Wigner, Phys. Rev. 138, B1576 (1965).
20. E. H. Kerner, J. Math. Phys. 6, 1218 (1965).
21. R. N. Hill, "Canonical Formulation of Relativistic Mechanics" (preprint).
22. D. G. Currie and E. J. Saletan, J. Math. Phys. 7, 967 (1966).
23. F. J. Kennedy, Jr., and E. H. Kerner, Am. J. Phys. 33, 463 (1965).

24. D. G. Currie and E. J. Saletan, Bull. Am. Phys. Soc. 11, 373 (1966).
25. D. G. Currie, Bull. Am. Phys. Soc. 11, 356 (1966).
26. D. G. Currie, to be published.

THE THIRRING MODEL†

Bernhard Klaiber
Palmer Physical Laboratory
Princeton University
Princeton, New Jersey

I. Introduction

The Thirring model describes a self-interaction of a massless spinor field in one-space dimension described by the Lagrangian

$$L_{int} = g j^{\mu} j_{\mu}.$$

The history of this model has been a dramatic one and revealed the Thirring model as being a rather tricky object. A comprehensive account of this history is given in Reference 1. We give here only a short sketch of this history. Thirring[2] showed that the model is exactly soluble and constructed the eigenstates of the Hamiltonian. Glaser[3] then solved the field equations. Both authors dealt with formal manipulations, and it was shown subsequently that these manipulations lead to contradictions. Johnson[4] then started from scratch, paid careful attention to the definition of the products of fields which occur in the definition of the currents and in the field equation and solved the system of coupled equations for the time-ordered functions. At that point one then asks for the unordered functions or even an operator solution. The latter was the aim of the work by Scarf and Wess.[5] This solution was again a formal one inasfar as nonexistent line integrals of the currents were used. The algebraic structure of this solution contained a lot of truth, and it was also possible to obtain the correct n-point functions (corresponding to Johnson's solution) in this framework.[7] What was too difficult to check was the positive definiteness condition.

The purpose of these lectures is to give an operator solution in a well-defined Hilbert space. We avoid the introduction of an indefinite metric space for the description of the massless scalar field. The positive definiteness condition is then automatically fulfilled. We

---

†Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

will obtain a two-parameter family of solutions in which the solutions of Johnson and Schwinger are contained as a subfamily. We will extensively discuss the properties of the solutions.

II. The Massless Dirac Field in Two Dimensions

In two-dimensional space-time, the zero mass Dirac field has some peculiar features which are the reason for the solubility of the Thirring model. Let us therefore begin with a discussion of these features.

The gamma matrices are realized by 2×2 matrices:

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \gamma^1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \gamma^5 = \gamma^0\gamma^1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{II.1}$$

We will also use the relations

$$\gamma^\mu\gamma^5 = e^{\mu\nu}\gamma_\nu \qquad \epsilon_{\mu\nu} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \epsilon_{\mu\nu}\epsilon^{\nu\rho} = g_\mu^\rho \tag{II.2}$$

One considers a two-component field whose equation of motion and anticommutation relation are

$$\gamma^\mu\partial_\mu\psi(x) = 0 \qquad \{\psi(x), \bar{\psi}(y)\} = \tfrac{1}{i}S(x-y) \tag{II.3}$$

Fourier representation:

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int dp^1\left\{a^*(p^1)e^{ipx} + b(p^1)e^{-ipx}\right\}u(p^1) \qquad p^0 = |p^1|$$

$$\{a(p^1),\ a^*(q^1)\} = \{b(p^1),\ b^*(q^1)\} = \delta(p^1-q^1). \tag{II.4}$$

$u(p^1)$ is the two-component fundamental spinor, satisfies

$$(\gamma p)u(p^1) = 0 \qquad u(p^1)\bar{u}(p^1) = \frac{1}{2p^0}\gamma p \tag{II.5}$$

and can be represented by

$$u(p^1) = \begin{pmatrix} \theta(-p^1) \\ \theta(p^1) \end{pmatrix}. \tag{II.6}$$

The interesting objects are the current $j^\mu(x) \equiv :\bar{\psi}(x)\gamma^\mu\psi(x):$ and the pseudocurrent $:\bar{\psi}(x)\gamma^\mu\gamma^5\psi(x): = \epsilon^{\mu\nu}j_\nu(x)$. Both are conserved in the zero mass case:

$$\partial^\mu j_\mu(x) = \partial^\mu \epsilon_{\mu\nu} j^\nu(x) = 0, \tag{II.7}$$

and this implies that $j^\mu(x)$ is a free field

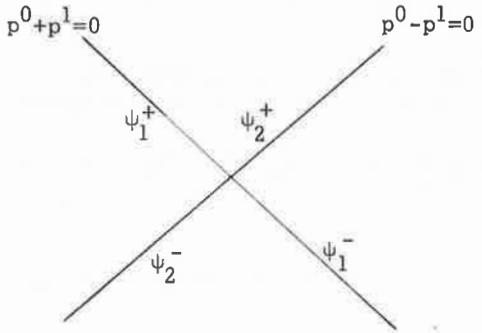
$$\Box j^\mu(x) = 0. \tag{II.8}$$

The situation becomes clear if one looks at the support properties of $\psi(x)$ in momentum space. First note that

$$j^0(x) = :\psi_1^*(x)\psi_1(x): + :\psi_2^*(x)\psi_2(x):$$

$$j^1(x) = -:\psi_1^*(x)\psi_1(x): + :\psi_2^*(x)\psi_2(x): . \tag{II.9}$$

If we decompose $\psi(x)$ into positive and negative frequency parts, the supports in momentum space can be characterized by the following diagram:



Now the convolution of ψ_1^* and ψ_1, for example, yields again something with the support $p^0 + p^1 = 0$.

But it turns out that $j^\mu(x)$ is even a canonical free field, i.e., its commutator is a c-number. Let us see, in x-space and computing formally, how this can happen. A straightforward calculation gives

$$[j^\mu(x), j^\nu(y)] = \frac{1}{i}:\bar\psi(x)\gamma^\mu S(x-y)\gamma^\nu\psi(y): - \frac{1}{i}:\bar\psi(y)\gamma^\nu S(y-x)\gamma^\mu\psi(x):$$

$$- \mathrm{Sp}\{\gamma^\mu S^-(x-y)\gamma^\nu S^+(y-x)\} + \mathrm{Sp}\{\gamma^\nu S^-(y-x)\gamma^\mu S^+(x-y)\}. \tag{II.10}$$

To show that the two operator parts cancel, one has to show that the first one is symmetric in x and y. (Symmetry in μ and ν is obvious

since $\gamma^\mu\gamma^\rho\gamma^\nu = \gamma^\nu\gamma^\rho\gamma^\mu$ in two dimensions.) To see this, use the explicit form of the S-function:

$$S(\xi) = i\begin{pmatrix} 0 & \delta(\xi^0+\xi^1) \\ \delta(\xi^0-\xi^1) & 0 \end{pmatrix} \tag{II.11}$$

and the fact that ψ_1, ψ_2 do not depend on x^0, x^1 separately but rather

$$\psi_1 = \psi_1(x^0+x^1) \qquad \psi_2 = \psi_2(x^0-x^1). \tag{II.12}$$

This follows from the equation of motion which, in components, reads

$$(\partial_0 - \partial_1)\psi_1 = 0 \qquad (\partial_0 + \partial_1)\psi_2 = 0. \tag{II.13}$$

So we have to consider terms of the type

$$:\bar{\psi}_2(x^0+x^1)\,\delta(x^0+x^1-y^0-y^1)\,\psi_1(y^0+y^1):,$$

and this is symmetric.

For a solid consideration, one best goes to momentum space. One then finds that the current has the representation

$$j^\mu(x) = -\frac{i}{\sqrt{2\pi}}\int \frac{dk^1}{\sqrt{2k^0}}\,k^\mu\left\{c(k^1)e^{-ikx} - c^*(k^1)e^{ikx}\right\} \quad k^0 = |k^1| \tag{II.14}$$

$$c(k^1) = \frac{i}{\sqrt{k^0}}\int dp^1\Big\{\theta(k^1p^1)\Big[b^*(p^1)b(p^1+k^1) - a^*(p^1)a(p^1+k^1)\Big]$$

$$+ \theta(p^1(k^1-p^1))a(k^1-p^1)b(p^1)\Big\}$$

$$[c(k^1), c^*(p^1)] = \delta(k^1-p^1) \qquad c(k^1)\Omega = 0 \tag{II.15}$$

if Ω is the original vacuum of the a and b.

Equation (II.14) anticipates yet another characteristic of the current: The quantity $e^{\mu\nu}\partial_\mu j_\nu$ is the two-dimensional analog of the curl of j_μ. Its vanishing implies that j_μ is a gradient. And (II.14) invites to the conclusion that it is even the gradient of a canonical free scalar field of mass zero. But here the trouble begins. In two dimensions there is no such object. The two-point function would be

$$\text{const.} \int \frac{dk^1}{2k^0}\,e^{-ikx},$$

thus divergent at $k=0$.

If one wants to solve the Thirring model, one has to overcome this problem. But before we have a look at this, let us list the commutators and singular functions of the zero mass spinor field:

$$(\Omega, \psi(x)\bar{\psi}(y)\Omega) = \frac{1}{i} S^-(x-y)$$

$$S^-(\xi) = \frac{i}{2\pi} \int dp^1 e^{-ip\xi} \begin{pmatrix} 0 & \theta(-p^1) \\ \theta(p^1) & 0 \end{pmatrix}$$

$$= \frac{1}{2\pi} \begin{pmatrix} 0 & 1/(\xi^0+\xi^1-i\epsilon) \\ 1/(\xi^0-\xi^1-i\epsilon) & 0 \end{pmatrix}$$

$$S(\xi) = S^-(\xi) + S^-(-\xi) = i \begin{pmatrix} 0 & \delta(\xi^0+\xi^1) \\ \delta(\xi^0-\xi^1) & 0 \end{pmatrix} \tag{II.16}$$

$$\{\psi(x), \bar{\psi}(y)\} = \frac{1}{i} S(x-y)$$

$$[j^\mu(x), \psi(y)] = -(g^{\mu\nu} + e^{\mu\nu}\gamma^5) D_\nu(x-y)\psi(y)$$

$$[j^\mu(x), j^\nu(y)] = \frac{i}{\pi} D^{\mu\nu}(x-y)$$

$$D_\nu(\xi) = \frac{1}{2\pi} \int \frac{dk^1}{2k^0} k_\nu (e^{-ikx} + e^{ikx}) = \tfrac{1}{2} \partial_\nu [\theta(\xi^0)\theta(\xi^2)]$$

$$D_{\mu\nu}(\xi) = \partial_\mu D_\nu(\xi)$$

For equal times, $x^0 = y^0$:

$$\begin{aligned} [j^0(x), \psi(y)] &= -\delta(x^1-y^1)\psi(y) \\ [j^1(x), \psi(y)] &= -\gamma^5 \delta(x^1-y^1)\psi(y). \end{aligned} \tag{II.17}$$

This means that $j^0(x)$ (resp. $j^1(x)$) is the generator of space-time dependent gauge transformations (resp. γ^5-transformations):

$$e^{iQ(\Lambda;y^0)}\psi(y)e^{-iQ(\Lambda;y^0)} = e^{i\Lambda(y)}\psi(y) \qquad Q(\Lambda;y^0) \equiv \int_{x^0=y^0} dx^1 \Lambda(x) j^0(x)$$

$$e^{i\tilde{Q}(\Lambda;y^0)}\psi(y)e^{-i\tilde{Q}(\Lambda;y^0)} = e^{i\gamma^5\Lambda(y)}\psi(y) \quad \tilde{Q}(\Lambda;y^0) \equiv \int_{x^0=y^0} dx^1 \Lambda(x) j^1(x).$$

For the interacting current $J^\mu(x)$ of the model we will also postulate relations of the form (II.17) (with general coefficients occurring), in addition to the conservation equations (II.7). Then $J^\mu(x)$ will again be a free field, and a relation of the type (II.17) defines a unique Cauchy problem (remember that $\partial^0 j^0 = \partial^1 j^1$, $\partial^0 j^1 = \partial^1 j^0$). Thus the commutator between the interacting field and current will essentially again have the simple structure (II.16).

III. A Substitute for the Massless Scalar Field in Two Dimensions

There is no massless scalar field in two dimensions. This boils down to the fact that in two dimensions there is no distribution (except, of course, $\delta(p)$) which has support given by $p^2 = 0$ $p^0 \geqq 0$ and which is positive definite and Lorentz invariant. The candidate would be, of course, $\theta(p^0)\delta(p^2)$. Written in the variables

$$u = p^0 + p^1 \qquad v = p^0 - p^1, \tag{III.1}$$

it is given by

$$\theta(u)\frac{1}{u}\delta(v) + \theta(v)\frac{1}{v}\delta(u),$$

and

$$\theta(u)\frac{1}{u},$$

for instance, is not a distribution. One way out of this difficulty is to redefine the two-point function as

$$\left(\frac{1}{u}\right)_+ \delta(v) + \left(\frac{1}{v}\right)_+ \delta(u)$$

where $(1/u)_+$ is a regularization of the function $\theta(u)(1/u)$. A regularization of a (nonintegrable) function is defined as a distribution which coincides with this function everywhere except in the neighbourhood of the singularities. In our case,

$$\left(\left(\frac{1}{u}\right)_+, f\right) = \int_0^\infty du \frac{1}{u} f(u) \tag{III.2}$$

for all test functions for which $f(u) = 0$ in a neighbourhood of the origin, $(T.f)$ denotes the application of the distribution T to the test function f. Such regularizations are furnished by

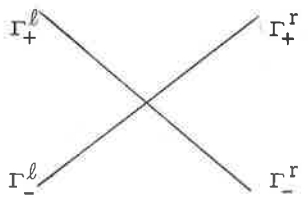
$$\left(\left(\frac{1}{u}\right)_+, f\right) = \int_0^\infty du\, \frac{1}{u}\Big(f(u) - \theta(2\kappa - u)f(0)\Big) \qquad \text{(III.3)}$$

with κ an arbitrary positive parameter (here with the dimension of a mass), and the factor 2 stands purely for later convenience.

This provides us with a 2-point function which is invariant but not positive definite. However, it has been shown[1] that a (indefinite) space of states and, in it, a field having this 2-point function, can actually be constructed.

This procedure has some drawbacks. If one constructs the solution of a model in such a space, one then has to show that the resulting n-point functions are positive definite, or, in other words, that the subspace generated by the resulting field operator is positive definite. For Schroer's derivative coupling[6] this has been done,[1] but for the Thirring model this would probably be more difficult. Then, there is another problem. We are not dealing with an independent scalar field (as one does in Schroer's model), but with the integrated current of a spinor field. A third field in the game is the integrated pseudocurrent. All three would be Lorentz invariant, but it turns out that they would transform with different unitary representations of the Lorentz group. The trouble comes in when one splits the integrated currents into positive and negative frequency parts. But to calculate the n-point functions one is seemingly forced to make this splitting. All this shows that there is no advantage in working with Lorentz invariant auxiliary fields. What is important is to ensure invariance of the final solution.

We will not use the indefinite metric, but the regularizations which characterize it will, at the end, come in in our solution as well (through the back door). Let us therefore discuss them in detail. The light cone consists of four invariant branches $\Gamma_\pm^r\, \Gamma_\pm^\ell$:



On these we can consider the following distributions:

$$\Gamma_+^r : G_+^r(u,v) = \left(\frac{1}{u}\right)_+ \delta(v) \quad (G_+^r, f) = \int_0^\infty du\, \frac{1}{u}\Big(f(u,o) - \theta(2\kappa-u)f(o,o)\Big)$$

$$\Gamma_+^\ell : G_+^\ell(u,v) = \left(\frac{1}{v}\right)_+ \delta(u) \quad (G_+^\ell, f) = \int_0^\infty dv\, \frac{1}{v}\Big(f(o,v) - \theta(2\kappa-v)f(o,o)\Big)$$

$$\Gamma_-^r : G_-^r(u,v) = \left(\frac{1}{v}\right)_- \delta(u) \quad (G_-^r, f) = \int_{-\infty}^0 dv\, \frac{1}{v}\Big(f(o,v) - \theta(2\kappa+v)f(o,o)\Big)$$

$$\Gamma_-^\ell : G_-^\ell(u,v) = \left(\frac{1}{u}\right)_- \delta(v) \quad (G_-^\ell, f) = \int_{-\infty}^0 du\, \frac{1}{u}\Big(f(u,o) - \theta(2\kappa+u)f(o,o)\Big). \tag{III.4}$$

Here we have also introduced the distribution

$$\left(\left(\frac{1}{u}\right)_-, f\right) = \int_{-\infty}^0 du\, \frac{1}{u}\Big(f(o) - \theta(2\kappa+u)f(o)\Big). \tag{III.5}$$

Let us first study the behaviour under Lorentz transformations:

$$x \to \Lambda x \qquad \Lambda^\mu{}_\nu = \begin{pmatrix} \cos\chi & \sin\chi \\ \sin\chi & \cos\chi \end{pmatrix} \qquad \Lambda u = e^{\chi} u \qquad \Lambda v = e^{-\chi} v. \tag{III.6}$$

By definition,

$$(G(\Lambda u, \Lambda v), f(u,v)) = (G(u,v), f(\Lambda^{-1}u, \Lambda^{-1}v)).$$

Thus, for example,

$$\begin{aligned}
(G_+^r(\Lambda u, \Lambda v), f(u,v)) &= \int_0^\infty du\, \frac{1}{u}\Big(f(e^{-\chi}u, o) - f(o,o)\theta(2\kappa-u)\Big) \\
&= \int_0^\infty du\, \frac{1}{u}\Big(f(u,o) - f(o,o)\theta(2\kappa e^{-\chi}-u)\Big) \\
&= \int_0^\infty du\, \frac{1}{u}\Big(f(u,o) - f(o,o)\theta(2\kappa-u)\Big) \\
&\quad + \int_0^\infty du\, \frac{1}{u}\Big(\theta(2\kappa-u) - \theta(2\kappa e^{-\chi}-u)\Big) f(o,o) \\
&= (G_+^r(u,v), f(u,v)) + \chi f(o,o))
\end{aligned}$$

and similarly with the other branches. We get the following table:

$$\begin{aligned} G_+^r(\Lambda u,\Lambda v) &= G_+^r(u,v) + \chi\delta(u)\delta(v) \\ G_+^\ell(\Lambda u,\Lambda v) &= G_+^\ell(u,v) - \chi\delta(u)\delta(v) \\ G_-^r(\Lambda u,\Lambda v) &= G_-^r(u,v) + \chi\delta(u)\delta(v) \\ G_-^\ell(\Lambda u,\Lambda v) &= G_-^\ell(u,v) - \chi\delta(u)\delta(v). \end{aligned} \tag{III. 7}$$

The most general invariant combination is

$$\alpha G_+^r + \beta G_+^\ell + \gamma G_-^r + \delta G_-^\ell \quad \text{with} \quad \alpha - \beta + \gamma - \delta = 0 \tag{III. 8}$$

Fourier transform:

$$\begin{aligned} \int dp e^{-ip\xi} G_+^r &= \frac{1}{2}\int du dv\, e^{-\frac{i}{2}u(\xi^0-\xi^1)-\frac{i}{2}v(\xi^0+\xi^1)} G_+^r(u,v) \\ &= \frac{1}{2}\int du \left(\frac{1}{u}\right)_+ e^{-\frac{i}{2}u(\xi^0-\xi^1)} \\ &= \frac{1}{2}\left(-\log|\kappa(\xi^0-\xi^1)| - \frac{i\pi}{2}\epsilon(\xi^0-\xi^1) + \Gamma'(1)\right). \end{aligned}$$

Introduce now a new mass $\mu = e^{-\Gamma'(1)}\kappa$. Then

$$\begin{aligned} \frac{i}{2\pi}\int dp\, e^{-ip\xi} G_+^r &= \frac{1}{4\pi i}\left(\log|\mu(\xi^0-\xi^1)| + \frac{i\pi}{2}\epsilon(\xi^0-\xi^1)\right) \\ &= \frac{1}{4\pi i}\left\{\log[\mu(\xi^0-\xi^1-i\epsilon)] + \frac{i\pi}{2}\right\} \\ \frac{i}{2\pi}\int dp\, e^{-ip\xi} G_-^r &= \frac{1}{4\pi i}\left(\log|\mu(\xi^0+\xi^1)| + \frac{i\pi}{2}\epsilon(\xi^0+\xi^1)\right) \\ &= \frac{1}{4\pi i}\left\{\log[\mu(\xi^0+\xi^1-i\epsilon)] + \frac{i\pi}{2}\right\}. \end{aligned} \tag{III. 9}$$

Both the sum and the difference of these functions will play a role. We have

$$D^-(\xi) \equiv \frac{i}{2\pi}\int dp\, e^{-ip\xi}(G_+^r + G_+^\ell) = \frac{1}{4\pi i}\left\{\log|\mu^2\xi^2| + i\pi\epsilon(\xi^0)\theta(\xi^2)\right\}$$

$$= \frac{1}{4\pi i}\log(-\mu^2\xi^2 + i\epsilon\xi^0)$$

$$\tilde{D}^-(\xi) \equiv \frac{i}{2\pi}\int dp\, e^{-ip\xi}(G_+^r - G_+^\ell) = \frac{1}{4\pi i}\log\frac{\xi^0 - \xi^1 - i\epsilon}{\xi^0 + \xi^1 - i\epsilon} \qquad \text{(III. 10)}$$

$$= \frac{1}{4\pi i}\log\left|\frac{\xi^0 - \xi^1}{\xi^0 + \xi^1}\right| - \frac{1}{4}\epsilon(\xi^1)\theta(-\xi^2).$$

Transformation properties:

$$D^-(\Lambda\xi) = D^-(\xi) \qquad \tilde{D}^-(\Lambda\xi) = \tilde{D}^-(\xi) + \frac{i}{2\pi}\chi \qquad \text{(III. 11)}$$

We get alternative representations for D^- and $\tilde{D}^-$ if we make the substitution $u = 2p^1$, $v = -2p^1$ in the Fourier integrals:

$$D^-(\xi) = \frac{i}{2\pi}\int \frac{dp^1}{2p^0}\left(e^{-ip\xi} - \theta(\kappa - p^0)\right) \qquad \tilde{D}^-(\xi) = \frac{i}{2\pi}\int \frac{dp^1}{2p^1}\left(e^{-ip\xi} - \theta(\kappa - p^0)\right). \qquad \text{(III. 12)}$$

Next, define

$$D(\xi) = \frac{i}{2\pi}\int dp\, e^{-ip\xi}\left(G_+^r + G_+^\ell + G_-^r + G_-^\ell\right) \qquad \text{(III. 13)}$$

and $\tilde{D}(\xi)$ as resulting from $D(\xi)$ by the operation $p^0 \leftrightarrow p^1$ or $x^0 \leftrightarrow -x^1$:

$$\tilde{D}(\xi) = \frac{i}{2\pi}\int dp\, e^{-ip\xi}\left(G_+^r - G_+^\ell - G_-^r + G_-^\ell\right) = D(-\tilde{\xi}) \qquad \tilde{\xi}^\mu = e^{\mu\nu}\xi_\nu. \qquad \text{(III. 14)}$$

$\tilde{D}^-(\xi)$ is then the negative frequency of $\tilde{D}(\xi)$. The positive frequency

$$\tilde{D}^+(\xi) = -\tilde{D}^-(-\xi). \qquad \text{(III. 15)}$$

It should be noted that we are using a somewhat sloppy terminology; D^-, $\tilde{D}^-$, et cetera, contain zero frequencies. The decomposition is not unique, and we have adopted the simplest possible definition of $G_\pm^{r,\ell}$ by choosing all four parameters κ to be the same. Finally,

$$D(\xi) = D^{-}(\xi) - D^{-}(-\xi) = \tfrac{1}{2}\epsilon(\xi^0)\theta(\xi^2)$$
$$\tilde{D}(\xi) = -\tfrac{1}{2}\epsilon(\xi^1)\theta(-\xi^2). \qquad \text{(III. 16)}$$

An alternative to the method of indefinite metric consists of regularizing not the 2-point function but the field: Define the zero mass scalar field by

$$\varphi^{-}(x) = \frac{1}{\sqrt{2\pi}}\int \frac{dk^1}{\sqrt{2k^0}}\, c(k^1)\left(e^{-ikx} - \theta(\kappa - k^0)\right)$$
$$\varphi^{+}(x) = \frac{1}{\sqrt{2\pi}}\int \frac{dk^1}{\sqrt{2k^0}}\, c^*(k^1)\left(e^{ikx} - \theta(\kappa - k^0)\right)$$
$$[c(k^1), c^*(p^1)] = \delta(k^1 - p^1). \qquad \text{(III. 17)}$$

Then

$$\left(\Omega, \varphi(x)\varphi(y)\Omega\right) = \frac{1}{i}D^{-}_{\kappa\kappa}(x,y)$$

$$D^{-}_{\kappa\kappa}(x,y) = \frac{i}{2\pi}\int \frac{dk^1}{2k^0}\left(e^{-ikx} - \theta(\kappa - k^0)\right)\left(e^{iky} - \theta(\kappa - k^0)\right) \qquad \text{(III. 18)}$$

$$[\varphi^{\mp}(x), \varphi^{\pm}(y)] = \frac{1}{i}D^{\mp}_{\kappa\kappa}(x,y) \qquad D^{+}_{\kappa\kappa}(x,y) = -D^{-}_{\kappa\kappa}(y,x) \qquad \text{(III. 19)}$$

$D^{-}_{\kappa\kappa}(x,y)$ is a well-defined distribution, but not translation invariant. To obtain a decomposition we will use later, observe that

$$(e^{-ikx} - \theta)(e^{iky} - \theta) = (e^{-ik(x-y)} - \theta) - \theta(e^{-ikx} - 1) - \theta(e^{iky} - 1). \qquad \text{(III. 20)}$$

Therefore,

$$D^{\mp}_{\kappa\kappa}(x,y) = D^{\mp}(x-y) - \Delta^{\mp}(x) + \Delta^{\pm}(y)$$

if

$$\Delta^{\mp}(x) = \pm\frac{i}{2\pi}\int \frac{dk^1}{2k^0}\,\theta(\kappa - k^0)(e^{\mp ikx} - 1) \qquad \text{(III. 21)}$$

and $D^{\mp}(\xi)$ is the regularized function defined above. This is the place where the functions of the indefinite metric again come in. As

to the unwanted Δ functions, we will have to find means to get rid of them.

IV. The Thirring Model

We consider first the classical version of the model. The equation of motion and conservation equations are:

$$\begin{aligned} i\gamma^\mu \partial_\mu \phi(x) &= -gJ^\mu(x)\gamma_\mu \phi(x) \\ J^\mu(x) &= \bar{\phi}(x)\gamma^\mu \phi(x) \\ \partial^\mu J_\mu(x) &= \partial^\mu \epsilon_{\mu\nu} J^\nu(x) = 0. \end{aligned} \tag{IV.1}$$

The current can again be written as a gradient:

$$J_\mu(x) = \frac{1}{\sqrt{\pi}} \partial_\mu J(x). \tag{IV.2}$$

Consider the expression

$$\phi_o(x) = e^{-i(g/\sqrt{\pi})J(x)} \phi(x)$$

and its current

$$j^\mu(x) = \phi_o(x)\gamma^\mu \phi_o(x).$$

They satisfy the relations

$$\begin{aligned} \gamma^\mu \partial_\mu \phi_o(x) &= 0 \\ j^\mu(x) &= J^\mu(x). \end{aligned} \tag{IV.3}$$

The most general solution of (IV.1) can therefore be written in the form

$$\phi(x) = e^{i(g/\sqrt{\pi})j(x)} \phi_o(x) \tag{IV.4}$$

where $\phi_o(x)$ obeys the free field equation and $j(x)$ is its integrated current.

From any solution $\psi(x)$ of the free field equation and its current $j(x)$, one can construct new ones by forming

$$\phi_o(x) = e^{ic\{j(x)+\gamma^5 \tilde{j}(x)\}} \psi(x), \tag{IV.5}$$

where we have introduced the integrated pseudo-current $\tilde{j}(x)$:

$$e^{\mu\nu}j_\nu(x) = \frac{1}{\sqrt{\pi}}\partial^\mu\tilde{j}(x). \tag{IV.6}$$

(The integrability of the pseudocurrent follows from the fact that the divergence of the current is the rotation of the pseudocurrent.) Note that $j(x)$ is the current of $\phi_0(x)$ as well as of $\psi(x)$. We find in that way that, starting from a free field $\psi(x)$, we get, for every g, a one-parameter family of solutions:

$$\phi(x) = e^{i\{\alpha j(x)+\beta\gamma^5\tilde{j}(x)\}}\psi(x) \tag{IV.7}$$

$$\alpha - \beta = \frac{g}{\sqrt{\pi}} \qquad \alpha,\beta \text{ real.} \tag{IV.8}$$

In quantizing the model, we assume that $\psi(x)$ is a canonical free field. This gives us a more general solution than if we chose $\phi_0(x)$ to be canonical. Now we define

$$\begin{aligned} j^-(x) &= \frac{1}{\sqrt{2\pi}}\int\frac{dk^1}{\sqrt{2k^0}}\,c(k^1)\left(e^{-ikx} - \theta(\kappa-k^0)\right) \\ j^+(x) &= \frac{1}{\sqrt{2\pi}}\int\frac{dk^1}{\sqrt{2k^0}}\,c^*(k^1)\left(e^{+ikx} - \phi(\kappa-k^0)\right) \end{aligned} \tag{IV.9}$$

where $c(k^1)$ is the variable occurring in (I.14). Then, clearly,

$$j_\mu(x) = \frac{1}{\sqrt{\pi}}\partial_\mu j(x). \tag{IV.10}$$

Observing that $e(k^1)k^\mu = e^{\mu\nu}k_\nu$ for $k^0 > 0$, we define

$$\begin{aligned} \tilde{j}^-(x) &= \frac{1}{\sqrt{2\pi}}\int\frac{dk^1}{\sqrt{2k^0}}\,e(k^1)c(k^1)\left(e^{-ikx} - \theta(\kappa-k^0)\right) \\ \tilde{j}^+(x) &= \frac{1}{\sqrt{2\pi}}\int\frac{dk^1}{\sqrt{2k^0}}\,e(k^1)c^*(k^1)\left(e^{ikx} - \theta(\kappa-k^0)\right) \end{aligned} \tag{IV.11}$$

$$e^{\mu\nu}j_\nu(x) = \frac{1}{\sqrt{\pi}}\partial^\mu\tilde{j}(x).$$

Another step we have to take to insure the existence of (IV.7) is to split the currents into positive and negative frequency parts and to rearrange the expression in the following way:

$$\phi(x) = e^{i\chi^+(x)}\psi(x)e^{i\chi^-(x)} \qquad \bar{\phi}(x) = e^{-i\bar{\chi}^+(x)}\bar{\psi}(x)e^{-i\bar{\chi}^-(x)}$$

$$\chi^{\pm}(x) = \alpha j^{\pm}(x) + \beta\gamma^5\tilde{j}^{\pm}(x) \qquad \bar{\chi}^{\pm}(x) = \alpha j^{\pm}(x) - \beta\gamma^5\tilde{j}^{\pm}(x). \tag{IV.12}$$

This object is an operator-valued distribution. To compute its n-point functions one has to take into account the commutation relations

$$\begin{aligned}
[j^-(x), j^+(y)] &= [\tilde{j}^-(x), \tilde{j}^+(y)] = \frac{1}{i}D^-_{\kappa\kappa}(x,y) \\
[j^-(x), \tilde{j}^+(y)] &= \frac{1}{i}\tilde{D}^-_{\kappa\kappa}(x,y) \\
[j^{\pm}(x), \psi(y)] &= -\sqrt{\pi}\left\{D^{\pm}_{\kappa}(x,y) + \gamma^5\tilde{D}^{\pm}_{\kappa}(x,y)\right\}\psi(y) \\
[\tilde{j}^{\pm}(x), \psi(y)] &= -\sqrt{\pi}\left\{\tilde{D}^{\pm}_{\kappa}(x,y) + \gamma^5 D^{\pm}_{\kappa}(x,y)\right\}\psi(y).
\end{aligned} \tag{IV.13}$$

The $\tilde{D}_{\kappa\kappa}$ function is obtained from $D^{\pm}_{\kappa\kappa}$ by writing a factor $\epsilon(k^1)$ in the Fourier representation (III.21):

$$\tilde{D}^-_{\kappa\kappa}(x,y) = \frac{i}{2\pi}\int\frac{dk^1}{2k^0}\,\epsilon(k^1)\left(e^{-ikx} - \theta(\kappa-k^0)\right)\left(e^{iky} - \theta(\kappa-k^0)\right)$$

$$\tilde{D}^+_{\kappa\kappa}(x,y) = -\tilde{D}^-_{\kappa\kappa}(y,x). \tag{IV.14}$$

The D_κ functions have only one subtraction,

$$\begin{aligned}
D^-_{\kappa}(x,y) &= \frac{i}{2\pi}\int\frac{dk^1}{2k^0}\left(e^{-ikx} - \theta(\kappa-k^0)\right)e^{iky} \\
D^+_{\kappa}(x,y) &= -\frac{i}{2\pi}\int\frac{dk^1}{2k^0}\left(e^{ikx} - \theta(\kappa-k^0)\right)e^{-iky}
\end{aligned} \tag{IV.15}$$

and $\tilde{D}^{\pm}_{\kappa}$ are again defined by an additional factor $\epsilon(k^1)$. The decomposition into wanted and unwanted parts analogous to (III.21) looks like

$$D^{\mp}_{\kappa}(x,y) = D^{\mp}(x-y) + \Delta^{\pm}(y). \tag{IV.16}$$

The decompositions of the $\tilde{D}$ functions have the same structure as those for the D functions.

Of course, the n-point functions of $\phi(x)$ turn out to be not translation invariant, since they are functions of the D_κ and $D_{\kappa\kappa}$.

(The only exception occurs for $\alpha=\beta=\sqrt{\pi}$, in which case all n-point functions are constants independent of x.) We use now a trick which makes that all the unwanted Δ terms cancel. The idea is to introduce operators $q^{\pm}(x)$, $\tilde{q}^{\pm}(x)$ which have the property that

$$[q^{\pm}(x), \psi(y)] = -\sqrt{\pi}\left\{\nu_1 \Delta^{\pm}(x) + \nu_2 \gamma^5 \tilde{\Delta}^{\pm}(x)\right\}\psi(y) \quad [q^{\pm}(x)]^* = q^{\mp}(x)$$

$$[\tilde{q}^{\pm}(x), \psi(y)] = -\sqrt{\pi}\left\{\mu_1 \tilde{\Delta}^{\pm}(x) + \mu_2 \gamma^5 \Delta^{\pm}(x)\right\}\psi(y) \quad [\tilde{q}^{\pm}(x)]^* = \tilde{q}^{\mp}(x). \tag{IV.17}$$

$q(x)$, $\tilde{q}(x)$ are, of course, given by

$$q^{\pm}(x) = \sqrt{\pi}\left\{\nu_1 Q\, \Delta^{\pm}(x) + \nu_2 \tilde{Q}\, \tilde{\Delta}^{\pm}(x)\right\} \tag{IV.18}$$

$$\tilde{q}^{\pm}(x) = \sqrt{\pi}\left\{\mu_1 Q\, \tilde{\Delta}^{\pm}(x) + \mu_2 \tilde{Q}\, \Delta^{\pm}(x)\right\}. \tag{IV.19}$$

Q and $\tilde{Q}$ are the charge and pseudocharge:

$$Q = \int dx^1 j^0(x) = \int dp^1 \left\{b^*(p^1)b(p^1) - a^*(p^1)a(p^1)\right\} \tag{IV.20}$$

$$\tilde{Q} = \int dx^1 j^1(x) = \int dp^1 \epsilon(p^1)\left\{b^*(p^1)b(p^1) - a^*(p^1)a(p^1)\right\}$$

$$[Q, \psi(x)] = -\psi(x)$$
$$[\tilde{Q}, \psi(x)] = -\gamma^5\psi(x). \tag{IV.21}$$

Now replace $\chi(x)$ in (IV.12) by

$$\chi^{\pm}(x) = \alpha j^{\pm}(x) + q^{\pm}(x) + \gamma^5\left[\beta \tilde{j}^{\pm}(x) + q^{\pm}(x)\right]. \tag{IV.22}$$

Let us now look at the n-point functions of $\phi(x)$. Think of a product of operators $\phi(x_j)$, $\phi(y_k)$ and pick two neighbouring factors, e.g.,

$$\phi(x_j)\phi(x_{j+1}) = e^{i\chi^+(x_j)}\psi(x_j)e^{i\chi^-(x_j)}e^{i\chi^+(x_{j+1})}\psi(x_{j+1})e^{i\chi^-(x_{j+1})}. \tag{IV.23}$$

One has to move the annihilation operators to the right and the creation operators to the left until they ultimately reach the vacuum. on

which they give zero. To commute $e^{iX^-(x_j)}$ and $e^{iX^+(x_{j+1})}$, one uses that $e^Ae^B = e^{[A,B]}e^Be^A$ if $[A,B]$ is a c-number. Then one commutes $\psi(x_j)$ and $e^{iX^+(x_{j+1})}$ (and $e^{iX^-(x_j)}$ and $\psi(x_{j+1})$) by using that, if $[A,B] = \lambda A$, λ a c-number, then $Ae^B = e^\lambda e^B A$. Having $X^-(x_j)$ and $X^-(x_{j+1})$ to the right of $\psi(x_{j+1})$, we commute them with the next factor, and so on. At the end of the process we have all e^{iX^+} to the left, all e^{iX^-} to the right, and the free operators ψ, $\bar\psi$ in between, and the n-point function is the free n-point function multiplied by an exponential of D_κ, $D_{\kappa\kappa}$ and Δ functions. Let us see how it is possible that for a certain value of $\mu_1, \mu_2, \nu_1, \nu_2$ we get a translation invariant exponent:

Commuting $X^-(x_j)$ and $X^+(x_{j+1}) \to D^-_{\kappa\kappa}(x_j, x_{j+1})$, contains $\Delta^-(x_j), \Delta^+(x_{j+1})$

Commuting $\psi(x_j)$ and $X^+(x_{j+1}) \to \begin{cases} D^+_\kappa(x_{j+1}, x_j), & \text{contains } \Delta^-(x_j) \\ \text{and } \Delta^+(x_{j+1}), \end{cases}$

Commuting $X^-(x_j)$ and $\psi(x_{j+1}) \to \begin{cases} D^-_\kappa(x_j, x_{j+1}), & \text{contains } \Delta^+(x_{j+1}) \\ \text{and } \Delta^-(x_j). \end{cases}$

(We have listed only the terms which are free of γ^5 matrices.) The condition that all $\Delta^-(x_j)$ and all $\Delta^+(x_{j+1})$ cancel each other gives us two equations for ν_1, which happen to be the same. Consideration of the terms proportional to $\gamma^5_{x_j}$, $\gamma^5_{x_{j+1}}$, $\gamma^5_{x_j}\gamma^5_{x_{j+1}}$ give one equation for each of ν_2, μ_1, μ_2. The result is

$$\nu_1 = \alpha\left(1 - \frac{\alpha}{\sqrt{\pi}}\right) \qquad \mu_1 = \alpha\left(1 - \frac{\beta}{\sqrt{\pi}}\right)$$
$$\nu_2 = \beta\left(1 - \frac{\alpha}{\sqrt{\pi}}\right) \qquad \mu_2 = \beta\left(1 - \frac{\beta}{\sqrt{\pi}}\right). \tag{IV.24}$$

Taking into account that some of the field operators in the n-point function are adjoints $\bar\phi(y)$ instead of $\phi(x)$ does not yield new conditions. With the choice (IV.23) for the μ and ν, the vacuum expectation values take on the form

$$\left(\Omega, \phi(x_1)\ldots\phi(x_n)\bar\phi(y_1)\ldots\bar\phi(y_n)\Omega\right)$$
$$= e^{iF(x,y)}\left(\Omega, \psi(x_1)\ldots\psi(x_n)\bar\psi(y_1)\ldots\bar\psi(y_n)\Omega\right)$$

$$F(x,y) = \sum_{j<k}\left\{\left(a+b\,\gamma^5_{x_j}\gamma^5_{x_k}\right)D^-(x_j-x_k) + \lambda\left(\gamma^5_{x_j}+\gamma^5_{x_k}\right)\tilde{D}^-(x_j-x_k)\right\}$$

$$+ \sum_{j<k}\left\{\left(a+b\,\gamma^5_{y_j}\gamma^5_{y_k}\right)D^-(y_j-y_k) - \lambda\left(\gamma^5_{y_j}+\gamma^5_{y_k}\right)\tilde{D}^-(y_j-y_k)\right\}$$

$$+ \sum_{j,k}\left\{\left(-a+b\,\gamma^5_{x_j}\gamma^5_{y_k}\right)D^-(x_j-y_k) + \lambda\left(-\gamma^5_{x_j}+\gamma^5_{y_k}\right)\tilde{D}^-(x_j-y_k)\right\}$$

$$a = \alpha^2 - 2\sqrt{\pi}\alpha \qquad b = \beta^2 - 2\sqrt{\pi}\beta \qquad \lambda = \alpha\beta - \sqrt{\pi}\alpha - \sqrt{\pi}\beta \tag{IV.25}$$

V. Lorentz Invariance

Having arrived at a solution which is translation invariant, we recognize that in general it is apparently not invariant with respect to homogeneous Lorentz transformations. This comes from the $\tilde{D}^-$ function which is not an invariant. Consequently, one would be tempted to put $\lambda = 0$. But this would mean losing one of the most interesting features of the model. For the free field $\psi(x)$ there exists a unitary transformation $V(\Lambda)$ with

$$V(\Lambda)^{-1}\psi(x)V(\Lambda) = e^{(\chi/2)\gamma^5}\psi(\Lambda^{-1}x)$$
$$V(\Lambda)^{-1}\bar{\psi}(x)V(\Lambda) = \bar{\psi}(\Lambda^{-1}x)e^{-(\chi/2)\gamma^5}. \tag{V.1}$$

$F(x,y)$ transforms in the following way:

$$F(\Lambda x,\Lambda y) = F(x,y) + \frac{i\lambda}{2\pi}\chi\left\{\sum_{j<k}\left(\gamma^5_{x_j}+\gamma^5_{x_k}\right) - \sum_{j<k}\left(\gamma^5_{y_j}+\gamma^5_{y_k}\right) + \right.$$

$$\left. + \sum_{j,k}\left(-\gamma^5_{x_j}+\gamma^5_{y_k}\right)\right\}$$

$$= F(x,y) + \frac{i\lambda}{2\pi}\chi\left\{-\sum_j\gamma^5_{x_j} + \sum_k\gamma^5_{y_k}\right\}.$$

Therefore, substituting $x_j \to \Lambda x_j$, $y_k \to \Lambda y_k$ in the expectation value of $\phi(x)$ is equivalent to multiplication with

$$\prod_{j=1}^{n}\prod_{k=1}^{n} e^{(\frac{1}{2}+(\lambda/2\pi))\chi\gamma^5_{x_j}} e^{-(\frac{1}{2}+(\lambda/2\pi))\chi\gamma^5_{y_k}},$$

and thus, in the subspace generated by ϕ and $\bar{\phi}$, there exists a unitary operator $U(\Lambda)$ which transforms $\phi(x)$ in the following way:

$$\begin{aligned} U(\Lambda)^{-1}\phi(x)U(\Lambda) &= e^{(\frac{1}{2}+(\lambda/2\pi)\chi\gamma^5}\phi(\Lambda^{-1}x) \\ U(\Lambda)^{-1}\bar{\phi}(x)U(\Lambda) &= \bar{\phi}(\Lambda^{-1}x)e^{-(\frac{1}{2}+(\lambda/2\pi)\chi\gamma^5}. \end{aligned} \tag{V.2}$$

In four-dimensional field theories (with finite number of field components) the way in which the field components transform under a Lorentz acceleration along the x'-axis is related to the spin in the following way: Since this is an abelian subgroup, its irreducible representations are one-dimensional and of the form $e^{\alpha\chi}$, and the maximum value of the α is the spin. For example, for a Dirac field, $\psi_i \to e^{\pm\frac{1}{2}\chi}\psi_i$ for an appropriate realisation of the γ-matrices, and for a vector field, $A_0 \pm A_1 \to e^{\pm\chi}(A_0 \pm A_1)$. In two dimensions there is, of course, no spin, since there are no rotations. Equation (V.2) shows that the solution (IV.24) interpolates the parameter α continuously. This situation cannot arise in four dimensions where the rotation subgroup gives periodic boundary conditions which make it integer or half integer. (It is interesting to note that this is not true for the infinite dimensional fully irreducible representations of SL(2,C) where the accelerations are independent from the rotations in the sense that the rotations are parametrized in a discrete way and the accelerations in a continuous way.[10])

The energy-momentum spectrum of the states $\phi(x)\Omega$, $\bar{\phi}(x)\Omega$ has a peculiar structure, as one sees if one looks at the two-point function:

$$\begin{aligned} \left(\Omega, \phi(x)\bar{\phi}(y)\Omega\right) &= \frac{1}{i} e^{i\{(-a+b\gamma^5_x\gamma^5_y)D^-(x-y)+\lambda(-\gamma^5_x+\gamma^5_y)\tilde{D}^-(x-y)\}} S^-(x-y) \\ &= \frac{1}{i} e^{-i\{(a+b)D^-(\xi)+2\lambda\gamma^5\tilde{D}^-(\xi)\}} S^-(\xi) \qquad (\xi = x-y) \end{aligned} \tag{V.3}$$

since $(\gamma^5_x+\gamma^5_y)S^-(x-y)=0$. The decomposition (III.9) gives

$$D^-(\xi) = \frac{1}{4\pi i}\left\{\log[\mu(\xi^0-\xi^1-i\epsilon)] + \log[\mu(\xi^0+\xi^1-i\epsilon)] + i\pi\right\}$$

$$\tilde{D}^-(\xi) = \frac{1}{4\pi i}\left\{\log[\mu(\xi^0-\xi^1-i\epsilon)] - \log[\mu(\xi^0+\xi^1-i\epsilon)]\right\} \qquad (V.4)$$

and therefore

$$(a+b)D^-(\xi) \mp 2\lambda\tilde{D}^-(\xi) = \frac{1}{4\pi i}\left\{(a+b\mp 2\lambda)\log[\mu(\xi^0-\xi^1-i\epsilon)] + (a+b\pm 2\lambda)\log[\mu(\xi^0+\xi^1-i\epsilon)] + i\pi(a+b)\right\}$$

$$e^{-i\{(a+b)D^-(\xi)+2\lambda\gamma^5\tilde{D}^-(\xi)\}} = e^{-i\frac{a+b}{4}} \times$$

$$\begin{bmatrix} \left[\frac{1}{\mu(\xi^0-\xi^1-i\epsilon)}\right]^{\frac{a+b-2\lambda}{4\pi}}\left[\frac{1}{\mu(\xi^0+\xi^1-i\epsilon)}\right]^{\frac{a+b+2\lambda}{4\pi}} & 0 \\ 0 & \left[\frac{1}{\mu(\xi^0-\xi^1-i\epsilon)}\right]^{\frac{a+b+2\lambda}{4\pi}}\left[\frac{1}{\mu(\xi^0+\xi^1-i\epsilon)}\right]^{\frac{a+b-2\lambda}{4\pi}} \end{bmatrix}$$

(V.5)

With the explicit form (II.16) for $S^-(\xi)$ this means that

$$\left(\Omega, \phi(x)\phi^*(y)\Omega\right) = \frac{\mu}{2\pi i} e^{-i\frac{a+b}{4}} \times$$

$$\begin{bmatrix} \left[\frac{1}{\mu(\xi^0-\xi^1-i\epsilon)}\right]^{\frac{a+b-2\lambda}{4\pi}} \left[\frac{1}{\mu(\xi^0+\xi^1-i\epsilon)}\right]^{\frac{a+b+2\lambda}{4\pi}+1} & 0 \\ 0 & \left[\frac{1}{\mu(\xi^0-\xi^1-i\epsilon)}\right]^{\frac{a+b+2\lambda}{4\pi}+1} \left[\frac{1}{\mu(\xi^0+\xi^1-i\epsilon)}\right]^{\frac{a+b-2\lambda}{4\pi}} \end{bmatrix}$$

$$= \frac{1}{2\pi}\int dp\, e^{-ip\xi} F(p). \qquad \text{(V.6)}$$

To get the Fourier transform F(p) take into account that

$$\int d\alpha(\alpha-i\epsilon)^{-\ell} e^{i\alpha\sigma} = \frac{2\pi e^{i(\pi/2)\ell}}{\Gamma(\ell)}\,\theta(\sigma)\sigma^{\ell-1}. \qquad \text{(V.7)}$$

Then, if $u=\mu(\xi^0+\xi^1)$, $v=\mu(\xi^0-\xi^1)$,

$$F(p) = \frac{1}{2\pi}\int d\xi\, e^{ip\xi}\left(\Omega, \phi(x)\phi^*(y)\Omega\right)$$

$$= \frac{1}{4\pi\mu^2}\int du\,dv\, e^{iu((p^0-p^1)/2\mu)}\, e^{iv((p^0+p^1)/2\mu)}\left(\Omega, \phi(x)\phi^*(y)\Omega\right)$$

$$= \frac{1}{2\mu}\,\frac{1}{\Gamma((a+b-2\lambda)/4\pi)\Gamma((a+b+2\lambda)/4\pi)+1)}\,\theta(p^0+p^1)\theta(p^0-p^1)\ \times$$

$$\begin{bmatrix} \left[\frac{p^0+p^1}{2\mu}\right]^{\frac{a+b-2\lambda}{4\pi}-1} \left[\frac{p^0-p^1}{2\mu}\right]^{\frac{a+b+2\lambda}{4\pi}} & 0 \\ 0 & \left[\frac{p^0+p^1}{2\mu}\right]^{\frac{a+b+2\lambda}{4\pi}} \left[\frac{p^0-p^1}{2\mu}\right]^{\frac{a+b-2\lambda}{4\pi}-1} \end{bmatrix}$$

(V.8)

For $\ell > 0$, the function

$$\frac{1}{\Gamma(\ell)}\,\theta(\sigma)\sigma^{\ell-1}$$

is integrable and defines a tempered distribution. For $\ell \to 0$, it converges to $\delta(\sigma)$ in the space $\mathcal{S}'$ of tempered distributions. If we reintroduce the original parameters α, β, then $a+b-2\lambda = (\alpha-\beta)^2$ and $a+b+2\lambda = (\alpha+\beta-2\sqrt{\pi})^2 - 4\pi$. Unless $\alpha=\beta$ or $\alpha+\beta=2\sqrt{\pi}$, the two-point function does not contain a δ-function contribution. The model then contains no one-particle states, no asymptotic fields, and the S-matrix cannot be defined. Finally, we note that the set of n-point functions (IV.24) satisfies the spectrum condition in an obvious way.

VI. Locality

Another property we have to check is the commutation behaviour for space-like separation. Look at the product

$$\phi(x)\bar{\phi}(y) = e^{i\chi^+(x)}\psi(x)e^{i\chi^-(x)}\,e^{-i\bar{\chi}^+(y)}\bar{\psi}(y)e^{-i\bar{\chi}^-(y)}.$$

We have to commute $\chi(x)$ with $\bar{\chi}(y)$, $\chi(x)$ with $\bar{\psi}(y)$ and $\psi(x)$ with $\bar{\chi}(y)$. The Δ terms again drop out and observing that, for x-y space-like, $D(x-y)=0$, $\tilde{D}(x-y) = -\frac{1}{2}\epsilon(x^1-y^1)$, one finds

$$\phi(x)\bar{\phi}(y) = -e^{iA(x,y)}\bar{\phi}(y)\phi(x) \qquad (x-y)^2 < 0$$

$$A(x,y) = \frac{\lambda}{2}\left(\gamma_x^5 - \gamma_y^5\right)\epsilon(x^1-y^1). \tag{VI.1}$$

In components, this means

$$\begin{aligned}
\phi_1(x)\phi_2^*(y) &= -\phi_2^*(y)\phi_1(x)\\
\phi_1(x)\phi_1^*(y) &= -e^{-i\lambda\epsilon(x^1-y^1)}\phi_1^*(y)\phi_1(x) \qquad \text{(VI.2)}\\
\phi_2(x)\phi_2^*(y) &= -e^{i\lambda\epsilon(x^1-y^1)}\phi_2^*(y)\phi_2(x).
\end{aligned}$$

For $\lambda = 2N\pi$ $(N=0,\pm1,\pm2\ldots)$, we have

$$\{\phi_i(x), \phi_i^*(y)\} = 0$$

and for $\lambda = (2N-1)\pi$,

$$[\phi_i(x), \phi_i^*(y)] = 0. \tag{VI.3}$$

The "spin" is $|\frac{1}{2} + \lambda/2\pi|$ and is, therefore, $|N + \frac{1}{2}|$ respectively $|N|$ for the two cases. The relation between "spin" and statistics is thus maintained and (VI.1) is a continuous interpolation of it. A similar structure is displayed by the relation

$$\begin{aligned}
\phi(x)\phi(y) &= -e^{iB(x,y)}\phi(y)\phi(x) \qquad (x-y)^2 < 0 \\
B(x,y) &= -\frac{\lambda}{2}\left(\gamma_x^5 + \gamma_y^5\right)\epsilon(x^1 - y^1) \\
\phi_1(x)\phi_2(y) &= -\phi_2(y)\phi_1(x) \qquad (VI.4) \\
\phi_1(x)\phi_1(y) &= -e^{i\lambda\epsilon(x^1-y^1)}\phi_1(y)\phi_1(x) \\
\phi_2(x)\phi_2(y) &= -e^{-i\lambda\epsilon(x^1-y^1)}\phi_2(y)\phi_2(x)
\end{aligned}$$

and again

$$\{\phi_i(x), \phi_i(y)\} = 0 \quad \text{for} \quad \lambda = 2N\pi \tag{VI.5}$$

$$[\phi_i(x), \phi_i(y)] = 0 \quad \lambda = (2N-1)\pi. \tag{VI.6}$$

VII. More Explicit Form of the n-Point Functions

Equation (IV.24) is a rather baroque way of characterizing an object which has a very simple structure. To see this, one has to realize that the free n-point function, which is a combinatorial sum of products of two-point functions, can be summed up to give a simple expression. For example,

$$\left(\Omega, \psi_1(x_1)\ldots\ {}_1(x_n)\psi_1^*(y_1)\ldots\psi_1^*(y_n)\Omega\right) = \frac{1}{(2\pi i)^n}\frac{\prod_{j<k}(u_j-u_k)\prod_{j<k}(U_j-U_k)}{\prod_{j,k}(u_j-U_k-i\epsilon)} \tag{VII.1}$$

$$u_j = x_j^0 + x_j^1 \qquad U_j = y_j^0 + y_j^1.$$

Look now at the n-point functions of $\phi_1(x)$. We have to put

$$\gamma^5_{x_j} = -1, \quad \gamma^5_{y_k} = 1$$

in $F(x,y)$ of (IV.24). With the form (V.4) for D^- and $\tilde{D}^-$, $e^{iF(x,y)}$ takes on the form

$$e^{iF(x,y)} = e^{-in\frac{a+b}{4}} \left(\frac{1}{\mu}\right)^{n\frac{a+b}{2\pi}} \prod_{j<k} (v_j - v_k - i\epsilon)^{\frac{a+b-2\lambda}{4\pi}} (u_j - u_k - i\epsilon)^{\frac{a+b+2\lambda}{4\pi}}$$

$$\cdot \prod_{j<k} (V_j - V_k - i\epsilon)^{\frac{a+b-2\lambda}{4\pi}} (U_j - U_k - i\epsilon)^{\frac{a+b+2\lambda}{4\pi}}$$

$$\cdot \prod_{j,k} \left(\frac{1}{v_j - V_k - i\epsilon}\right)^{\frac{a+b-2\lambda}{4\pi}} \left(\frac{1}{u_j - U_k - i\epsilon}\right)^{\frac{a+b+2\lambda}{4\pi}} \tag{VII.2}$$

and therefore

$$\left(\Omega, \phi_1(x_1)\ldots\phi_1(x_n)\phi_1^*(y_1)\ldots\phi_1^*(y_n)\Omega\right)$$

$$= \left(\frac{1}{2\pi i}\right)^n e^{-in(\pi/2)(\delta_1+\delta_2-1)} \left(\frac{1}{\mu}\right)^{n(\delta_1+\delta_2-1)}$$

$$\cdot \prod_{j<k} (u_j - u_k - i\epsilon)^{\delta_1} (U_j - U_k - i\epsilon)^{\delta_1} \prod_{j,k} \left(\frac{1}{u_j - U_k - i\epsilon}\right)^{\delta_1} \tag{VII.3}$$

$$\cdot \prod_{j<k} (v_j - v_k - i\epsilon)^{\delta_2} (V_j - V_k - i\epsilon)^{\delta_2} \prod_{j,k} \left(\frac{1}{v_j - V_k - i\epsilon}\right)^{\delta_2}$$

$$\delta_1 = \frac{a+b+2\lambda}{4\pi} + 1 \qquad \delta_2 = \frac{a+b-2\lambda}{4\pi}.$$

The n-point functions of $\psi_2(x)$ and $\phi_2(x)$ follow from those of $\psi_1(x)$ and $\phi_1(x)$ by the substitutions $u \leftrightarrow v$, $U \leftrightarrow V$. For δ_1 and δ_2 integer,

$$\left(\Omega,\phi_1(x_1)\dots\phi_1(x_n)\phi_1^*(y_1)\dots\phi_1^*(y)\Omega\right) = \left(\frac{2\pi}{\mu}\right)^{n(\delta_1+\delta_2-1)}$$

$$\left(\Omega,\psi_1(x_1)\dots\psi_1(x_n)\psi_1^*(y_1)\dots\psi_1^*(y_n)\Omega\right)^{\delta_1} \qquad \text{(VII. 4)}$$

$$\left(\Omega,\psi_2(x_1)\dots\psi_2(x_n)\psi_2^*(y_1)\dots\psi_2^*(y_n)\Omega\right)^{\delta_2}$$

and the n-point function of $\phi_1(x)$ has the structure of a direct product of free n-point functions. It is well known that, if A(x) and B(x) are two fields, one can construct from their n-point functions a new sequence of n-point functions by forming

$$W_n^C(x_1\dots x_n) = W_n^A(x_1\dots x_n)W_n^B(x_1\dots x_n).$$

In terms of operators, this corresponds to forming the direct product $\mathfrak{H}^C=\mathfrak{H}^A\otimes\mathfrak{H}^B$ of the Hilbert spaces and computing the n-point functions of the direct product of the fields: $C(x) = A(x) \otimes B(x)$. The structure (VII. 2) is a continuous interpolation (with respect to the exponent) of such multiple products. For a field with non-vanishing spin in four dimensions such interpolations cannot occur, since they would lead to interpolation of the spin. For a scalar field the question is open (the critical point is, of course, the positive definiteness).

For the mixed functions

$$\left(\Omega,\prod_1^{\ell}\phi_1(x_j)\prod_{\ell+1}^{n}\phi_2(x_j)\prod_1^{\ell}\phi_1^*(y_j)\prod_{\ell+1}^{n}\phi_2^*(y_j)\Omega\right)$$

use that

$$\left(\Omega,\prod_1^{\ell}\psi_1(x_j)\prod_{\ell+1}^{n}\psi_2(x_j)\prod_1^{\ell}\psi_1^*(y_j)\prod_{\ell+1}^{n}\psi_2^*(y_j)\Omega\right) = (-1)^{\ell(n-\ell)}$$

$$\left(\Omega,\prod_1^{\ell}\psi_1(x_j)\prod_1^{\ell}\psi_1^*(y_j)\Omega\right)\left(\Omega,\prod_{\ell+1}^{n}\psi_2(x_j)\prod_{\ell+1}^{n}\psi_2^*(y_j)\Omega\right).$$

Therefore, the mixed functions of ϕ split into a product of functions of ϕ_1 and of ϕ_2, which is, however, multiplied by a factor which comes from the summations with $1\leqq j\leqq\ell$, $\ell+1\leqq k\leqq n$ and $\ell+1\leqq j\leqq n$, $1\leqq k\leqq\ell$ in F(x,y) in (IV.24):

$$\left(\Omega,\prod_{1}^{\ell}\phi_1(x_j)\prod_{\ell+1}^{n}\phi_2(x_j)\prod_{1}^{\ell}\phi_1^*(y_j)\prod_{\ell+1}^{n}\phi_2^*(y_j)\Omega\right) = (-1)^{\ell(n-\ell)}$$

$$\cdot\frac{\prod_{j=1}^{\ell}\prod_{k=\ell+1}^{n}(u_j-u_k-i\epsilon)^{\delta}(v_j-v_k-i\epsilon)^{\delta}(U_j-U_k-i\epsilon)^{\delta}(V_j-V_k-i\epsilon)^{\delta}}{\prod_{j=1}^{\ell}\prod_{k=\ell+1}^{n}(u_j-U_k-i\epsilon)^{\delta}(v_j-V_k-i\epsilon)^{\delta}\prod_{j=\ell+1}^{n}\prod_{k=1}^{\ell}(u_j-U_k-i\epsilon)^{\delta}(v_j-V_k-i\epsilon)^{\delta}} \tag{VII. 5}$$

$$\cdot\left(\Omega,\prod_{1}^{\ell}\phi_1(x_j)\prod_{1}^{\ell}\phi_1^*(y_j)\Omega\right)\left(\Omega,\prod_{\ell+1}^{n}\phi_2(x_j)\prod_{\ell+1}^{n}\phi_2^*(y_j)\Omega\right)$$

$$\delta = \frac{a-b}{4\pi}.$$

VIII. Cluster Properties

In a field theory which satisfies the conditions of relativistic invariance, spectrum and locality, the vacuum is unique if and only if the n-point functions have the cluster decomposition property, in our case:

$$\lim_{\rho\to\infty}\Big(\Omega,\psi(x_1+\rho\eta)\ldots\psi(x_p+\rho\eta)\bar{\psi}(y_1+\rho\eta)\ldots\bar{\psi}(y_q+\rho\eta)\psi(x_{p+1})\ldots\psi(x_n)\bar{\psi}(y_{q+1})$$

$$\ldots\bar{\psi}(y_n)\Omega\Big) = \Big(\Omega,\psi(x_1)\ldots\psi(x_p)\bar{\psi}(y_1)\ldots\bar{\psi}(y_q)\Omega\Big)$$

$$\cdot\Big(\Omega,\psi(x_{p+1})\ldots\psi(x_n)\bar{\psi}(y_{q+1})\ldots\bar{\psi}(y_n)\Omega\Big), \tag{VIII. 1}$$

if η is a space-like vector, $\eta^2 = -1$.

For the two-point function, we must have

$$\lim_{\rho\to\infty}\Big(\Omega,\psi(x)\bar{\psi}(y+\rho\eta)\Omega\Big) = 0. \tag{VIII. 2}$$

With formula (V.5) this gives the necessary condition

$$a + b + 2\pi > 0. \tag{VIII. 3}$$

Higher functions: The expression

$$\left(\Omega, \prod_{1}^{p} \psi_i(x_j+\rho\eta) \prod_{p+1}^{n} \psi_i(x_j) \prod_{1}^{q} \psi_i^*(y_j+\rho\eta) \prod_{q+1}^{n} \psi_i^*(y_j)\Omega\right)$$

goes like

$$\left(\frac{1}{\rho}\right)^{(p-q)^2}$$

for large ρ. The corresponding object for ϕ_1 goes, therefore, like

$$\left(\frac{1}{\rho}\right)^{(p-q)^2((a+b)/2\pi)+1)},$$

and for $p \neq q$ this converges to zero as it should, provided (VIII. 3) holds. The mixed functions give stronger conditions. Look at

$$\left(\Omega, \prod_{1}^{p_1} \phi_1(x_j+\rho\eta) \prod_{p_1+1}^{\ell} \phi_1(x_j) \prod_{\ell+1}^{\ell+p_2} \phi_2(x_j+\rho\eta) \prod_{\ell+p_2+1}^{n} \phi_2(x_j) \prod_{1}^{q_1} \phi_1^*(y_j+\rho\eta)\right.$$

$$\left. \prod_{q_1+1}^{\ell} \phi_1^*(y_j) \prod_{\ell+1}^{\ell+q_2} \phi_2^*(y_j+\rho\eta) \prod_{\ell+q_2+1}^{n} \phi_2^*(y_j)\Omega\right).$$

The second and third factor in (VII. 4) then go like

$$\left(\frac{1}{\rho}\right)^{(p_1-q_1)^2((a+b)/2\pi)+1)}$$

and

$$\left(\frac{1}{\rho}\right)^{(p_2-q_2)^2((a+b)/2\pi)+1)}.$$

The first factor goes like

$$\left(\frac{1}{\rho}\right)^{(p_1-q_1)(p_2-q_2)(a-b)/\pi}.$$

The whole object must tend to zero unless simultaneously $p_1 = q_1$, $p_2 = q_2$. This gives the condition

$$\left[(p_1-q_1)^2+(p_2-q_2)^2\right]\left(\frac{a+b}{2\pi}+1\right)+(p_1-q_1)(p_2-q_2)\,\frac{a-b}{\pi}>0$$

$$p_i, q_k = 0,1,2,\ldots$$

unless $p_1=q_1$ and $p_2=q_2$. This is equivalent to

$$a+\pi>0, \qquad b+\pi>0. \tag{VIII.4}$$

There remains the case $p_1=q_1$, $p_2=q_2$. Consider first

$$\left(\Omega, \prod_1^p \phi_1(x_j+\rho\eta) \prod_{p+1}^n \phi_1(x_j) \prod_1^p \phi_1^*(y_j+\rho\eta) \prod_{p+1}^n \phi_1^*(y_j)\Omega\right).$$

There is no loss in generality in assuming $\eta^0=0$, $|\eta^1|=1$. If u_j is translated and u_k is not, then $u_j-u_k \to u_j+\rho\eta^1-u_k-i\epsilon \sim \rho(\eta^1-i\epsilon)$ and

$$(u_j-u_k+\rho\eta^1-i\epsilon)^{\delta_1} \sim \rho^{\delta_1} e^{-i\pi\delta_1\theta(-\eta^1)}.$$

Therefore,

$$\prod_{j<k}(u_j-u_k-i\epsilon)^{\delta_1}(U_j-U_k-i\epsilon)^{\delta_1}(v_j-v_k-i\epsilon)^{\delta_2}(V_j-V_k-i\epsilon)^{\delta_2}$$

$$\cdot\prod_{j,k}\left(\frac{1}{u_j-U_k-i\epsilon}\right)^{\delta_1}\left(\frac{1}{v_j-V_k-i\epsilon}\right)^{\delta_2}$$

$$\to\left(\prod_{1\le j,k\le p}\cdots\right)\left(\prod_{p+1\le j,k\le n}\cdots\right)e^{-i\pi\delta_1\theta(-\eta^1)p(n-p)}e^{i\pi\delta_1\theta(\eta^1)p(n-p)}$$

$$\cdot e^{-i\pi\delta_2\theta(\eta^1)p(n-p)}e^{i\pi\delta_2\theta(-\eta^1)p(n-p)}$$

$$=(\quad)(\quad)\,e^{i\pi(\delta_1-\delta_2)\epsilon(\eta^1)p(n-p)}$$

$$=(\quad)(\quad)\,(-1)^{p(n-p)}e^{i\lambda\epsilon(\eta^1)p(n-p)}$$

which implies that

$$\lim_{\rho\to\infty}\Big(\Omega,\prod_{1}^{p}\phi_1(x_j+\rho\eta)\prod_{p+1}^{n}\phi_1(x_j)\prod_{1}^{p}\phi_1^*(y_j+\rho\eta)\prod_{p+1}^{n}\phi_1^*(y_j)\Omega\Big)=(-1)^{p(n-p)}$$

$$\cdot e^{i\lambda\epsilon(\eta^1)p(n-p)}\Big(\Omega,\prod_{1}^{p}\phi_1(x_j)\prod_{1}^{p}\phi_1^*(y_j)\Omega\Big)\Big(\Omega,\prod_{p+1}^{n}\phi_1(x_j)\prod_{p+1}^{n}\phi_1^*(y_j)\Omega\Big).$$

Mixed functions: The first factor in (VII.5) just splits into two factors relating to the two clusters, and so

$$\lim_{\rho\to\infty}\Big(\Omega,\prod_{1}^{p_1}\phi_1(x_j+\rho\eta)\prod_{p_1+1}^{\ell}\phi_1(x_j)\prod_{\ell+1}^{\ell+p_2}\phi_2(x_j+\rho\eta)\prod_{\ell+p_2+1}^{n}\phi_2(x_j)\prod_{1}^{p_1}\phi_1^*(y_j+\rho\eta)$$

$$\cdot\prod_{p_1+1}^{\ell}\phi_1^*(y_j)\prod_{\ell+1}^{\ell+p_2}\phi_2^*(y_j+\rho\eta)\prod_{\ell+p_2+1}^{n}\phi_2^*(y_j)\Omega\Big)$$

$$=(-1)^{\ell(n-\ell)}\frac{\prod_{j=1}^{p_1}\prod_{k=\ell+1}^{\ell+p_2}\cdots\prod_{j=p_1+1}^{\ell}\prod_{k=\ell+p_2+1}^{n}\cdots\cdots}{\prod_{j=1}^{p_1}\prod_{k=\ell+1}^{\ell+p_2}\cdots\prod_{j=\ell+1}^{\ell+p_2}\prod_{k=1}^{p_1}\cdots\prod_{j=p_1+1}^{\ell}\prod_{k=\ell+p_2+1}^{n}\cdots\prod_{j=\ell+p_2+1}^{n}\prod_{k=p_1+1}^{\ell}}$$

$$\cdot(-1)^{p_1(\ell-p_1)}e^{i\lambda\epsilon(\eta^1)p_1(\ell-p_1)}\Big(\Omega,\prod_{1}^{p_1}\phi_1(x_j)\prod_{1}^{p_1}\phi_1^*(y_j)\Omega\Big)$$

$$\Big(\Omega,\prod_{p_1+1}^{\ell}\phi_1(x_j)\prod_{p_1+1}^{\ell}\phi_1^*(y_j)\Omega\Big)$$

$$\cdot(-1)^{p_2(n-\ell-p_2)}e^{-i\lambda\epsilon(\eta^1)p_2(n-\ell-p_2)}\Big(\Omega,\prod_{\ell+1}^{\ell+p_2}\phi_2(x_j)\prod_{\ell+1}^{\ell+p_2}\phi_2^*(y_j)\Omega\Big)$$

$$\Big(\Omega,\prod_{\ell+p_2+1}^{n}\phi_2(x_j)\prod_{\ell+p_2+1}^{n}\phi_2^*(y_j)\Omega\Big)\quad=$$

$$= (-1)^{p_1(n-p_1)+p_2(n-p_2)} e^{i\lambda\epsilon(\eta^1)[p^1(\ell-p^1)-p_2(n-\ell-p_2)]}$$

$$\cdot\Big(\Omega, \prod_1^{p_1}\phi_1(x_j)\prod_{\ell+1}^{\ell+p_2}\phi_2(x_j)\prod_1^{p_1}\phi_1^*(y_j)\prod_{\ell+1}^{\ell+p_2}\phi_2^*(y_j)\Omega\Big)$$

$$\cdot\Big(\Omega, \prod_{p_1+1}^{\ell}\phi_1(x_j)\prod_{\ell+p_2+1}^{n}\phi_2(x_j)\prod_{p_1+1}^{\ell}\phi_1^*(y_j)\prod_{\ell+p_2+1}^{n}\phi_2^*(y_j)\Omega\Big).$$

The phase appearing here is just the one which is produced when one commutes all the translated fields to the left. We conclude that the cluster property (VIII. 1) is fulfilled provided the coefficients a and b satisfy the condition (VIII. 4). Since

$$a + \pi = (\alpha - \sqrt{\pi})^2, \quad b + \pi = (\beta - \sqrt{\pi})^2,$$

we obtain the necessary and sufficient condition

$$\alpha \neq \sqrt{\pi}, \qquad \beta \neq \sqrt{\pi}. \tag{VIII. 5}$$

IX. Current and Equation of Motion

The current $J^\mu(x)$ of the field $\phi(x)$ shall be defined by the properties

$$[J^\mu(x), \phi(y)] = -\Big(c_1 g^{\mu\nu} + c_2 e^{\mu\nu}\gamma^5\Big)D_\nu(x-y)\phi(y)$$

$$(\Omega, J^\mu(x)\Omega) = 0. \tag{IX. 1}$$

We make the Ansatz

$$J^\mu(x) = \frac{1}{\sqrt{\pi}}\partial^\mu J(x)$$

$$J(x) = j(x) + q(x;\sigma_1,\sigma_2)$$

$$q(x) = \sqrt{\pi}\Big\{\rho_1\Delta(x)Q + \rho_2\tilde{\Delta}(x)\tilde{Q}\Big\} \tag{IX. 2}$$

with ρ_1, ρ_2 to be determined. The commutator of J and ϕ is

$$[J(x), \phi(y)] = \Big\{\alpha D_{\kappa\kappa}(x,y) + \beta\gamma^5 \tilde{D}_{\kappa\kappa}(x,y) - \sqrt{\pi} D_\kappa(x,y)$$
$$-\sqrt{\pi}\gamma^5 \tilde{D}_\kappa(x,y) - \sqrt{\pi}\rho_1 \Delta(x) - \sqrt{\pi}\rho_2 \gamma^5 \tilde{\Delta}(x)\Big\}\phi(y)$$
$$= \Big\{(\alpha - \sqrt{\pi}) D(x-y) + (\beta - \sqrt{\pi})\gamma^5 \tilde{D}(x-y) + (\alpha - \sqrt{\pi})\Delta(y)$$
$$+ (\beta - \sqrt{\pi})\gamma^5 \tilde{\Delta}(y) - (\alpha + \sqrt{\pi}\rho_1)\Delta(x) - (\beta + \sqrt{\pi}\rho_2)\tilde{\Delta}(x)\Big\}\phi(y), \tag{IX.3}$$

and so we have to put

$$\rho_1 = -\frac{\alpha}{\sqrt{\pi}} \qquad \rho_2 = -\frac{\beta}{\sqrt{\pi}}.$$

Then we have

$$[J^\mu(x), \phi(y)] = -\Big\{\Big(1 - \frac{\alpha}{\sqrt{\pi}}\Big) g^{\mu\nu} + \Big(1 - \frac{\beta}{\sqrt{\pi}}\Big) \epsilon^{\mu\nu}\gamma^5\Big\} D_\nu(x-y)$$

$$J^\mu(x) = \frac{1}{\sqrt{\pi}} \partial^\mu J(x) \tag{IX.4}$$

$$J(x) = j(x) + q\Big(x; -\frac{\alpha}{\sqrt{\pi}}, -\frac{\beta}{\sqrt{\pi}}\Big).$$

The charge of the interacting field is then

$$\int dx^1 J^0(x) = \int dx^1 j^0(x) - \frac{\alpha}{\sqrt{\pi}} \int dx^1 \partial^0 \Delta(x) \cdot Q = \Big(1 - \frac{\alpha}{\sqrt{\pi}}\Big) Q, \tag{IX.5}$$

differing only by a numerical factor from the free charge.

We have obtained the solution of the model by quantizing the classical solution (IV.7) instead of quantizing the equation of motion (IV.1). Our next aim is to determine what the precise form of the equation of motion is. Differentiation of (IV.12), (IV.22) yields

$$i\gamma^\mu \partial_\mu \phi(x) = -\gamma^\mu \partial_\mu \chi^+(x)\phi(x) - \phi(x)\gamma^\mu \partial_\mu \chi^-(x). \tag{IX.6}$$

The projection onto positive or negative frequencies is a nonlocal operation, and therefore the right hand side of (IX.6) is not in an obviously covariant form. We pretend now that (IX.6) can be written in the following form:

$$i\gamma^\mu \partial_\mu \phi(x) = -\tfrac{1}{2} \lim_{\epsilon \to 0} \left\{ \gamma^\mu \partial_\mu \chi(x+\epsilon)\phi(x) + \phi(x)\gamma^\mu \partial_\mu \chi(x-\epsilon) \right\}. \quad (\epsilon^2 \neq 0). \quad \text{(IX. 7)}$$

To prove this, one has to note that (IX. 6) can be written as

$$i\gamma^\mu \partial_\mu \phi(x) = -\tfrac{1}{2} \lim_{\epsilon \to 0} \Big\{ \gamma^\mu \partial_\mu \chi^+(x+\epsilon)\phi(x) + \phi(x)\gamma^\mu \partial_\mu \chi^-(x+\epsilon) + \gamma^\mu \partial_\mu \chi^+(x-\epsilon)\phi(x) + \phi(x)\gamma^\mu \partial_\mu \chi^-(x-\epsilon) \Big\},$$

and we have to show that

$$[\phi(x), \gamma^\mu \partial_\mu \chi^-(x+\epsilon)] + [\gamma^\mu \partial_\mu \chi^+(x-\epsilon), \phi(x)] = 0.$$

But this follows from the relation

$$[\gamma^\mu \partial_\mu \chi^\pm(y), \phi(x)] = (\alpha - \beta)^2 \gamma^\mu \partial_\mu D^\pm(y-x)\phi(x).$$

Finally, $\gamma^\mu \partial_\mu \chi(x)$ can be expressed by the current

$$\gamma^\mu \partial_\mu \gamma^5 \left[\beta \tilde{j}(x) + \tilde{q}(x;\mu_1,\mu_2) \right] = -\gamma^\mu \left[\sqrt{\pi}\beta j_\mu(x) + \partial_\mu q(x;\mu_1,\mu_2) \right]$$

and

$$\gamma^\mu \partial_\mu \chi(x) = \gamma^\mu \left[\sqrt{\pi}(\alpha-\beta) j_\mu(x) + \partial_\mu q(x;\nu_1-\mu_1,\nu_2-\mu_2) \right]$$

$$= \sqrt{\pi}(\alpha-\beta)\gamma^\mu J_\mu(x).$$

The equation of motion assumes then the form

$$i\gamma^\mu \partial_\mu \phi(x) = -\sqrt{\pi}(\alpha-\beta)\gamma^\mu \lim_{\epsilon \to 0} \tfrac{1}{2} \left\{ J_\mu(x+\epsilon)\phi(x) + \phi(x)J_\mu(x-\epsilon) \right\}. \quad \text{(IX. 8)}$$

It is possible to express J_μ in terms of ϕ, but we will do this only for the case when $\lambda = 0$.

X. The Case $\lambda = 0$ and the Currents of Johnson and Schwinger

If one postulates that $\phi(x)$ has the same Lorentz transformation properties as the free field $\psi(x)$, one arrives at the solution found by Johnson.[4] The condition $\lambda = 0$ is equivalent to

$$\left(\frac{\alpha}{\sqrt{\pi}} - 1\right)\left(\frac{\beta}{\sqrt{\pi}} - 1\right) = 1. \tag{X.1}$$

This gives the parametric representation

$$\alpha = \sqrt{\pi}(1+\rho) \qquad \beta = \sqrt{\pi}\left(1 + \frac{1}{\rho}\right) \qquad (-\infty < \rho < +\infty,\ \rho \neq 0)$$
$$a = \pi(\rho^2 - 1) \qquad b = \pi\left(\frac{1}{\rho^2} - 1\right). \tag{X.2}$$

We now want to express the current in terms of the field $\phi(x)$. One way to do this is to define it as a limit of a nonlocal expression. For a free Dirac field with arbitrary mass,

$$j^\mu(x) = \tfrac{1}{2} \lim_{\epsilon \to 0} \left\{ \bar{\psi}(x+\epsilon)\gamma^\mu\psi(x) - \gamma^\mu\psi(x+\epsilon)\bar{\psi}(x) \right\} \tag{X.3}$$

in two and four dimensions. ϵ may be time- or space-like. In the Thirring model one has either to average over a spacelike and a timelike direction[4] or to modify the right-hand side of Eq. (X.3).[8] The introduction of timelike vectors, however, means that one is outside the framework of the canonical formalism. We give here a spacelike current definition which contains a free parameter and which reproduces the results of both Johnson[4] and Schwinger.[8),9)] Look first at the expression

$$\bar{\phi}(x+\epsilon)\gamma^\mu\phi(x) = e^{-i\bar{\chi}^+(x+\epsilon)}\bar{\psi}(x+\epsilon)e^{-i\bar{\chi}^-(x+\epsilon)}\gamma^\mu e^{i\chi^+(x)}\psi(x)e^{i\chi^-(x)}$$
$$= e^{-i(\alpha-\beta)^2 D^-(\epsilon)} e^{-i[\bar{\chi}^+(x+\epsilon)-\bar{\chi}^+(x)]}\bar{\psi}(x+\epsilon)\gamma^\mu\psi(x)$$
$$e^{-i[\chi^-(x+\epsilon)-\chi^-(x)]}.$$

Separate the singular part of $\bar{\psi}_\alpha(x+\epsilon)\psi_\beta(x)$:

$$\bar{\psi}_\alpha(x+\epsilon)\psi_\beta(x) = :\bar{\psi}_\alpha(x+\epsilon)\psi_\beta(x): + \frac{1}{i} S^+_{\alpha\beta}(-\epsilon)$$
$$= :\bar{\psi}_\alpha(x+\epsilon)\psi_\beta(x): + \frac{1}{2\pi i}\frac{(\gamma\epsilon)_{\beta\alpha}}{\epsilon^2}$$

and develop

$$e^{-i[\chi^-(x+\epsilon)-\chi^-(x)]} = 1 - i\epsilon^\rho \partial_\rho \chi^-(x) + \mathcal{O}(\epsilon^2).$$

Then,

$$e^{i(\alpha-\beta)^2 D^-(\epsilon)} \bar{\psi}(x+\epsilon)\gamma^\mu \psi(x) = :\bar{\psi}(x)\gamma^\mu\psi(x): + \frac{1}{2\pi i}\frac{\mathrm{Sp}\{\gamma^\mu(\gamma\epsilon)\}}{\epsilon^2}$$

$$- \frac{1}{2\pi}\mathrm{Sp}\left\{\frac{\gamma\epsilon}{\epsilon^2}\,\epsilon^\rho \partial_\rho \bar{\chi}^+(x)\gamma^\mu\right\}$$

$$- \frac{1}{2\pi}\mathrm{Sp}\left\{\frac{\gamma\epsilon}{\epsilon^2}\gamma^\mu \epsilon^\rho \partial_\rho \chi^-(x)\right\} + \mathcal{O}(\epsilon)$$

$$= :\bar{\psi}(x)\gamma^\mu\psi(x): + \frac{1}{i\pi}\frac{\epsilon^\mu}{\epsilon^2} - \frac{1}{2\pi}\frac{\epsilon_\nu \epsilon_\rho}{\epsilon^2}\mathrm{Sp}\left\{\gamma^\nu\gamma^\mu\partial^\rho\chi(x)\right\} + \mathcal{O}(\epsilon)$$

with

$$\mathrm{Sp}\left\{\gamma^\nu\gamma^\mu\partial^\rho\chi(x)\right\} = 2\left\{g^{\mu\nu}\left[\sqrt{\pi}\alpha j^\rho(x) + \partial^\rho q(x;\nu_1,\nu_2)\right]\right.$$

$$\left. + e^{\mu\nu}e^{\rho\lambda}\left[\sqrt{\pi}\beta j_\lambda(x) + \partial_\lambda q(x;\mu_1,\mu_2)\right]\right\}.$$

If we use

$$e^{\mu\nu}e^{\rho\lambda} = g^{\mu\lambda}g^{\nu\rho} - g^{\mu\rho}g^{\nu\lambda},$$

we obtain

$$|e^2|^{(\alpha-\beta)^2/4\pi}\bar{\psi}(x+\epsilon)\gamma^\mu\psi(x)$$

$$= j^\mu(x) - \frac{1}{\pi}\left[\sqrt{\pi}\beta j^\mu(x) + \partial^\mu q(x;\mu_1,\mu_2)\right]$$

$$- \frac{1}{\pi}\frac{\epsilon^\mu\epsilon^\lambda}{\epsilon^2}\left[\sqrt{\pi}(\alpha-\beta)j_\lambda(x) + \partial_\lambda q\left(x;\nu_1-\mu_1,\nu_2-\mu_2\right)\right]$$

$$= \left(1 - \frac{\beta}{\sqrt{\pi}}\right)J^\mu(x) - \frac{\alpha-\beta}{\sqrt{\pi}}\frac{\epsilon^\mu\epsilon^\lambda}{\epsilon^2}J_\lambda(x) + \frac{1}{i\pi}\frac{\epsilon^\mu}{\epsilon^2} + \mathcal{O}(\epsilon). \qquad (X.4)$$

Since $\epsilon^2 < 0$, we also have

$$-|\epsilon^2|^{(\alpha-\beta)^2/4\pi}\gamma^\mu\phi(x+\epsilon)\bar{\phi}(x)$$

$$= |\epsilon^2|^{(\alpha-\beta)^2/4\pi}\bar{\phi}(x)\gamma^\mu\phi(x+\epsilon) \tag{X.5}$$

$$= \left(1 - \frac{\beta}{\sqrt{\pi}}\right) J^\mu(x) - \frac{\alpha-\beta}{\sqrt{\pi}} \frac{\epsilon^\mu\epsilon^\lambda}{\epsilon^2} J_\lambda(x) - \frac{1}{i\pi} \frac{\epsilon^\mu}{\epsilon^2} + \mathcal{O}(\epsilon).$$

Averaging over (X.4) and (X.5) eliminates the singular part but leaves us with a non-covariant expression. One gets rid of the non-covariant terms if one defines the current $K^\mu(x)$ as (g, σ are coefficients to be determined)

$$\begin{aligned} K^\mu(x) \equiv \lim_{\delta\to 0}\lim_{\epsilon\to 0} |\epsilon^2|^{(\alpha-\beta)^2/4\pi} \frac{1}{2}\Big\{&\bar{\phi}(x+\epsilon)\gamma_\nu\phi(x)\Big[g^{\mu\nu}(1+ig\epsilon^\lambda K_\lambda(x+\delta)) + \\ &+ i\sigma\epsilon^\nu K^\mu(x+\delta)\Big] - \gamma_\nu\phi(x+\epsilon)\bar{\phi}(x)\Big[g^{\mu\nu}(1-ig\epsilon^\lambda K_\lambda(x+\delta)) \\ &- i\sigma\epsilon^\nu K^\mu(x+\delta)\Big]\Big\} \\ = \left(1 - \frac{\beta}{\sqrt{\pi}}\right) J^\mu(x) + \frac{\sigma}{\pi} K^\mu(x) + &\left\{-\frac{\alpha-\beta}{\sqrt{\pi}} J_\lambda(x) + \frac{g}{\pi} K_\lambda(x)\right\} \lim_{\epsilon\to 0} \frac{\epsilon^\mu\epsilon^\lambda}{\epsilon^2}. \end{aligned} \tag{X.6}$$

If the non-covariant term is to vanish we must have

$$K_\lambda(x) = \frac{\sqrt{\pi}(\alpha-\beta)}{g} J_\lambda(x), \tag{X.7}$$

and Eq. (X.6) then tells us

$$K^\mu(x) = \left(1 - \frac{\beta}{\sqrt{\pi}} + \frac{\sigma}{g}\frac{\alpha-\beta}{\sqrt{\pi}}\right) J^\mu(x). \tag{X.8}$$

g and σ must satisfy the relation (X.9)

$$\left(1 - \frac{\beta}{\sqrt{\pi}}\right) g + \frac{\alpha-\beta}{\sqrt{\pi}}\sigma = \sqrt{\pi}(\alpha-\beta). \tag{X.10}$$

The equation of motion can be written

$$i\gamma^\mu\partial_\mu\phi(x) = -g\gamma^\mu \lim_{\epsilon\to 0} \frac{1}{2}\left\{K_\mu(x+\epsilon)\phi(x) + \phi(x)K_\mu(x-\epsilon)\right\}. \tag{X.11}$$

We have obtained the result that to any given solution (i.e., to given α, β, $\alpha-\beta\neq 0$) and any $g\neq 0$ we can find a current definition (i.e., find a σ from Eq. (X.9)) such that g is the coupling constant occurring in the equation of motion. (If $\alpha=\beta$, then $\gamma^\mu\partial_\mu\phi(x) = 0$, and if $g=0$, $\alpha-\beta\neq 0$, $K_\rho(x)$ is infinite.)

Solving (X.1) and (X.10) yields a parametric representation of α and β:

$$\alpha = \sqrt{\pi}\left(1 \pm \sqrt{\frac{g+\sigma-\pi}{\sigma-\pi}}\right) \qquad \left(\frac{g}{\sigma-\pi} > -1\right)$$
$$\beta = \sqrt{\pi}\left(1 \pm \sqrt{\frac{\sigma-\pi}{g+\sigma-\pi}}\right) \tag{X.12}$$

The parameter ρ in (X.2) is therefore

$$\rho = \pm\sqrt{\frac{g+\sigma-\pi}{\sigma-\pi}}\,. \tag{X.13}$$

The occurrence of the $\pm$ sign means that there exist two different operator solutions which satisfy the equation of motion with the same g and σ. If one considers only n-point functions, there is no difference between the two solutions. We have

$$a = \pi\frac{g}{\sigma-\pi} \qquad b = -\pi\frac{g}{g+\sigma-\pi}\,. \tag{X.14}$$

The solution of Schwinger corresponds to $\sigma=0$:

$$\alpha = \pi(1 \pm \sqrt{1-g/\pi}) \qquad (g < \pi)$$
$$\beta = \pi\left(1 \pm \sqrt{\frac{1}{1-g/\pi}}\right) \tag{X.15}$$
$$a = -g \qquad b = \frac{g}{1-g/\pi}$$

Johnson defined the current as

$$\frac{1}{4}\lim \sum_{\epsilon,\tilde{\epsilon}} e^{i(\alpha-\beta)^2 D^-(\epsilon)}\left\{\bar{\phi}(x+\epsilon)\gamma^\mu\phi(x) - \gamma^\mu\phi(x+\epsilon)\bar{\phi}(x)\right\} \qquad \tilde{\epsilon}^\mu \equiv \epsilon^{\mu\nu}\epsilon_\nu\,.$$

This leads to

$$\alpha = \sqrt{\pi}\left(1 \pm \sqrt{\frac{1 - g/2\pi}{1 + g/2\pi}}\right) \qquad \beta = \sqrt{\pi}\left(1 \pm \sqrt{\frac{1 + g/2\pi}{1 - g/2\pi}}\right)$$

$$a = -\frac{g}{1 - g/2\pi} \qquad b = \frac{g}{1 + g/2\pi} . \tag{X.16}$$

The Johnson current can be obtained with our definition with $\sigma = -g/2$. It should be noted that these apparently different solutions are only different parametrizations of the same family of solutions defined by (X.2). The Johnson solution results from the Schwinger solution by the replacement

$$g \to \frac{g}{1 + g/2\pi} .$$

The coupling constant is undetermined in the sense that any value can be produced with an appropriate current definition. Thus, being given two solutions, not even the statement that the coupling constant of the first one is larger than of the second one is an invariant one. We arrive at the conclusion that the coupling constant should not be defined in terms of the equation of motion.

References

1. A. S. Wightman, Lectures at the Summer School of Theoretical Physics, Cargèse (1964).
2. W. Thirring, Annals of Physics 3, 91 (1958).
3. V. Glaser, Nuovo Cimento 9, 990 (1958).
4. K. Johnson, Nuovo Cimento 20, 773 (1961).
5. F. Scarf, J. Wess, Nuovo Cimento 26, 150 (1952).
6. B. Schroer, Fortschr. der Physik 11, 1(1963).
7. B. Klaiber, Helv. Phys. Acta. 37, 554 (1964).
8. J. Schwinger, Phys. Rev. Letters 3, 296 (1959).
9. A current definition which interpolates the Johnson and Schwinger currents has been given in C. R. Hagen, IAEA preprint (1967). We use a different definition.
10. M. A. Neumark, Lineare Darstellungen der Lorentzgruppe (Berlin 1963), Kap. III, Sec. 15.

GROUP THEORETIC APPROACH TO QUANTUM FIELDS[†‡]

P. T. Matthews
Imperial College
London, England

Notes by J. G. Wray

Contents

* * *

I. Introduction

We wish to set up a formalism for relativistic quantum systems which can incorporate simply internal symmetries such as SU(2) or SU(3). The latter are normally treated in an explicitly group theoretic

---

[†]Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.
[‡]The research reported in this document has been sponsored in part by the Air Force Office of Scientific Research OAR through the European Office Aerospace Research, United States Air Force.

manner. The relativistic invariance has traditionally been developed through Lagrangian field theory. The two approaches appear to have little in common and this has lead to some confusion in attempts to combine internal and external symmetries into larger groups. Our first objective is to consider relativistic invariance from a purely group theoretic point of view, closely parallel to the usual treatment of SU(2) and SU(3).

One of the most important products of Lagrangian quantum field theory is the notion of causality, which leads to anti-particles, the relation between spin and statistics, the substitution law and the C.T.P. theorem. These appear to be experimentally correct. Thus, although we do not require Lagrangians, or even equations of motion, we require causality.

From the two requirements of relativistic invariance and causality, we arrive at local causal fields from which we construct S-matrix elements. The formalism can be trivially extended to include internal symmetries such as SU(2) or SU(3). It provides a framework in which one can understand, criticize and hopefully develop a great deal of work done particularly in the last three years on the combination of internal and external symmetries, and the use of non-compact "dynamical" groups.

In this connection we are led to consider generalised infinite component fields which are unitary representations of the homogeneous Lorentz group. We show that for such fields there is a complete breakdown of the Pauli Theorem, which for finite fields establishes, through causality, the experimentally observed connection between spin and statistics.

II. Groups

A. Definitions.

In this section we summarise the rudiments of group theory, particularly Lie groups and their representations. Theorems will be stated without proof but can be found in standard group theory texts some of which are listed in the references.[1),2),3)]

We begin with the definition of a group. Elements $a,b\ldots$ define a set, $\{a,b,\ldots c\}$, which becomes a group, G, if we impose the conditions

(i) Product: $ab=c \in G$ (i.e., contained in G) for $a,b, \in G$.
(ii) Associativity: $(ab)c = a(bc)$,
(iii) Unit element: There exists $e \in G$ such that $ae = ea$ for every $a \in G$. (II. 1)
(iv) Inverse: To each $a \in G$ there exists $a^{-1} \in G$ such that $aa^{-1} = a^{-1}a = e$.

Further we have the following definitions. A group is abelian if ab = ba. Discrete groups have a discrete set of elements (e.g., reflections). Continuous groups can be defined in terms of continuous, ϵ^i (e.g., the rotation group where the parameters can be taken as the Euler angles). The order of a continuous group is the number, N, of essential parameters, ϵ^i. A continuous group is finite if N is finite. A continuous group is compact if the range of the ϵ^i is finite (e.g., rotations), and non-compact if the range is infinite (e.g., displacements, Lorentz group (£)). A subgroup, S, is a subset, $S \subset G$, is such that for any $s \in S$, $asa^{-1} \in S$ for every $a \in G$. A group is simple if it contains no invariant subgroup. A group is semi-simple if it contains no abelian invariant subgroup. (The Lorentz group is semi-simple, but the Poincare group—or inhomogeneous Lorentz group—is not).

Representations of continuous groups by linear transformations entail the association of each element of the group with a matrix,

$$a: \quad \hat{M}(a) \quad \text{for every} \quad a \in G, \tag{II.2}$$

such that the group properties, (II.1), are preserved by the matrices,

$$\hat{M}(a)\,\hat{M}(b) = \hat{M}(c) \quad \text{if} \quad ab = c,$$

$$\hat{M}(a^{-1}) = \hat{M}^{-1}(a), \tag{II.3}$$

$$\hat{M}(e) = \hat{1}.$$

Suppose the vector space in question is spanned by the vectors $|\xi\rangle$ such that each element of the group induces a transformation

$$a: \quad |\xi\rangle \rightarrow |\xi'\rangle = \hat{M}(a)\,|\xi\rangle. \tag{II.4}$$

The adjoint space, spanned by $\langle\xi|$, transforms contravariantly such that each $a \in G$ induces the transformation

$$a: \quad \langle\xi| \rightarrow \langle\xi'| = \langle\xi|\hat{M}^{-1}(a). \tag{II.5}$$

A representation is said to be reducible if the matrices can be block diagonalized, e.g.,

$$\hat{M}(a) = \left(\begin{array}{c|c} \hat{A}(a) & \hat{0} \\ \hline \hat{0} & \hat{B}(a) \end{array}\right) \quad \text{for every} \quad a \in G. \tag{II.6}$$

An irreducible representation is one which cannot be reduced. In

terms of the vectors—rather than the matrices—an irreducible representation (IR) is a complete orthonormal set of vectors that spans the representation and transforms into itself under the $\hat{M}(a)$'s. (For example, for the rotation group the vectors $|j,j_3\rangle$, $j \geqq j_3 \geqq -j$, form such a set.) The dimension of the IR characterized by j is $(2j+1)$. It will turn out that the specification and dimensionality of the IR's of the relevant groups is of vital importance in physical applications.

The following theorems are stated without proof.

Theorem 1. A finite representation of a compact group can always be made unitary by a similarity transformation. That is, if the $\hat{M}(a)$'s are $n \times n$ matrices with n finite then we can choose $\hat{M}^{-1}(a) = \hat{M}^{+}(a)$.

Theorem 2. Any unitary representation of a non-compact group is either one or infinite dimensional. That is, if $\hat{M}^{-1}(a) = \hat{M}^{+}(a)$, then the matrices are 1×1 or $\infty \times \infty$. (The exceptional finite case is called the trivial representation.)

B. Lie Groups.

A continuous group is a Lie Group if

(i) given real ϵ_i $(i = 1, \ldots N)$ there exists an $\bar{\epsilon}$ such that

$$\hat{M}(\epsilon)\ \hat{M}(\bar{\epsilon}) = \hat{M}(\bar{\epsilon})\ \hat{M}(\epsilon) = \hat{1}$$

and $\bar{\epsilon} = f(\epsilon)$ is analytic (Note: $\hat{M}(\epsilon) \equiv \hat{M}(\epsilon_1 \ldots \epsilon_N)$ for N the order of the continuous group.); and

(ii) for $\hat{M}(\epsilon) = \hat{M}(\epsilon')\,\hat{M}(\epsilon'')$ then $\epsilon_k = \phi_k(\epsilon_1' \ldots \epsilon_N' \epsilon_1'' \ldots \epsilon_N'')$ is analytic.

Examples of the classical simple Lie Groups are the following:

Group		Representation
1. Compact.		
$O(3)$ - Rotation Group	:	All linear transformations that preserve $x_1^2 + x_2^2 + x_3^2$. (Orthogonal 3×3 matrices.)
$O(n)$	:	All linear transformations that preserve $x_1^2 + x_2^2 + \ldots + x_n^2$. (Orthogonal $n \times n$ matrices.)
$U(n)$	:	All linear transformations that preserve $\lvert\xi_1\rvert^2 + \lvert\xi_2\rvert^2 + \ldots + \lvert\xi_n\rvert^2$. (Unitary $n \times n$ matrices.)
$S_p(n)$	:	All linear transformations that preserve $\tilde{X}GY = \sum_{i,j=1}^{n} g_{ij} x_i y_j$

where the metric G satisfies $\tilde{G} = -G$. (Unimodular matrices, A, satisfying $\tilde{A}GA = G$.)

2. Non-compact.

$O(3,1)$ - Lorentz group	:	All linear transformations that preserve $x_1^2 + x_2^2 + x^2 - x^2$.
$O(n,m)$	:	All linear transformations that preserve $x_1^2 + x_2^2 + \ldots + x_n^2 - x_{n+1}^2 - x_{n+2}^2 - \cdots - x_{n+m}^2$.
$U(n,m)$	:	All linear transformations that preserve $\lvert\xi_1\rvert^2 + \lvert\xi_2\rvert^2 + \ldots + \lvert\xi_n\rvert^2 - \lvert\xi_{n+1}\rvert^2 - \ldots - \lvert\xi_{n+m}\rvert^2$.
$U(2,2)$	:	All linear transformations that preserve $\bar{\Psi}\psi = \psi^+ \gamma_0 \psi = \lvert\psi_1\rvert^2 + \lvert\psi_2\rvert^2 - \lvert\psi_3\rvert^2 - \lvert\psi_4\rvert^2$.

For a Lie group with N parameters, ϵ_i $(i = 1, \ldots N)$, the matrices $M(\epsilon)$ can be written

$$\hat{M}(\epsilon) = e^{i\epsilon^i \hat{F}_i}, \tag{II.7}$$

where the $\hat{F}^i$ are the infinitesimal generators. If the representation is unitary then

$$\hat{M}^{-1}(\epsilon) = \hat{M}^{\dagger}(\epsilon) \quad \text{and} \quad \hat{F}^i = \hat{F}^{i\dagger}. \tag{II.8}$$

For example, for the rotation group

$$\hat{R}(\epsilon) = e^{i\hat{J}^i \epsilon_i}, \qquad i = 1,2,3. \tag{II.9}$$

The unimodular condition, $\det \hat{M}(\epsilon) = 1$, implies

$$\mathrm{Tr}[\hat{F}_i] = 0 \quad i = 1\ 2, \ldots N. \tag{II.10}$$

If the groups listed above are restricted to be unimodular they are distinguished by an S; e.g., SO(n), SU(n). For an infinitesimal transformation we have

$$\hat{M}(\epsilon) = \hat{1} + i\epsilon_i \hat{F}^i + O(\epsilon^2). \tag{II.11}$$

The local structure of the group is defined by the algebra of infinitesimal generators

$$[\hat{F}_i, \hat{F}_j] = i f_{ij}^k \hat{F}_k, \tag{II.12}$$

where the f_{ij}^k are the structure constants. Two groups are said to be

locally isomorphic if they have the same algebra (e.g., SU(2) and SO(3)).

The rank, r, is the number of $\hat{F}_i$'s that can be simultaneously diagonalized; e.g., for the rotation group r=1 and of the three $\hat{J}$'s only $\hat{J}_3$, say, can be taken diagonal.

The Casimir operators are constructed from the $\hat{F}^i$ and commute with all of them. A standard method of construction is to take all the forms,

$$\hat{C}^{(n)} = f_{i_1 j_1}^{\ \ j_2} f_{i_2 j_2}^{\ \ j_3} \cdots f_{i_{n-1} j_{n-1}}^{\ \ j_n} f_{i_n j_n}^{\ \ j_1} \hat{F}^{i_1} \cdots \hat{F}^{i_n}. \qquad \text{(II. 13)}$$

Within an IR the Casimir operators are proportional to the unit operator (Schur's Lemma); e.g., for the rotation group the Casimir operator is

$$\hat{J}^2 = \hat{J}_1^2 + \hat{J}_2^2 + \hat{J}_3^2, \qquad \text{(II. 14)}$$

and its matrix representation in a given IR,

$$\hat{J}^2 = j(j+1)\hat{1}. \qquad \text{(II. 15)}$$

Thus the Casimir operators provide convenient labels to specify the IR.

For simple Lie groups we have the following theorems.

Theorem 3. The number of independent $C^{(n)}$ is equal to the rank r.

Theorem 4. The number of labels needed to specify an IR is also the rank r.

To specify a component of a representation on additional $\frac{1}{2}(N-r)$ parameters are needed. These can be provided by the eigenvalues of the diagonal $\hat{F}^i$'s and/or Casimir operators of subgroups. Thus, for example, the components (vectors) of an IR may be the state $|c_i, f_j\rangle$ where

$$\begin{aligned} \hat{c}_i |c_i, f_j\rangle &= c_i |c_i, f_j\rangle, \\ \hat{F}_j |c_i, f_j\rangle &= f_j |c_i, f_j\rangle, \end{aligned} \qquad \text{(II. 16)}$$

where the c's label the IR and the f's distinguish the components. For example, for the rotation group, r=1, N=3 and $\frac{1}{2}(N-r)=1$. The representation is specified by the eigenvalue of $\hat{J}^2$ and the components by the eigenvalue of $\hat{J}_3$, so that the vectors of an IR are given by $|j, j_3\rangle$ with

$$\hat{J}^2 |j,j_3\rangle = j(j+1)|j,j_3\rangle,$$
$$\hat{J}_3 |j,j_3\rangle = j_3 |j,j_3\rangle. \qquad \text{(II. 17)}$$

The elements of the algebra have the matrices,

$$\langle j,j_3|\hat{J}^i|j,j_3'\rangle, \qquad i = 1,2,3.$$

III. Poincaré Invariance

A. Representations and Quantum Mechanics.

In the quantum mechanical description of a physical system the set of observables are represented by hermitian operators $\hat{\alpha}$, which can have the values α, specified by the eigenvalue equation

$$\hat{\alpha}|\alpha = \alpha|\alpha\rangle \qquad (\alpha \text{ real}), \qquad \text{(III. 1)}$$

for physical states. If $[\hat{\alpha}_i,\hat{\alpha}_j]=0$ then $\hat{\alpha}_i$ and $\hat{\alpha}_j$ are simultaneously observable and have simultaneous eigenstates. A unique quantum state is defined by the eigenvalues of a complete set of commuting operators $\hat{\beta}_1 \ldots \hat{\beta}_M$. The states

$$|\beta_1 \ldots \beta_M\rangle \equiv |\ell\rangle \qquad \text{(III. 2)}$$

are: (i) orthonormal,

$$\langle \ell|m\rangle = \delta_{\ell m} \quad (\delta(\ell-m) \text{ for continuum}), \qquad \text{(III. 3)}$$

and (ii) complete,

$$\sum_{\ell} |\ell\rangle\langle \ell| = \hat{1}. \qquad \text{(III. 4)}$$

A transformation of coordinate base, such as a displacement of the origin or a rotation of the axes, produces a change in the states and observables which is induced by an operator of the corresponding group:

$$|\ell\rangle \to \hat{U}_\epsilon|\ell\rangle, \quad \hat{\alpha}_i \to \hat{U}_\epsilon \hat{\alpha}_i \hat{U}_\epsilon^{-1}. \qquad \text{(III. 5)}$$

To preserve (III. 1)-(III. 4), $\hat{U}_\epsilon$ must be unitary,

$$\langle \ell|\hat{U}_\epsilon^{-1}|m\rangle = \langle \ell|\hat{U}_\epsilon^{\dagger}|m\rangle, \qquad \text{(III. 6)}$$

and the representation $|\ell\rangle$ is to be called a physical representation. If the group is continuous then

$$\hat{U}_\epsilon = e^{i\epsilon^j \hat{F}_j}, \tag{III. 7}$$

where $\hat{F}_i$ are the infinitesimal generators of the group. In the physical representations

$$\hat{U}_\epsilon^\dagger = \hat{U}_\epsilon^{-1} \quad \text{and} \quad \hat{F}_i^\dagger = \hat{F}_i. \tag{III. 8}$$

For a discrete group, such as space reflection or particle-antiparticle conjugation, a double application takes us back to the starting point, so we have

$$\hat{U}^2 |\ell\rangle = |\ell\rangle. \tag{III. 9}$$

Hence for eigenstates of $\hat{U}$

$$\hat{U}|\ell\rangle = \pm|\ell\rangle. \tag{III. 10}$$

By considering the diagonal form with eigenvalues ± 1, we see

$$\hat{U} = \hat{U}^{-1} = \hat{U}^\dagger. \tag{III. 11}$$

So $\hat{U}$ is both unitary and hermitian and therefore an observable of the system.

The S-matrix, which specifies the scattering amplitudes of the system, is $\langle m|\hat{S}|\ell\rangle$. Under a change of coordinate bases, induced by $\hat{U}_\epsilon$, both states and operators transform, and one obtains the trivial result

$$\langle m|S|\ell\rangle \to \langle m|U^{-1}(USU^{-1})U|\ell\rangle = \langle m|S|\ell\rangle. \tag{III. 12}$$

However, if S is invariant with respect to $\hat{U}_\epsilon$, we have the nontrivial statement that

$$\langle m|U_\epsilon^{-1} S U_\epsilon|\ell\rangle = \langle m|S|\ell\rangle, \tag{III. 13}$$

or

$$[U_\epsilon, S] = 0, \tag{III. 14}$$

the algebraic statement of invariance. If $\hat{U}_\epsilon$ is a continuous transformation, then we also have for the generators

$$[\hat{S}, \hat{F}_i] = 0, \tag{III. 15}$$

and a fortiori for the Casimir operators

$$[\hat{S}, \hat{C}^{(n)}] = 0.$$

Thus the eigenvalues of the hermitian operators, $\hat{F}_i$ and $\hat{C}^{(n)}$ for continuous groups, and $\hat{U}$ for discrete groups, can be measured simultaneously with $\hat{S}$ when $\hat{S}$ is invariant under the corresponding group of transformations. We can therefore include these operators in the characterization of the state $|\ell\rangle$:

$$\hat{F}|\ell\rangle = f_\ell|\ell\rangle. \tag{III. 16}$$

Taking an expectation value of (III. 15)

$$\langle\ell|[\hat{F},\hat{S}]|m\rangle = 0 \rightarrow (f_\ell - f_m)\ \langle\ell|\hat{S}|m\rangle = 0. \tag{III. 17}$$

Equation (III. 17) implies that either $\langle\ell|\hat{S}|m\rangle = 0$ (i.e., the $\ell \rightarrow m$ transition is forbidden) or $f_\ell = f_m$. We have thus established the important result that the observables $\hat{F}$—generators of groups with respect to which $\hat{S}$ is invariant—are conserved in allowed transitions. Clearly the same argument applies to the Casimir operators $\hat{C}^{(n)}$, and for discrete groups, to $\hat{U}$ (parity operators).

For example the generator of displacements is the hermitian operator for the momentum $\hat{D}_x = \hat{P}_x/\hbar$ where

$$\hat{U}_{\vec{x}} = e^{i\vec{x}\cdot\hat{\vec{D}}_x} = e^{(i/\hbar)\vec{x}\cdot\hat{\vec{P}}} \tag{III. 18}$$

and momentum is conserved for any system invariant under displacements.

Further, we see that the operators which label the IR's and components of the groups with respect to which S is invariant are just the simultaneously observable, (conserved) physical properties of the system described by S. So these operators satisfy the requirements of a commuting set and supplemented, if necessary, for completeness, provide a possible (and usually extremely convenient) choice for the operators $\hat{\beta}_1 \ldots \hat{\beta}_m$, which define a physical representation of the states of the system. Hence the extreme importance of the IR's. (For those more familiar with quantum mechanics than group theory, it is often convenient to invert this above argument and determine the irreducible representations of the group by finding a complete commuting set from among the infinitesimal generators and Casimir operators.)

B. The Poincare Group.

In discussing the Poincare group $\mathcal{P}$ the following conventions will be used.

$$\begin{aligned}
&\hbar = c = 1,\\
&x_\mu = (x_0;\vec{x}); \quad \mu = 0,1,2,3,\\
&x^\mu = g^{\mu\nu}x_\nu = (x_0,-\vec{x}) \quad \text{for} \quad g = (1,-1,-1,-1),\\
&\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{\partial}{\partial x_0}; -\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}, -\frac{\partial}{\partial x_3}\right)\\
&a\cdot b = a^\mu b_\mu = a_0 b_0 - \vec{a}\cdot\vec{b},\\
&\epsilon_{0123} = +1, \quad \epsilon^{0123} = -1,
\end{aligned} \tag{III.19}$$

so, e.g.,

$$\epsilon_{\mu\nu\rho\pi}\epsilon^{\mu\nu\rho'\pi'} = -2!\left[\delta_\rho^{\ \rho'}\delta_\pi^{\ \pi'} - \delta_\rho^{\ \pi'}\delta_\pi^{\ \rho'}\right].$$

$\mathcal{P}$ can be represented by the group of linear transformations that preserve the form

$$(x-y)^\mu(x-y)_\mu = (x_0-y_0)(x_0-y_0) - (\vec{x}-\vec{y})\cdot(\vec{x}-\vec{y}). \tag{III.20}$$

More precisely

$$\mathcal{P} = T(4) \sim \mathcal{L}, \tag{III.21}$$

the semi-direct product of translations in four dimension and the Lorentz group. A transformation under $\mathcal{P}$ parameterised by $(a^\mu, \eta^{\mu\nu})$

$$x_\mu \to x'_\mu = x_\mu + a_\mu + \Lambda_{\mu\nu}(\eta)x^\nu \cong (1+\delta)x, \tag{III.22}$$

is generated by (see (IV.24) for the relation between Λ and η)

$$\hat{U}(a,\eta) = e^{i(a^\mu\hat{P}_\mu + \frac{1}{2}\eta^{\mu\nu}\hat{J}_{\mu\nu})}, \tag{III.23}$$

where

$$\eta_{\mu\nu} = -\eta_{\nu\mu}.$$

The algebra of the generators of $\mathcal{P}$ can be obtained by considering how a scalar field, $\phi(x)$, behaves under an infinitesimal transformation of variables

$$x \to x' = Ax = (1+\delta)x. \tag{III.24}$$

For $\phi(x)$ a scalar invariant with respect to $\mathcal{P}$, the transformed field is

$$\phi'(x') = \phi(x) = \phi(A^{-1}x') = \phi[(1-\delta)x']. \tag{III.25}$$

Further we have the generators, $\hat{f}(x_\mu, \partial_\mu)$, which are defined by the property that

$$(1 + i\delta\hat{f})\,\phi(x) \equiv \phi'(x) = \phi[(1-\delta)x]. \tag{III.26}$$

In the final equality we have used (III.25) dropping the dash. For displacements, for example,

$$\begin{aligned}(1 + ia^\mu \hat{f}_\mu + \ldots)\,\phi(x) &= \phi(x_\mu - a_\mu) \\ &= \phi(x) - a^\mu \partial_\mu\, \phi(x) + \ldots,\end{aligned} \tag{III.27}$$

which implies

$$\hat{f}_\mu = i\partial_\mu \equiv \frac{\hat{P}_\mu}{\hbar}\ (\hbar = 1).$$

The commutators of $\hat{f}_\mu$ (and hence of $\hat{P}_\mu$) can then be found trivially from the properties of partial derivatives, i.e.,

$$[\hat{f}_\mu, \hat{f}_\nu] = 0 \quad \text{or} \quad [\hat{P}_\mu, \hat{P}_\nu] = 0. \tag{III.28}$$

Similarly for Lorentz rotations

$$\hat{f}_{\mu\nu} = i(x_\mu \partial_\nu - x_\nu \partial_\mu) \equiv \hat{J}_{\mu\nu}, \tag{III.29}$$

and the resultant algebra of $\mathcal{P}$ is

$$\left.\begin{aligned}[\hat{P}_\mu, \hat{P}_\nu] &= 0, \\ [\hat{P}_\lambda, \hat{J}_{\mu\nu}] &= i[g_{\lambda\mu}\hat{P}_\nu - g_{\lambda\nu}\hat{P}_\mu],\end{aligned}\right\}$$

$$[\hat{J}_{\mu\nu},\hat{J}_{\rho\pi}] = i[g_{\mu\rho}\hat{J}_{\nu\pi} + g_{\nu\pi}\hat{J}_{\mu\rho} - g_{\mu\pi}\hat{J}_{\nu\rho} - g_{\nu\rho}\hat{J}_{\mu\pi}]. \quad \text{(III. 30)}$$

The second of these equations is equivalent to the statement that under Lorentz transformations $\hat{P}_\mu$ transforms, like x_μ, as a four vector.

The IR's of $\mathcal{P}$ are to be labelled by a complete commuting set of operators constructed from $\hat{P}_\mu$ and $\hat{J}_{\mu\nu}$. For momentum we have the Casimir operator

$$\hat{P}^2 = \hat{P}^\mu\hat{P}_\mu = t. \quad \text{(III. 31)}$$

Further, since the $\hat{P}_\mu$'s commute with each other and with $\hat{P}^2$, we can specify their eigenvalues, p_μ, simultaneously with t. Additional operators available for specifying the IR's are those components of $\hat{J}_{\mu\nu}$ which do not alter p_μ. These, by definition, are the generators of the little group (LG) of p_μ.

A covariant specification of the generators of the LG are the three independent components of the Pauli-Lubanski vector

$$\hat{W}^\mu \equiv -\tfrac{1}{2}\epsilon^{\mu\nu\lambda\rho}\hat{P}_\nu\hat{J}_{\lambda\rho}, \quad \text{(III. 32)}$$

satisfying

$$\hat{P}^\mu\hat{W}_\mu = 0 \quad \text{(III. 33)}$$

and

$$[\hat{P}^\mu,\hat{W}_\nu] = 0, \quad \text{(III. 34)}$$

$$[\hat{W}_\mu,\hat{W}_\nu] = i\epsilon_{\mu\nu\pi\rho}W^\pi P^\rho. \quad \text{(III. 35)}$$

The various possibilities for the LG are the following.

1) $t = m^2 > 0$. (t corresponds to the energy variable in two-body scattering, or the mass of a single particle.) In the C. M. frame

$$p_\mu = (m; \vec{O}),$$

and

$$\hat{W}_\mu = m\left(O; \hat{J}_{23}, \hat{J}_{31}, \hat{J}_{12}\right), \quad \text{(III. 36)}$$

and the LG is O(3).

2) $t = -q^2 < 0$ (momentum transfer in two-body scattering). In the brick-wall frame, p_μ can be cast in the form $p_\mu = (O;O,O,q)$ and

$$\hat{W}^{\mu} = q\left(\hat{J}_{12}; \hat{J}_{20}, \hat{J}_{03}, O\right), \tag{III. 37}$$

and the LG is O(2,1). The algebra of the LG is

$$\begin{aligned} [\hat{J}_{12}, \hat{J}_{20}] &= +i\hat{J}_{01}, \\ [\hat{J}_{20}, \hat{J}_{01}] &= -i\hat{J}_{12}, \\ [\hat{J}_{01}, \hat{J}_{12}] &= +i\hat{J}_{20}. \end{aligned} \tag{III. 38}$$

3) $t = 0$ (the light-like case for a single particle). One can choose a frame so that p_{μ} (w, o, o, w).

Now with p_{μ} null,

$$\hat{W}^{\mu} = w\left(-\hat{J}_{12}; -\hat{\Pi}_2, \hat{\Pi}_1, \hat{J}_{12}\right) \tag{III. 39}$$

where

$$\begin{aligned} \hat{\Pi}_1 &= \hat{J}_{10} - \hat{J}_{13}, \\ \hat{\Pi}_2 &= \hat{J}_{20} - \hat{J}_{23}, \end{aligned} \tag{III. 40}$$

$$\left[\hat{J}_{12}, \hat{\Pi}_1\right] = i\hat{\Pi}_2, \left[\hat{J}_{12}, \hat{\Pi}_2\right] = -i\hat{\Pi}_1, \left[\hat{\Pi}_1, \hat{\Pi}_2\right] = 0,$$

and the LG is $T(2) \sim O(2)$—the Euclidean group in two dimensions.

4) $p_{\mu} = 0$. (Momentum transfer in elastic forward scattering.) For this case, LG = £. This group is discussed, in a different connection, in more detail below.

C. The Single Particle States.

Single particle states of non-zero rest mass belong to Case 1) and can be labelled utilizing the IR's of that LG. The Casimir operators for the LG are

$$\hat{W}^2 = \hat{W}^{\mu}\hat{W}_{\mu} = -\tfrac{1}{2}\hat{P}^{\lambda}\hat{P}_{\lambda}\hat{J}^{\mu\nu}\hat{J}_{\mu\nu} + \hat{J}^{\rho\mu}\hat{J}_{\rho\nu}\hat{P}_{\mu}\hat{P}^{\nu}, \tag{III. 41}$$

and

$$\frac{\hat{\vec{J}} \cdot \hat{\vec{P}}}{|\hat{\vec{P}}|} = \lambda \equiv \text{helicity}. \tag{III. 42}$$

In the rest frame, $\hat{W}^2$ becomes

$$\hat{W}^2 = -m^2\hat{J}^2, \tag{III. 43}$$

which, for single particles, has eigenvalues

$$-m^2 s(s+1), \tag{III. 44}$$

where s is the spin. Instead of the helicity, λ, we will use s_3 which is related to λ by a pure rotation. Thus an IR of the Poincare group is specified by m^2 and s, while the components of an IR are labelled by $\vec{p}$ and s_3. A single particle state is, therefore,

$$|m^2 s; \vec{p}, s_3\rangle \equiv |p_\mu, s, s_3\rangle. \tag{III. 45}$$

The non-compactness of the group manifests itself through the infinite range of $\vec{p}$. These states are orthogonal and can be covariantly normalised:

$$\langle p_\mu, s, s_3 | p'_\mu, s', s'_3\rangle \Delta^+(p) = (2\pi)^4 \delta^{(4)}(p-p')\delta_{ss'}\delta_{s_3 s'_3}, \tag{III. 46}$$

where

$$\Delta^+(p) = 2\pi\theta(p_o)\delta(p^2-m^2).$$

Parity can be introduced through the operator, $\hat{R}$, satisfying the physically motivated commutation relations

$$\begin{aligned} [\hat{R}, \hat{P}_o] &= 0, & \{\hat{R}, \hat{P}_i\} &= 0, \\ [\hat{R}, \hat{J}_{ij}] &= 0, & \{\hat{R}, \hat{J}_{oi}\} &= 0, \\ [\hat{R}, \hat{P}^\mu \hat{P}_\mu] &= 0, & [\hat{R}, \hat{W}^\mu \hat{W}_\mu] &= 0. \end{aligned} \tag{III. 47}$$

Now, since $\hat{R}$ commutes with the other operators labelling the rest frame state and is a discrete transformation such that its unitary representation is also hermitian, it can be included as a state label through its eigenvalue, π.

$$\text{State:} \quad |m^2, s, \pi; \vec{p} s_3\rangle.$$

Additional labels can be incorporated provided their operators commute with those operators labelling the state. Thus, if the system is invariant under SU(3) each particle is also a component of some IR of SU(3), the multiplet being specified by two Casimir operators $C^{(1)}$

and $C^{(2)}$, and the component by the diagonal generators i_3, y and the Casimir operator of the isotopic subgroup i^2:

$$|m^2, s, \pi; \vec{p}, s_3; C^{(1)}, C^{(2)}; i^2, i_3, y\rangle.$$

(This is a simple application of Theorem 4 et seq. for the case $N=8$, $r=2$.)

D. The Lorentz Group.[4),5)]

Though it is not obvious at this stage it turns out that Poincare invariance of single particle states requires an analysis in terms of the Lorentz group, $\mathfrak{L}$ or $O(3,1)$ and its covering group, $SL(2,C)$. So far $\mathfrak{L}$ has only been shown to be relevant in forward scattering where it was identified as the little group (LG). However, $\mathfrak{L}$ will eventually play a much larger role, so we include here a brief discussion of its representation.

The algebra of $\mathfrak{L}$ is a subalgebra of $\mathfrak{P}$ given in the final equation of (III. 30).

Introducing

$$\hat{J}_{ij} \equiv \epsilon_{ijk}\hat{J}_k \quad \text{and} \quad \hat{J}_{oi} \equiv \hat{K}_i, \tag{III. 48}$$

these commutation relations can be re-written in the form

$$\begin{aligned} [\hat{J}_i, \hat{J}_j] &= i\epsilon_{ijk}\hat{J}_k, \\ [\hat{J}_i, \hat{K}_j] &= i\epsilon_{ijk}\hat{K}_k, \\ [\hat{K}_i, \hat{K}_j] &= -i\epsilon_{ijk}\hat{J}_k. \end{aligned} \tag{III. 49}$$

The most familiar specification of IR's uses as a basis for the algebra

$$\begin{aligned} \hat{L}_i &\equiv \tfrac{1}{2}(\hat{J}_i + i\hat{K}_i), \\ \hat{M}_i &\equiv \tfrac{1}{2}(\hat{J}_i - i\hat{K}_i); \end{aligned} \tag{III. 50}$$

with the commutation relations of $SO(3) \otimes SO(3)$ or $SU(2) \otimes SU(2)$,

$$\begin{aligned} &[\hat{L}_i, \hat{L}_j] = i\epsilon_{ijk}\hat{L}_k, \quad [\hat{M}_i, \hat{M}_j] = i\epsilon_{ijk}\hat{M}_k, \\ &[\hat{L}_i, \hat{M}_j] = 0. \end{aligned} \tag{III. 51}$$

(Thus O(3,1) is locally isomorphic to SL(2,C) with basic generators $\hat{J}_i \sim \hat{\sigma}_i$, $\hat{K}_i \sim i\hat{\sigma}_i$.)

The Casimir operators $\hat{L}^2$ and $\hat{M}^2$ have eigenvalues $\ell(\ell+1)$ and $m(m+1)$, ℓ, m being integer or half integer and providing labels for the I.R. From standard theory of angular momentum, it is clear that components are labelled by (j,j_3) with the range,

$$|\ell - m| \leqq j \leqq \ell + m, \quad -j \leqq j_3 \leqq j. \tag{III.52}$$

Since j and j_3 take on only integer (or half integer) values, these representations are finite dimensional. Since the matrices $\hat{L}$, $\hat{M}$ in these representations are hermitian, the matrices $\hat{K}$ are anti-hermitian, so these finite representations, (ℓ,m), of this non-compact group are non-unitary, in accordance with Theorem 2.

If one is interested also in the (infinite) unitary representations it is more convenient, following Naimark,[4] to specify the eigenvalues of the Casimir operators, and hence the I.R.'s by (k_o,c) where

$$\tfrac{1}{2}\hat{J}_{\mu\nu}\hat{J}^{\mu\nu}|k_o c\rangle = (\hat{J}^2 - \hat{K}^2)|k_o,c\rangle$$

$$= (k_o^2 + c^2 - 1)|k_o,c\rangle,$$

$$-\frac{1}{8}\epsilon^{\mu\nu\pi\rho}J_{\mu\nu}J_{\pi\rho}|k_o,c\rangle = \hat{\vec{J}}\cdot\hat{\vec{K}}|k_o c\rangle = -ik_o c|k_o,c\rangle. \tag{III.53}$$

It is clear from the definition of k_o and c that $(-k_o,-c)$ specify an equivalent representation. By convention we will restrict ourselves to representations having $k_o \geqq 0$.

For any finite or unitary irreducible representation k_o can take on one of the values

$$k_o = 0,\ 1/2,\ 1,\ 3/2,\ \ldots. \tag{III.54}$$

In all cases components of the representation are labelled by (j,j_3) integer (or half integer), satisfying

$$k_o \leqq j, \tag{III.55}$$

$$-j \leqq j_3 \leqq j. \tag{III.56}$$

The two alternative types of representation are distinguished by:

(i) Finite Dimensional (non-unitary).

$$|c| = k_o + n \qquad (n \text{ positive integer}). \tag{III. 57}$$

Then j has the finite range

$$k_o \leqq j \leqq |c| - 1. \tag{III. 58}$$

(ii) Unitary (infinite dimensional).

Either

(a) Principle series

$$c \text{ pure imaginary;} \tag{III. 59}$$

or

(b) Supplementary series

$$k_o = 0, \quad c \text{ real} \quad 0 < c \leqq 1. \tag{III. 60}$$

In both of these latter cases there is no upper limit to j which thus runs over an infinite range of integer (or half-integer) values. This is in accordance with Theorem 2. (Note that the infinite dimensionality of the unitary representations of $\mathcal{L}$ involve this infinite "tower" of spins, whereas for the Poincare Group a particle of single definite spin belongs to a unitary representation because it can have an infinite range of momentum.)

For both finite and unitary representations the states are written

$$|k_o, c; j, j_3\rangle. \tag{III. 61}$$

We introduce the abbreviation

$$\begin{aligned} \tau &\equiv (k_o, c) = (-k_o, -c), \\ \dot{\tau} &\equiv (k_o, -c), \\ \alpha &\equiv (j j_3). \end{aligned} \tag{III. 62}$$

Then if the matrix

$$\langle \tau\alpha | \hat{K}_i | \tau\beta \rangle \equiv \hat{K}_{i\alpha\beta},$$

the matrix

$$\langle \dot{\tau},\alpha|\hat{K}_i|\dot{\tau},\beta\rangle = -(-1)^{[j]+[j']}\hat{K}_{i\alpha\beta}, \tag{III. 63}$$

where

$$[j] = \begin{cases} j & \text{(integers)} \\ j-\frac{1}{2} & (\tfrac{1}{2}\text{ integers}). \end{cases} \tag{III. 64}$$

The parity operator R satisfies

$$\begin{aligned} \hat{R}(\hat{\vec{J}}\cdot\hat{\vec{K}})|\tau\alpha\rangle &= -(\hat{\vec{J}}\cdot\hat{\vec{K}})\hat{R}|\tau\alpha\rangle, \\ \hat{R}(\hat{J}^2-\hat{K}^2)|\tau\alpha\rangle &= (\hat{J}^2-\hat{K}^2)\hat{R}|\tau\alpha\rangle. \end{aligned} \tag{III. 65}$$

Now suppose

$$\hat{R}|\tau\alpha\rangle = c(\tau,\alpha)|\dot{\tau}\alpha\rangle; \tag{III. 66}$$

then consistency with (III. 53) and (III. 65) implies

$$\hat{R}|\tau\alpha\rangle = (-1)^j|\dot{\tau}\alpha\rangle. \tag{III. 67}$$

Thus the parity self-conjugate representations are given by (k_0, O) and (O,c); for example,

$$\begin{aligned} &(O,1), && \text{scalar,} \\ &(\tfrac{1}{2},O), && j = 1/2, 3/2, \ldots \ (\tfrac{1}{2}\text{ integer spin tower}), \\ &(O,\tfrac{1}{2}), && j = 0, 1, \ldots\ldots \ \text{(integer spin tower).} \end{aligned} \tag{III. 68}$$

Otherwise the inclusion of the space reflection operation with $\mathfrak{L}$ requires a reducible representation (τ and $\dot{\tau}$).

IV. Poincare Invariant S-Matrix Elements

A. Statement of the Problem.

Free particle creation-annihilation operators which transform according to IR's of $\mathcal{P}$ can be used to generate the scattering states,

$$\begin{aligned} &\text{1 particle:} && |p_\mu, s\rangle, \\ &\text{n particle:} && |p_\mu^{(1)}s^{(1)}; p_\mu^{(2)}s^{(2)}; \ldots p_\mu^{(n)}s^{(n)}\rangle_{in} \equiv |\ell\rangle_{in}, \end{aligned} \tag{IV. 1}$$

when acting on the $\mathcal{P}$ invariant vacuum $|\rangle_0$. The suffix "in" means by definition a state with plane beams coming into the scattering

region. We assume the vacuum and n-particle states are complete,

$$\sum_{\ell} |\ell\rangle_{\text{in in}}\langle\ell| = \hat{1}. \tag{IV.2}$$

The set of 'out' states are defined analogously as states having plane beams coming out of the scattering region, satisfying completeness

$$\sum_{\ell} |\ell\rangle_{\text{out out}}\langle\ell| = \hat{1}. \tag{IV.3}$$

The S-operator is defined

$$\hat{S} \equiv \sum_{\ell} |\ell\rangle_{\text{in out}}\langle\ell|, \tag{IV.4}$$

so that the S-amplitude is

$${}_{\text{in}}\langle \text{final}|\hat{S}|\text{initial}\rangle_{\text{in}} = {}_{\text{out}}\langle \text{final}|\text{initial}\rangle_{\text{in}}. \tag{IV.5}$$

The construction of a $\mathcal{P}$ invariant $\hat{S}$, i.e. such that

$$[\hat{S},\hat{F}^i] = 0 \quad \text{for } \hat{F} \ (i=1,\ldots 10) \text{ the generators of } \mathcal{P}, \tag{IV.6}$$

entails a particular difficulty that is best illustrated by comparison with invariance with respect to an internal symmetry group. Take the invariant group to be that of I-spin, SU(2), with the basis representation defined by

$$|N_a\rangle = \begin{matrix}|p\rangle \\ |n\rangle\end{matrix}. \tag{IV.7}$$

It is convenient to introduce annihilation operators labelled by the particles on which they operate.

$$\hat{N}_a = \begin{pmatrix}\hat{p} \\ \hat{n}\end{pmatrix}, \quad \hat{\bar{N}}_a = (\hat{\bar{p}},\hat{\bar{n}}),$$

which transform under I-spin rotations parameterised by $\vec{d}$ according to

$$\hat{N}_a \to \left(e^{i(\vec{\tau}/2)\cdot\vec{d}}\right)_a^{a'} \hat{N}_{a'}, \quad \hat{\bar{N}}_a \to \hat{\bar{N}}^{a'}\left(e^{-i(\vec{\tau}/2)\cdot\vec{d}}\right)_{a'}^{a}; \tag{IV.8}$$

and

$$\left(\hat{\Pi}_a^b\right) = \begin{pmatrix} \hat{\pi}^o/\sqrt{2} & \hat{\pi}^+ \\ \hat{\pi}^- & -\hat{\pi}^o/\sqrt{2} \end{pmatrix},$$

which transform according to

$$\hat{\pi}_a^b \to \left(e^{i(\tau/2)\cdot\vec{d}}\right)_a^{a'} \quad \hat{\Pi}_{a'}^{b'}\left(e^{-i(\tau/2)\cdot\vec{d}}\right)_{b'}^{b}. \tag{IV.9}$$

The I-spin invariant vertex part for the $\bar{N}N\Pi$ interaction is

$$\hat{s} = g\hat{\bar{N}}^a\hat{\Pi}_a^b\hat{N}_b = g\left\{(\hat{p}\hat{p} - \hat{n}\hat{n})\frac{\hat{\pi}^o}{\sqrt{2}} + \hat{p}\hat{n}\,\hat{\Pi}^+ + \hat{n}\hat{p}\,\hat{\Pi}^-\right\}. \tag{IV.10}$$

In general I-spin invariant amplitudes are very simply formed by saturating suffixes and it is trivial to check that the exponential factors arising from an arbitrary I-spin transformation on an expression such as (IV.10) then cancel in pairs.

Now consider the same problem for Poincare invariance. For creation-annihilation operators we have

$$\hat{a}(p,s)\rangle_o = 0, \quad \hat{a}^+(p,s)\rangle_o = |p_\mu, s\rangle, \tag{IV.11}$$

with the (anti) commutation relations

$$\left\{\hat{a}(p,s),\ \hat{a}^+(p',s')\right\}_\pm \Delta^+(p) = (2\pi)^4\delta^{(4)}(p-p')\delta_{ss'}. \tag{IV.12}$$

Analogous to Eq. (IV.8), the $\hat{a}$'s transform under $\mathcal{L}$ like

$$p \to p' : \hat{a}(p,s)^\eta \to e^{i\eta^{\mu\nu}\hat{J}_{\mu\nu}}\hat{a}(p,s)e^{-i\eta^{\rho\pi}J_{\rho\pi}} = (e^{i\vec{\theta}\cdot\vec{J}})_s^{s'}\hat{a}(p',s'), \tag{IV.13}$$

where (as will be shown explicitly below—(IV.26)) the spin transformation is a pure rotation, but

$$\vec{\theta} = \vec{\theta}(\eta_{\mu\nu}, p, p'). \tag{IV.14}$$

The difficulty arises from the dependence of $\vec{\theta}$ on the initial momentum, p, in (IV.14). The fact that the momenta associated with different particles at the vertex are different prevents the simple construction of a scalar invariant by a naive saturation of spin indices, in the manner of the I-spin example. The standard solution to this

problem, well-known from Lagrangian field theory, is to replace spin by a related "spinor" variable so that Lorentz transformations on momentum and the new "spinor" variable factorize. From Lagrangian theory we know that the "spinor" labels specify representation of the homogeneous Lorentz group $\mathcal{L}$, e.g.,

$$\bar{\psi}^{\alpha}(\gamma_{\mu})_{\alpha}^{\beta}\,\psi_{\beta}A^{\mu}.$$

From a purely group theoretic point of view, this development is surprising since the physically relevant symmetry is Poincare and not Lorentz, and we have seen that the IR's of the two groups have very different structure. Specifically:

(1) Physical particle states belong to unitary representations of $\mathcal{P}$ which in general are not reduced with respect to the subgroup $\mathcal{L}$. Instead, one specifies m^2 and p and then reduces $\mathcal{P}$ with respect to the LG (little group), e.g., O(3) (spin).

(2) Furthermore, the finite representations of $\mathcal{L}$ appearing in the Lagrangian program are necessarily non-initary, whereas physics is normally concerned only with unitary representations.

B. Auxiliary Group Solution.[6),7),8)]

We will now show how the homogeneous Lorentz representations get into the theory, from a purely group theoretic point of view.

We define the single particle state at rest as

$$|m,s,s_3\rangle \equiv |m,s_3\rangle,$$

with the normalization

$$\langle m,s,s_3|m,s',s_3'\rangle = \delta_{ss'}\delta_{s_3s_3'}. \tag{IV.15}$$

Particles not at rest can be described with the aid of the following theorem.

Theorem. The moving states can be obtained from the rest frame states by the Lorentz transformation

$$|p_{\mu},s_3\rangle = Ne^{-i\vec{\epsilon}(p)\cdot\vec{K}}|m,s_3\rangle, \tag{IV.16}$$

if $\cosh|\epsilon(p)| = p_o/m$ and $\sinh|\vec{\epsilon}(p)| = |\vec{p}|/m$, with

$$\frac{\vec{\epsilon}(p)}{|\vec{\epsilon}(p)|} = \frac{\vec{p}}{|\vec{p}|}. \tag{IV.17}$$

Proof. The proof consists of showing that the right hand side of Eq. (IV.16) is an eigenstate of P_μ with eigenvalue p_μ. Explicitly,

$$\hat{P}_\mu e^{-i\vec{\epsilon}(p)\cdot\vec{K}}|ms_3\rangle = e^{-i\vec{\epsilon}(p)\cdot\hat{\vec{K}}} e^{i\vec{\epsilon}(p)\cdot\hat{\vec{K}}} \hat{P}_\mu e^{-i\vec{\epsilon}(p)\cdot\hat{\vec{K}}}|ms_3\rangle. \quad \text{(IV.19)}$$

Now, take $p=(0,0,p_3)$ so that the relevant factor of (IV.19) becomes

$$e^{-i\epsilon_3\hat{K}_3}\hat{P}_\mu e^{i\epsilon_3\hat{K}_3} = \sum_{n=0}^{\infty} \frac{1}{n!}(-i\epsilon_3)^n\Big[\ldots\Big[\hat{P},\hat{K}_3\Big],\hat{K}_3\Big],\ldots\hat{K}_3\Big]. \quad \text{(IV.20)}$$

By (III.30),

$$\Big[\hat{P}_\mu,\hat{K}_3\Big] = i\left(g_{\mu 0}\hat{P}_3 - g_{\mu 3}\hat{P}_0\right).$$

Hence, with $|p_\mu,s_3\rangle$ defined by (IV.16),

$$\begin{aligned}\hat{P}_\mu|p_\mu,s_3\rangle &= e^{-i\epsilon_3\hat{K}_3}p_\mu|m,s_3\rangle \\ &= p_\mu|p_\mu,s_3\rangle\,.\end{aligned} \quad \text{(IV.21)}$$

N.B. $\vec{\epsilon}(p)$ boosts $|m,s_3\rangle$ to $|p_\mu s_3\rangle$.

Consider an arbitrary Lorentz transformation, $\eta_{0i}(v)\equiv\eta_i(v)$, specified by the velocity, $\vec{v}$, with

$$\cosh\eta = \gamma, \quad \sinh\gamma = \gamma|\vec{v}|, \quad \gamma = (1-\vec{v}^2)^{-\frac{1}{2}} \quad \text{(IV.22)}$$

such that

$$p_\mu \to p'_\mu = \Lambda_\mu^{\ \nu}p_\nu. \quad \text{(IV.23)}$$

If $\vec{v}$ is in the z-direction

$$\Lambda_\mu^{\ \nu}(\eta) = \begin{pmatrix}\cosh\eta & 0 & 0 & \sinh\eta\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ \sinh\eta & 0 & 0 & \cosh\eta\end{pmatrix}. \quad \text{(IV.24)}$$

N.B. η takes $p\to p'$.

The Lorentz transformation of single particle states is

$$|p,s_3\rangle \xrightarrow{\eta} |p,s_3\rangle' = e^{-i\vec{\eta}\cdot\vec{K}}|p,s_3\rangle.$$

Using (IV.16) and

$$\vec{\epsilon}' \equiv \vec{\epsilon}(p'),$$

we can write

$$|p,s_3\rangle' = e^{-i\vec{\epsilon}'\cdot\hat{\vec{K}}}e^{i\vec{\epsilon}'\cdot\hat{\vec{K}}}e^{-i\vec{\eta}\cdot\hat{\vec{K}}}e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}}|m,s_3\rangle, \tag{IV.25}$$

$$p' \longleftarrow m \longleftarrow p' \longleftarrow p \longleftarrow m;$$

the effect of the various exponential factors (pure Lorentz transformations) being represented schematically below the equation. We see that the three Lorentz transformations

$$e^{i\vec{\epsilon}'\cdot\hat{\vec{K}}}e^{-i\vec{\eta}\cdot\hat{\vec{K}}}e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} \qquad \text{take} \quad m \to m,$$

and together generate a pure rotation—the Wigner rotation. An effective complete set of states between these and the remaining exponential factor is

$$\sum_{s_3'} |m,s_3'\rangle\langle m,s_3'|$$

yielding

$$\begin{aligned} |p,s_3\rangle' &= \sum_{s_3'} e^{-i\vec{\epsilon}'\cdot\hat{\vec{K}}}|m,s_3'\rangle\langle m,s_3'|e^{i\vec{\epsilon}'\cdot\hat{\vec{K}}}e^{-i\vec{\eta}\cdot\hat{\vec{K}}}e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}}|m,s_3\rangle \\ &\equiv \sum_{s_3'} |p',s_3'\rangle\langle m,s_3'|e^{i\vec{\theta}(\eta,p,p')\cdot\hat{\vec{J}}}|m,s_3\rangle, \end{aligned} \tag{IV.26}$$

which establishes the result anticipated in (IV.13). The Wigner rotation, $e^{i\vec{\theta}\cdot\vec{J}}$, is analogous to the $e^{(i/2)\vec{\tau}\cdot\vec{\Phi}}$ rotation in SU(2). Now, however, simple cancellations of phase factors are not achieved by saturating spin indices because of the momentum dependence of $\theta(\eta,p,p')$.

Rather than using states, it is convenient to work with the Fock space creation (annihilation) operators. These transform

according to

$$a(ps) \xrightarrow{\eta} [a(ps)]' = e^{i\vec{\eta}\cdot\hat{\vec{K}}} a(ps) e^{-i\vec{\eta}\cdot\hat{\vec{K}}}$$

$$= \sum_{s_3'} \langle m, s_3 | e^{i\vec{\epsilon}\cdot\hat{\vec{K}}} e^{i\vec{\eta}\cdot\hat{\vec{K}}} e^{-i\vec{\epsilon}'\cdot\hat{\vec{K}}} | m, s_3' \rangle a(p', s_3'). \quad \text{(IV.27)}$$

This is like the single particle state, with $p \to p'$ and the spin index undergoing a pure momentum dependent (Wigner) rotation.

To avoid the difficulty we introduce an auxiliary group, $\mathcal{A}$, which is chosen to provide an explicit representation for the three separate exponential factors of the Wigner rotation. Any group which contains the generators $\hat{\vec{K}}$ and $\hat{\vec{J}}$ will be sufficient. The simplest possibility for $\mathcal{A}$ is $\mathcal{L}$, the homogeneous Lorentz group (O(3,1) or SL(2,C)) with basic generators $\vec{\sigma}$ and $i\vec{\sigma}$.

Another possibility is to use U(2,2)—(locally isomorphic to O(4,2)). This is the space time part of $\tilde{U}(12)$—with basic generators the sixteen Dirac matrices.

Having chosen $\mathcal{A}$, we must choose a representation $|\tau, \alpha\rangle$, the only requirement being that the j, j_3, contained in α, include the physical spins s, s_3 of the particles to be described. For $\mathcal{A} = \mathcal{L}$ the representations τ can be either unitary or non-unitary. If we desire $|\tau, \alpha\rangle$ to be finite dimensional ($j_{min} = k_0 \leqq j \leqq |c| - 1 = j_{max}$) then the $|\tau, \alpha\rangle$ representation is non-unitary. Thus $\langle \tau, \alpha | \hat{K} | \tau, \beta \rangle$ is not hermitian, and in fact for the examples under consideration will be anti-hermitian,

$$\langle \alpha | \hat{K} | \beta \rangle = -\langle \alpha | \hat{K}^+ | \beta \rangle. \quad \text{(IV.28)}$$

On the other hand if $|\tau, \alpha\rangle$ is infinite dimensional (unitary) then we have a spin tower $k_0 \leqq j$. Most of quantum field theory has been concerned with the first choice.

If parity is to be included we have seen that we must use both $|\tau, \alpha\rangle$ and $|\dot{\tau}, \alpha\rangle$. The most familiar example of an auxiliary group representation is the Dirac spinor label for which $\mathcal{A} = \mathcal{L}$ and $\tau = (1/2, 3/2)$, $\dot{\tau} = (1/2, -3/2)$.

We drop the label τ and specify the representation by $|\alpha\rangle$. For a Lorentz transformation $\hat{a}(p, s_3) \to \hat{a}(p, s_3)'$ where

$$\hat{a}(p, s_3)' = \sum_{s_3'} \langle m, s_3 | \rho \rangle \langle \rho | e^{i\vec{\epsilon}\cdot\hat{\vec{K}}} | \delta \rangle \langle \delta | e^{i\vec{\eta}\cdot\hat{\vec{K}}} | \beta \rangle \langle \beta | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} | \alpha \rangle \langle \alpha | m, s_3' \rangle$$

$$\hat{a}(p', s_3'). \quad \text{(IV.29)}$$

Now define the auxiliary operator

$$\hat{A}_\alpha(p) \equiv \sum_{s_3'} \langle\alpha|e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}}|\beta\rangle\langle\beta|ms_3'\rangle\hat{a}(p,s_3')f(ms) \qquad \text{(IV.30)}$$

$$\equiv \sum_{s_3'} U_\alpha(p)^{s_3'}\hat{a}(p,s_3'). \qquad \text{(IV.31)}$$

This has the important property that under a Lorentz transformation

$$\begin{aligned}\hat{A}_\alpha(p) \xrightarrow{\vec{\eta}} [\hat{A}_\alpha(p)]' &= U_\alpha(p)^{s_3}\hat{a}'(p,s_3)\\ &= U_\alpha(p)^{s_3}\langle m,s_3|\rho\rangle\langle\rho|e^{i\vec{\epsilon}\cdot\hat{\vec{K}}}|\gamma\rangle\langle\gamma|e^{i\vec{\eta}\cdot\hat{\vec{K}}}|\beta\rangle\hat{A}_\beta(p')\\ &= \delta_\alpha^\gamma\langle\gamma|e^{i\vec{\eta}\cdot\hat{\vec{K}}}|\beta\rangle\hat{A}_\beta(p')\\ &= \langle\alpha|e^{i\vec{\eta}\cdot\hat{\vec{K}}}|\beta\rangle\hat{A}_\beta(p'). \qquad \text{(IV.32)}\end{aligned}$$

In (IV.32) the "spinor" indices α,β have replaced (s,s_3) as the spin variables. Further we see that, under Lorentz transformations:

(i) The new "spinor" label (β) undergoes a pure matrix transformation independent of p, parameterised simply by $\vec{\eta}$.

(ii) $U_\alpha(p)^{s_3}$ is a generalised spinor. Equation (IV.31) which expresses the generalised spinor explicitly as the auxiliary group matrix element of a Lorentz transformation (boost) is the crucial link between the field theoretic and group theoretic approaches.

The dual operator $A^\alpha(p)$ is defined as

$$\hat{A}^\alpha(p) = \sum_{s_3'}\hat{a}^+(p,s_3')\bar{f}(ms')\langle m,s_3'|e^{i\vec{\epsilon}\cdot\hat{\vec{K}}}|\alpha\rangle \qquad \text{(IV.33)}$$

(with the sign of the boost opposite to that in (IV.30)).

Under a Lorentz transformation,

$$\hat{A}^\alpha(p) \xrightarrow{\vec{\eta}} [\hat{A}^{\alpha'}(p)] = \hat{A}^\beta(p')\langle\beta|e^{-i\vec{\eta}\cdot\hat{\vec{K}}}|\alpha\rangle. \qquad \text{(IV.34)}$$

If $|\alpha\rangle$ is finite dimensional (non-unitary), then

$$\hat{A}^\alpha(p) \neq (\hat{A}_\alpha(p))^+ \quad \text{but} \quad \hat{A}^\alpha(p) = \bar{\hat{A}}_\alpha(p). \qquad \text{(IV.35)}$$

Also for spinors, defined by

$$\hat{A}^{\alpha}(p) \equiv \hat{a}^{+}(p,s_3)U^{\alpha}(p)_{s_3}; \tag{IV.36}$$

for non-unitary representations,

$$U^{\alpha} \neq (U_{\alpha})^{+}. \tag{IV.37}$$

For example, in the Dirac representation $U^{\alpha}U_{\alpha} = \overline{U}U$ (and not $(U_{\alpha})^{+}U_{\alpha}$) is a scalar.

Under displacements,

$$\begin{aligned} a^{+}(p,s) &\xrightarrow{a} e^{i\hat{P}_{\mu}a^{\mu}}\hat{a}^{+}(p,s_3), \\ \hat{a}(p,s) &\xrightarrow{a} e^{-i\hat{P}_{\mu}a^{\mu}}\hat{a}(p,s_3); \end{aligned} \tag{IV.38}$$

and the auxiliary field operators transform in the same way,

$$\begin{aligned} \hat{A}^{\alpha}(p) &\xrightarrow{a} e^{iP_{\mu}a^{\mu}}\hat{A}^{\alpha}(p), \\ \hat{A}_{\alpha}(p) &\xrightarrow{a} e^{-iP_{\mu}a^{\mu}}\hat{A}_{\alpha}(p). \end{aligned} \tag{IV.39}$$

The introduction of the auxiliary group has solved the problem of separating "spin" and momentum variables in the Lorentz transformation of field operators. Now invariant factors can be constructed in analogy to the isotopic spin by saturating indices.

Thus, for example, for Dirac operators (which we discuss in more detail below), a four-point sub-matrix of the S-matrix could have the form

$$\begin{aligned} \hat{S} \sim \int \hat{A}^{\alpha}(p_1)\,\hat{B}^{\pi}(p_2)\,t\left(\gamma_{\mu};q^{\mu}\right)_{\alpha}^{\beta}\hat{B}_{\beta}(p_3)\hat{A}_{\pi}(p_A) \\ (2\pi)^4\delta\left(p_1+p_2-p_3-p_4\right)\prod_{i=1}^{4}2\pi\theta(p_o)\,\delta\left(p_i^2-m_i^2\right)\frac{d^4p_i}{(2\pi)^4}. \end{aligned} \tag{IV.40}$$

Under displacements, $\hat{S}$ transforms

$$\hat{S} \xrightarrow{a} \int e^{i(p_1^\mu + p_2^\mu - p_3^\mu - p_4^\mu)a_\mu} \hat{A}^\alpha_{(p_1)} \hat{B}^\pi_{(p_2)} t\left(\gamma_\mu; q^\mu\right)^\beta_\alpha$$

$$\hat{B}_\beta(p_3)\hat{A}_\pi(p_4)\delta\left(p_1+p_2-p_3-p_4\right)\prod_{i=1}^{4}\Delta_+(p_i)\frac{d^4p_i}{(2\pi)^4}, \qquad \text{(IV,41)}$$

$$= \hat{S},$$

showing very explicitly how translation invariance leads to energy momentum conservation in $\hat{S}$-matrix elements. Under Lorentz transformations of $\hat{S}$, it is most important that only the auxiliary operators transform, and that these transform as densities, i.e., that for $A^\alpha(p)$, $p \to p'$ in addition to the matrix transformation on α. Consider the case

$$t\left(\gamma_\mu, q^\mu\right)^\beta_\alpha = \left(q^\mu \gamma_\mu\right)^\beta_\alpha, \qquad \text{(IV.42)}$$

$$\hat{S} \xrightarrow{\vec{\eta}} \hat{S}' = \int \hat{A}^{\alpha'}(p_1')\left(e^{-i\vec{\eta}\cdot\vec{K}}\right)^\alpha_{\alpha'} B^{\pi'}(p_2')\left(e^{-i\vec{\eta}\cdot\vec{K}}\right)^\pi_{\pi'}$$

$$\left(q^\mu \gamma_\mu\right)^\beta_\alpha \left(e^{i\vec{\eta}\cdot\vec{K}}\right)^{\beta'}_\beta B_{\beta'}(p_3')\left(e^{i\vec{\eta}\cdot\vec{K}}\right)^\gamma_\pi A_\gamma(p_4') \qquad \text{(IV.43)}$$

$$\delta\left(p_1+p_2-p_3-p_4\right)\prod_{i=1}^{4}\Delta_+(p_i)\frac{d^4p_i}{(2\pi)^4}.$$

Now,

$$e^{-i\vec{\eta}\cdot\vec{K}}\gamma_\mu e^{i\vec{\eta}\cdot\vec{K}} = \Lambda_\mu(\eta)^\nu \gamma_\nu, \qquad \text{(IV.44)}$$

and

$$q^\mu \Lambda_\mu{}^\nu(\eta) \equiv q'^\nu. \qquad \text{(IV.45)}$$

One can now make a change from the undashed to the dashed variables. Since the Jacobian is unity, the invariance of $\hat{S}$ is established.

We will have occasion to consider below pure matrix transformations of the auxiliary operators

$$\hat{A}_\alpha(p) \to \left(e^{i\vec{\eta}\cdot\vec{K}}\right)^\beta_\alpha \hat{A}_\beta(p) \qquad \text{(IV.46)}$$

(p unchanged). For invariance under such transformations it is evident from the example above that the saturation must take place between suffixes on auxiliary operators only, and factors such as $(q^\mu \gamma_\mu)$ are excluded from S-matrix elements.

C. Physical Content of the Auxiliary Field Operators.

The correspondence of the auxiliary operators to physical particles depends on the choice of the constant spinors

$$\langle \beta | m, s_3 \rangle \equiv \langle k_o c, jj_3 | m, s, s_3 \rangle, \tag{IV.47}$$

where (s, s_3) are the physical spins and (j, j_3) are components of the auxiliary representation. For example, take $s = \frac{1}{2}$ and the auxiliary group to be $O(3,1)$. The simplest representation $|\alpha\rangle$ containing spin $\frac{1}{2}$ is

$$|\alpha\rangle = |k_o c; jj_3\rangle = |\tfrac{1}{2}, \pm 3/2; \tfrac{1}{2}, \pm \tfrac{1}{2}\rangle, \tag{IV.48}$$

which are the four Dirac spinor labels. In this representation the boosts are just Dirac matrices,

$$\langle \alpha | \hat{R}_i | \beta \rangle = (\hat{\sigma}_{oi/2})_\alpha^\beta = -(\hat{\sigma}^+_{oi/2})_\alpha^\beta, \tag{IV.49}$$

where

$$\sigma_{\mu\nu} = \tfrac{i}{2}[\gamma_\mu, \gamma_\nu], \{\gamma_\mu \gamma_\nu\} = 2g_{\mu\nu}. \tag{IV.50}$$

We have two physical states in the rest frame, $|m, s_3\rangle (s_3 = \pm \frac{1}{2})$, and there are four auxiliary states, $|\alpha\rangle$. This redundancy can be resolved by additional conditions on $\langle \alpha | m, s_3 \rangle$.

The operator γ_o satisfies

$$[\gamma_o, \sigma_{ij}] = 0, \tag{IV.51}$$

and can therefore be simultaneously specified with spin only in the rest frame. In fact, for this representation γ_o plays the role of the parity operator [i.e., has the correct commutation relations with $\sigma_{\mu\nu}/2 = J_{\mu\nu}$]. Thus specifying the eigenvalue of γ_o specifies intrinsic particle parity.

$$\gamma_o^2 = 1, \tag{IV.52}$$

so γ_o has eigenvalues ± 1. We may impose the conditions

$$\text{(i)} \quad \langle\alpha|\gamma_o|\beta\rangle\langle\beta|m,s_3\rangle = +\langle\alpha|m,s_3\rangle, \quad (s=\tfrac{1}{2}), \tag{IV.53}$$

or

$$\gamma_o|m,s_3\rangle = +|m,s_3\rangle, \tag{IV.54}$$

defining positive intrincis parity for particles; or

$$\text{(ii)} \quad \gamma_o|m,\bar{s}_3\rangle = -|m,\bar{s}_3\rangle, \quad \bar{s} = \tfrac{1}{2}, \tag{IV.55}$$

defining negative intrinsic parity for particles.

Either of the conditions reduces the number of independent auxiliary states to that of the physical states. By boosting the γ_o equations, (IV.54) or (IV.55), we have

$$\langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{K}}\gamma_o e^{i\vec{\epsilon}\cdot\vec{K}}e^{-i\vec{\epsilon}\cdot\vec{K}}|m,(s/\bar{s})\rangle = \pm\langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{K}}|m,(s/\bar{s})\rangle,$$

or

$$\langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{K}}\gamma_o e^{i\vec{\epsilon}\cdot\vec{K}}|\beta\rangle\langle\beta|e^{-i\vec{\epsilon}\cdot\vec{K}}|m,(s/\bar{s})\rangle = \pm\langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{K}}|m,(s/\bar{s})\rangle,$$

or

$$\left(\frac{p^\mu\gamma_\mu}{m}\right)_\alpha^\beta U_\beta(p,(s/\bar{s})) = \pm U_\alpha(p,(s/\bar{s})), \tag{IV.56}$$

i.e., the Dirac equation. Equations of motion (apart from the Klein-Gordan equation) always play this role of eliminating redundant components in the auxiliary representation $|\alpha\rangle$. In this formalism, equations of motion may or may not exist. In any case, they are not needed to calculate cross-sections since we know the spinors explicitly.

Note that in terms of these states we can write the unit matrix in $|\alpha\rangle$ space as

$$\hat{1} = \sum_{s_3} |s,s_3\rangle\langle s,s_3| + |\bar{s},\bar{s}_3\rangle\langle\bar{s},\bar{s}_3|, \tag{IV.57}$$

and

$$\gamma_o = \sum_{s_3} |s,s_3\rangle\langle s,s_3| - |\bar{s},\bar{s}_3\rangle\langle\bar{s}_3,\bar{s}_3|. \tag{IV.58}$$

So far we have started with the physical particles and proceeded to the auxiliary representation, with the aid of supplementary

conditions to eliminate redundancy. However, we could have started with the auxiliary representation and decided from that point of view what particles should be physical. For example, physical particles $|m,s,s_3\rangle$ could be specified to correspond to all of the $|j,j_3\rangle$ in a particular representation $|\alpha\rangle$. For example, for the Dirac case $A_\alpha(p)$ we can have parity doubling with four physical states, $|s,s_3\rangle$ and $|\bar{s},\bar{s}_3\rangle$. In this case there is no redundancy and $\hat{A}_\alpha(p)$ satisfies no equation of motion, but the spin sum over physical spin states is then

$$\sum_{s_3} \langle\alpha|s,s_3\rangle\langle s,s_3|\beta\rangle + \sum_{\bar{s}_3} \langle\alpha|\bar{s},\bar{s}_3\rangle\langle\bar{s},\bar{s}_3|\beta\rangle \equiv \delta_\alpha^\beta. \quad \text{(IV.59)}$$

The sum over spinors is extremely important for calculating cross sections

$$\sigma = \sum_{\text{spin}} SS^+.$$

To illustrate the general features we need below we can consider special cases

$$g_\alpha^{\ \beta} = \sum_{\substack{\text{spins} \\ (s,s_3)}} U_\alpha(p,s,s_3)U^\beta(p,s,s_3). \quad \text{(IV.60)}$$

Formal proofs of the statements made below are to be found in Feldman and Matthews.[9]

(1) $|\alpha\rangle$ finite dimensional (non-unitary); e.g., for the Dirac case with parity doubling we have (using particularly (IV.49))

$$\begin{aligned} g_\alpha^{\ \beta} &= \sum_{s\bar{s}s_3} U_\alpha(p,s,s_3)U_\pi^+(p,s,s_3)(\gamma_o)_\pi^{\ \beta}, \\ &= \sum_{s\bar{s}s_3} \langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{\hat{K}}}|\gamma\rangle\langle\gamma|m,s,s_3\rangle\langle m,s,s_3|\sigma\rangle\langle\sigma|e^{i\vec{\epsilon}\cdot\vec{\hat{K}}^+}|\pi\rangle\langle\pi|\gamma_o|\beta\rangle, \\ &= \langle\alpha|e^{-i\epsilon_i(\sigma_{oi}/2)}e^{+i\epsilon_i(\sigma_{oi}^+/2)}\gamma_o|\beta\rangle, \\ &= \langle\alpha|e^{-i\epsilon_i\sigma_{oi}}\gamma_o|\beta\rangle = \left(\frac{\gamma^\mu p_\mu}{m}\right)_\alpha^{\ \beta}. \end{aligned} \quad \text{(IV.61)}$$

This is a slight variation of the more familiar case with two physical states (no parity doubling) where

$$U_\alpha \overline{U}^\beta = \frac{1}{2m}(\not{p}+m)_\alpha^\beta, \quad [+ \text{ parity}],$$

and

$$U_\alpha \overline{U}^\beta = \frac{1}{2m}(\not{p}-m)_\alpha^\beta, \quad [- \text{ parity}].$$

The important general point is that for finite (non-unitary) representations the $\hat{K}$ matrices are anti-hermitian, so the momentum dependent exponentials in the spinors do not cancel but combine to give a momentum dependent factor like $\not{p}$.

(2) $|\alpha\rangle$ infinite dimensional and unitary. We assume a one to one correspondence between physical states and auxiliary states so that the constant spinors are

$$\langle k_o c; j, j_3 | m, s, s_3 \rangle = \delta_{js}\delta_{j_3 s_3}, \tag{IV.62}$$

implying an infinite spin tower of physical particles. Unitarity of the representation implies

$$\begin{aligned} \langle\alpha|K_i|\beta\rangle &= \langle\alpha|K_i^+|\beta\rangle, \\ U^\alpha(p) &= U_\alpha^+(p), \end{aligned} \tag{IV.63}$$

and the sum over spinors becomes

$$\begin{aligned} g_\alpha^\beta &\equiv \sum_{\text{spin}} U_\alpha(p, s_3)U^\beta(p, s_3) \\ &= \langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{\hat{K}}}|\pi\rangle\langle\pi|m, s_3\rangle\langle m, s_3|\rho\rangle\langle\rho|e^{i\vec{\epsilon}\cdot\vec{\hat{K}}^+}|\beta\rangle \\ &= \langle\alpha|e^{-i\vec{\epsilon}\cdot\vec{\hat{K}}}e^{i\vec{\epsilon}\cdot\vec{\hat{K}}}|\beta\rangle \\ &= \langle\alpha|\beta\rangle = \delta_\alpha^\beta. \end{aligned} \tag{IV.64}$$

Thus we see that for this case the spin sum has no momentum dependence.

V. Poincare Invariance and Internal Symmetry

We now consider the problem of combining external with internal symmetry. The simplest possibility for an internal symmetry such as U(3) is to assume invariance of the theory with respect to the direct product $\mathcal{P} \otimes U(3)$. Let the representations and components of U(3) be labelled by the suffix a, and the generators by λ_i (i=0,1,... 8) including the unit matrix λ_0. Thus the Fock space annihilation operation is now $\hat{a}(p,s;a)$ and the auxiliary operator $A_{\alpha,a}(p)$.

Invariant S-matrix elements are again constructed by saturating independently all the indices of both groups.

$$\hat{S} = \int A^{\alpha,a}(p_1) B^{\pi,b}(p_2)\, t(\not{q})_{\alpha}^{\beta} B_{\beta,a}(p_3) A_{\pi,b}(p_4)$$

$$(2\pi)^4 \delta(p_1+p_2-p_3-p_4) \prod_{i=1}^{4} \frac{d^4 p_i}{(2\pi)^4} \Delta^{+}(p_i). \qquad (V.1)$$

Lorentz transformations operate both on the indices α and the arguments p_μ as discussed in (IV.45)-(IV.45), so momentum dependent terms $t(\not{q})$ are allowed in the $\hat{S}$. For this it is essential that both A_α and p_μ transform according to representations of O(3,1).

If we attempt a more intricate connection between the Poincare group and U(3), we can take the simplest representations of the Lorentz group SL(2,C), for which (k_0, c) are (3/2, 1/2) and (3/2, -1/2) (dotted and undotted spinors). These 2×2 representations give rise to four vectors

$$\sigma_\mu = (1, \vec{\sigma}), \qquad (V.2)$$

and

$$\tilde{\sigma}_\mu = (1, -\sigma). \qquad (V.3)$$

The simplest possible non-trivial extension of the physical Lorentz group $\mathcal{L}$, which contains $\mathcal{L} \otimes U(3)$ as a subgroup, is to take the group SL(6,C) with basic generators given by the outer product

$$\sigma_\mu \otimes \lambda_i \qquad 36 \text{ generators}, \qquad (V.4)$$

or

$$\tilde{\sigma}_\mu \otimes \lambda_i \qquad 36 \text{ generators}. \qquad (V.5)$$

Both have to be included for a theory which involves space-reflection invariance. If this is to be generalised to an extension of the

Poincare group, the vector p_μ must be a representation of the new larger group SL(6,C). But the simplest representation which includes four components which transform as a four-vector under Lorentz transformation is the 72-component multiplet, transforming like the basic generators. One is thus stuck with a 72-component energy-momentum vector.[10),11),12)]

This is an example of a general theorem established by Michel and Sakita.[13)]

Theorem. Any generalisation of $\mathcal{P} \otimes U(3)$ which contains $\mathcal{P} \otimes U(3)$ as a subgroup necessarily involves an energy-momentum vector with more than four components.

The way around this difficulty is to extend, not the Poincare group, but the Auxiliary group. For example,[14),15)] we can take the latter to be U(6,6) or $\tilde{U}(12)$—with basic generators given by the outer product of the 16 Dirac matrices with λ_i:

$$\left(1, i\gamma_5, \gamma_\mu, \gamma_\mu\gamma_5, \tfrac{1}{2}\sigma_{\mu\nu}\right) \otimes \lambda_i. \qquad (V.6)$$

The basic auxiliary operator is twelve component

$$A_A(p) \text{ where } A = \alpha, a$$
$$\alpha = 1,2,3,4 \quad a = 1,2,3.$$

Since the transformations on both indices α and a are pure index transformations, we can require the S-matrix to be index invariant under the pure index transformations of the U(6,6) auxiliary group. This gives rise to a subset of the invariants allowed by invariance under $\mathcal{P} \otimes U(3)$. They have the property of being SU(6) invariant, in the extreme static limit, in which the masses of all particles tend to infinity, since then momentum dependence is frozen out and the auxiliary group spinor labels are equivalent to the spin labels. However, since the index transformations do not operate on the momentum p_μ, index invariance excludes factors like $\not{q}$ in S-matrix elements (see (IV.43)-(IV.45)).

We must now consider whether index invariance is consistent with the unitarity of the S-matrix. Expressed in terms of T where

$$S = 1 + iT, \qquad (V.7)$$

this requires that

$$2\,\mathrm{Im}T = TT^+. \qquad (V.8)$$

Suppose that T and T^+ are index invariant and that ImT is defined by (V.8). The product on the right hand side involves a sum over

spinors. Writing one such sum explicitly we can put

$$TT^+ \equiv \sum_{s_3} t^\alpha U_\alpha(p,s_3)U^\beta(p,s_3)\bar{t}_\beta.$$

If $|\alpha\rangle$ is a finite (non-unitary) representation of the auxiliary group, we have seen that this gives momentum dependent factors—typically (IV.61)—

$$TT^+ = t^\alpha(\not{p})^\beta_\alpha \bar{t}_\beta.$$

But factors $\not{p}$ are excluded by index invariance. Thus even though T and T^+ are index invariant, unitarity of the S-matrix implies that ImT is not, if the particle multiplets are finite dimensional. This is the so-called "conflict with unitarity." The escape route has been anticipated in (IV.64). If the physical multiplets are taken in one-to-one correspondence with unitary representations of the (extended) auxiliary group, the spin sum is unity, and consistency between index invariance and unitarity of the S-matrix is restored. The price paid is that we now have infinite particle multiplets.[16),17),18)]

For a more general and formal statement of the argument of this chapter, see Feldman and Matthews.[9)]

VI. Causality—Spin and Statistics

Notice that so far we have made absolutely no mention of anti-particles, the relation between spin and statistics, or CTP invariance. All these concepts arise from the requirement of causality, which goes beyond Poincare invariance. We turn now to this problem and find that it leads to further complications for unitary, index invariant theories.

Causality is best discussed in configuration space, but before taking the Fourier transform of the auxiliary operators we develop what will turn out to be the anti-particle formalism. Following Weinberg,[7)] we note that for any rotation there exists a matrix B, such that

$$\langle s,s_3|e^{i\hat{\vec{J}}\cdot\vec{\theta}}|s,s_3'\rangle = \langle s,s_3'|\hat{B}^{-1}e^{-i\hat{\vec{J}}\cdot\vec{\theta}}\hat{B}|s,s_3\rangle. \qquad \text{(VI.1)}$$

Thus to construct a creation operator that transforms under the auxiliary group like an annihilation operator, we make use of (VI.1) to obtain

$$\hat{b}^+(p,s_3) \xrightarrow{\eta} \langle s_3|\hat{B}^{-1}e^{i\vec{\epsilon}\cdot\hat{\vec{K}}}e^{i\vec{\eta}\cdot\hat{\vec{K}}}e^{-i\vec{\epsilon}'\cdot\hat{\vec{K}}}\hat{B}|s_3'\rangle\hat{b}^+(p',s_3'). \qquad \text{(VI.2)}$$

Analogous to the auxiliary field $\hat{A}_\alpha(p)$, we introduce

$$\hat{\tilde{B}}_\alpha(p) = \sum_{s,s_3} \langle \alpha | e^{-i\vec{\epsilon}\cdot\vec{\hat{K}}} \hat{B} | m,s,s_3 \rangle g(m,s) \hat{b}^+(p,s,s_3) \qquad \text{(VI.3)}$$

$$\equiv \sum_{s,s_3} \tilde{U}_\alpha(p,s,s_3) \hat{b}^+(p,s,s_3). \qquad \text{(VI.4)}$$

As we have seen, there can be more than one spin occurring in the sum and the factor $g(m,s)$ allows for a spin dependent mass,[20] though we will not consider such fields here. Under a Lorentz transformation

$$\hat{\tilde{B}}_\alpha(p) \xrightarrow{\eta} \langle \alpha | e^{i\vec{\eta}\cdot\vec{\hat{K}}} | \beta \rangle \hat{\tilde{B}}_\beta(p'), \qquad \text{(VI.5)}$$

i.e., $\hat{\tilde{B}}_\alpha$ is a creation operator transforming like the annihilation operator $\hat{A}_\alpha(p)$.

Under translations, however, $\hat{\tilde{B}}_\alpha(p)$ transforms with the opposite sign from $\hat{A}_\alpha(p)$,

$$\hat{\tilde{B}}_\alpha(p) \xrightarrow{a} \hat{\tilde{B}}'_\alpha(p) = e^{i\hat{P}_\mu a^\mu} \hat{\tilde{B}}_\alpha(p). \qquad \text{(VI.6)}$$

Thus $\hat{\tilde{B}}_\alpha(p)$ transforms like a particle creation operator.

We now define the free field in configuration space

$$\hat{\psi}_\alpha(x) = \int \left\{ \hat{A}_\alpha(p) e^{-ip\cdot x} + \tilde{B}_\alpha(p) e^{ip\cdot x} \right\} \Delta^+(p) \frac{d^4p}{(2\pi)^4}. \qquad \text{(VI.7)}$$

Under Poincare transformations, we have

$$\hat{\psi}_\alpha(x) \xrightarrow{\eta} \left(e^{i\vec{\eta}\cdot\vec{\hat{K}}} \right)_\alpha^\beta \hat{\psi}_\beta(x'), \qquad \text{(VI.8)}$$

$$\hat{\psi}_\alpha(x) \xrightarrow{a} \hat{\psi}_\alpha(x+a), \qquad \text{(VI.9)}$$

for quite arbitrary choice of $f(m,s)$ and $g(m,s)$ (including zero). The factorization of transformations on spinor and space time labels follows from the factorization of transformations on spinor and p labels in momentum space. In this way we arrive at a local field operator normally taken as the starting point in Lagrangian field theory.

The causality condition imposed upon such fields is

$$\left\{ \hat{\psi}_\alpha(x), \hat{\Phi}_\beta(y) \right\}_\pm = 0 \quad \text{for} \quad (x-y)^2 < 0 \quad \text{(spacelike)}. \qquad \text{(VI.10)}$$

This implies that no mutual distrubance from observation of the fields travels faster than the speed of light.

This is a stringent requirement which may be stronger than is physically necessary since it is not clear that one need observe the fields in this sense. The condition does, however, have the enormous advantage of being simple and precise. Applied to finite component fields, it has the important consequences, mentioned above, concerning anti-particles and statistics, which are well borne out by experiment.

It is the basis of the analytic properties of the S-matrix. In particular, the procedure for introducing electromagnetic interactions through the substitution

$$\partial_\mu \to \partial_\mu - ieA_\mu$$

only leads to causal currents, when it is applied to causal fields. Only for such fields does the time-ordered product—and hence the standard Feynman-Dyson S-matrix expansion—have a well-defined covariant meaning. It is thus a property which one does not lightly give up.

We have been lead to consider infinite component fields corresponding to unitary representations of the Lorentz group. We now consider what survives of the Pauli Theorem concerning spin and statistics if the causality condition is applied to these generalised unitary fields.

If a theory is constructed around the field $\hat{\psi}$, hermitian operators (e.g., currents) have to be constructed, and the theory necessarily also involves $\hat{\psi}^+$. Thus a minimum condition for a causal theory is that $\hat{\psi}$ and $\hat{\psi}^+$ satisfy (VI.10). One must also make sure that the condition is satisfied by any other pair of operators in the theory, but this is usually simple.

The causality condition is satisfied by the free field commutator (anti-commutator) if it can be expressed in the form

$$\left\{\hat{\psi}_\alpha(x), \hat{\psi}^+_\beta(y)\right\}_\pm = f_{\alpha\beta}(\partial) \int \left\{e^{-ip\cdot(x-y)} - e^{ip\cdot(x-y)}\right\} \Delta^+(p) d^4p, \tag{VI.11}$$

where the crucial feature is the relative minus sign between the positive and negative frequency parts. Now substitute (VI.7) into the left hand side using (VI.4), (IV.31) and the (anti) commutation relations (IV.12) where

$\{\ \}_+$ denotes (Fermi statistics) anti-commutator,

and

$\{\ \}_-$ denotes (Bose statistics) commutator.

Then

$$\left\{\hat{\psi}_\alpha(x), \hat{\psi}^+_\beta(y)\right\}_\pm = \sum_{s s_3} \int \Big[U_\alpha(p,s,s_3) U^*_\beta(p,s,s_3) e^{-ip(x-y)}$$

$$\pm \tilde{U}_\alpha(p,s,s_3) \tilde{U}^*_\beta(p,s,s_3) e^{ip(x-y)} \Big] \Delta^+(p) \frac{d^4p}{(2\pi)^4} \qquad \text{(VI. 12)}$$

$$= \int \Big[\langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}^+} | \delta \rangle \langle \delta | m,s,s_3 \rangle |f(s)|^2 \langle m,s,s_3 | \pi \rangle \cdot$$

$$\cdot \langle \pi | e^{i\vec{\epsilon}\cdot\hat{\vec{K}}^+} | \beta \rangle e^{-ip(x-y)} \pm \langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} B | m,\bar{s},s_3 \rangle$$

$$|g(\bar{s})|^2 \langle m,\bar{s},s_3 | B^{-1} e^{i\vec{\epsilon}\cdot\vec{K}^+} | \beta \rangle e^{ip(x-y)} \Big] \Delta^+(p) \frac{d^4p}{(2\pi)^4}, \qquad \text{(VI. 13)}$$

the upper sign referring to Fermi statistics. This is the basis for the Pauli spin-statistics theorem in that, if finite non-unitary representations are used, then (VI. 10) is satified only with Bose statistics for integer spin and Fermi statistics for half integer spin particles. To illustrate this, we consider the spin O and $\frac{1}{2}$ cases.

(1) S=O. For the scalar case,

$$\hat{K} = \hat{O} \quad \text{and} \quad \hat{U} = \hat{U}^+ = \hat{I}, \qquad \text{(VI. 14)}$$

so that from (VI. 13) we obtain

$$\left\{\hat{\psi}_{(x)}, \hat{\psi}^+_{(y)}\right\}_\pm = \int \Big[|f|^2 e^{-ip\cdot(x-y)} \pm |g|^2 e^{ip(x-y)} \Big] \Delta^+(p) \frac{d^4p}{(2\pi)^4}. \qquad \text{(VI. 15)}$$

Thus for agreement with (VI. 10) we must choose Bose statistics (lower sign) and

$$|f|^2 = |g|^2 = 1, \qquad \text{(VI. 16)}$$

i.e., anti-particles must be included. It is further noted that if we require the parity transform of the scalar field

$$\hat{R}\hat{\psi}(x_o,\vec{x})\hat{R}^{-1} = \pm\hat{\psi}(x_o,-\vec{x}) \,. \tag{VI.17}$$

it implies that particle and anti-particle have the same parity; i.e.,

$$\hat{R}\hat{a}^+(\underset{\sim}{p},s)\rangle_o = \pm\hat{a}^+(-\underset{\sim}{p},s)\rangle_o, \quad \hat{R}\hat{b}^+(\underset{\sim}{p},s)\rangle_o = \pm\hat{b}^+(-\underset{\sim}{p},s)\rangle_o,$$

taking either both upper or both lower signs.

(2) $s = \frac{1}{2}$. For the spinor case,

$$\hat{K}_{oi} = \frac{\hat{\sigma}_{oi}}{2} = -\frac{\hat{\sigma}^+_{oi}}{2}. \tag{VI.18}$$

Particle parity is chosen by convention to be positive,

$$\hat{R}\hat{a}^+(\underset{\sim}{p},s_3)\rangle_o = +\hat{a}^+(-\underset{\sim}{p},s)\rangle_o, \tag{VI.19}$$

and will be determined for the anti-particle by the causality condition

$$\hat{R}\hat{b}^+(\underset{\sim}{p},s_3)\rangle_o = \pm\hat{b}^+(-\underset{\sim}{p},s_3)\rangle_o. \tag{VI.20}$$

The spin sums in (VI.13) are (see (IV.62))

$$\begin{aligned}
&\sum_{\substack{\text{spin}\\ \text{parity}(+)}} U_\alpha(p,s,s_3)\,U^*_\beta(p,s,s_3) = \frac{1}{2m}\Big[(\not{p}+m)(\gamma_o)\Big]^\beta_\alpha |f|^2, \\
&\sum_{\substack{\text{spin}\\ \text{parity}(\pm)}} \tilde{U}_\alpha(p,s,s_3)\,\tilde{U}^*_\beta(p,s,s_3) = \frac{1}{2m}\Big[(\not{p}\pm m)\gamma_o\Big]^\beta_\alpha |g|^2,
\end{aligned} \tag{VI.21}$$

so that we obtain

$$\begin{aligned}
\Big\{\hat{\psi}_\alpha(x),\hat{\psi}^+_\beta(y)\Big\}_\pm = \int\Big\{&\frac{1}{2m}\Big[(\not{p}+m)\gamma_o\Big]^\beta_\alpha |f(s)|^2 e^{-ip(x-y)} \\
&\pm \frac{1}{2m}\Big[(\not{p}\pm m)\gamma_o\Big]^\beta_\alpha |g(s)|^2 e^{ip(x-y)}\Big\}\Delta^+(p)\,\frac{d^4p}{(2\pi)^4}\,.
\end{aligned} \tag{VI.22}$$

The causality statement (VI.10) then requires:

(i) Fermi statistics,
(ii) (-) parity for the anti-particles,
(iii) $|f|^2 = |g|^2 = 1$, (VI.23)

so that

$$\left\{\hat{\psi}_\alpha(x)\hat{\psi}^+_\beta(y)\right\}_+ = \frac{1}{2m}\left[(i\not{\partial}+m)\gamma_o\right]^\beta_\alpha \int \left(e^{-ip(x-y)} - e^{ip(x-y)}\right)\Delta^+(p)\,\frac{d^4p}{(2\pi)^4}. \tag{VI.24}$$

As is well-known, these arguments generalize for all finite non-unitary fields with the result that:

(i) Pauli spin-statistics theorem is valid;
(ii) Substitution law [$p_\mu \to -p_\mu$: particle in $\to$ anti-particle out] is valid;
(iii) CPT theorem is valid.

Index invariant theories such as SL(6,C) or U(6,6) with finite particle multiplets work with local causal fields, which preserves all of these features and give SU(6) in the extreme static limit [all particles have $m \to \infty$]. However, as we have seen, they violate unitarity. This was the original motivation for discussing infinite component fields, considered at the end of V, which allow for index invariance consistent with the unitarity of the S-matrix.

(3) Unitary Fields of index invariant theories. For these index invariant fields we have $|\alpha\rangle$ unitary and $K = K^+$. In addition to satisfy unitarity we required that

(i) $f = g = 1$,
(ii) $\langle j, j_3 | s, s_3 \rangle = \delta_{js}\delta_{j_3 s_3}$,

so that

$$U_\alpha U^*_\beta = \delta^\beta_\alpha$$

and

$$\tilde{U}_\alpha \tilde{U}^*_\beta = \delta^\beta_\alpha. \tag{VI.25}$$

Thus

$$\left\{\hat{\psi}_\alpha(x), \hat{\psi}^+_\beta(y)\right\}_\pm = \delta^\beta_\alpha \int \left[e^{-ip(x-y)} \pm e^{ip(x-y)}\right]\Delta^+(p)\,\frac{d^4p}{(2\pi)^4} \tag{VI.26}$$

so causality (VI.10) can only be satisfied for Bose statistics for all such fields, whether of half interger or integer spins.

Notice that there is a very direct conflict between causality and unitarity with index invariance, since the latter demands

$$U_\alpha U^*_\beta = \delta^\beta_\alpha,$$

while causality for consistence with Fermi statistics requires

$$U_\alpha U^*_\beta \neq \delta^\beta_\alpha,$$

in order to produce the crucial minus sign in (VI.10).

Fronsdal[16] and Dao and Nguyen[16] have suggested that one can construct infinite fields of the auxiliary group to satisfy index invariance, but quantise each separate spin component in the conventional manner to preserve the correct spin-statistics relation. This requires that both a unitary (infinite) and a finite auxiliary representation are associated with each physical spin. The relation between the two is wildly non-local, so that local currents constructed from the unitary fields are very a-causal in the quantised (finite) fields. This brute force approach does not appear to us to provide a solution to the problem.

Apart from its implications for index invariant theories, this result clearly shows that the Pauli Theorem connecting spin and statistics is not valid for unitary fields. We may wonder if any such connection survives. To this end we consider:

(4) The self-conjugate unitary fields. The two unitary fields specified by either

$$(k_o, c) = (0, \tfrac{1}{2}) \; : \; j = 0, 1, 2, \ldots$$

or

$$(k_o, c) = (\tfrac{1}{2}, 0) \; : \; j = \tfrac{1}{2}, 3/2, \ldots \qquad \text{(VI.27)}$$

are parity self-conjugate fields and we now show <u>both</u> can be made causal using Fermi, instead of the Bose, statistics derived above. This completes the collapse of the Pauli theorem in that it demonstrates that not even the wrong spin-statistics correlation results. In general for unitary causal fields there is no spin-statistics correlation.

For these special representations there exist operators[4),5)] $(\Gamma_\mu)^\beta_\alpha$ that transform like four vectors, where

$$\langle j, j_3 | \Gamma_o | j', j'_3 \rangle = (j + \tfrac{1}{2}) \delta_{jj'} \delta_{j_3 j'_3}. \qquad \text{(VI.28)}$$

The meaning of Γ_μ as a 4-vector is that under a boost,

$$\Gamma_o \xrightarrow{\vec{\epsilon}} e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}}\Gamma_o e^{i\vec{\epsilon}\cdot\hat{\vec{K}}} = \frac{p^\mu \Gamma_\mu}{m}. \tag{VI.29}$$

Now take the physical states to correspond to the infinite tower such that

$$\langle k_o, c, j, j_3 | m, s, s_3 \rangle = \delta_{js}\delta_{j_3 s_3}. \tag{VI.30}$$

so that

$$\Gamma_o | m, s, s_3 \rangle = (s + \tfrac{1}{2}) | m, s, s_3 \rangle, \tag{VI.31}$$

and Γ_o can be written

$$\Gamma_o = \sum_{s s_3} | s, s_3 \rangle (s + \tfrac{1}{2}) \langle s, s_3 |. \tag{VI.32}$$

The trick that enables one to satisfy causality with Fermi statistics is to take

$$|f(s)|^2 = |g(s)|^2 = (s + \tfrac{1}{2}). \tag{VI.33}$$

Then the (anti) commutator (VI.13) is, using (VI.32),

$$\left\{\hat{\psi}_\alpha(x), \hat{\psi}^+_\beta(y)\right\}_\pm = \int \Big\{ \langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} | s, s_3 \rangle (s+\tfrac{1}{2}) \langle s, s_3 | e^{i\vec{\epsilon}\cdot\hat{\vec{K}}^+} | \beta \rangle e^{-ip(x-y)}$$

$$\pm \langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} B | s, s_3 \rangle (s+\tfrac{1}{2}) \langle s, s_3 | B^{-1} e^{i\vec{\epsilon}\cdot\hat{\vec{K}}^+} | \beta \rangle e^{ip(x-y)} \Big\}$$

$$\Delta^+(p) \frac{d^4p}{(2\pi)^4}$$

$$= \int \Big\{ \langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} \Gamma_o e^{i\vec{\epsilon}\cdot\hat{\vec{K}}} | \beta \rangle e^{-ip(x-y)}$$

$$\pm \langle \alpha | e^{-i\vec{\epsilon}\cdot\hat{\vec{K}}} B \Gamma_o B^{-1} e^{+i\vec{\epsilon}\cdot\hat{\vec{K}}} | \beta \rangle e^{ip(x-y)} \Big\} \Delta^+(p) \frac{d^4p}{(2\pi)^4};$$

and, with (VI.29), becomes

$$\left\{\hat{\psi}_\alpha(x), \hat{\psi}_\beta^{+}(y)\right\}_{\pm} = \left(\frac{\Gamma_\mu i\partial^\mu}{m}\right)_\alpha^\beta \int \left(e^{-ip(x-y)} \mp e^{ip(x-y)}\right)\Delta^{+}(p)\,\frac{d^4p}{(2\pi)^4}\,. \tag{VI.34}$$

Thus causality (VI.10) requires Fermi statistics (upper sign) for both the $(0,\frac{1}{2})$—integer spin—and $(\frac{1}{2},0)$—one-half integer spin—unitary representations. Thus there is no spin statistics correlation for unitary representations. (Note that since the sum over spinors is not unity, these Fermi fields cannot be used to construct unitary index invariant theories.)

We have confined this discussion to multiplets of equal mass. Multiplets with different masses for the different spin have been considered by Feldman and Matthews.[20] If a unitary field is assumed to satisfy a linear equation

$$\left(i\partial^\mu\Gamma_\mu + K\right)\psi = 0,$$

giving a mass spectrum

$$m_s = \frac{K}{s+\frac{1}{2}},$$

it has been shown by Abers, Grodsky and Norton[21] that one can satisfy causality with either Fermi or Bose statistics, without introducing any anti-particles. This demonstrates the loss for unitary fields of one more of the physically attractive features of finite component (non-unitary) fields.

We close with the very negative comment that the difficulties encountered in combining internal and external symmetries in a non-trivial way through index invariance are of a subtle nature, involving as they do problems of causality, CTP and the relation between particles and anti-particles. It has only been possible to see them because these theories have been formulated in a very clear and precise manner.

Similarly, attempts to find reasonably realistic model theories for current commutators, which have an equally precise basis, lead to the highly non-local Lagrangians. No light can be thrown on these fundamental problems by phenomenological non-relativistic quark models, and the whole question of the connection between internal and external symmetries remains extremely obscure.

Summary

Starting from the notion of particles and Poincare invariance, we have arrived at local fields. In order that these should have simple transformation properties the spin variable, specifying a representation of the Little Group O(3), is replaced by a spinor, specifying a representation of the auxiliary group. This must contain the homogeneous Lorentz group and is most simply chosen to be isomorphic to the homogeneous Lorentz group (Chapters III and IV).

The combination of internal, SU(3), and external, $\mathcal{P}$, symmetries is considered in Chapter V. The non-trivial extension of $\mathcal{P} \otimes$ SU(3) leads—by the Michel-Sakita Theorem—to an energy momentum vector of more than four components. This can be avoided by extending the auxiliary group to include the internal symmetry—SL(6,C) or U(6,6)—and requiring "index invariance" of the S-matrix under the purely index transformations of this larger group. This provides a relativistic theory with SU(6) as its static limit, but is in conflict with the unitarity of the S-matrix unless the particle multiplets are infinite unitary "towers."

Causality is considered in Chapter VI. It is shown that index invariant unitary theories can only be made causal if all particles satisfy Bose statistics. In general, it is shown that for (infinite component) unitary causal fields there is no connection between spin and statistics.

References

1. M. Hamermesh, "Group Theory," Addison Wesley Pub. Co., Inc. (1962).
2. R. E. Behrends, J. Dreitlein, C. Fronsdal, B. W. Lee, Rev. Mod. Phys. 34, 1 (1962).
3. Abdus Salam, "Theoretical Physics," Trieste Summer School, p. 173. IAEA, Vienna (1963).
4. M. A. Naimark, "Linear Representations of the Lorentz Group," Pergamon Press, New York (1964).
5. I. M. Gelfand, R. A. Minlos and Z. Ya Shapiro, Representations of the Rotation and Lorentz Groups, Pergamon Press, New York (1964).
6. H. Joos, Fortschritte der Physik. 10, 65 (1962).
7. S. Weinberg, Phys. Rev. 133B, 1318 (1964).
8. G. Feldman and P. T. Matthews, Ann. of Physics 40, 19 (1966).
9. G. Feldman and P. T. Matthews, Phys. Rev. 151, 1176 (1966).
10. T. Fulton and J. Wess, Phys. Lett. 14, 57 (1965); 14, 334 (1965).
11. W. Rühl, Nuovo Cim. 37, 301 (1965); 37, 319 (1965).
12. J. M. Charap, P. T. Matthews and R. F. Streater, Proc. Roy. Soc. A290, 24 (1965).

13. L. Michel and B. Sakita, Ann. Ins. Henri Poincare, 1965. See also L. O'Raifeartaigh, Phys. Rev. 139, B1052 (1965).
14. A. Salam, R. Delbourgo and J. Strathdee, Proc. Roy. Soc. A284, 146 (1965).
15. B. Sakita and K. C. Wali, Phys. Rev. 139, B1355 (1965).
16. C. Fronsdal, Proc. 1965 Trieste Seminar, IAEA, Vienna (1965); Phys. Rev. 156, 1657 (1967).
17. W. Rühl, Nuovo Cim. 42, 619; 43, 171; 44, 572, 44, 659 (1966).
18. Dao Van Duc and Nguyen Van Hieu, Dubna preprints (1966).
19. K. Koller, Imperial College Preprint, ICTP/67/25: (Feynman diagrams for unitary representations of SL(2,C)).
20. G. Feldman and P. T. Matthews, Phys. Rev. 154, 1241 (1967).
21. E. Abers, I. T. Grodsky and R. E. Norton, Phys. Rev. 159, 1222 (1967).

BOOTSTRAPS, FIELDS AND GENERALISED GROUPS†

John G. Taylor
Mathematical Institute
Oxford, England

Contents

I. A Survey of N/D Bootstraps

In this series of lectures, I want to set up a field theory of bootstraps and discuss briefly some of the mathematical problems arising in such a theory.[1] As an introduction to this I want to start by giving a survey of the theory and results of the N/D approach to bootstraps. By doing this, I can tell you in a very simple way what the bootstrap idea is, at the same time following the historical method of development, since the bootstrap idea was first developed in the N/D framework. At the same time you will see how the calculations indicate the need for a more complete approach to bootstraps which includes many particles in both direct and crossed channels. This can be achieved, at least in principle, by means of field theory, so we will naturally be lead to setting up a field theory of bootstraps. In the process of doing this we will have first to set up a field theory of composites, and then make all particles composite. These problems and their resolution will be discussed in more detail in later lectures.

So let me begin, then, with the N/D method of bootstrapping particles.[2] The basic idea here is that a particle 'bootstraps' itself by being its own potential—that potential caused by the exchange of the particle. This potential then acts between two other (elementary or composite) particles to produce the bootstrapped particle as

---

†Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

a bound state. To set this up in detail let me consider the elastic scattering amplitude for two equal neutral scalar a-particles of mass m. The ℓ-th partial wave amplitude is a function $t_\ell(s)$ of the invariant energy s of the incoming particles; in the centre of mass system for the incoming particles, $s = 4(m^2 + \vec{k}^2)$, where $\vec{k}$ is the centre of mass momentum of the particles. Also the elastic scattering phase shift $\delta_\ell(s)$ is related to $t_\ell(s)$ by

$$k \cdot t_\ell(s) = s^{\frac{1}{2}} \cdot e^{i\delta_\ell} \sin\delta_\ell .$$

It is to be expected[3] that $t_\ell(s)$ map be continued analytically in s to the whole complex plane except for two real branch cuts: the right hand cut running from $s = 4m^2$ to $+\infty$ and the left hand cut, running from a value $s_c (< 4m^2)$ to $-\infty$. The right hand cut is sometimes called the unitarity cut, since for $s < 4m^2$ the unitarity condition is[3]

$$t_\ell(s) - t_\ell^*(s) = 2ik|t_\ell(s)|^2/s^{\frac{1}{2}} + \text{inelastic contributions}$$

$$= 2i\rho(s)|t_\ell(s)|^2. \qquad (I.1)$$

The physical value $t_\ell(s)$ is the value of the analytic function $t_\ell(z)$ on the upper side of the right hand cut, while $t_\ell^*(z^*)$ will be the value on the lower side of the cut. Thus the right hand side of (I.1) denotes the discontinuity of t_ℓ across the right hand cut; this contribution may be regarded as arising from rescattering through two or more particle intermediate states.

The left hand cut has a more complicated origin, being the place where the dynamics of the system becomes evident. If we consider the simplest particle exchange, that of a single c particle of mass μ, coupled with strength g to two a particles of mass m, then the Feynman diagram for this gives a contribution to the total scattering amplitude equal to

$$g^2/[(p_1 - p_3)^2 - \mu^2] . \qquad (I.2)$$

In the centre of mass system, with scattering angle θ, so that

$$p_1 = (w, \vec{k}), \quad p_2 = (w, -\vec{k}), \quad p_3 = (w, \vec{k}'), \quad p_4 = (w, -\vec{k}'), \quad |\vec{k}| = |\vec{k}'|,$$

with

$$w = (\vec{k}^2 + m^2)^{\frac{1}{2}}, \qquad \cos\theta = \vec{k} \cdot \vec{k}' / |\vec{k}|\,|\vec{k}'| ,$$

then $t = -2\vec{k}^2(1 - \cos\theta)$ and the contribution of (I.2) to $t_\ell(s)$ will be

$$g^2 \int_{-1}^{+1} P_\ell(\cos\theta)\left[-2\vec{k}^2(1-\cos\theta) - \mu^2\right]^{-1} d(\cos\theta). \qquad (I.3)$$

Evidently (I.3) is singular in $\vec{k}^2$(or s) for $\vec{k}^2 \leq -\mu^2/4$, and $\vec{k}^2 = -\mu^2/4$ is a logarithmic branch point. There will be higher branch points at $\vec{k}^2 = -n^2\mu^2/4$, with $n = 2,3,\ldots$, due to higher numbers of particles being exchanged; these will lie along the negative real axis in the s-plane. Thus the highest branch point s_c of the left hand cut will be at $s_c = 4m^2 - \mu^2$.

There may also be poles in t_ℓ along the real axis between s_c and $4m^2$; these poles will correspond to bound states. For the bootstrap situation we are especially interested in showing that the single c particle exchange can produce a c-particle bound state, i.e., a pole in $t_0(s)$ at $s = \mu^2$ with the correct residue g^2.

In order to see if such a bootstrap situation is possible, and more generally to obtain $t_\ell(s)$ in terms of its discontinuity across the left-hand cut, we write

$$t_\ell = N_\ell / D_\ell$$

where we choose N_ℓ and D_ℓ only to have left- and right-hand branch cuts in s, respectively. We now show constructively how this may be done.

We apply Cauchy's theorem for $D_\ell(z)$ to the contour C_1 in the s-plane, which is a large circle together with a contour encircling the right hand cut, all taken counter-clockwise, z is any point not on the right hand cut; then

$$D(z) = \frac{1}{2\pi i}\int_{C_1} \frac{D_\ell(z')dz'}{(z'-z)} = \frac{1}{2\pi i}\int_{4m^2}^{\infty} \frac{[D_\ell(s') - D_\ell^*(s')]ds'}{(s'-z)} \qquad (I.4)$$

where we are assuming that the contribution from the circular part of C_1 to the middle part of (I.4) vanishes as the radius of the circle is made infinite. If we denote $(D_\ell - D_\ell^*)$ by 2idisc D_ℓ, then

$$D_\ell(z) = \frac{1}{\pi}\int_{4m^2}^{\infty} \frac{\text{disc}\, D_\ell(s')ds'}{(s'-z)}.$$

Since we cannot determine each of N_ℓ and D_ℓ to within a common arbitrary constant, we assume arbitrarily that $D_\ell(s_0) = 1$, where $s_c < s_0 < 4m^2$. Then

$$D_\ell(z) = 1 + \frac{(z - s_0)}{\pi} \int_{4m^2}^{\infty} \frac{\text{disc } D_\ell(s')ds'}{(s'-z)(s'-s_0)} \quad . \tag{I.5}$$

The physical value of D_ℓ is obtained by taking

$$\lim_{\epsilon \to 0+} D_\ell(s + i\epsilon) .$$

Similarly, by using the contour C_2 (which is identical to C_1 except that the left hand cut is encircled in place of the right hand cut) for N_ℓ, we obtain

$$N_\ell(z) = \frac{1}{\pi} \int_{-\infty}^{s_c} \frac{\text{disc } N_\ell(s')ds'}{(s' - z)} \quad .$$

Now

$$\text{disc } D_\ell(s') = N_\ell \, \text{disc } t_\ell^{-1} = -\rho(s')N_\ell(s'),$$

where $\rho(s')$ is given by the unitarity equation (I.1), so is $k'/(s')^{\frac{1}{2}}$ for $4m^2 < s' < 9m^2$. Also $\text{disc } N_\ell = D_\ell \cdot \text{disc } t_\ell = D_\ell f_\ell$, so

$$D_\ell(s) = 1 - \frac{(s - s_0)}{\pi} \int_{4m^2}^{\infty} \frac{\rho(s')N_\ell(s')ds'}{(s'-s-i\epsilon)(s'-s_0)} \tag{I.6}$$

$$N_\ell(s) = \frac{1}{\pi} \int_{-\infty}^{s_c} \frac{f_\ell(s')D_\ell(s')ds'}{(s' - s - i\epsilon)} \quad . \tag{I.7}$$

Thus if $f_\ell(s')$ is known for $s' < s_c$ we may solve the pair of coupled linear integral equations (I.6), (I.7) to obtain N_ℓ and D_ℓ explicitly.

The bound states arising in t_ℓ may be due to poles of N_ℓ or zeros of D_ℓ. The form of (I.7) has excluded such poles in N_ℓ; if we had included them we would not have been able to determine their position or residue. In other words, they would not be dynamical bound states; as we switch off the interaction represented by the exchange of particles and described quantitatively by the function f_ℓ such poles would have remained fixed, contrary to the behaviour expected from a dynamical bound state. Since we are using the bootstrap approach we require our bound states to be dynamical, so they correspond to (positive) zeros of D_ℓ between s_c and $4m^2$.

Let me first describe how we may determine general conditions on f_ℓ so that a zero of D_ℓ may or may not occur. Before doing that it is necessary to remark that we are not only interested in bound states but also in resonances. These will correspond to complex zeros of D_ℓ. We may approximately describe these resonances as arising from zeros of the real part $\mathrm{Re}D_\ell$ of D_ℓ, since at such a zero, say s_r, we have

$$t_\ell(s) \sim N_\ell(s_r) \Big/ \left[(s - s_r)\mathrm{Re}\, D'_\ell(s_r) + i\, \mathrm{Im}D_\ell(s_r) \right] \quad (I.8)$$

which is a Breit-Wigner form for a resonance. For the rest of this section we will treat resonances as if they were bound states.

Returning now to (I.6) and (I.7) we see that if f_ℓ is negative on the left hand cut, N_ℓ will be positive on the right hand cut. Then D_ℓ will vary between the value +1 at s_o to $-\infty$ as s approaches $4m^2$, so must have a zero between s_o and $4m^2$. On the other hand, if f_ℓ is positive on the left hand cut then N_ℓ will be negative on the right hand cut, and D_ℓ will lie between +1 and $+\infty$ when s lies between s_o and $4m^2$. It is possible that D_ℓ has a zero to the left of s_o; if we choose s_o close to s_c then this will be unlikely. Hence we conclude that for negative (positive) values of f_ℓ on the left hand cut there is (is not) a dynamical bound state. This result will enable us to discuss the bootstrap in a quantitative manner, since we will be able to determine directly whether the exchange of the c-particle gives an attractive potential (negative f_ℓ) or a repulsive one (positive f_ℓ).

In order to go beyond these purely qualitative results, and also to be able to discuss cases in which f_ℓ is neither strictly positive nor strictly negative on the left hand cut, we may set up simple approximate solutions to (I.6) and (I.7). Such an approximation is to set D_ℓ to be unitary on the left hand cut, so that if f_ℓ is the discontinuity arising from a set of Feynman graphs with contribution T_1 (which only has the left hand cut) then $N_\ell(s) = T_\ell(s)$ for all s, and taking two-particle unitarity,

$$D_\ell(s) = 1 - \frac{(s - s_o)}{\pi} \int_{4m^2}^{\infty} \frac{k'}{(s')^{\frac{1}{2}}}\, ds' \cdot \frac{T_\ell(s')}{(s' - s)(s' - s_o)} . \quad (I.9)$$

The condition for a zero of $D_\ell(s)$ at $s = \mu^2$ (a zero of the real part of $D_\ell(s)$ if $\mu^2 > 4m^2$) may now be easily written down from (I.9).

We may even approximate (I.8) further by taking $T_\ell(s) = T_\ell(\mu^2)$ on the right hand cut, so that

$$D_\ell(s) = 1 - T_\ell(\mu^2)\ \frac{(s-s_0)}{\pi}\int_{4m^2}^{\infty} \frac{k'}{(s')^{\frac{1}{2}}}\ \frac{ds'}{(s'-s)(s'-s_0)}. \qquad (I.10)$$

We now return to the bootstrap. We wish to obtain a scalar bound state at $s=\mu^2$, when f_ℓ arises from single particle exchange with value determined by (I.2). Thus we need

$$D_0(\mu^2) = 0. \qquad (I.11)$$

We also require that the residue at this bound state be g^2, or from (I.8) that

$$g^2 = N_0(\mu^2)\Big/D_0'(\mu^2). \qquad (I.12)$$

If we combine (I.6) and (I.7) for $\ell=0$ with (I.11), (I.12) and the fact that f_0 is the discontinuity on the left hand cut arising from (I.2), we get a set of non-linear equations for g and μ which in principle should determine them (though not necessarily uniquely). In the approximation (I.10) we see that (I.12) becomes

$$\frac{g^2}{\pi}\int_{4m^2}^{\infty} \frac{ds'}{(s'-\mu^2)^2}\cdot\frac{k'}{(s')^{\frac{1}{2}}} = 1 \qquad (I.13)$$

with (I.11) given by an evident equation. If we denote the second order c-particle self-energy bubble by $\pi(s)$, with $s=p^2$, then

$$\pi(s) = \frac{g^2}{(2\pi)^2}\int \frac{d^4k}{[k^2-m^2][(p-k)^2-m^2]}$$

$$= \frac{g^2}{\pi}\int_{4m^2}^{\infty}\frac{ds'}{(s'-s)}\cdot\frac{k'}{(s')^{\frac{1}{2}}}.$$

Thus (I.13) becomes

$$\left(\frac{d\pi}{ds}\right)_{s=\mu^2} = 1. \qquad (I.14)$$

To summarise our position, we may discuss the possibility of bootstrapping a c-particle of mass μ by exchanging it between two a-particles of mass m to produce the c-particle as a bound state with mass μ and correct residue by

(a) qualitatively discussing the sign of the discontinuity function $f_0(s)$ arising from the single particle exchange term (I.2);

(b) quantitatively by solving (I.6), (I.7), (I.11) and (I.12) with f_ℓ arising from (I.2), or in a weaker approximation replacing (I.6) by (I.9), or even weaker replacing (I.12) by (I.13).

We may evidently generalise the above analysis to:

(a) the bootstrap of a particle of higher spin than zero;

(b) the bootstrap of a set of particles belonging, say, to an SU_3 multiplet;

(c) the bootstrap of a set of multiplets, having different mass and spin for each multiplet.

As an example of a bootstrap, let me consider the ρ bootstrap in pion-pion scattering. The 2π system has $I=0$, 1, or 2, and we suppose that the scattering is caused by the exchange of a single $I=J=1$ ρ-meson. We use the spin 1, isospin 1 analogue of (I.2) and the crossing matrix $\alpha_{II'}$ which gives the contribution to the channel with isotopic spin I due to exchange of a particle of isotopic spin I'. We further use the approximation (I.9). Since single exchange gives a positive value for T_0 and T_2 in the positive region (the suffix denoting spin only), and α_{21} is negative, we expect no bound state or resonance in the isospin 2 channel (due to no zero of the relevant D-functions). On the other hand α_{01} is positive, so that resonances or bound states are expected in the isospin zero channel with spins 0 and 2. The scalar state is dubious, due to the lack of a repulsive angular momentum barrier in which a resonance can be trapped; the spin 2 state may be identified with the f_0 meson of mass 1250 MeV.

The channel of interest, with isotopic spin 1, has α_{11} positive, and since single ρ exchange gives a positive T_1 in the physical region, then the ρ-meson can appear as a resonance or bound state in this channel.

If we now turn to a quantitative discussion of the ρ bootstrap we meet the difficulty that the ρ has spin 1, so $T_1(s)$ increases with s. In general the exchange of a particle of spin ℓ produces a discontinuity function $f_L(s)$ which for large s behaves as $s^{\ell-1}$. Since $D_L(s)$ behaves at least as a constant for large s then the integral in (I.7) will require at least ℓ subtractions to achieve convergence. For the ρ we will need to have 1 subtraction, and the results will depend on this subtraction constant. Due to this we cannot expect to obtain separate values for both the mass and width of the ρ.

Besides this divergence difficulty there are other difficulties:

(a) The threshold behaviour of t_ℓ should be $t_\ell(s) \sim q^{2\ell}$ as $q \sim 0$. This can be achieved by writing down dispersion relations for $q^{-2\ell} t_\ell$, but this introduces difficulties of behaviour at infinity. To avoid this, one may use $s^\ell q^{-2\ell} t_\ell$, though this introduces a pole at $s=0$.

(b) The results depend on the value of the normalisation point s_0.

(c) States (bound or resonant) are found with the wrong sign of residue. These are called ghost states.

Difficulties (a) and (b) have been avoided by Shaw,[4] to which I refer you for details. I will ignore the difficulty of the ghost states. The final numerical results for the ρ bootstrap may be given either as self-consistent solutions for a given value of the subtraction constant, or alternatively as the value of the width of the resonance in the $I=J=1$ channel by exchange of a ρ meson with the experimental mass and width when the subtraction constant has been chosen so that this resonance has mass equal to that of the ρ meson. We use the latter form, so that the exchange of a ρ meson of the experimentally correct mass and width of 760 MeV and 108 MeV respectively gives rise to a ρ meson of mass 760 MeV provided a cut-off Λ is taken on the integration range in (I.7) of $72\, m_\pi$ (m_π = pion mass), the width of the produced ρ being 600 MeV. This is more than a factor of 5 larger than the experimental value.

We should note that this is only a partial bootstrap. The pion still has to be bootstrapped, and in particular could be considered as a bound state in $\rho\pi$ scattering caused by single pion exchange.

The general qualitative agreement but quantitative disagreement by a factor of 5 or so occurs in other partial bootstraps. An example of this is the reciprocal bootstrap in which nucleon exchange in $\pi - N$ scattering generates the $N^*(3,3)$ resonance, whilst N^* exchange in $\pi - N$ scattering generates the N. In the static limit, N exchange in $\pi - N$ scattering is most attractive in the isospin 3/2 spin 3/2 channel; N^* exchange is most attractive in the $I=\frac{1}{2}$, $J=\frac{1}{2}$ channel, so there is qualitative agreement with the reciprocal bootstrap requirement. However, there is about a factor of two difference between the πN coupling constant and the N^* width as predicted and known experimentally, assuming a cut-off at about the nucleon mass.

There is qualitative agreement for other bootstraps. Thus:

(a) The reciprocal bootstrap for the SU_3 octet P of 0^- mesons, the baryon octet B of $\frac{1}{2}^+$ baryons, and the baryon decuplet Δ of $(3/2)^+$ baryons has the property that B exchange in PB scattering is attractive in the $(3/2)^+$ decuplet, and Δ exchange is attractive in the $\frac{1}{2}^+$ octet.

(b) The $K^*(891)$ with $I=\frac{1}{2}$ may be bootstrapped by ρ and K^* exchange in πK scattering, provided the ρ coupling is stronger than the K^*.

(c) P,P scattering with the exchange of a vector meson octet V is attractive in the $I=J=1$ octet, so bootstrapping the V.

In order to improve the quantitative agreement with experiment, attempts have been made to include both further direct channels and

exchanged channels. The addition of further direct channels has been investigated[5] by adding in further closed channels, e.g., in $\pi\pi$ scattering the $\pi\omega$ and $K\bar{K}$ channels have been added in, using ρ, K^* and ϕ exchange. Using known coupling constants and adjusting the cut-off to reproduce a ρ with mass 760 MeV, it was found that the ρ width was reduced to 500 MeV. The addition of the $K\Sigma$ channel in the reciprocal bootstrap, with Λ, Σ and Ξ exchange, however, does not improve the agreement for the N^* width.

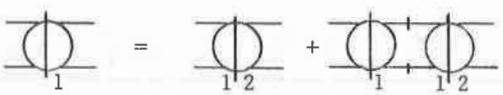
Additional exchanged particles have been treated by Pran Nath and collaborators,[6] who have attempted to bootstrap the ρ by exchange of the complete Regge trajectory on which the ρ lies and reproduce it in the direct channel. This has produced a ρ width of 125 MeV (again the ρ mass cannot be predicted), though by means of a number of approximations.

Both of these results indicate the need for a more complete treatment of the direct and exchange channels. In other words, inelastic effects and multi-particle exchange must be handled more satisfactorily before we expect to obtain satisfactory numerical agreement for bootstraps.

II. Off-Mass Shell Bootstraps

As a first step to handling many-particle states we use the Bethe-Salpeter equation instead of the N/D equation. The ladder approximation to the former equation will enable us to take some account of many-particle exchange, though we will only be able to satisfy two-particle unitarity exactly. The bootstrap equations resulting from this were first set up by Cutkosky;[7] we will see how to generalise these later so as to satisfy higher particle unitarity, and investigate the resulting equations. I should also mention that the Bethe-Salpeter equation allows higher spin composite particles to be considered without introducing further divergences, since the off-mass shell vertex functions have then a suitable high energy damping (in some sense a Regge behaviour).

The Bethe-Salpeter (B.S.) equation for two-particle scattering may be written graphically as



(II.1)

We denote by $M_1(p_1p_2p_3p_4)$ the off-mass shell S-matrix element for the scattering of particles of momenta p_1, p_2 to particles of momenta p_3, p_4 without a single particle intermediate state, this being the left hand side of (II.1), and $V(p_1p_2p_3p_4)$ the similar quantity with no one- or two-particle intermediate states, this being the first term in

the right hand side of (II.1). Also $D_F'(p)$ denotes the complete propagator for a single particle. Then (II.1) may be written in the form

$$M_1(p_1p_2p_3p_4) = V(p_1p_2p_3p_4) + \frac{1}{2}\int M_1(p_1p_2p_5p_6)\delta^4(p_1+p_2-p_5-p_6)$$

$$\times V(p_5p_6p_3p_4)D_F'(p_5)D_F'(p_6)\,d^4p_5\,d^4p_6. \qquad \text{(II.2)}$$

When the total energy $s = (p_1+p_2)^2$ is near the value for a resonance or bound state, say $s = M^2$, then M_1 has a simple pole in s, and we may write

$$M_1(p_1p_2p_3p_4) \sim F(p_1p_2)F(p_3p_4)(s-M^2)^{-1}. \qquad \text{(II.3)}$$

Then for s near M^2, (II.2) reduces to

$$F(p_3p_4) = \frac{1}{2}\int F(p_5p_6)\delta^4(p_3+p_4-p_5-p_6)\,D_F'(p_5)D_F'(p_6)V(p_5p_6,p_3,p_4). \qquad \text{(II.4)}$$

We may approximate (II.4) further by taking $D_F'(p) = (p^2-m^2)^{-1}$, and also take V to be described by exchange of the single particle of mass M, so

$$V(p_5p_6p_3p_4) = F(p_5p_3)F(p_6p_4)\left[(p_3-p_5)^2 - M^2\right]^{-1}$$

$$+ F(p_6p_3)F(p_5p_4)\left[(p_3-p_6)^2 - M^2\right]^{-1}. \qquad \text{(II.5)}$$

Then for s near M^2 we obtain from (II.4) and (II.5)

$$F(p_3,p_4) = \int F(p_5p_6)F(p_5p_3)F(p_6p_4)\delta^4(p_3+p_4-p_5-p_6)(p_5^2-m^2)^{-1}$$

$$\times (p_6^2-m^2)^{-1}\left[(p_5-p_3)^2 - M^2\right]^{-1} d^4p_5\,d^4p_6 \qquad \text{(II.6)}$$

which is a non-linear 'bootstrap' equation for the vertex function F. If we now approximate F by a constant, and take (II.6) to be valid other than at $(p_3+p_4)^2 = M^2$, we obtain the simplest form of B.S. bootstrap.[7]

Before we discuss this bootstrap and its extension to take account of corrections to (II.6) when $s \neq M^2$, let me first describe how

the B.S. equation reduces to the Schrödinger equation,[8] and precisely what the wave function is. We will do this for the ladder approximation to the bound state form of the B.S. equation, the exchanged particle having mass μ. We replace the F's in (II.5) by a constant, gm, (so g is dimensionless), and substitute in (II.4). Inserting the correct factors of (2π) and i, and taking $p_3 = p+q$, $p_4 = p-q$, with $p = (p_o, \vec{o})$ in the centre of mass system, and $F(p+q, p-q) = \phi(q)$, (II.4) becomes

$$\phi(q) = \frac{ig^2m^2}{(2\pi)^4} \int \frac{d^4k\, \phi(k)}{[(q-k)^2-\mu^2][(p+k)^2-m^2][(p-k)^2-m^2]} . \qquad (II.7)$$

We neglect the retardation effects in the exchange denominator in (II.7), which means replacing $[(q-k)^2 - M^2]$ in the denominator by $-[(\vec{q}-\vec{k})^2 + M^2]$. Then $\phi(q)$ is independent of q_o, and so the integration over k_o may be done exactly in (II.7). We close the integration along the real k_o axis by a semi-circle in the upper half k_o-plane, and pick up the contribution from the poles at $k_o = -p_o - \omega, p_o - \omega$, with $\omega = (\vec{k}^2 + m^2)^{\frac{1}{2}}$. Then (II.7) becomes

$$\phi(\vec{q}) = -\frac{g^2m^2}{4(2\pi)^3} \int \frac{d^3\vec{k}\, \phi(\vec{k})}{[(\vec{q}-\vec{k})^2+\mu^2]\omega(p_o^2-\omega^2)} . \qquad (II.8)$$

The binding energy B of the bound state is defined as

$$2p_o = 2m - B.$$

We assume that $\phi(\vec{k})$ is only appreciable if $|\vec{k}| \ll m$, and also that $B \ll m$, so $p_o^2 - \omega^2 \sim -Bm - \vec{k}^2$, and we replace ω by m in the denominator of (II.8), so it now becomes

$$\phi(\vec{q}) = \frac{g^2}{4(2\pi)^3} \int \frac{d^3\vec{k}\, \phi(\vec{k})}{[(\vec{q}-\vec{k})^2+\mu^2](B+\vec{k}^2/m)} . \qquad (II.9)$$

If we define $\psi(\vec{q})$ as

$$\phi(\vec{q}) = (B+\vec{q}^2/m)\psi(\vec{q})$$

then

$$(B+\vec{q}^2/m)\psi(\vec{q}) = \frac{g^2}{4(2\pi)^3} \int \frac{d^3\vec{k}\, \psi(\vec{k})}{[(\vec{q}-\vec{k})^2+\mu^2]} . \qquad (II.10)$$

In position space, (II.10) becomes

$$(-\nabla^2/m + B)\tilde{\psi}(\vec{r}) = \frac{g^2}{4m^2} \cdot \frac{e^{-\mu r}}{r} \cdot \tilde{\psi}(\vec{r}) \qquad \text{(II.11)}$$

where $\tilde{\psi}$ is the Fourier transform of ψ. Since m is twice the reduced mass, then (II.11) is the correct Schrödinger equation for two particles of mass m interacting through a Yukawa potential of range μ^{-1}. We now return to the wave function. Using the same technique of contour integration, and the approximations preceding (II.9),

$$\psi(\vec{q}) = \frac{2m^2}{\pi i} \int \frac{dq_o \phi(q)}{[(p+q)^2 - m^2][(p-q)^2 - m^2]} \qquad \text{(II.12)}$$

under the assumption that $\phi(q)$ is independent of q_o. If we define the co-ordinate space wave function

$$f(x_1 - x_3, x_2 - x_3) = \int \exp\left[i(p_1 x_1 + p_2 x_2 + p_3 x_3)\right] \cdot \prod_{i=1}^{3} d^4 p_i \delta^4(p_1 + p_2 + p_3) D_F(p_1) D_F(p_2) F(p_1, p_2), \qquad \text{(II.13)}$$

then (II.13) may be written in terms of $\phi(q)$ by means of the equal time wave function as

$$f(\vec{r}_1, t; \vec{r}_2 t) = \exp i\big(2p_o t + \vec{p} \cdot (\vec{r}_1 + \vec{r}_2) + \vec{q} \cdot (\vec{r}_1 - \vec{r}_2)\big) D_F(p+q) D_F(p-q) \times F(p+q, p-q) dp_o dq_o d^3\vec{p}\, d^3\vec{q}.$$

We use (II.12) to obtain

$$f(\vec{r}_1, t; \vec{r}_2 t) = (\pi i/2m^2) \int \exp\left[i\big(2p_o t + \vec{p} \cdot (\vec{r}_1 + \vec{r}_2) + \vec{q} \cdot (\vec{r}_1 - \vec{r}_2)\big) \psi(\vec{q}; \vec{p}, p_o)\right] \times d^3\vec{q}\, dp_o\, d^3\vec{p} \qquad \text{(II.14)}$$

where we have included the explicit dependence of ψ on the total energy and momentum $p_o, \vec{p}$. Then since $\tilde{\psi}$ is the Fourier transform of ψ in the variable $\vec{q}$, we have

$$\tilde{\psi}(\vec{r}_1-\vec{r}_2;\vec{p},p_o) = \frac{2m^2}{\pi i}\int f(\vec{r}_1,t;\vec{r}_2,t)e^{-2ip_o t - i\vec{p}.(\vec{r}_1+\vec{r}_2)}\,dt\,d^3(\vec{r}_1+\vec{r}_2) \tag{II.15}$$

which is as expected: the projection with respect to the centre of mass variables of the equal time Green's function is equal to the Schrödinger wave function in co-ordinate space.

Let me now return to the simplest bootstrap derived from the B.S. equation—the Cutkosky bootstrap.[7] I will consider a set of N vector mesons with equal mass. The potential of (II.5) can be split up into a term symmetric in the space variables and one which is antisymmetric, under the approximation that the vertex function F is the constant G_{ijk}, where the i, j, k refer to the particular mesons considered, so take the values 1,2,...N and G_{ijk} is completely antisymmetric in its indices. The antisymmetric part of the potential is

$$\left(G_{\ell rj}G_{mrk} - G_{\ell rk}G_{mrj}\right)\left[(q-k)^2 - m^2\right]^{-1} - \left[(q+k)^2 - m^2\right]^{-1}.$$

We take all external momenta to be on the mass shell so that (II.5) becomes in this case

$$G_{ijk} = \lambda^{-1} V_{ij,\ell m} G_{\ell mk} \tag{II.16}$$

where $V_{ij,\ell m} = (G_{ir\ell}G_{jrm} - G_{irm}G_{jr\ell})$ and λ^{-1} is the triangle function evaluated on the mass shell:

$$\lambda^{-1} = \int d^4k\left[(p-k)^2 - m^2\right]^{-1}\left[(p+k)^2 - m^2\right]^{-1},$$

$$p^2 = (p-q)^2 = (p+q)^2 = m^2.$$

Strictly speaking, the triangle function should involve particles of spin one on its internal lines and also involve a spin one projection on its external lines; I have dropped these complications, since the essential part of the argument involves the way the indices on G are related by (II.16).

A further relation on the coupling constants G_{ijk} is obtained by requiring that self-energy effects due to meson pair formation do not alter the meson masses differently. Since these self-energy effects are proportional to $\Pi_{ab} = G_{acd}G_{cdb}$, we require (with suitable normalisation)

$$G_{acd}G_{cdb} = \delta_{ab}. \qquad \text{(II.17)}$$

If we consider V as a matrix with respect to its initial and final pair of indices, (II.16) requires that V have a set of N degenerate eigenvectors belonging to the eigenvalue λ. Since there is a total of $\frac{1}{2}N(N-1)$ independent asymmetric tensors of rank 2, then there must be $\frac{1}{2}N(N-3)$ orthogonal eigenvectors. Let their eigenvalues be λ_i, with degeneracy d_i, so that

$$\mathrm{Tr}V^2 = N\lambda^2 + \sum_i d_i \lambda_i^2.$$

But by direct computation from the definition of V and (II.16), (II.17),

$$\mathrm{Tr}V^2 = 2N - N\lambda$$

so

$$2 - \sum d_i \lambda_i^2 N^{-1} = \lambda(1+\lambda)$$

and thus $\lambda \leq 1$, with $\lambda = 1$ if and only if all $\lambda_i = 0$. Suppose there is a gauge group with generators $\bar{G}^A_{ab}$ (A running over a finite set, a,b taking the values $1,2,\ldots N$) under which the G_{ijk} are invariant, so

$$G_{xbc}\bar{G}^{\alpha}_{xa} + G_{axc}\bar{G}^{\alpha}_{xb} + G_{abx}\bar{G}^{\alpha}_{xc} = 0. \qquad \text{(II.18)}$$

If we multiply (II.18) by G_{bad} and sum on a and b, then

$$\bar{G}^A_{cd} = V_{cd,ab}\bar{G}^A_{ab}$$

so that $\bar{G}^A_{cd}$ is an eigenvector of V belonging to the eigenvalue 1. Then either

(a) $\bar{G}$ is orthogonal to all the G's, $\bar{G}^A_{ab}G_{abc} = 0$, $1 \leq c \leq N$. Then at least one of the λ_i is one, and $\lambda < 1$. If we assume that increasing λ causes a decrease in the mass of the bootstrapped mesons, we see that there will be a set of more massive bootstrapped mesons belonging to the eigenvalue $\lambda_i = 1$. This is inconsistent with the idea of bootstrapping the lowest masses first, and extending to higher masses, so we reject this possibility.

(b) $\bar{G}$ is not orthogonal to all the G's, so $\lambda = 1$, and all the λ_i's are zero.

Then we have that V is the sum

$$V = \sum \lambda_n P_n$$

where λ_n run over the eigenvectors of V, P_n the projection onto the eigensubspace belonging to λ_n, so

$$V = G_{abr}G_{cdr}$$

or

$$G_{abr}G_{cdr} + G_{bcr}G_{adr} + G_{car}G_{bdr} = 0. \qquad \text{(II.19)}$$

Equation (II.19) is the Jacobi identity satisfied by the structure constants of a Lie group and, combined with (II.17), we see that the N mesons belong to the regular representation of a compact semi-simple Lie group.

We may extend this approach to consider symmetry breaking, but I do not want to go into this here, but restrict myself to bootstraps.

The next step beyond the Cutkosky bootstrap is to treat the non-linear vertex equation (II.6) more completely. It is also of interest to consider more general bootstraps, which include more than two particles in the direct channel, one in the exchanged channel. It is easy to see how to write down these more general bootstraps by taking suitable Feynman diagrams and replacing point vertices by complete vertices, where a vertex now may have three or more legs. I will return to the study of these bootstraps later, under the heading of generalised groups. But first we must obtain a consistent method to continue the B.S. bootstrap (II.6) off the mass shell.

III. A Field Theory for Composite Particles

In order to continue off-mass shell for bootstraps we will set up a field theory of bootstraps. We do this first for a composite particle, and obtain a bootstrapped theory by taking all particles to be composites.

We obtain an indication of a basic condition on a particle to be a composite if we note that the wave function renormalisation constant Z_c for a particle c is defined by

$$Z_c = |\langle \text{bare c} | \text{physical c} \rangle|^2 \qquad \text{(III.1)}$$

where $|\text{bare}\rangle$ and $|\text{physical}\rangle$ denote the bare and physical one-particle states; it is the interaction with other particles which clothes the bare c-particle, making it physical. Then we have a

composite c-particle evidently when there is no bare c-particle, i.e.,

$$Z_c = 0. \tag{III.2}$$

This condition (III.2) is an implicit condition relating renormalised coupling constants and masses for the c-particle and other elementary and composite particles with which it interacts. We expect all renormalisation constants to be unobservable, though we do not expect that the consequences of imposing (III.2) will also be unobservable. To see what these consequences are on the field operators describing the various particles, let me take the simple model of a scalar neutral c-particle interacting with a scalar neutral a-particle. If the bare fields and masses for the particles are ϕ_o, ψ_o and m_o, M_o for the a- and c-particle respectively, we take the Lagrangian density to be

$$L = \tfrac{1}{2}(\partial_\mu \phi_o)^2 + \tfrac{1}{2}(\partial_\mu \psi_o)^2 - \tfrac{1}{2}m_o^2\phi_o^2 - \tfrac{1}{2}M_o^2\psi_o^2 + g_o\phi_o^2\psi_o + f(\phi_o) \tag{III.3}$$

where f is any polynomial in the field ϕ_o (it is essentially the self-interaction which will bind the c-particle in the limit $Z_c = 0$). The equations of motion arising from (III.3) are

$$\left(\Box + m_o^2\right)\phi_o = 2g_o\phi_o\psi_o + f'(\phi_o)$$

$$\left(\Box + M_o^2\right)\psi_o = g\phi_o^2.$$

We renormalise these equations in standard fashion, introducing the renormalised fields ϕ, ψ and wave function renormalisation constants Z_a, Z_c by

$$\phi_o = Z_a^{\frac{1}{2}}\phi, \quad \psi_o = Z_c^{\frac{1}{2}}\psi$$

and the explicit counter terms in the Lagrangian by

$$L = \tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 + \tfrac{1}{2}(\partial_\mu\psi)^2 - \tfrac{1}{2}M^2\psi^2 + \tfrac{1}{2}(Z_a-1)\left[(\partial_\mu\phi)^2 - m^2\phi^2\right]$$

$$+ \tfrac{1}{2}(Z_c-1)\left[(\partial_\mu\psi)^2 - M^2\psi^2\right] + g_oZ_aZ_c^{\frac{1}{2}}\phi^2\psi + \tfrac{1}{2}m^2Z_a\phi^2 + \tfrac{1}{2}M^2Z_c\psi^2$$

$$+ F(\phi) \tag{III.4}$$

where $m^2 = m_o^2 + \delta m^2$, $M^2 = M_o^2 + \delta M^2$, $F(\phi) = f(\phi_o)$. The resulting

field equation for ψ is

$$(\Box + M^2)\psi = (1-Z_c)(\Box + M^2)\psi + \delta M^2 Z_c \psi + g_o Z_a Z_c^{\frac{1}{2}} \phi^2 . \quad \text{(III.5)}$$

If we let $Z_c \to 0$ in (III.4), this equation becomes

$$\psi(x) = \lambda \phi^2(x) \quad \text{(III.6)}$$

where

$$\lambda = \lim_{Z_c \to 0} \left(-g_o Z_a / \delta M^2 Z_c^{\frac{1}{2}} \right),$$

provided this limit exists. In other words, $\psi(x)$ is a local function of $\phi(x)$, evaluated at the same point x, in the limit $Z_c \to 0$.

We have derived (III.6) in a very sloppy manner, with no discussion of the topology to be imposed on the operator equation (III.5) or on the operators. Since we do not even know that a solution to (III.5) exists, it is not possible at the present time to give such a discussion. There are, however, two non-rigorous approaches which we may follow to justify (III.6) with more believability than the discussion we just gave. One of these follows Zimmerman,[9] who proved that a composite c-particle may be described by an interpolating Heisenberg field ψ which is quasi-local in the elementary particle field ϕ; we have to impose the further requirement that ψ be a local function of ϕ to obtain (III.6), provided this limit is defined with suitable counter-terms.[9] Then (III.6) may be derived from the Lagrangian

$$L = \frac{a}{2} (\psi - \lambda \phi^2)^2 + L_o(\phi) \quad \text{(III.7)}$$

where L_o is the part of the Lagrangian depending on ϕ only, and assumedly binding the two a-particles to make a composite c-particle. L of (III.7) is the limit of L of (III.5) as $Z_c \to 0$, in some sense. We have not bridged the gap of how to define this limit of Z_c approaching zero by the arguments of Zimmerman, but only justified the limiting result (III.6). We will see later that this limiting process is a very delicate one, and is not yet understood.

The second approach which I remarked on above is to discuss the $Z_c \to 0$ limit by means of the Green's functions equations (GFE's) which may be written down from (III.4).[10] The approach through G.F.E.'s has the added advantage that it allows us to set up in a simple manner bootstraps which generalise the B.S. bootstrap which we discussed in the previous chapter.

To use the G.F.E.'s, let me first define them for a field alone as

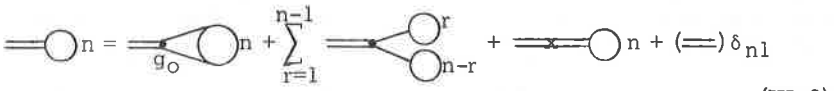
$$\langle 0|T\big(\phi(x_1)\ldots\phi(x_n)\big)|0\rangle = G\big(x_1,\ldots,x_n\big)$$

and in Fourier space

$$\tilde{G}(p_1\ldots p_n)\delta^4\Big(\sum_{j=1}^{n} p_j\Big) = \int \prod_{i=1}^{n} d^4x \left[\prod_{i=1}^{n}\big(\Box_i^2+m^2\big)\right] G(x_1\ldots x_n)$$

$$\times\ \exp\left[i\sum_{j=1}^{n} p_j x_j\right].$$

By the asymptotic condition the S-matrix elements for a process involving a total of n particles is $\tilde{G}(p_1\ldots p_n)$, for $p_j^2=m^2$, $j=1,\ldots,n$. We now extend the definition of the functions $\tilde{G}$ to include the c-particles as well. We now wish to consider the equations which relate the various Green's functions as follows from (III.5). We will write these down in graphical form, since this gives the most transparent way of seeing what is happening, and also enables a considerable saving of space. We denote the connected part of $\tilde{G}(p_1\ldots p_n)$ (that part having no δ^4 functions involving a subset of the $p_1\ldots p_n$) by ◯n, $iD_F^{(a)}(p)$ by — and $iD_F^{(c)}(p)$ by ═. The way a c-particle double line is joined to any graph is, by (III.5):[10]



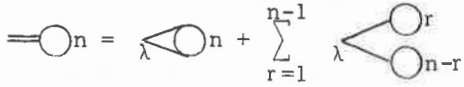
(III.8)

where each line carries a momentum four-vector, which is integrated over on internal lines, with energy-momentum conservation at each vertex, the dotted vertex denotes $i(2\pi)^4 g_o$, there is an extra factor of $(2\pi)^{-4}$ on each internal propagator, and x denotes

$$[(1-Z_c)(p^2-M^2) + Z_c\delta M^2].$$

It is easy to see that iteration of (III.8), together with the analogous equation which indicated how a single a-particle propagator is attached to a graph, produces exactly the Feynman graphs of perturbation theory, with all mass and wave function renormalisation counter terms correctly inserted.

If we let $Z_c \to 0$ in (III.8) we obtain, for $n \neq 1$,



(III.9)

which is just a restatement of (III.6), since (III.6) as a relation between Green's functions is

$$G_c(x_1, \ldots) = G_{aa}(x, x, \ldots) \tag{III.10}$$

where the subscripts c and a denote which fields are involved and the number of times these subscripts appear equals the number of fields. Then (III.10) results in (III.9), if we use that for a function $f(x,y)$ with

$$\tilde{f}(p) = \int e^{ipx} f(x,x)dx$$

then

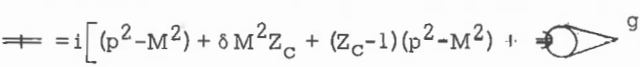
$$\tilde{f}(p) = \int \tilde{f}(p-k,k)dk$$

where

$$\tilde{f}(p,k) = \int e^{ipx+iky} f(x,y)dx\,dy.$$

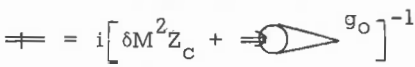
So far we have not shown that any composite c-particle actually occurs in the a-particle scattering amplitude. To do this we have to show that a pole exists in the a-particle scattering amplitude at $s = M^2$. We will regard M^2, g_r^2 (the renormalised coupling constant) as the basic variables, and approach the curve $Z_c(M^2, g_r^2) = 0$ in such a way that λ is finite and non-zero.

For the complete c-propagator, which we denote by ⟹, it follows from (III.9) that



(III.11)

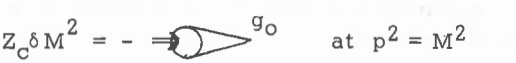
where ⟹ n = ⟹ n. Then as $Z_c \to 0$, (III.11) becomes



or

(III.12)

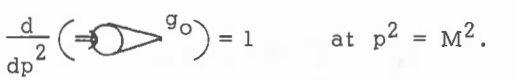
At $p^2 = M^2$ we require $\sim i(p^2 - M^2)^{-1}$, so achieving both mass and charge renormalisation. Thus we require



or

(III.13)

and also



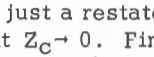
(III.14)

Since the proper self-energy function $\Pi(p^2) =$  (correct to all orders in g_0), then (III.14) may be rewritten as

$$1 - \Pi'(M^2) = 0. \qquad \text{(III.15)}$$

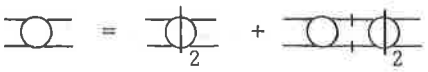
But we see from (III.11) that in order that the residue of the right hand side is i at $p^2 = M^2$ it is necessary that $Z_c = 1 - \Pi(M^2)$, so that (III.15) is just a restatement of $Z_c = 0$, and is automatically satisfied in the limit $Z_c \to 0$. Finally (III.13), in the non-relativistic limit discussed earlier in (III.13) et seq., becomes

$$(\pi i/2m^2\lambda) = \tilde{\psi}(\vec{0}) = \int \psi(\vec{q})\, d^3\vec{q}\,. \qquad \text{(III.16)}$$

For λ finite and non-zero then (III.13) requires that $F(p+q, p-q)$ does not vanish identically in q at $p^2 = M^2$, so that the pole term  has non-zero residue at $p^2 = M^2$ in the a-particle scattering amplitude. Thus the composite c-particle is actually present.

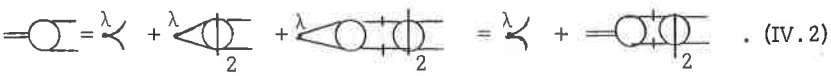
IV. Composite Potential and the Jin-MacDowell Cancellation

We now turn to a more detailed discussion of the manner in which the limit $Z_c \to 0$ is achieved. In particular we wish to determine what the 'potential' is for $Z_c \neq 0$. For $Z_c = 0$ this potential is defined by means of the two-particle exposure



(IV.1)

where $\dashv\!\vdash$ denotes the complete a-particle propagator, and the relativistic potential is the first term on the right hand side of (IV.1), having no intermediate state involving two a-particles. It is through scattering through this potential that the composite c-particle arises, i.e., iteration of (IV.1) diverges at $s = M^2$. If we now use (III.6) or its G.F.E. form (III.9) we will obtain from (IV.1)

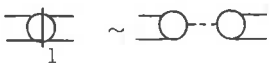


. (IV.2)

We see from (IV.2), by comparison with (III.8) for $n = 2$, that the effective coupling 'constant' for the c-particle in interaction with two a-particles is $\lambda(p^2 - M^2)$, and is actually zero on the mass shell. This is in fact necessary so that (IV.2) becomes a homogeneous integral equation for the composite particle wave function on the mass shell, so is an eigenvalue equation for the bound state mass M.

We now enquire into the possible form of the potential when $Z_c \neq 0$. We wish to choose it as part of the a-particle scattering amplitude which is constructed in a straightforward fashion, and which has no c-particle pole in it. The only such quantity seems to be the potential arising in the B.S. equation of (II.1), being the first term on the right hand side of that equation (the suffix 1 now denoting no c-particle state). We will now see that such a natural choice of potential leads to the impossibility of continuing off-mass-shell in a non-trivial fashion.

To show this, we suppose that the one c-particle irreducible amplitude shown on the left of (II.1) possesses a single particle pole at a point $s = \mu^2$, when $Z_c \neq 0$ (we do not consider the pathological behaviour corresponding to no pole at all in the amplitude for $Z_c \neq 0$, this pole only arising at $Z_c = 0$). Then if we denote $(p^2 - \mu^2)^{-1}$ by a dashed line, we will have near $s = \mu^2$ that

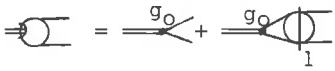


(IV.3)

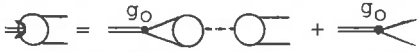
so from (II.1) and (IV.3) at $s = \mu^2$

.

Now the mass-renormalised vertex function equation arising from (III.8) with $n = 2$ is

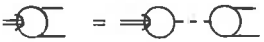


so near $s=\mu^2$ this becomes



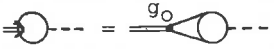
(IV.4)

and we may neglect the last term on the right hand side of (IV.4). Thus the vertex function has a pole at $s=\mu^2$, and if we take there

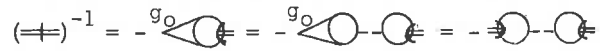


(IV.5)

then

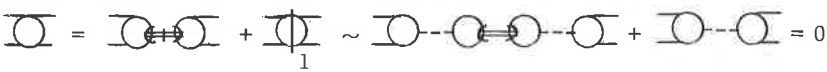


Further, (III.11) at $p^2=\mu^2$ becomes



(IV.6)

so that the c-particle propagator is zero at $p^2=\mu^2$. If we combine this zero with the poles in the c-particle vertex function, then we find[11] in the a-particle scattering amplitude at $s=\mu^2$

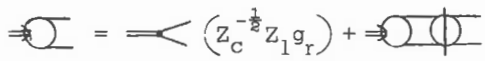


and the pole at $s=\mu^2$ does not appear physically. This was first shown in the two-particle unitarity approximation, and is known as the Jin-MacDowell cancellation.[12]

So far we have kept $Z_c \neq 0$, and have an elementary c-particle. We now want to make the composite pole at μ^2 to coincide with the elementary pole at M^2, and finally have the residue at this composite pole equal to that corresponding to the physical coupling constant g_r.

To obtain $M^2=\mu^2$, we use that g_o and g_r are related by $g_r Z_1 = g_o Z_c^{\frac{1}{2}}$, where Z_1 is the vertex function renormalisation constant. Then we see that near $s=M^2$ and on mass shell for the a-particles, the left hand side of (IV.4) is $Z_1^{-1} g_r$ (by the definition of g_r as the value of the vertex function when all its external particles are on their mass shells), while the right hand side is proportional to $(M^2-\mu^2)^{-1}$. Thus to obtain $M=\mu$, and keep g_r finite, we need to

take $Z_1 = 0$. If we take now the charge renormalised vertex function equation



(IV.7)

and set $Z_1 = 0$, we see that the only solution to (IV.7) for the renormalised vertex function is zero everywhere off the c-particle mass shell, and non-zero only on it. Thus there is no non-trivial off-mass shell vertex function for the composite particle, in contradistinction to the non-trivial off-mass shell continuation obtained from (IV.2).

We finally turn to the residue condition. We see that this is obtained by choosing $Z_c = 0$, since then from (IV.6) we have  at $s = M^2$, so that since  at $s = M^2$ (and the a-particles are on their mass shell) then from (IV.5) the composite wave function takes the value g_r on the mass shell. Thus we need both $Z_1 = 0$ and $Z_c = 0$. There has been a great deal of discussion of the need for $Z_1 = 0$ in composite particle theory.[1] However, we see that if we require $Z_1 = 0$ then there is no non-trivial off mass shell continuation. Since we wish to set up a field theory of composites based on (III.6) we cannot impose this extra condition in addition to $Z_c = 0$.

Beyond this we notice that the correct limiting form of vertex function equation (IV.2) has a coupling constant $\lambda(p^2 - M^2)$ which is momentum dependent. Evidently to achieve this by a suitable limiting process from a theory which, before the limit is achieved, has a momentum independent coupling constant, will require a non-uniform limiting process. The evident non-uniformity and subtlety of this limiting process requires much further discussion before it is better understood. It may be better to avoid this subtlety by taking (III.6) directly and attempting to quantise the composite without starting from an elementary particle. We will return to this point after an example of a relativistic model which has a composite obtained by taking $Z_c = 0$ but $Z_1 \neq 0$.

V. A Composite in the Relativistic Lee Model

We consider the relativistic Lee model discussed recently by Yndurain.[13] This model is identical with the original Lee model,[14] except that in the free Hamiltonian the energies of the V and N particles are

$$E_V(\vec{p}) = \left(m_V^2 + \vec{p}^{\,2}\right)^{\frac{1}{2}}$$

and

$$E_N(\vec{p}) = \left(m_N^2 + \vec{p}^{\,2}\right)^{\frac{1}{2}},$$

respectively, so that the total Hamiltonian now becomes

$$H = H_o + H_{int},$$

$$H_o = \int d^3\vec{p}\, V_{\vec{p}}^{+} V_{\vec{p}} E_V(\vec{p}) + \int d^3\vec{p}\, N_{\vec{p}}^{+} N_{\vec{p}} E_N(\vec{p}) + \int d^3\vec{k}\, a_{\vec{k}}^{+} a_{\vec{k}} \omega(\vec{k})$$

$$H_{int} = \int d^3\vec{p}\, \delta E_V(\vec{p}) V_{\vec{p}}^{+} V_{\vec{p}} + \int \lambda \frac{d^3\vec{p}\, d^3\vec{k}}{(E_V(\vec{p}) E_N(\vec{p}-\vec{k})\, \omega(\vec{k}))^{\frac{1}{2}}} \times$$

$$\left[f(\vec{p},\vec{k}) \cdot V_{\vec{p}}^{+} N_{\vec{p}-\vec{k}} a_{\vec{k}} + \text{h.c.} \right].$$

The local limit of H is for $f \equiv 1$. The physical one-particle N and θ states are identical to the bare ones, while the V state will be of form

$$|V,p\rangle = Z_V^{\frac{1}{2}}(p) \left\{ V_{\vec{p}}^{+} |0\rangle + \int d^3\vec{k}\, \phi(\vec{k},\vec{p})\, N_{\vec{p}-\vec{k}}^{+} a_{\vec{k}}^{+} |0\rangle \right\} . \qquad (V.1)$$

If we solve the eigenvalue equation

$$H|V,p\rangle = E_V(p)|V,p\rangle$$

by taking scalar products with $V_{\vec{p}}^{+}|0\rangle$ and $N_{\vec{p}-\vec{k}}^{+} a_{\vec{k}}^{+}|0\rangle$ we find that

$$\phi(\vec{k},\vec{p}) = g^*(\vec{p},\vec{k}) \left[E_V(\vec{p}) - E_N(\vec{p}-\vec{k}) - \omega(\vec{k}) \right]^{-1}$$

where

$$g(\vec{p},\vec{k}) = \lambda\, f(\vec{p},\vec{k}) \left[8 E_V(\vec{p})\, E_N(\vec{p}-\vec{k})\, \omega(\vec{k}) \right]^{-\frac{1}{2}}$$

and

$$\delta E_V(\vec{p}) = \int d^3\vec{k}\, |g(\vec{p},\vec{k})|^2 \left[E_N(\vec{p}-\vec{k}) + \omega(\vec{k}) - E_V(\vec{p}) \right]^{-1}.$$

Further, the normalisation condition on $|V,p\rangle$ requires

$$Z_V^{-1} = 1+\lambda^2 \int \frac{|f(\vec{p},\vec{k})|^2 d^3\vec{k}}{8E_V(\vec{p})E_N(\vec{p}-\vec{k})\omega(\vec{k})[E_V(\vec{p}) - E_N(\vec{p}-\vec{k}) - \omega(\vec{k})]^2} . \qquad (V.2)$$

We see that in the local limit, or more generally if $f(\vec{p},\vec{k})$ does not depend on $\vec{k}$, then Z_V is finite, while δE_V is logarithmically divergent. We see, however, that in general Z_V will be a function of p unless $f(\vec{p},\vec{k})$ is chosen suitably. We will return to the implications of this later; it is still true that Z_V is the probability of finding a bare V-particle in the single V-particle state (V.1), since

$$Z_V = |\langle 0|V_{\vec{p}}^{+}|V,p\rangle|^2 .$$

The scattering matrix may be computed as in the standard Lee model,[14] and the resulting s-wave scattering amplitude A(s) (the only partial wave in which there is scattering) is

$$A(s) = \frac{\pi\lambda^2|f(\vec{o},\vec{k})|^2}{2M_V(s^{\frac{1}{2}} - M_V)} \times \left\{1 - \int \frac{\lambda^2\pi^2 p^2 dp|f(\vec{o},\vec{p})|^2}{2M_V E_N(\vec{p})\omega(\vec{p})[s^{\frac{1}{2}} - E_N(\vec{p}) - \omega(p) + i\varepsilon][E_N(\vec{p}) + \omega(\vec{p}) - M_V]}\right\}^{-1} \qquad (V.3)$$

where $s^{\frac{1}{2}} = \omega(\vec{k}) + E_N(\vec{k})$. There are no ghosts in this theory, since for $s^{\frac{1}{2}} < M_N + \mu$ the integral in the denominator of (V.3) is negative, so the denominator can never vanish. There is only one pole in A(s), at $s = M_V^2$, being the expected V-particle pole. We define the residue at this pole to be the square of the renormalised coupling constant so that

$$\lambda_r = \lambda^2|f(\vec{o},\vec{k}_o)|^2\left\{1 + \frac{\pi\lambda^2}{2M}\int_o^\infty \frac{p^2dp|f(\vec{o},\vec{p})|^2}{E_N(\vec{p})\omega(\vec{p})[E_N(\vec{p})+\omega(\vec{p})-M_V]^2}\right\}^{-1}$$

where

$$\vec{k}_o^{\,2} = \left(M_V^2 + \mu^2 - M_N^2\right)\Big/ 2M_V^2 - \mu^2 ,$$

and we suppose that $f(\vec{o},\vec{k}_o)$ is defined by suitable analytic continuation in $\vec{k}$ to $\vec{k}_o$. We take from now on $f(\vec{p},\vec{k}) = f(\vec{p})$, so this continuation is trivial, and A(s) will be independent of f(p). Further A(s) will be analytic in the cut s-plane, cut from $(M_N+\mu)^2$ to $+\infty$ and

from 0 to $-\infty$ (from the square-root function), with a pole at M_V^2. Thus except for crossing, A(s) has all the correct properties of relativistic invariance and analyticity (in spite of the non-local form factor f(p)). Further the renormalised coupling constant λ_r is finite, even for finite bare coupling constant λ.

In order to relate to what was discussed in the preceding section, I will take

$$f(p) = \left(h(p)a^2\right)^{-1} \tag{V.4}$$

where

$$h(p) = \int d^3\vec{k} \left\{ 8E_V(\vec{p})E_N(\vec{p}-\vec{k})\omega(\vec{k})\left[E_V(\vec{p}) - E_N(\vec{p}-\vec{k}) - \omega(\vec{k})\right]^2 \right\}^{-1}$$

is a finite function of p, and a is an arbitrary constant. Then

$$Z_V^{-1} = 1 + \lambda^2/a^2$$

and Z_V is independent of p, while

$$\lambda_r^2 = \lambda^2 Z_V(o) = \lambda^2\left(1+\lambda^2/a^2\right)^{-1}.$$

If we define the vertex function renormalisation constant Z_1 as

$$\lambda_r^2 = Z_1^{-1} Z_V \lambda^2, \tag{V.5}$$

then

$$Z_1 = 1. \tag{V.6}$$

On the other hand, if we do not choose (V.4) then $Z_V = Z_V(p)$ depends on p, and Z_1, as defined by (V.5) will also depend on p. In particular, in the local limit $f \equiv 1$ then

$$Z_1(p) = \left[Z_V(p)/Z_V(o)\right]^{\frac{1}{2}}. \tag{V.7}$$

We now make the physical V-particle into a composite V-particle by taking $Z_V \to 0$. In order to achieve this, even when Z_V depends on p, we take

$$\lambda \to \infty. \tag{V.8}$$

In this limit,

$$|V,p\rangle = h^{-\frac{1}{2}}(p)\int d^3\vec{k}N^{+}_{\vec{p}-\vec{k}}a^{+}_{\vec{k}}|o\rangle\left[8E_V(\vec{p})E_N(\vec{p}-\vec{k})\omega(\vec{k})\right]^{-\frac{1}{2}}$$

$$\times\left[E_V(\vec{p})-E_N(\vec{p}-\vec{k})-\omega(\vec{k})\right]^{-1} \qquad (V.9)$$

and is a linear combination of (N,θ) states, as it should be for a composite V state. At the same time, either (V.6) persists in the composite limit, if we choose (V.4), or in the local limit $f\equiv 1$, we have

$$Z_1 \to \left[h(o)/h(p)\right]$$

which again is non-zero, though p-dependent. The model does not possess crossing symmetry (which is automatically violated in any model with the Lee model selection rules), though crossing was not explicitly used in our discussion in the last chapter and, in fact, could have been dispensed with entirely. Thus this model is a satisfactory counter-example to the condition $Z_1 = 0$ for a composite.

We see that the $N\theta$ scattering amplitude in the composite limit is

$$A(s) = \left[2M_V(s^{\frac{1}{2}}-M_V)\right]^{-1}$$

$$\times\left\{\int p^2dp\left[2M_VE_N(\vec{p})\left(s^{\frac{1}{2}}-E_N(\vec{p})-\omega(\vec{p})\right)\left(E_N(\vec{p})+\omega(\vec{p})-M_V\right)\right]^{-1}\right\}^{-1}$$

so that $A(s)\sim(\ln s)^{-1}$ as $s\sim\infty$. For the elementary V-particle, from (V.3), $A(s)\sim s^{-\frac{1}{2}}$, so the model has the usual property that the high energy behaviour becomes worse if the V-particle is made composite.

We can attempt to define a composite operator as

$$V^{+}(\vec{p})|o\rangle = |V,\vec{p}\rangle$$

and

$$V(\vec{x}) = \int e^{i\vec{p}\vec{x}}V(p)\left[2E_V(\vec{p})\right]^{-\frac{1}{2}}d^3\vec{p} + \text{herm. conj.}$$

However, it is evident that $V(\vec{x})$ will be a non-local function of the $\vec{x}$-space operators $N(\vec{x})$, $\theta(\vec{x})$, due to the appearance of the factor $[E_V(\vec{p})-E_N(\vec{p}-\vec{k})-\omega(\vec{k})]^{-1}$ in the denominator of (V.9).

VI. Classification of Particles

We now ask the question: can we determine which particles are composite, which elementary, by evaluation of their wave function renormalisation constants (and using suitable experimental quantities, if necessary)? As remarked earlier, the Z's are usually considered to be unobservable; however, we noted for the relativistic Lee model that the value of Z_V determines the high energy behaviour of N-θ scattering. A difficulty in this is that for any realistic relativistic field theory all the Z's appear to be zero due to the divergence of high energy behaviour. For such theories, and assuming the Z's to be zero due to high energy behaviour even outside perturbation theory, we may define a composite theory in this case as that obtained by imposing a cut-off Λ on the theory and requiring the wave function renormalisation constant $Z_C(\Lambda, \ldots)$ to be zero, regarded as a function of Λ and the renormalised masses and coupling constants. We then let $\Lambda \to \infty$ always keeping $Z_C(\Lambda, \ldots) = 0$. If the theory has a limit with a particle of mass M, then we term this theory a composite theory. This double limiting process is likely to be even more difficult to use than the single limiting process, some of whose difficulties we discussed in Chapter IV. This is a further reason to turn to the composite particle defined directly through (III.6), and avoid the limiting procedure through an elementary particle.

It is possible to attempt to determine the value of Z_C for particles in which non-relativistic models may be satisfactory. This has been done for the deuteron by Weinberg.[15] He showed that the wave-function renormalisation constant Z_t for the deuteron may be determined from the n-p triplet scattering length a_t and effective range r_t in the limit of zero binding by

$$\begin{aligned} 2(1-Z_t)/(2-Z_t) &= a_t/R \\ -Z_t/(1-Z_t) &= r_t/R \end{aligned} \qquad \text{(VI.1)}$$

where R is the deuteron radius, related to the deuteron binding by $R = (mB)^{-\frac{1}{2}}$, m = nucleon mass. The corrections to (VI.1) are of order $(1/m_\pi)$, where m_π is the pion mass. The experimental values $r_t = 1.75$ fermis, $a_t = 5.41$ fermis do not allow Z_1 to be very large, $|Z_t| < 0.4$.

An alternative evaluation of Z_t has been made by Amado and co-workers,[16] who showed that even a small non-zero value for Z_t alters the triton binding energy considerably (using n-p scattering in length approximation to calculate three nucleon scattering). The results favor $Z_t = 0.05$, though the effect of a small admixture of tensor forces in the deuteron may affect this result.

We may apply (VI.1) to the anti-bound singlet n-p state which has a 'binding' energy equal to one-thirtieth that of the deuteron. In this case, the singlet effective range and scattering length are $r_2 = 2.7F$, $a_s = -23.7F$, so that from (VI.1), $Z_s < 0.1$; since the approximation under which (VI.1) is obtained is more nearly satisfied in this case, this anti-bound state is a better candidate than the deuteron for a composite particle.

It should be possible to extend (VI.1) to heavier nuclei.

As I remarked earlier in this chapter, it is difficult to discuss the usual so-called "elementary particles" in such a fashion. However, we may deduce certain results from our field theory, albeit in a non-rigorous fashion.

(a) The photon is not composite.[17]

(b) For other particles we do not know the particular form of interaction which binds them. But we can discuss whether or not a particular interaction can do so. Thus for the $\pi - N$ system, if we suppose the system is interacting through a non-derivative Yukawa coupling $g\overline{N}\Gamma N\pi$, $\Gamma = 1$ or γ_5, then the extension of (III.6) for the nucleon is

$$N = \lambda \pi \cdot \Gamma N \tag{VI.2}$$

and for the pion is

$$\pi = \lambda' \overline{N} \Gamma N. \tag{VI.3}$$

So far, we have considered a system of particles in which at least one is elementary, so we take either (VI.2) or (VI.3) and not both (we will take both in the next chapter, and so get a bootstrap).

We see that if we take Γ as diagonal in (VI.2) we would be able to cancel the field N on both sides so that π would be the constant λ^{-1}. Thus if cancellation is allowed in the field theoretic case, then π would only take a single value at all points of space-time. Evidently the Fourier transform of π could not have a singularity on its mass shell so it could not describe a particle. (This lack of singularity in the Fourier transform would still be so if π only took a finite number of values in co-ordinate space.) Thus if the cancellation hypothesis is valid then (VI.2) can never support a composite nucleon, whatever the pion field.

On the other hand, if we take a derivative Yukawa interaction $g(\partial_\mu \pi)(\overline{N}\gamma_\mu \Gamma N)$, we now have

$$N = \lambda (\partial_\mu \pi)(\gamma_\mu \Gamma N) \tag{VI.4}$$

and it is not possible to cancel N since we cannot diagonalise the $\gamma_\mu \Gamma$'s simultaneously. However, we now expect subtraction

constants or other parameters to take account of the high energy behaviour introduced by the derivative coupling.

I do not want to justify the cancellation hypothesis for composite fields, but turn to the field-theoretic bootstrap and the problem of quantising it. We will then return to the cancellation hypothesis for the bootstrap and see how the conclusions derived from it may be justified.

VII. A Field Theory of Bootstraps

After the difficulties we have met with in understanding the $Z \to 0$ limit, it is natural to define a field-theoretic bootstrap as a set of field equations for fields $\psi_1, \ldots, \psi_N$, arising from some Lagrangian density L by taking all the wave function renormalisation constants $Z_1, \ldots, Z_N$ to be zero; the field equations being of local polynomial form:

$$\psi_i(x) = F_i(\vec{\psi}(x)), \quad \vec{\psi}(x), \ldots). \tag{VII.1}$$

Here $\vec{\psi}$ denotes the vector with components $\psi_1, \ldots, \psi_N$. We choose a Lagrangian L as starting point so that unitarity of the resulting theory is assured.

Our earlier discussion in Chapter III shows that suitable approximations to the G.F.E.'s arising from (VII.1) will lead to the N/D bootstrap equations or the B.S. bootstraps which were discussed in Chapters I and II. However, if we do not approximate to (VII.1), we have hopes of determining bootstrap parameters with inclusion of all intermediate particle states; we had indications in Chapter I that this might improve numerical agreement with experiment.

So far, it has not proved possible to obtain such improved numerical results, due to the complex operator structure of (VII.1). It is evidently possible to make better approximations than those described in Chapters I and II, but they require development of methods for dealing with functions of large numbers of variables; such has not yet been achieved. Instead of discussing in more detail the difficulties arising here, I would like to consider the general problem of quantising (VII.1), and then use the results to rule out certain bootstrap systems as not being possible. In the process we will see how the results of the cancellation hypothesis are justified.

The quantisation of the composite field ψ defined by (III.6) is straightforward if ϕ is an elementary particle field. For then the commutation relation between ψ and its time derivative will be

$$\left[\psi(\vec{x},t), \dot{\psi}(\vec{y},t)\right]_- = 4i\lambda Z_a^{-1}\psi(\vec{x},t)\delta^3(\vec{x}-\vec{y}) \tag{VII.2}$$

which is the correct limiting form as $Z_c \to 0$ of the elementary particle

canonical commutation relation for the ψ field:

$$\left[\psi(\vec{x},t),\ \dot{\psi}(\vec{y},t)\right]_- = iZ_c^{-1}\delta^3(\vec{x}-\vec{y}). \qquad \text{(VII.3)}$$

Comparing (VII.2) and (VII.3) we see that as $Z_c \to 0$ we must interpret Z_c^{-1} as taking the limiting value $4\lambda Z_a^{-1}\psi(\vec{x},t)$. Thus Z_c^{-1} acquires both co-ordinate space and operator-valued dependence; this is another aspect of the subtlety of the limiting process $Z_c \to 0$. In spite of this subtlety, (VII.2) is well-defined, provided we do not also take $Z_a \to 0$. But in the bootstrap situation, with all particles composite, we must do precisely that, so we cannot use (VII.2) without great care.

It might be possible to develop suitable canonical transformations which still allow the $Z_i \to 0$ limit for all particles to be discussed without too much difficulty.[18] I will not follow this approach here but will attempt to quantise (VII.1) in a direct manner, though one which should agree with a $Z \to 0$ prescription if this can be obtained. The quantisation method we use will be the Feynman 'history integral' approach to quantum field theory.[19] We obtain the Green's functions $G(x_1,\ldots,x_n)$ for a system of fields $\vec{\psi}$ with Lagrangian density $L(x)$ by taking

$$G(x_1,\ldots,x_n) = \int d\mu(\vec{\psi})\exp\left[i\int L(x)d^4x\right]\psi(x_1)\ldots\psi(x_n). \qquad \text{(VII.4)}$$

The measure $\mu(\vec{\psi})$ may be obtained by means of lattice-space integration as follows. Space-time is split into a large number N of cells, each of volume ε, together with a remainder. We take a representative point x_k in the k-th cell, and take the measure

$$\prod_{i,k} d\psi_i(x_k).$$

In the limit $\varepsilon \to 0$, $N \to \infty$, with the cells filling the whole of space time, this measure becomes $d\mu(\vec{\psi})$.

Let me apply this quantisation to the case of a non-derivative bootstrap. By this I mean a set of bootstrap equations of the type of (VII.1) with no derivatives of the fields entering on the right hand side. For the case of bosons we require $L(x)$ to be even in each boson field since, otherwise, we expect the energy to be unbounded below, following an argument of Baym.[20] From this evenness and the independence of different lattice points, we see that $G(x_1,\ldots,x_n)$ is zero unless the x_i's are equal in pairs. Since we require at least that $G(x_1,\ldots,x_n)$ be a distribution in its variables we need that it be a sum of covariant derivatives of δ-functions of the differences of

pairs of variables x_i, x_j. In order that G be a distribution, the order of the derivatives must be finite so that the Fourier transform of G will have no particle singularities. Thus there is no bootstrap.

When we include spinor particles we have to be more careful, since it is necessary to use anti-commutative integration. This has not been completely worked out for the history integral; we will define the functional integration over commuting fields, and then perform an anti-symmetrisation of the result over the space and spin variables of the spinor fields. If we use this method, then the arguments of Heber and co-workers[21] again lead to Green's functions which are δ-functions in the co-ordinate differences of the spinor fields. Thus if the spinor variables ψ and $\bar{\psi}$ are being integrated over (Majorana fields are thus excluded but do not occur in any case among the known particles), we define

$$\psi_\alpha(x_k) = r_{\alpha k} \exp(i\phi_{\alpha k}), \quad \psi^*_\alpha(x_k) = r_{\alpha k} \exp(-i\phi_{\alpha k}). \qquad \text{(VII.5)}$$

In a term such as

$$\int d\mu(\psi, \bar{\psi}) \cdot \bar{\psi}(x_1)\psi(x_2) \exp\Big[i \int L(x) d^4x\Big]$$

we will have the term

$$\int_0^{2\pi} d\phi_{\alpha_1} \int_0^{2\pi} d\phi_{\alpha_2} \exp i\Big[\phi_{\alpha_1} - \phi_{\alpha_2}\Big]$$

and this will be zero unless $x_1 = x_2$, if L is independent of the arguments ϕ_{α_k}. We may always choose a suitable angle variable so that L is independent of it and it enters in the manner of (VII.5), so that the integration over this angle variable will again introduce δ-functions of the differences of pairs of variables x_i, x_j. Thus we do not have any bootstrap in a theory with spin $\frac{1}{2}$ fields; we may extend this argument to higher spin fields by using the Rarita-Schwinger form of higher spin wave functions.[22]

If we start with a bootstrap theory of particles, some being bosons, some fermions, we see that the only non-trivial possibility is if the S-matrix involving bosons is unitary on the boson subspace, after the history integrals over the internal fermion fields is performed. In order for this to be so, we expect the resulting boson bootstrap to arise from a set of equations of the type of (VII.1), being derived from some Lagrangian. If those new bootstrap equations involve derivatives then we have a derivative bootstrap, and expect to have to introduce new parameters. If the new equations involve no derivatives then we have already seen that they cannot bootstrap themselves. Thus we have reached the conclusion that it is not

possible to bootstrap a system of particles without either an explicit or implicit derivative interaction. Thus the purest form of bootstrap, in which all free (dimensionless) parameters are prescribed, is not possible.

This gives the same result as the cancellation hypothesis for non-derivative bootstrap equations. It does not actually justify this hypothesis; however, since it has justified the result we wished to prove by means of the hypothesis, we will not consider it further here.

It might be possible to obtain particles from non-derivative bootstraps by introducing a measure in function space which also depends on the derivatives of the fields. For example, we might use the measure

$$\prod_k d\left((\Box + m^2)\phi(x_k)\right).$$

However, choosing the new variable

$$\phi' = (\Box + m^2)\phi$$

leads to terms such as

$$\int \prod_k d\phi'(x_k) \exp\left[i \int d^4x \, \phi' (\Box^2 + m^2)^{-2} \phi'\right]$$

and momentum space Green's functions will have a zero at $p^2 = m^2$ due to this. It will only be possible to introduce poles in the propagator by using a measure such as

$$\prod_k d\,(\Box + m^2)^{-\frac{1}{2}}\phi(x_k);$$

this is non-local, so no longer satisfies the property of statistical independence, and so will not be suitable as a measure on function space. Thus a different choice of measure does not seem to help.

Thus we are faced with considering derivative bootstraps. I want to talk about a general property of such bootstraps. Let me consider the derivative interaction $g\bar{\psi}\gamma_\mu\gamma_5\psi\cdot\partial_\mu\phi$, with bootstrap equations:

$$\psi = \lambda(\partial_\mu\phi)(\gamma_\mu\gamma_5\psi) \qquad \text{(VII.6)}$$

$$\phi = \lambda'\partial_\mu(\bar{\psi}\gamma_\mu\gamma_5\psi). \qquad \text{(VII.7)}$$

We see that (VII.7) is just the PCAC model of the pseudoscalar meson

(the vector interaction $g\bar{\psi}\gamma_\mu\psi\partial_\mu\phi$ would give $\phi = \partial_\mu(\bar{\psi}\gamma_\mu\psi)$, so would vanish due to nucleon conservation).

What can we say about the solutions to (VII.6), (VII.7)? The history integral quantisation can be used, but the lattice space measure does not separate out independent contributions from separate lattice points. Thus we cannot perform the lattice space integration in the trivial fashion we used for the non-derivative case.[23] But there are certain properties of solutions to the quantised form of (VII.6), (VII.7) which can be derived under the assumption that the solutions are suitably well-behaved. In particular, we can derive the existence of additively conserved quantum numbers, known as kinks.[24]

To see how these arise, consider (VII.6) in the form

$$(I - \lambda\gamma_\mu\gamma_5\partial_\mu\phi)\psi = 0. \qquad \text{(VII.8)}$$

At points with $\psi \neq 0$ we require $\det(I - \lambda\gamma_\mu\gamma_5\partial_\mu\phi) = 0$, or

$$(\partial_\mu\phi)^2 = -\lambda^{-2}. \qquad \text{(VII.9)}$$

Thus the four-vector $\partial_\mu\phi$ lies on the non-linear manifold (VII.9), which, for real λ, is a hyperboloid of one sheet. This manifold is doubly connected: there is a loop ℓ on the manifold, in the plane $\partial_0\phi = 0$, which cannot be continuously deformed into a point.

To see in what way this is important, let us consider the evolution in time of the system given by the history integral quantisation. The state at time $t = 0$ is a functional of the classical field quantities $\psi_i(x,o)$ measured at time $t = 0$; we denote it by $\Psi_o[\psi_i(x,o)]$. Then the state at time t is

$$\Psi_t[\psi_i(x,t)] = \int d\mu(\vec{\psi})\exp\Big[i\int L(x)d^4x\Big]\Psi_o[\psi_i(x,o)]. \qquad \text{(VII.10)}$$

In (VII.10) the integration is performed over all histories beginning with the initial configuration $\psi_i(\ ,o)$ and ending with $\psi_i(\ ,t)$. We assume that only the histories with non-zero measure are continuous, i.e., only $\vec{\psi}(\vec{x},t)$ continuous in all variables are integrated over in (VII.10) (this is true for Wiener measure, and if we consider the measure μ in (VII.10) as some suitable limit of Wiener measures, this result should still be true).

Let me return to the non-linear manifold (VII.9). More generally let me suppose that the field variables $\psi_i(\vec{x},t)$, for a given t, lie on a manifold Ψ which may be covered by local co-ordinates, so has the structure of, say, a differential manifold of dimension n. We suppose that the field variables are required to be equal to a

given field $\psi_i^{(o)}(\vec{x})$ in the limit as $|\vec{x}| \to \infty$. Since the time-development of the field variables for a classical system is continuous then this time development will not be able to change the connectivity of the field variables, i.e., the number of holes in the field variables manifold Ψ. As an example, if the fields variables are ψ_1, ψ_2 and lie on the unit circle $\psi_1^2 + \psi_2^2 = 1$, and we take 1 space dimension, then at time $t = 0$ there may be n 'twists' of the field variable in going from $x = -\infty$ to $x = +\infty$. As the time t increases, the continuous classical development of the system cannot break the twists, so n is conserved. This situation is best described by means of homotopy groups; we say that two fields $\vec{\psi}^{(1)}(\vec{x})$ and $\vec{\psi}^{(2)}(\vec{x})$ are homotopic if there exists a homotopy (or deformation) between them, this being a set of continuous function $\vec{\psi}(\vec{x}, \alpha)$ for $\alpha \in [0,1]$, with $\vec{\psi}(\vec{x}, o) = \vec{\psi}^{(1)}(\vec{x})$, and $\vec{\psi}(\vec{x}, 1) = \vec{\psi}^{(2)}(\vec{x})$, with

$$\lim_{|\vec{x}| \to \infty} |\vec{\psi}(\vec{x}, \alpha) - \vec{\psi}^{(0)}(\vec{x})| = 0,$$

all $\alpha \in [0,1]$. Then the relation of homotopy between two functions is an equivalence relation, and we may divide the field variables at t into equivalence classes following this relation. We may further define a multiplication on the set of homotopy classes $\{\vec{\psi}\}$ by saying that the product of $\{\vec{\psi}^{(1)}\}$ and $\{\vec{\psi}^{(2)}\}$ is the homotopy class which has as a member the function $\vec{\psi}^{(3)}$ which is obtained by 'sewing' $\vec{\psi}^{(1)}$ and $\vec{\psi}^{(2)}$ together as follows. $\vec{\psi}^{(1)}$ is homotopied to a function which is zero for, say, $x_1 > 0$, $\vec{\psi}^{(2)}$ to a function zero for $x_1 < 0$, and $\vec{\psi}^{(3)}$ is the function equal to $\vec{\psi}^{(1)}$ for $x_1 < 0$, $\vec{\psi}^{(2)}$ for $x_1 > 0$. This multiplication turns the set of homotopy classes into a group which, in our case, is the 1-st homotopy group $\pi_1(\Psi)$ (in 1-space dimension), and in n-dimensional-space is the n-th homotopy group $\pi_n(\Psi)$. ($\pi_1(\Psi)$ = fundamental group of Ψ). Evidently $\pi_1(\Psi) = \infty$ (additive group of integers) with generator equal to a single twist. The case we are interested in has similarly

$$\pi_3(\Psi) = \infty$$

with the generator being the single loop ℓ.

We now turn to the quantisation of the system, going back to Eq. (VII.10). If we suppose that $\Psi_o[\psi_i(\ ,o)]$ is non-zero only for $\psi_i(\ ,o)$ in some homotopy class $C \in \pi_3(\Psi)$, then since we have continuous time development the only homotopy class to which $\psi_i(\ ,t)$ can belong to have a non-zero value in (VII.10) is C, since the only continuous histories we sum over in (VII.10) will lie in C, at each time t. Thus the state preserves the homotopy class of the field variables, and thus the homotopy number n (the number of 'twists' in C). This number is called the 'kink' number by Finkelstein.[24]

We see that our meson field thus carries this conserved quantum number, and we expect it to be transferred to the fermion by the trilinear interaction; it will be an additively conserved quantum number in any elementary particle process. We may also have further conserved quantum numbers arising from the nucleon field manifold

$$\psi = \lambda' \lambda \gamma_5 \gamma_\mu \psi \cdot \partial_\mu \partial_\nu (\bar{\psi} \gamma_\nu \gamma_5 \psi) \tag{VII.11}$$

for which

$$(\lambda \lambda')^{-2} = -\left[\partial_\mu \partial_\nu (\bar{\psi} \gamma_\nu \gamma_5 \psi)\right]^2. \tag{VII.12}$$

It is much more difficult to analyse the homotopy structure of (VII.12). However, there does not seem to be any conserved quantum number arising which may be interpreted as a fermion number, such as nucleon or lepton number. The conserved number obtained already may be interpreted as strangeness, though we will have to break it, possibly by expanding about $Z=0$.[23] There are other interactions of a slightly more general form, which will also give kinks, e.g., if we do not want to allow characteristics for ϕ depending on ϕ (in the $Z=0$ limit) we can still take

$$L_{int} = F(\bar{\psi}, \psi) \cdot (\bar{\psi} \gamma_\mu \gamma_5 \psi) \partial_\mu \phi$$

to obtain

$$\phi = \lambda \partial_\mu \left[F(\bar{\psi}, \psi) \bar{\psi} \gamma_\mu \gamma_5 \psi \right]$$

$$\psi = \lambda \left\{ (\partial F / \partial \bar{\psi}) (\bar{\psi} \gamma_\mu \gamma_5 \psi) + F(\bar{\psi}, \psi) \gamma_\mu \gamma_5 \psi \right\} \partial_\mu \phi.$$

For $F = a\bar{\psi}\psi$, say, we obtain

$$\left[1 - \lambda a (\bar{\psi} \gamma_5 \gamma_\mu \psi) \partial_\mu \phi\right]^2 = -a^2 (\partial_\mu \phi)^2$$

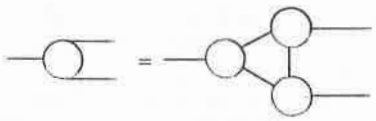
which still has $(\partial_\mu \phi)^2 \leq 0$, so giving a non-trivial third homotopy group (unless $1 = \lambda a (\bar{\psi} \gamma_5 \gamma_\mu \psi) \partial_\mu \phi$, which can be avoided by a suitable choice of a). Thus we cannot use the existence (or not) of a kink to single out a specific form of interaction.

We remark here that we may regard the bootstrap in our field theory approach as the 0-th approximation to a strong coupling solution; if all the Z's ~ 0 we could expand in the Z's about this approximation. If this expansion was good this would explain the successes of SU_6 (for when $Z=0$ the kinematic terms $\gamma_\mu \partial_\mu$ are absent, and one

can easily write down interaction which are SU_6 invariant but not kinematic terms which are also SU_6 invariant).

VIII. Generalised Groups

I now want to turn to a different direction than that of field theory to develop bootstraps. I want to return to the approximate form of bootstrap embodied in the Cutkosky bootstrap equations.



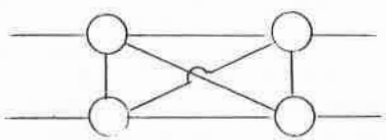
(VIII.1)

We consider functions of three variables f(a,b,c); the right hand side of (VIII.1) suggests the triple product

$$(fgh)_{ijk} = \sum_{\ell,m,n} f_{imn} g_{\ell jn} h_{\ell mk} \qquad \text{(VIII.2)}$$

where Σ denotes summation over discrete variables or integration over continuous variables. If we denote by G the vector space of the functions f(a,b,c) (with suitable smoothness properties with respect to the variables a,b,c if they take a continuous range of values) then (VIII.2) defines a mapping of $G \times G \times G \to G : (f,g,h) \to (fgh)$. We note that we cannot have a binary map $G \times G \to G$, but only a three-fold map; we thus meet a generalisation of a group (or ring or algebra, et cetera) which we may denote as three-group (three-ring, et cetera).[25]

We may generate a four-fold map if we take functions of four variables, and consider the map defined similarly to (VIII.2) with the Feynman-type of diagram



We may generalise this to an n-fold map or an n-ary operation in an obvious fashion. The Feynman-like form of these mappings shows how they would arise in more general bootstrap equations than (VIII.1), and a number of them may occur simultaneously. Thus it is useful to analyse the structure of sets G with n-ary operations on them which generalise the group structure. In particular, the

representation theory of such objects may enable us to diagonalise the n-ary operation and so enable the solution of equations like (VIII.1) to be found in a simpler fashion on the irreducible subspaces. A three-group also arises in calculating higher moments of two-body nuclear operators, and a reduction theory would possibly enable much simpler calculations to be performed.[26]

Let me now make some definitions.

Definition 1. An n-groupoid is a set A of objects $a_1, a_2, a_3, \ldots$, together with an n-ary operation of A^n to A which I will, as a product, denote by $a_1 a_2 \ldots a_n$.

Definition 2. An n semi-group is an n-groupoid with an associative n-ary operation, so that all the possible products of (2n-1) objects are equal, i.e.,

$$(a_1 \ldots a_n) a_{n+1} \ldots a_{2n-1} = a_1 (a_2 \ldots a_{n+1}) a_{n+2} \ldots a_{2n-1} = \ldots$$

$$= a_1 a_2 \ldots a_{n-1} (a_n \ldots a_{2n-1}).$$

Definition 3. An object a belonging to an n groupoid A has a left (right) inverse $a_l (a_r)$ if

$$(a_l a)^{n-1} b = b, \quad b \in G$$

$$(a a_r)^{n-1} b = b, \quad b \in G.$$

If $a_l = a_r$ we say a has an inverse.

Definition 4. An n semi-group for which every element has an inverse is an n-group.

Such objects have been considered in a very general manner elsewhere[27),28)] though no work on representation theory seems to have been done. As examples of these systems:

(1) The odd integers form a 3-group under addition, but have no identity (which would be zero).

(2) The negative non-zero numbers form a 3-group under multiplication, again having no identity.

(3) The tensors of rank 3, $a_{ijk} (1 \; i,j,k \; N)$ form a 3-groupoid, with

$$(abc)_{ijk} = \sum_{\ell,m,n}{}' a_{imn} b_{\ell jn} c_{\ell mk} .$$

However, the 3-product is not associative, $(abc)de \neq ab(cde)$. Neither is there an inverse to a general element a. In spite of this we may try to develop an analysis for a subset of this 3-groupoid closely paralleling that for tensors of rank 2 when regarded as matrices. We may regard the tensors of rank 3 (3-tensors) as elements of a 3-dimensional cubical array. This will enable us to generalise many concepts holding for matrices to the case of 3-tensors, one example being that of a determinant. We will not consider this further here.

(4) The set of n-tensors of dimension N form an n-groupoid, with

$$(a_1 a_2 \ldots a_n)_{i_1 i_2 \ldots i_n} = \sum_{1 \le j_1 \ldots j_n \le N} a_{1\, j_1 j_2 \ldots j_n} \times a_{2\, j_1 i_2 j_3 \ldots j_n} \ldots a_{n\, j_1 \ldots j_{n-1} i_n}.$$

As for 3-tensors this n-groupoid is non-associative, nor does a general element have an inverse.

Embedding Theorem.[29] Every n-groupoid G may be isomorphically embedded in a semi-group (that is, a 2-semi-group).

For if ω denotes the n-ary operation on G, let R be the row-algebra of $G \cup \{\omega\}$, that is, the set of rows $R = (a_1, \ldots, a_s)$ with $a_i \in G \cup \{\omega\}$, the product of any two elements of R being just that element of R made up of the two sets of rows joined together:

$$(a_1, \ldots, a_s)(b_1, \ldots, b_t) = (a_1, \ldots, a_s, b_1, \ldots, b_t).$$

To each $a \in G$ we associate the operator ρ_a on R by

$$\rho_a(a_1, \ldots, a_s) = (a, a_1, \ldots, a_s)$$

and to ω we associate σ_ω defined by

$$\sigma_\omega(a_1, \ldots, a_s) = (\omega(a_1, \ldots, a_n), a_{n+1}, \ldots, a_s) \quad \text{if } s = m+n, \; a_i \in G \text{ for } 1 \le i \le n$$

$$= (\omega, a_1, \ldots, a_s) \quad \text{otherwise.}$$

Then

$$\sigma_\omega \rho_{a_1} \ldots \rho_{a_n} = \rho_{\omega(a_1 \ldots a_n)}$$

if $a_i \in G$, $(1 \leq i \leq n)$, and $\rho_a = \rho_b$ if and only if $a = b$. Thus G is isomorphic to a semi-group of 1-ary operators on R. But we have

$$\sigma_\omega\left(\sigma_\omega\left(\rho_{a_1} \ldots \rho_{a_n}\right)\rho_{a_{n+1}} \ldots \rho_{a_{2n-1}}\right) =$$

$$= \rho_\omega\left(\omega(a_1 \ldots a_n)a_{n+1} \ldots a_{2n-1}\right) = \rho_A$$

$$\sigma_\omega\left(\rho_{a_1}\sigma_\omega\left(\rho_{a_2} \ldots \rho_{a_{2n-1}}\right)\rho_{a_{n+2}} \ldots \rho_{a_{2n-1}}\right)$$

$$= \rho_\omega\left(a_1\omega(a_2 \ldots a_{n+1})a_{n+2} \ldots a_{2n-1}\right) = \rho_B$$

and $A \neq B$, due to lack of associativity. Thus the associative product on R is not the (generally) non-associative n-ary product on G in which we are interested. We mainly wish to set up a theory of n-groupoids which allows us to simplify the non-linear bootstrap equations of type (VIII.1); since the binary product in R cannot in general be directly related to the n-ary product in which we are interested, it would seem necessary to analyse the structure and representation theory of n-groupoids by methods which do not depend on the embedding theorem.

We will turn, then, to a direct analysis of the representation theory, and start with the very simplest type of n-groupoid, which will be a finite Abelian 3-group. We will see very similar results to those for finite Abelian groups, as is to be expected from the embedding theorem in the associative case; these results may possibly be proved directly by this method. However, we hope that it may be possible to extend our methods of proof to the more general non-associative case.

Definition 5. A representation of an Abelian 3-group is a map of $G \times G$ into the set $\mathfrak{L}(V \to V)$ of linear operators of a vector space V into itself; if this map is denoted by $L_{a,b}$, for $a, b \in G$, then we require

$$L_{a,b}L_{c,d} = L_{abc,d} = L_{a,bcd}. \qquad \text{(VIII.3)}$$

We have used (VIII.3) since we want to have the regular representation included in our Definition 5. This is obtained when we take $V = G$ (when G also has a vector space structure compatible with the 3-group structure), and

$$L_{a,b}c = abc.$$

Let me consider the one-dimensional representations

$$L_{a,b} = \chi(a,b)I \tag{VIII.4}$$

where I is the identity operator on V, and V is one-dimensional. Then $\chi(a,b)$ is a complex-valued function on $G \times G$ with

$$\chi(a,b)\chi(c,d) = \chi(abc,d) = \chi(a,bcd). \tag{VIII.5}$$

Any such complex-valued function on $G \times G$ will be called a character of G; the set of all such functions is the character group $\hat{G}$ of G.

Definition 6. In any finite Abelian n-group G, the least integer r for which, for a given $a \in G$, $a^{r(n-1)+1} = a$, is called the order of the element a.

Definition 7. In any finite Abelian n-group G the number of distinct elements in G is the order of G, o(G).

Then in any finite Abelian 3-group, every element has finite order, and if the order of G is denoted by n, for any $a \in G$

$$a^{2n+1} = a.$$

Then from (VIII.5) we have, for any $a \in G$,

$$\chi(a,a)^{n+1} = \chi(a^{2n+1},a) = \chi(a,a)$$

so $\chi(a,a)$ is an n-th root of unity. Also for any $a,b \in G$, from (VIII.5)

$$\chi(a,a)^{2n+2} = \chi\left((ab)^{2n+1}a,b\right) = \chi(a.b)^2$$

so

$$\chi(a,b)^{2n} = 1$$

and $\chi(a,b)$ is a 2n-th root of unity. Then $|\chi(a,b)\chi = 1$ for any $\chi \in \hat{G}$, so that $\hat{G}$ is a group, with the inverse of each χ being the complex conjugate function χ^*, and multiplication of two characters χ_1, χ_2 defined by the evident rule

$$\chi_1\chi_2(a,b) = \chi_1(a,b)\chi_2(a,b).$$

Definition 8. An n-group G is generated by the elements $x_1, \ldots, x_N$ if every element of G may be written as a product of the elements $x_1, \ldots, x_N$, each element taken any number (including zero or negative numbers) of times.

Definition 9. A cyclic n-group G is that generated by a single element.

If G is a cyclic 3-group, generated by the element a of order n, then for any integers r,s,

$$\chi\left(a^{2r+1},a^{2s+1}\right) = \chi(a,a)^{r+s+1}.$$

Hence the values of χ on G are just the possible n-th roots of unity. Also the set of all possible values of G on a are all the possible n-th roots of unity, so that there are n different functions in $\hat{G}$, and $o(G) = o(\hat{G})$.

We may now extend this above result to all possible Abelian 3-groups which are the direct products of cyclic 3-groups, in the following sense.

Definition 10. An n-group G is the direct product of n-groups $G_1,\ldots G_N$ (N odd) if these n-groups are sub-n-groups of G, and any element x of G may be written as

$$x = \prod_{i=1}^{N} x_i, \qquad x_i \in G_i.$$

Then for any Abelian 3-group G which is the direct product of a finite number of cyclic 3-groups G_i, generated by the element x_i, any element $x \in G$ is of the form

$$x = \prod_{i=1}^{N} x_i^{\alpha_i}$$

where the α_i are odd. Then any character on G can be factorised by the relation

$$\chi(x,y) = \chi\left(\pi x_i^{\alpha_i},y\right) = \chi\left(x_j,x_j\right)\chi\left(\prod_{i\neq j} x_i^{\alpha_i}.x_j^{\alpha_j-2},y\right)$$

so we may reduce $\chi(x,y)$ to a product of powers of $\chi(x_j,x_j)$, for various j, times

$$\chi\left(\prod_{i=1}^{N} x_i, \prod_{i=1}^{N} x_i\right),$$

which may be easily seen to equal

$$\prod_{i=1}^{N} \chi(x_i,x_i).$$

Thus the character group $\hat{G}$ of G will be the direct product of the character groups of the factor groups of G, and so we still have $o(G) = o(\hat{G})$. It is not known whether the basis theorem for finite Abelian 3-groups is true (this being that every such 3-group is the direct product of cyclic 3-groups); the proof for the 2-group case is well known.[30]

Now that we know the possible 1-dimensional representations for any direct product of cyclic 3-groups, we turn to the analysis of any general representation V on the complex vector space H(V). We first derive an elementary property of the characters. Consider

$$\chi(a,b) \sum_{c,d \in G} \chi(c,d) = \sum_{c,d \in G} \chi(abc,d) = \sum_{c',d \in G} \chi(c',d) \quad \text{(VIII.6)}$$

where we use the 3-group structure in the last step of (VIII.6) to show that for any $c' \in G$ there is a c in G so that $abc = c'$, this c being $a^{-1}b^{-1}c^{1}$. Thus

$$[1 - \chi(a,b)] \sum_{c,d \in G} \chi(c,d) = 0.$$

If $\chi \neq 1$ then

$$\sum_{c,d \in G} \chi(c,d) = 0. \quad \text{(VIII.7)}$$

For two distinct characters χ_1, χ_2, then $\chi_1 \chi_2^*$ is a character, and $\chi_1(a,b)\chi_2^*(a,b) = \chi_1 \chi_2^{-1}(a,b) \neq 1$, so

$$\sum_{a,b \in G} \chi_1(a,b)\chi_2^*(a,b) = 0. \quad \text{(VIII.8)}$$

Evidently (VIII.8) is an orthogonality property of the characters, so they form a basis for the vector space of complex valued functions on $G \times G$. We also note that since

$$\chi(a,b) = \chi(a,bcc^{-1}) = \chi(a,b)\chi(c,c^{-1})$$

then $\chi(c,c^{-1}) = 1$ for any $c \in G$.

Finally we note the further identity $\chi^*(b^{-1},a^{-1}) = \chi(a,b)$. To analyse the general representation V, we define for each $\chi \in \hat{G}$,

$$P_\chi = [o(G)]^{-2} \sum_{a,b \in G} \chi^*(a,b) V_{a,b}.$$

Then we have the following properties of P:

(1)
$$P_\chi^2 = P_\chi, \quad \text{any} \quad \chi \in \hat{G} \tag{VIII.9}$$

(2)
$$P_{\chi_1} P_{\chi_2} = P_{\chi_2} P_{\chi_1} = 0 \quad (\chi_1 \neq \chi_2) \tag{VIII.10}$$

(3)
$$\sum_{\chi \in \hat{G}} P_\chi = I \quad \text{(the identity on } H(V)) \tag{VIII.11}$$

(4)
$$V_{a,b} P_\chi = \chi^*(a,b) P_\chi, \quad \text{all } a,b \in G, \ \chi \in \hat{G}. \tag{VIII.12}$$

To prove (VIII.9) and (VIII.10), we have

$$P_{\chi_1} P_{\chi_2} = [o(G)]^{-4} \sum_{a,b,c,d \in G} \chi_1^*(a,b) \chi_2^*(c,d) V_{a,b} V_{c,d}$$

$$= [o(G)]^{-4} \sum_{a,b \in G} \chi_1^*(a,b) \chi_2(a,b) \sum_{c',d \in G} \chi_2^*(c',d) V_{c',d}$$

$$= [o(G)]^{-2} \sum_{a,b \in G} \chi_1^*(a,b) \chi_2(a,b) P_\chi \tag{VIII.13}$$

where $c' = abc$ in (VIII.13). If we use (VIII.8) when $\chi_1 \neq \chi_2$ we obtain (VIII.10), and if $\chi_1 = \chi_2$, since $|\chi| = 1$, and

$$\sum_{a,b \in G} 1 = [o(G)]^2,$$

then we obtain (VIII.9).

To obtain (VIII.11) we have

$$\sum_{\chi \in \hat{G}} P_\chi = [o(G)]^{-2} \sum_{a,b \in G} V_{a,b} \sum_{\chi \in \hat{G}} \chi^*(a,b).$$

Now for given a,b, $\chi(a,b)$ is a mapping of $\hat{G}$ to the complex numbers with absolute value 1 and satisfying the property of a character, so is in $\hat{\hat{G}}$. Thus using (VIII.7) (true also for $\hat{G}$ as for G),

$$\sum_{\chi \in \hat{G}} \chi^*(a,b) = 0 \quad \text{unless} \quad \chi(a,b) \equiv 1, \quad \text{all} \quad \chi \in \hat{G}.$$

But when G is a direct product of cyclic 3-groups,

$$\chi(a,b) \equiv 1 \qquad \text{if and only if} \quad a = b^{-1}$$

so

$$\sum_{\chi \in \hat{G}} P_\chi = [o(G)]^{-2} \sum_{a \in G} V_{a,a^{-1}} \sum_{\chi \in \hat{G}} 1. \tag{VIII.14}$$

Further $V_{a,a^{-1}} = I$ (the identity on H(V)), since for any $b, c \in G$,

$$V_{a,a^{-1}} V_{b,c} = V_{a,a^{-1}b,c} = V_{b,c},$$

and we multiply both sides by the inverse of $V_{b,c}$. Thus the first summation in (VIII.14) gives o(G), the second summation gives $o(\hat{G}) = o(G)$, so proving (VIII.11). Finally

$$V_{a,b} P_\chi = [o(G)]^{-2} \sum_{c,d} \chi^*(c.d) V_{a,b} V_{c,d}$$

$$= [o(G)]^{-2} \sum_{c,d \in G} \chi^*(a.b)\chi^*(abc,d) V_{abc,d}$$

$$= \chi^*(a,b) P$$

so proving (VIII.12).

Let H_χ be the range of P_χ. Then from (VIII.9), (VIII.10) and (VIII.11) every vector in H is uniquely expressible as a sum of vectors, one from each H_χ:

$$\phi = \sum_{\chi \in \hat{G}} P_\chi \phi.$$

Also for each $\phi \in H_\chi$,

$$V_{a,b} \phi = \chi^*(a,b) \phi.$$

Then each H defines a sub-representation V^χ of the form

$a, b \to \chi^*(a, b)I$; since

$$V = \sum_{\chi \in \hat{G}} V^{\chi},$$

we obtain the primary decomposition of V. The irreducible representations are obtained by taking any basis in each V^{χ}, so that each basis element defines a one-dimensional irreducible representation (though this decomposition into irreducibles is not unique).

Further, the irreducible subspaces of H may be made up by taking any subspace of each V^{χ} and forming the union over χ.

These results are identical with those for finite Abelian 2-groups. I do not want here to go into details about the extension of these results to infinite non-Abelian non-associative n-groupoids. The removal of the finiteness and Abelian characters can be done, provided suitable topological properties (at least local compactness) are added. These and applications to bootstraps and nuclear physics will be considered elsewhere.

Acknowledgments

I would like to thank W. Brittin for the stimulating hospitality of the Summer Institute while these lectures were delivered; I would also like to thank M. Broido, H. Osborn, A. Pagnamenta, F. Schroeck and J. G. Valatin for helpful discussions at various stages of the work.

References

1. For an initial discussion of this see M. M. Broido and J. G. Taylor, Phys. Rev. 147, 993 (1966). See also M. M. Broido, "A Field Theory of Bootstraps," Lectures in Theoretical Physics, Vol. IX (1966), Boulder, Colorado.
2. For other surveys see, for example, S. Gasiorowicz, "Elementary Particle Physics," Wiley, New York (1966), Chapter 25; B. Diu, Ecole Internationale de la Physique des Particules Elementaires, Herceg Novi, Jugoslavia (1966); see also articles in Book III of High Energy Physics and Elementary Particles, International Atomic Energy Authority, Vienna (1965).
3. See S. Gasiorowicz, Reference 2, Chapters 22 and 24.
4. G. L. Shaw, Phys. Rev. Lett. 12, 345 (1964).
5. R. Fulco, G. Shaw and D. Wong, Phys. Rev. 137B, 1242 (1965).
6. W. Abbe, P. Kaus, P. Nath and Y. Srivastava, Phys. Rev. 154, 1515 (1967). For an alternative approach, see P. D. B. Collins, Phys. Rev. 157, 1432 (1967).
7. R. E. Cutkosky, Phys. Rev. 131, 1888 (1963).
8. See also A. Pagnamenta, CERN preprint "A Bootstrap Theory of Quarks" (1967).

9. W. Zimmerman, Nuovo Cimento 10, 597 (1958).
10. For a full discussion of this see M. M. Broido and J. G. Taylor, Reference 1, and also J. G. Taylor, Nuovo Cimento Suppl. 1, 857 (1964), or J. G. Taylor, "Fixed Point Methods in Field Theory," Lectures in Theoretical Physics, Vol. IX (1966), Boulder, Colorado.
11. Private Communication from H. Osborn, University College, London. I would like to thank Mr. Osborn for a stimulating discussion on this problem.
12. See, for example, N. Nakanishi, J. Math. Phys. 7, 698 (1965), and references quoted therein.
13. Yndurain, J. Math. Phys. 7, 1133 (1966). We would like to thank Dr. Broido for pointing out this reference to us and letting us see his unpublished manuscript on composite particles in this model; we, however, have different results, and in particular take seriously the momentum dependence of $Z_V(p)$, as discussed in the text.
14. For a full discussion of the non-relativistic Lee model, see, for example, S. Schweber, "Relativistic Quantum Field Theory," Row, Peterson (1961), Chapter 12, Section 6.
15. S. Weinberg, Phys. Rev. 137, 672 (1965).
16. R. Aaron, R. Amado and Y. Yam, Phys. Rev. Letts. 13, 574 (1964).
17. See Reference 1, and also M. M. Broido, Phys. Rev. 157, 1444 (1967).
18. Private communication from M. M. Broido.
19. See, for example, accompanying lectures by J. Tarski in this volume.
20. G. Baym, Phys. Rev. 117, 886 (1960).
21. G. Heber and A. Kühnel, Zeits. für Naturforsch. 199, 1245 (1964).
22. See, for example, the accompanying lectures by Y. Takahashi in this volume.
23. See, however, the approach by Heber and co-workers described by G. Heber, "Lattice Space Methods in Functional Integration," Winter School in Theoretical Physics, Karpacz, Poland (1967), University of Warsaw press.
24. D. Finkelstein, J. Math. Phys. 7, 1218 (1966).
25. Such a system was already considered in this context in J. G. Taylor, J. Math. Phys. 6, 1148 (1965).
26. J. B. French, "Combinatorial Methods for Spectroscopic Averaging Spectral Moments," Univ. of Rochester preprint, and F. E. Schroeck, Jr. (unpublished). The use of 3-groups in this connection was pointed out by F. E. Schroeck, Jr. in private discussion.

27. R. H. Bruck, "A Survey of Binary Systems," Springer (1958).
28. P. M. Cohn, "Universal Algebra," Harper (1965).
29. Reference 26, page 37; Reference 27, page 185; see also the original paper by E. L. Post, Trans. Amer. Math. Soc. 48, 208 (1940) where finite n-groups are discussed in some length, and the embedding theorem was originally proved though only for an n-group.
30. See, for example, E. Schenkman, "Group Theory," van Nostrand (1965), or any standard text on group theory.

FUNCTIONAL INTEGRALS IN BROWNIAN MOTION†

George J. Papadopoulos
Department of Mathematics
University of Leeds
England

I. Introduction and Summary

In these lectures we propose to show how the phenomenological theory of Brownian motion could be developed via the techniques of functional calculus. The formalism refers to all situations in which we are seeking the distribution of a dynamic variable, related through an equation of motion to a stochastic variable of given distribution. However, we shall base these lectures on a particular situation—that of a particle in a liquid.

It was Einstein's idea that particles in liquid environment suffer perpetually collisions from the molecules of the surrounding medium due to the thermal agitation of the latter.[1] As a result of the thermal kicks, a particle of approximately colloidal size is continuously kinking; we say that it executes Brownian motion or simply it is a Brownian particle.

Langevin postulated the following equation of motion for the Brownian particle:

$$\frac{d\vec{p}}{d\tau} = -B\vec{p} + \vec{F}(\vec{r},\vec{p},\tau) + \vec{f}(\tau) \tag{I.1}$$

where

$$\vec{p} = m\frac{d\vec{r}}{d\tau} . \tag{I.2}$$

B^{-1} is a relaxation time matrix which is symmetric positive definite and depends on the viscous properties of the medium and the geometry of the particle.[2] (The orientation dependence of B will be considered averaged.) $-B\vec{p}$ is the Stokes resistance of the medium to the particle. $\vec{F}(\vec{r},\vec{p},\tau)$ is the external force on the particle, assumed to be slowly varying and containing no memory. $\vec{f}(\tau)$ is the force of

† Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

collisions assumed random and independent of the kinetic state of the particle.[3] The forces $-B\vec{p}$ and $\vec{F}$ are called systematic, whereas $\vec{f}$ is called thermal force.

Our problem is: Given a functional distribution for $\vec{f}(\tau)$ over a time interval $[t_o,t]$, what is the distribution of the position $\vec{r}$ and momentum $\vec{p}$ of the particle at time t, if at an earlier time t_o the particle occupied the phase point $(\vec{r}_o,\vec{p}_o)$? We make for the thermal force $\vec{f}$ the stochastic assumption that its functional distribution over $[t_o,t]$ is the following continual Gaussian distribution:

$$W[\vec{f}(\tau)]_{t_o}^{t} = \Big[\prod_{t_o \le \tau < t} \det \pi^{-1} g(\tau) d\tau \Big]^{\frac{1}{2}} \exp\Big\{ \int_{t_o}^{t} g_{\alpha\beta}(\tau) f_\alpha(\tau) f_\beta(\tau) d\tau \Big\} . \tag{I.3}$$

Summation convention from 1 to 3 will be understood for repeated indices throughout this text. $g_{\alpha\beta}$ is a positive definite symmetric matrix. It will be further specified later on in such a way that for a free Brownian particle (i.e., $\vec{F}=0$) the distribution of the momenta after infinite time goes over to the Maxwellian distribution:

$$[2\pi\, m\kappa T]^{-3/2} \exp\Big\{ -\frac{\vec{p}^{\,2}}{2m\kappa T} \Big\} . \tag{I.4}$$

Alternatively we may demand of the equation governing the distribution of the particle momentum to admit (I.4) as a solution in the case $\vec{F}=0$. g thus determined is found:

$$g^{-1} = 4\kappa T m B. \tag{I.5}$$

The functional distribution (I.3) represents the distribution at each time $\tau \in [t_o,t]$ of the collision forces $\vec{f}(\tau)$ on the particles of an ensemble of identical Brownian particles. It is easy to deduce that the thermal forces at two different times are not correlated. This fact is very important for the derivation of the integral Smoluchowski equation.

The statistical description of the Brownian particle is effected through the ensemble average conditional probability distribution (ECPD) $G(\vec{r}|\vec{r}_o,\vec{p}|\vec{p}_o;t|t_o)$, $t>t_o$, of finding the particle in the vicinity of the phase point $(\vec{r},\vec{p})$ at time t, if at time t_o it occupied the phase point $(\vec{r}_o,\vec{p}_o)$. The object of the subsequent sections is to develop techniques for obtaining the ECPD. We shall demonstrate the method by concentrating attention to the case of momentum space. For this purpose we shall consider only external forces of the form $\vec{F}(\vec{p},t)$, i.e., depending on momentum and time. In this case we can ask what is the ECPD $G(\vec{p}|\vec{p}_o;t|t_o)$ of

finding the particle at time t with momentum in the vicinity of $\vec{p}$ if at time t_o its momentum were $\vec{p}_o$. The generalization to the case of phase space is effected in a similar manner.

In Section II we derive the distribution G as a functional average of the conditional probability distribution (CPD) in Liouville (or deterministic) sense. Furthermore, we exemplify the method by (a) calculating the ECPD in the case the external force is a prescribed function of time, and (b) by finding G when $\vec{F} = \vec{F}(\vec{p},\tau)$, but for $(t - t_o)$ very short.

In Section III, using the method for the construction of ECPD G, we show that G obeys the Smoluchowski (-Kolmogorov-Chapman) equation. Also we show that G is a Green function of the Fokker-Planck equation.

In Section IV we represent G as a conditional functional integral over Wiener measure in the space of momentum functions.

In Section V we obtain from the formal representation of G in Section IV a compact approximate expression for G. The method can provide solutions to non-linear problems in Brownian motion and related topics. It is analogous to the WKB approximation.

II. Construction of the ECPD $G(\vec{p}\,|\,\vec{p}_o;t|t_o)$

We deal with the case $\vec{F} = \vec{F}(\vec{p},\tau)$. The Langevin equation is

$$\frac{d\vec{p}}{d\tau} = -B\vec{p} + \vec{F}(\vec{p},\tau) + \vec{f}(\tau). \tag{II.1}$$

Let

$$\vec{P}(t) = \vec{P}\Big(\vec{p}_o,[\vec{f}(\tau)]_{t_o}^{t}\Big)$$

be the solution of the Langevin equation (II.1) which satisfies the initial condition

$$\vec{P}(t_o) = \vec{p}_o.$$

This solution obviously depends on all the values of $\vec{f}(\tau)$ with $\tau \in [t_o,t]$.

The CPD in deterministic sense of finding the particle with momentum in the vicinity of $\vec{p}$ at time t, if its momentum at time t_o were $\vec{p}_o$, is given by the δ-functional:

$$\delta\left\{\vec{p} - \vec{P}\Big(\vec{p}_o,[f(\tau)]_{t_o}^{t}\Big)\right\}. \tag{II.2}$$

The ECPD is obtained as the functional average of the deterministic CPD (II.2) with respect to the thermal distribution measure (I.3),[4] i.e.,

$$G(\vec{p}\,|\,\vec{p}_o;t|\,t_o) = \int \delta\left\{\vec{p} - \vec{P}\Big(p_o,[\vec{f}(\tau)]_{t_o}^{t}\Big)\right\} W[\vec{f}(\tau)]_{t_o}^{t} \prod_{t_o\le\tau<t} d\vec{f}(\tau). \quad \text{(II.3)}$$

Let us now, before we exemplify the above procedure for the construction of the ECPD, state the formula:

$$\int \delta\left\{\vec{x} - \int_{t_o}^{t} Q(\tau)\vec{f}(\tau)d\tau\right\} W[\vec{f}(\tau)]_{t_o}^{t} \prod_{t_o\le\tau<t} d\vec{f}(\tau) =$$

$$= \Big[\det \pi \int_{t_o}^{t} Q(\tau)g^{-1}(\tau)\widetilde{Q}(\tau)d\tau\Big]^{-\frac{1}{2}} \exp\Big\{-\Big[\int_{t_o}^{t} Q(\tau)g^{-1}(\tau)d\tau\Big]^{-1}_{\alpha\beta} x_\alpha x_\beta\Big\} \quad \text{(II.4)}$$

where $\vec{x}$ is a three-dimensional vector and Q a 3×3 matrix. $\widetilde{Q}$ stands for the transposed of Q. This formula is obtained by writing the δ-functional in (II.4) as a Fourier integral, and so the problem reduces to the functional integration of a linear exponential functional. (Appendix.)

As a first example we consider a Brownian particle under the influence of a time-prescribed external force $\vec{F}(\tau)$. Furthermore, we shall take the matrix B constant, as this is the usual case with applications. In this case the Langevin equation is

$$\frac{d\vec{p}}{d\tau} = -B\vec{p} + \vec{F}(\tau) + \vec{f}(\tau). \quad \text{(II.5)}$$

Its solution satisfying the condition $\vec{P}(t_o) = \vec{p}_o$ is

$$\vec{P}(t) = \vec{U} + \int_{t_o}^{t} \exp[-B(t-\tau)]\,\vec{f}(\tau)d\tau \quad \text{(II.6)}$$

where

$$\vec{U} = \exp[-B(t-t_o)]\,\vec{p}_o + \int_{t_o}^{t} \exp[-B(t-\tau)]\,\vec{F}(\tau)d\tau.$$

Then using formula (II.4) we find for the ECPD in momentum space the result:

$$G(\vec{p}\,|\,\vec{p}_o;t|\,t_o) = \int \delta\left\{\vec{p} - \vec{U} - \int_{t_o}^{t} \exp[-B(t-\tau)]\,\vec{f}(\tau)\,d\tau\right\} W[\vec{f}(\tau)]_{t_o}^{t} \prod_{t_o \le \tau < t} d\vec{f}(\tau)$$

$$= \left[\det 2\pi m\kappa T\left(\exp[-2B(t-t_o)] - I\right)\right]^{-\frac{1}{2}} \times$$

$$\times \exp\left\{-\frac{1}{2m\kappa T}\left(\exp[-B(t-t_o)] - I\right)^{-1}_{\alpha\beta}\left(p_\alpha - U_\alpha\right)\left(p_\beta - U_\beta\right)\right\} \tag{II.7}$$

where we have used (I.5) (to be established later) for the matrix g and the symmetry property of B. For $\vec{F}(\tau) = 0$, formula (II.7) shows how the Brownian particles settle to equilibrium through the frictional force $-B\vec{p}$. The particle dissipates energy to the medium through the frictional force $-B\vec{p}$, whereas through the thermal force $\vec{f}(\tau)$ the medium does work on the particle. In the state of equilibrium the two forces balance each other.

As a final application, let us calculate the expression for the ECPD of finding the particle with momentum $\vec{p}$ at time $t+\Delta t$, if at time t it had momentum $\vec{p}\,'$. Δt is taken short enough so that the systematic forces do not change appreciably during this interval.

From the Langevin equation (II.5) we have for $\vec{P}(t) = \vec{p}\,'$

$$\vec{P}(t+\Delta t) = \vec{p}\,' + \int_t^{t+\Delta t} (-B\vec{p} + \vec{F}(\vec{p}\,,\tau))\,d\tau + \int_t^{t+\Delta t} \vec{f}(\tau)\,d\tau \simeq$$

$$\simeq \vec{p}\,' + \left[-B\vec{p}\,' + \vec{F}(\vec{p}\,',t)\right]\Delta t + \int_t^{t+\Delta t} \vec{f}(\tau)\,d\tau \tag{II.8}$$

where we have replaced the time integral of the systematic forces by the first non-vanishing term of its Taylor expansion. This is done due to the slow variation of the systematic forces. We do not do this for the thermal force due to its rapid variation.[3)]

Employing (II.3) for the construction of the ECPD and using (II.8) and (II.4), we obtain:

$$G(\vec{p}\,|\,\vec{p}\,';t+\Delta t|t) = \left[\det \pi\, g^{-1}(t)\,\Delta t\right]^{-\frac{1}{2}} \times$$

$$\times \exp\left\{-g(t)\left(\frac{\vec{p}-\vec{p}\,'}{\Delta t} + B\vec{p}\,' - \vec{F}(\vec{p}\,',t)\right)_\alpha \cdot \left(\frac{\vec{p}-\vec{p}\,'}{\Delta t} + B\vec{p}\,' - \vec{F}(\vec{p}\,',t)\right)_\beta \Delta t\right\} \tag{II.9}$$

where in (II.9) we made the simplification

$$\int_t^{t+\Delta t} g^{-1}(\tau)\,d\tau = g^{-1}(t)\Delta t.$$

Introducing the transformation

$$\vec{p}\,' = \vec{p} - \Delta\vec{p}, \tag{II.10}$$

formula (II.9) takes the form of a function $T(\vec{p}-\Delta\vec{p};\Delta\vec{p})$ and represents the transition probability density for the Brownian particle having momentum $\vec{p}-\Delta\vec{p}$ at time t to change by $\Delta\vec{p}$ in the short time Δt. We are interested in the transition probability distribution for the momentum to change from $\vec{p}$ at time t by $\Delta\vec{p}$ in the short time Δt. Replacing in (II.9) $\vec{p}\,'$ by $\vec{p}$ and $\vec{p}-\vec{p}\,'$ by $\Delta\vec{p}$ we obtain

$$T(\vec{p};\Delta\vec{p}) = \left[\det \pi\, g^{-1}(t)\Delta t\right]^{-\frac{1}{2}} \times$$

$$\times \exp\left\{-g(t)\left(\frac{\Delta\vec{p}}{\Delta t} + B\vec{p} - \vec{F}(\vec{p},t)\right)_\alpha \cdot \left(\frac{\Delta\vec{p}}{\Delta t} + B\vec{p} - \vec{F}(\vec{p},t)\right)_\beta \Delta t\right\}. \tag{II.11}$$

Defining the average value of a function $\varphi(\Delta\vec{p})$ by

$$\langle\varphi(\Delta\vec{p})\rangle = \int \varphi(\Delta\vec{p})\,T(\vec{p},\Delta\vec{p})\,d(\Delta\vec{p}), \tag{II.12}$$

we obtain, for later reference, utilizing (II.12), the results:

$$\begin{aligned} \langle 1\rangle &= 1 \\ \langle \Delta p_\alpha\rangle &= \left[F_\alpha(\vec{p},t) - (B\vec{p})_\alpha\right]\Delta t \\ \langle \Delta p_\alpha \Delta p_\beta\rangle &= \tfrac{1}{2}\,(g^{-1}(t))_{\alpha\beta}\Delta t \end{aligned} \tag{II.13}$$

$\langle \Delta p_\alpha \Delta p_\beta \Delta p_\gamma\rangle$ and so on are of higher order in Δt.

where $\alpha, \beta, \gamma, \ldots$ run from 1 to 3.

III. The Smoluchowski Equation and the Fokker-Planck Equation

Since the Langevin equation of motion (II.1) has no memory, there follows that for $t_o < t' < t$ we have

$$\vec{P}\Big(\vec{p}_o, [\vec{f}(\tau)]_{t_o}^{t}\Big) = \vec{P}\Big(\vec{P}', [\vec{f}(\tau)]_{t'}^{t}\Big) \tag{III.1}$$

where

$$\vec{P}' = \vec{P}\Big(\vec{p}_o, [f(\tau)]_{t_o}^{t'}\Big) .$$

Furthermore, using (III.1), it is easy to verify that:

$$\delta\Big\{\vec{p} - \vec{P}\Big(\vec{p}_o, [f(\tau)]_{t_o}^{t}\Big)\Big\} =$$

$$= \int \delta\Big\{\vec{p} - \vec{P}\Big(\vec{p}', [\vec{f}(\tau)]_{t'}^{t}\Big)\Big\} \cdot \delta\Big\{\vec{p}' - \vec{P}\Big(\vec{p}_o, [\vec{f}(\tau)]_{t_o}^{t'}\Big)\Big\} d\vec{p}' . \tag{III.2}$$

The thermal distribution (I.3) factorizes for any pair of disjoint subintervals covering $[t_o, t]$, i.e.,

$$W[\vec{f}(\tau)]_{t_o}^{t} = W[\vec{f}(\tau)]_{t_o}^{t'} \cdot W[\vec{f}(\tau)]_{t'}^{t} . \tag{III.3}$$

Multiplying (III.2) and (III.3) by members and integrating both sides over all $f(\tau)$ with $\tau \in [t_o, t]$, we have

$$G(\vec{p} \,|\, \vec{p}_o; t \,|\, t_o) = \int G(\vec{p} \,|\, \vec{p}'; t \,|\, t') \, G(\vec{p}' \,|\, \vec{p}_o; t' \,|\, t_o) d\vec{p}' , \tag{III.4}$$

where we have utilized formula (II.3) for the construction of the ECPD. Equation (III.4) is the Smoluchowski (-Kolmogorov-Chapman) equation for the ECPD. We note that should either the Langevin equation of motion contain memory or the thermal distribution does not factorize, then the relation (III.4) would break down.

On the Smoluchowski equation we can base all the calculations of the Brownian motion. In particular, we shall go into a differential equation—the Fokker-Planck equation.

Let us now replace in (III.4) t' by t and t by $t+\Delta t$ and introduce the transformation (II.10). Then, utilizing (II.9), we obtain

$$G(\vec{p}\,|\,\vec{p}_o;t+\Delta t|\,t_o) = \int T(\vec{p}-\Delta\vec{p};\Delta\vec{p})\,G(\vec{p}-\Delta\vec{p};t|\,t_o)d(\Delta\vec{p})\,. \qquad \text{(III.5)}$$

The Fokker-Planck equation for the probability distribution $\Phi(\vec{p},t)$ is obtained through the same integral equation (III.5), i.e.,[3]

$$\Phi(\vec{p}\,,t+\Delta t) = \int T(\vec{p}-\Delta\vec{p};\Delta\vec{p})\,\Phi(\vec{p}-\Delta\vec{p},t)d(\Delta\vec{p})\,. \qquad \text{(III.5a)}$$

Expanding the left hand side of (III.5a) in power series of Δt and the right hand side in power series of Δp_α by Taylor's theorem and making some rearrangements, then we have

$$\Phi + \frac{\partial\Phi}{\partial t}\Delta t + 0((\Delta t)^2) =$$

$$= \int\left\{\Phi T - \frac{\partial}{\partial p_\alpha}\Phi T\,\Delta p_\alpha + \frac{1}{2}\frac{\partial^2}{\partial p_\alpha \partial p_\beta}\Phi T\,\Delta p_\alpha \Delta p_\beta + \cdots\right\}d(\Delta\vec{p}) \qquad \text{(III.6)}$$

where we have denoted by Φ and T the functions $\Phi(\vec{p},t)$ and $T(\vec{p};\Delta\vec{p})$. Dividing (III.6) by Δt and taking into account the results (II.13), and passing to the limit as $\Delta t \to 0$, we obtain the Fokker-Planck equation in momentum space:

$$\left\{\frac{\partial}{\partial t} + \frac{\partial}{\partial p_\alpha}\left[F_\alpha(\vec{p},t) - (B\vec{p})_\alpha\right] - \frac{1}{4}\frac{\partial^2}{\partial p_\alpha \partial p_\beta}(g^{-1})_{\alpha\beta}\right\}\Phi = 0. \qquad \text{(III.7)}$$

Demanding of this equation to admit the Maxwellian distribution for the momenta (I.4) in the case of the free Brownian particle (i.e., $\vec{F} = 0$), we establish for g the relation (I.5): $g^{-1} = 4\kappa T m B$.

Since for $t > t_o$ the ECPD G satisfies (III.5a), there follows that (for $t > t_o$) G solves the Fokker-Planck equation (III.7). Furthermore, by replacing t by t_o, $t+\Delta t$ by t, and $\vec{p}\,'$ by $\vec{p}_o$ in (II.9), it is easy to see that as $\Delta t \to 0$ we have:

$$G(\vec{p}\,|\,\vec{p}_o;t|\,t_o) \longrightarrow \delta(\vec{p}-\vec{p}_o) \quad \text{as} \quad t \longrightarrow t_o + 0. \qquad \text{(III.8)}$$

From property (III.8) and the fact that G satisfies the Fokker-Planck equation (III.7), there follows that the ECPD G defined in (II.3) is a Green function of the Fokker-Planck equation. In particular, it is the Green function satisfying the integrability condition.

The ECPD G has the property to propagate the solutions of the Fokker-Planck equation. In other words, given the distribution for the momenta $\Phi_o(\vec{p})$ at time t_o, the distribution $\Phi(\vec{p},t)$ at a later time t, which solves the Fokker-Planck equation and satisfies the same boundary conditions as G w.r.t. $\vec{p}$ is given by:

$$\Phi(\vec{p},t) = \int G(\vec{p}\,|\,\vec{p}_o;t|\,t_o)\,\Phi(\vec{p}_o)d\vec{p}_o. \qquad \text{(III.9)}$$

This is the propagation equation. That Φ defined in (III.9) solves the Fokker-Planck equation is easily seen, since G w.r.t. $\vec{p},t$ satisfies this equation. The same applies for the boundary conditions. The initial condition requirement follows from property (III.8). We have

$$\Phi(\vec{p},t) \longrightarrow \int \delta(\vec{p}-\vec{p}_o)\,\Phi_o(\vec{p}_o)d\vec{p}_o = \Phi_o(\vec{p}) \quad \text{as} \quad t \longrightarrow t_o+0.$$

IV. The ECPD $G(\vec{p}\,|\,\vec{p}_o;t|\,t_o)$ as a Conditional Functional Integral in Momentum Space

Consider a fine subdivision of the interval $[t_o,t]$:

$$\mathcal{D}_n = \{t_0, t_1, t_2, \cdots, t_n\}$$

with

$$t_0 < t_1 < t_2 < \cdots < t_n = t.$$

Repeated application of Smoluchowski's integral equation (III.4) gives

$$G(\vec{p}\,|\,\vec{p}_o;t|\,t_o) =$$

$$= \int G\left(\vec{p}\,|\,\vec{p}(n-1);t|\,t_{n-1}\right) G\left(\vec{p}(n-1)\,|\,\vec{p}(n-2);t_{n-1}|\,t_{n-2}\right) \cdots \times$$

$$G\left(\vec{p}(2)\,|\,\vec{p}(1);t_2|\,t_1\right) \times G\left(\vec{p}(1)\,|\,\vec{p}_o;t_1|\,t_o\right) d\vec{p}(n-1)d\vec{p}(n-2)\cdots \times$$

$$d\vec{p}(2)d\vec{p}(1). \qquad \text{(IV.1)}$$

Employing formula (II.9) for the ECPD between two neighboring times, we have

$$G(\vec{p}|\vec{p}_o;t|t_o) \approx G^{(n)} = \Big[\prod_{j=0}^{n-1} \det \pi\, g^{-1}(t_j)\Delta t_j\Big]^{-\frac{1}{2}} \int \exp\Big\{-\sum_{j=0}^{n-1} g_{\alpha\beta}(j) \times$$

$$\Big(\frac{\vec{p}(j+1)-\vec{p}(j)}{\Delta t_j} - \vec{F}(j) + B(j)\vec{p}(j)\Big)_\alpha \cdot \Big(\frac{\vec{p}(j+1)-\vec{p}(j)}{\Delta t_j} - \vec{F}(j) + B(j)\vec{p}(j)\Big)_\beta \Delta t_j\Big\} \times$$

$$\delta(\vec{p}(0) - \vec{p}_o)\,\delta(\vec{p}(n) - \vec{p}) \prod_{j=0}^{n} d\vec{p}(j), \qquad \text{(IV.2)}$$

where we have denoted

$$\Delta t_j = t_{j+1} - t_j, \quad \vec{p}(j) = \vec{p}(t_j), \quad \vec{F}(j) = \vec{F}(\vec{p}(t_j), t_j).$$

Passing to the limit as $n \to \infty$ provided $\max.\Delta t_j \to 0$, this limit (when it exists independently of the choice of the sequence of subdivisions $\{\mathcal{D}_n\}$) defines a conditional functional integral in momentum space, representing G, which we denote in the following suggestive manner:

$$G(\vec{p}\,|\,\vec{p}_o;t|\,t_o) = \Big[\prod_{t_o \le \tau < t} \det \pi\, g^{-1}(\tau)d\tau\Big]^{-\frac{1}{2}} \int_{\vec{p}(t_o)=\vec{p}_o}^{\vec{p}(t)=\vec{p}} \exp\Big\{-\int_{t_o}^{t} g_{\alpha\beta}(\tau) \times$$

$$\Big(\frac{d\vec{p}(\tau)}{d\tau} - \vec{F}(\vec{p}(\tau),\tau) + B(\tau)\vec{p}(\tau)\Big)_\alpha \Big(\frac{d\vec{p}(\tau)}{d\tau} - \vec{F}(\vec{p}(\tau),\tau) + B(\tau)\vec{p}(\tau)\Big)_\beta d\tau\Big\} \times$$

$$\prod_{t_o \le \tau < t} d\vec{p}(\tau). \qquad \text{(IV.3)}$$

Notice the difference in the normalization factor from that of (I.3) for the thermal force distribution. The normalization factor in (IV.3) is such that one could express this integral as a conditional integral over the Wiener measure:

$$\hbar \exp\Big\{-\int_{t_o}^{t} g_{\alpha\beta}(\tau)\frac{dp_\alpha}{d\tau}\frac{dp_\beta}{d\tau}d\tau\Big\}$$

where $\hbar$ is the symbolic quantity in front of the integral sign in (IV.3).

We also remark that one could arrive at the representation (IV.3) for the ECPD G through transformation theory by employing the original definition of G (II.3) and the transformation (II.5). It is also possible to express G as a conditional functional integral in the coordinate space.

V. Approximation Methods for the Green Function

The functional integral representation (IV.3) of the Green function of the Fokker-Planck equation (III.7) is not only a formal expression, but its practical importance lies in the fact that it is offered to various approximation procedures. By employing (IV.3), one could devise these techniques by introducing various transformations of $\vec{p}(\tau)$, which transform part of the integral into a functionally-known integrable form. The rest is treated as a perturbation.

One particular procedure, which usually picks up most of the Green function in the zero order approximation, is analogous to the WKB approximation in Quantum Mechanics.[5] We wish to demonstrate this technique by treating in detail the case of one-dimensional momentum. The extension to three-dimensions is trivially effected by replacing the scalar quantities involved with the corresponding matrices and vectors. Furthermore we shall take g constant as this is the usual case for applications.

We begin with writing down (IV.3) in one-dimensional form in the momentum.

$$G(p\,|\,p_o;t\,|\,t_o) = \Big[\prod_{t_o\le\tau<t} \pi g^{-1} d\tau\Big]^{-\frac{1}{2}} \times$$

$$\times \int_{p(t_o)=p_o}^{p(t)=p} \exp\Big\{-\int_{t_o}^{t} g\Big[\frac{dp(\tau)}{d\tau} - F(p(\tau),\tau) + Bp(\tau)\Big]^2 d\tau\Big\} \prod_{t_o\le\tau<t} dp(\tau). \quad (V.1)$$

Since g is positive, there follows that the definite integral in the exponent of (V.1) is positive for every path $p(\tau)$. For smooth $F(p,\tau)$ there exists a certain path through (t_o,p_o) and (t,p) for which the definite integral is minimized and hence the exponential functional is maximized. Then most of the contribution, from the integrations over $p(\tau)$ to the functional integral (V.1), comes from a neighborhood around this path.

To find this path we apply the usual methods of the calculus of variations. We have for the required path:

$$\delta \int_{t_o}^{t} g\Big[\frac{dp}{d\tau} - F(p,\tau) + Bp\Big]^2 d\tau = 0 \quad (V.2a)$$

together with the conditions

$$p(t_o) = p_o, \quad p(t) = p \tag{V.2b}$$

from which there follows that the required path is the solution of the second order ordinary differential equation:

$$\frac{d^2p}{d\tau^2} - \left(\frac{\partial F}{\partial \tau} + B^2 p\right) + \left(Bp - F\right)\frac{\partial F}{\partial p} = 0 \tag{V.3}$$

which passes through (t_o, p_o) and (t, p). This equation, although in general non-linear, nevertheless is an ordinary one and it is easier to handle numerically than a partial differential equation.

Let $p^*(\tau)$ be the path solving (V.3) and passing through (t_o, p_o); (t, p). Upon introducing the transformation

$$p(\tau) = p^*(\tau) + k(\tau), \tag{V.4}$$

we have

$$\int_{t_o}^{t} \left(\dot{p}(\tau) - F(p,\tau) + Bp(\tau)\right)^2 d\tau = A + \int_{t_o}^{t} \left(\dot{k}^2(\tau) + \alpha(\tau)k(\tau)\dot{k}(\tau) + \gamma(\tau)k^2(\tau)\right)d\tau$$

$$+ \text{ terms of higher power in } k \tag{V.5}$$

where we have denoted differentiation w.r.t. τ by a dot. Notice that the term first order in k vanishes since the first variation (V.2) of the left hand side of (V.5) is taken zero. We have denoted by

$$A = \int_{t_o}^{t} \left(\frac{dp^*}{d\tau} - F(p^*,\tau) + Bp^*\right)^2 d\tau = A(p, p_o; t, t_o)$$

$$\alpha(\tau) = 2\left[B - \left(\frac{\partial F}{\partial p}\right)_{p=p^*}\right] = \alpha(p, p_o; \tau) \tag{V.6}$$

$$\gamma(\tau) = \left[\alpha^2(\tau) - \left(\frac{dp^*}{d\tau} - F(p^*,\tau) + Bp^*\right)\cdot\left(\frac{\partial^2 F}{\partial p^2}\right)_{p=p^*}\right] = \gamma(p, p_o; \tau).$$

The Jacobian of the transformation $p \to k$ is: $J(p \to k) = 1$ and due to (V.2b) it follows

$$k(t_o) = k(t) = 0. \tag{V.7}$$

Therefore we have for the Green function:

$$G(p|p_o;t|t_o) \approx G_o(p|p_o;t|t_o) =$$

$$= e^{-gA}\Big[\prod_{t_o\le\tau<t} \pi\, g^{-1} d\tau\Big]^{-\frac{1}{2}} \int_{k(t_o)=0}^{k(t)=0} \exp\Big\{-g\int_{t_o}^{t}\Big(\dot{k}^2 + \tag{V.8}$$

$$+ \alpha(\tau)k\dot{k} + \gamma(\tau)k^2\Big)\,d\tau\Big\} \prod_{t_o\le\tau<t} dk(\tau).$$

The approximation sign in (V.8) is due to the omission of powers higher than 2 in k.

To calculate the continual Gaussian integral in (V.8) we pass to the discrete form. Let us consider a subdivision of the interval $[t_o, t]$:

$$\mathfrak{D}_N = \{t_o; t_1, t_2, \ldots, t_N\}$$

with

$$t_o < t_1 < t_2 < \ldots < t_N = t.$$

For simplicity we take $\mathfrak{D}_N$ isometric, i.e., $t_{j+1} - t_j = \Delta t$ for $(j = 0, 1, 2, \ldots, N-1)$. Then we form the expression $G_o^{(N)}$ by replacing the integrals and symbolic products in (V.8) by sums and products over the points of the subdivision $\mathfrak{D}_N$. Denoting by S the integral in the exponent of (V.8), we have for the discrete form the sum:

$$S_N = \sum_{j=o}^{N-1} \Big\{\frac{[k(j+1)-k(j)]^2}{\Delta t} + \alpha_j k(j)[k(j+1) - k(j)] + \gamma_j k^2(j)\Delta t\Big\} = \tag{V.9}$$

$$= \frac{1}{\Delta t}\sum_{j=1}^{N-1}\Big\{\Big[2 - \alpha_j\Delta t + \gamma_j(\Delta t)^2\Big]k^2(j) + 2\Big[1 - \tfrac{1}{2}\alpha_j(\Delta t)^2\Big]k(j)k(j+1)\Big\}$$

where in (V.9) we have taken into account the condition (V.7), which in the discrete case reads

$$k(0) = K(N) = 0.$$

Here again we have denoted $\alpha(t_j)$, $\gamma(t_j)$ and $k(t_j)$ by α_j, γ_j and $k(j)$.

Let us now transform (V.9) through a principal axis transformation as follows:[6] We cast (V.9) in the form

$$S_N = \frac{1}{\Delta t} \sum_{j=1}^{N-1} \Big[k(j) - b_{j+1} k(j+1) \Big] \Lambda_j \Big[k(j) - b_{j+1} k(j+1) \Big] \quad (V.10)$$

where $k(N) = 0$. In case k is a three-dimensional vector then b and Λ are 3×3 matrices and the quantity in the first square brackets of (V.10) is transposed.

Now comparing the right hand side of (V.9) and (V.10) we obtain

$$\begin{aligned} \Lambda_{j+1} + \Lambda_j b_{j+1}^2 &= 2 - \alpha_{j+1} \Delta t + \gamma_{j+1} (\Delta t)^2 \\ \Lambda_j b_{j+1} &= 1 - \tfrac{1}{2} \alpha_j \Delta t \\ (j &= 1, 2, \ldots, N-1) \end{aligned} \quad (V.11)$$

with

$$\Lambda_1 = 2 - \alpha_1 \Delta t + \gamma_1 (\Delta t)^2 .$$

From (V.11) we obtain

$$\begin{aligned} b_{j+1} &= \Lambda_j^{-1} (1 - \tfrac{1}{2} \alpha_j \Delta t) \\ \Lambda_{j+1} &= -\Lambda_j^{-1} (1 - \tfrac{1}{2} \alpha_j \Delta t)^2 + \Big[2 - \alpha_{j+1} \Delta t + \gamma_{j+1} (\Delta t)^2 \Big] . \end{aligned} \quad (V.12)$$

We make a further substitution:

$$\begin{aligned} \xi_j &= k(j) - b_{j+1} k(j+1) \\ (j &= 1, 2, \ldots, N-2) \\ \xi_{N-1} &= k(N-1) . \end{aligned} \quad (V.13)$$

Again the Jacobian $J(k \to \xi) = 1$. Employing (V.13), we write for (V.10):

$$S_N = \frac{1}{\Delta t} \sum_{j=1}^{N-1} \Lambda_j \xi_j^2 . \quad (V.14)$$

With the aid of (V.14) the discrete form of (V.8) becomes:

$$G_o^{(N)} = e^{-gA}\left[\prod_{j=o}^{N-1} \pi g^{-1}\Delta t\right]^{-\frac{1}{2}} \int \exp\left\{-\frac{g}{\Delta t}\sum_{j=1}^{N-1}\Lambda_j \xi_j^2\right\} \prod_{j=1}^{N-1} d\xi_j$$

$$= e^{-gA}\left[\frac{g}{\pi\Delta t \prod_{j=1}^{N-1}\Lambda_j}\right]^{\frac{1}{2}}. \tag{V.15}$$

What we need now for passing to the limit as $N\to\infty$ is to find:

$$\lim_{N\to\infty} \Delta t \prod_{j=1}^{N-1} \Lambda_j (= \text{let } D(t)) = D(p,p_o;t,t_o). \tag{V.16}$$

To do this let

$$D_n = \Delta t \prod_{j=1}^{n} \Lambda_j \tag{V.17}$$

then

$$D_{n+1} = D_n \Lambda_{n+1}. \tag{V.18}$$

Employing (V.18) and the second relation of (V.12), we obtain the finite difference equation for D_n.

$$\frac{D_{n+1}-2D_n+D_{n-1}}{(\Delta t)^2} = -\frac{D_n\alpha_{n+1}-D_{n-1}\alpha_n}{\Delta t} - \frac{1}{4}D_{n-1}\alpha_n^2 + D_n\gamma_{n+1}. \tag{V.19}$$

Passing to the limit as $\Delta t\to 0$, we obtain for $D(t)$ defined in (V.16) the differential equation:

$$\frac{d^2D}{d\tau^2} + \frac{d}{d\tau}(\alpha(\tau)D) + \left[\frac{1}{4}\alpha^2(\tau) - \gamma(\tau)\right]D = 0. \tag{V.20}$$

The required $D(t)$ is the solution of (V.20) satisfying the initial conditions:

$$D(0) = 0, \quad \dot{D}(0) = 1 \tag{V.21}$$

since

$$D_1 = \Delta t\left[2 - \alpha_1\Delta t + \gamma_1(\Delta t)^2\right] \to 0 \quad \text{as} \quad \Delta t\to 0$$

and

$$\frac{D_2 - D_1}{\Delta t} \to 1 \quad \text{as} \quad \Delta t \to 0.$$

From (V.15) and (V.16) we obtain the following approximate result for G:

$$\lim_{N\to\infty} G_o^{(N)} = G_o(p|\, p_o;t|\, t_o) = \left[\frac{g}{\pi D(p, p_o;t, t_o)}\right]^{\frac{1}{2}} e^{-gA(p, p_o;t, t_o)}. \tag{V.22}$$

This result can be taken as zero order approximation to G. We can further improve the approximation by expanding the rest of the exponential functional in (V.1) in power series of k and subsequently in ξ and employ the measure in (V.15) for the averaging over ξ. Dr. Tarski has given in his lectures formulae for such averages.

The present work originated from a series of lectures on functional integration which I gave at the Mathematics Department of the University of Salford in October, 1966. I would like to thank Professor S. F. Edwards, F.R.S., Manchester University, and Dr. S. Sampanthar, University of Salford for many useful discussions.

References

1. A. Einstein, "Investigations on the Theory of Brownian Movement," Page 1, Dover Publications, Inc. (1956).
2. L. D. Landau and E. M. Lifshitz, "Fluid Mechanics," Page 67, Pergamon Press, London (1959).
3. S. Chandrasekhar, R.M.P. 15, No. 1, 1-96 (1943).
4. S. F. Edwards, J. Fl. Mec. 18, Part 2, 239-273 (1964).
5. R. P. Feynman and A. R. Hibbs, "Quantum Mechanics and Path Integrals," Pages 58-61, McGraw-Hill Co. (1965).
6. R. Abé, Busseiron Kenkyu 79, 101 (1954).

Appendix

We wish to evaluate the functional integral:

$$J = \int \delta\left\{\vec{x} - \int_{t_o}^{t} Q(\tau)\vec{f}(\tau)d\tau\right\}\left[\prod_{t_o \le \tau < t} \det\pi^{-1} g(\tau)d\tau\right]^{\frac{1}{2}} \tag{A.1}$$

$$\exp\left\{-\int_{t_o}^{t} g_{\alpha\beta}(\tau)f_\alpha(\tau)f_\beta(\tau)d\tau\right\} \prod_{t_o \le \tau < t} d\vec{f}(\tau).$$

$\vec{x}$, $\vec{f}$ can be n-dimensional vectors but for our purposes will be taken

as three-dimensional. For the same reason $Q(\tau)$ is a 3×3 matrix. We recall that $g(\tau)$ is a 3×3 positive definite symmetric matrix.

For the evaluation of the functional integral (A.1), it is convenient to express the δ-functional involved in Fourier form. We have

$$J = \int\Big[\frac{1}{(2\pi)^3}\int \exp\Big\{i\lambda_\alpha x_\alpha - i\lambda_\alpha\int_{t_o}^{t} Q_{\alpha\gamma}(\tau)f_\gamma(\tau)d\tau\Big\}\prod_{\alpha=1}^{3} d\lambda_\alpha\Big]\times$$

$$\times\Big[\prod_{t_o\le\tau<t}\det\pi^{-1}g(\tau)d\tau\Big]^{\frac{1}{2}}\exp\Big\{-\int_{t_o}^{t} g_{\alpha\beta}(\tau)f_\alpha(\tau)f_\beta(\tau)d\tau\Big\}\prod_{t_o\le\tau<t} d\vec{f}(\tau). \tag{A.2}$$

To evaluate (A.2) we consider a sequence of subdivisions $\{\mathfrak{D}_N\}$ of the interval $[t_o,t]$

$$\mathfrak{D}_N = \{\tau_o, \tau_1, \tau_2, \ldots, \tau_N\}$$

with the property: $t_o<\tau_1<\tau_2<\ldots<\tau_N=t$ and $\max\Delta\tau_j\to0$ as $N\to\infty$. Let us associate with each $\mathfrak{D}_N$ the N-tuble integral:

$$J_{\mathfrak{D}_N} = \int_{-\infty}^{+\infty} 1\ \text{to}\ N\int\Big[\frac{1}{(2\pi)^3}\int\exp\Big\{i\lambda_\gamma x_\gamma - i\sum_{j=o}^{N-1}\lambda_\gamma Q_{\gamma\alpha}(j)f_\alpha(j)\Delta\tau_j\Big\}\prod_{\gamma=1}^{3} d\lambda_\gamma\Big]\times$$

$$\times\Big[\prod_{j=o}^{N-1}\det\pi^{-1}g(j)\Delta\tau_j\Big]^{\frac{1}{2}}\exp\Big\{-\sum_{j=o}^{N-1} g_{\alpha\beta}(j)f_\alpha(j)f_\beta(j)\Delta\tau_j\Big\}\prod_{\substack{j=o\\ \alpha=1,2,3}}^{N-1} df_\alpha(j) \tag{A.3}$$

where we have denoted $Q(j)=Q(\tau_j)$, $f_\alpha(j)=f_\alpha(\tau_j)$, $g(j)=g(\tau_j)$. Let us introduce at this stage the formulae

$$\int\exp\Big\{-g'_{\alpha\beta}f_\alpha f_\beta\Big\}\prod_{\alpha=1}^{3} df_\alpha = \Big[\det\pi^{-1}g'\Big]^{-\frac{1}{2}}$$

$$\int\exp\Big\{\pm ib_\alpha f_\alpha - g'_{\alpha\beta}f_\alpha f_\beta\Big\}\prod_{\alpha=1}^{3} df_\alpha = \Big[\det\pi^{-1}g'\Big]^{-\frac{1}{2}}\exp\Big\{-\frac{1}{4}(g'^{-1})_{\alpha\beta}b_\alpha b_\beta\Big\} \tag{A.4}$$

where g' is a 3×3 positive definite matrix. With the aid of formulae (A.4) we perform the integration over all $f_\alpha(j)$ in (A.3) and obtain:

$$J_{\mathfrak{S}_N} = \frac{1}{(2\pi)^3} \int \exp\Big\{ i\lambda_\gamma x_\gamma - \frac{1}{4} \sum_{j=0}^{N-1} \Big[g^{-1}_{(j)} \Big]_{\alpha\beta} Q_{\gamma\alpha}(j) Q_{\gamma\beta}(j) \lambda_\gamma \lambda_\delta \Delta\tau_j \Big\} \prod_{\gamma=1}^{3} d\lambda_\gamma . \tag{A.5}$$

Utilizing the fact

$$Q_{\gamma\alpha} \Big[g^{-1} \Big]_{\alpha\beta} Q_{\delta\beta} = \Big[Q g^{-1} \widetilde{Q} \Big]_{\gamma\delta}$$

and passing to the limit as $N \to +\infty$, we find for the functional integral (A.1) the result

$$J = \frac{1}{(2\pi)^3} \int \exp\Big\{ i\lambda_\gamma x_\gamma - \frac{1}{4} \int_{t_o}^{t} \Big[Q(\tau) g^{-1}(\tau) \widetilde{Q}(\tau) \Big]_{\gamma\delta} \lambda_\gamma \lambda_\delta \Big\} \prod_{\gamma=1}^{3} d\lambda_\gamma . \tag{A.6}$$

It is desirable to get rid of the parameter λ_γ. Upon performing the integration over λ_γ in (A.6), utilizing the second formula of (A.4), we establish formula (II.4), i.e.,

$$J = \Big[\det \pi \int_{t_o}^{t} Q(\tau) g^{-1}(\tau) \widetilde{Q}(\tau) d\tau \Big]^{-\frac{1}{2}} \exp\Big\{ -\Big[\int_{t_o}^{t} Q(\tau) g^{-1}(\tau) \widetilde{Q}(\tau) d\tau \Big]^{-1}_{\alpha\beta} x_\alpha x_\beta \Big\} . \tag{A.7}$$

We shall close the appendix by stating the result of a generalization of formula (A.7) for the functional integral of products of δ-functionals of the form appearing in (A.1). We have

$$\int \Big(\prod_{j=1}^{n} \delta \Big\{ \vec{x}^{(j)} - \int_{t_o}^{t} Q^{(j)}(\tau) \vec{f}(\tau) d\tau \Big\} \Big) \Big[\prod_{t_o \le \tau < t} \det \pi^{-1} g(\tau) d\tau \Big]^{\frac{1}{2}} \exp\Big\{ -$$

$$- \int_{t_o}^{t} g_{\alpha\beta}(\tau) f_\alpha(\tau) f_\beta(\tau) d\tau \Big\} \prod_{t_o \le \tau < t} d\vec{f}(\tau) =$$

$$= \Big[\det \pi \int_{t_o}^{t} G(\tau) d\tau \Big]^{-\frac{1}{2}} \exp\Big\{ - \sum_{r,s=1}^{3n} \Big[\int_{t_o}^{t} G(\tau) d\tau \Big]^{-1}_{rs} x'_r x'_s \Big\} \tag{A.8}$$

where

$$x'_{\alpha+3(j-1)} = x_{\alpha}^{(j)} \quad (\alpha = 1,2,3); \quad (j = 1, 2, \ldots, n)$$

and $G(\tau)$ is a $3n \times 3n$ (positive definite matrix given in the partitioned form:

$$G(\tau) = \begin{bmatrix} Q^{(1)}{}_g{}^{-1}\tilde{Q}^{(1)} & \cdots & Q^{(1)}{}_g{}^{-1}\tilde{Q}^{(n)} \\ \cdot & \cdots & \cdot \\ \cdot & \cdots & \cdot \\ Q^{(n)}{}_g{}^{-1}\tilde{Q}^{(1)} & \cdots & Q^{(n)}{}_g{}^{-1}\tilde{Q}^{(n)} \end{bmatrix}$$

Formula (A.8) is particularly useful if one is looking for joint conditional probability distributions.

NON-EQUILIBRIUM STATISTICAL MECHANICS: IRREVERSIBILITY AND MACROSCOPIC CAUSALITY*

G. L. Sewell
Department of Physics
Queen Mary College
London

Contents

* * *

I. Introduction and Discussion

The subject-matter of this course will be concerned with a statistical mechanical formulation of macroscopic dynamical laws. The existence of such laws (e.g., hydrodynamics, heat conduction) is an empirical fact. In general the laws are expressible in terms of a set of macroscopic (extensive) variables, $\{a\} = (a_1, \ldots, a_m)$, appropriate to the system under consideration. These variables then constitute a description of the macrostate. Their dynamics correspond to a self-contained set of equations of motion, describing an irreversible approach to equilibrium:

$$\frac{da_j}{dt} = \varphi_j(a) \tag{I.1}$$

These equations represent a causal macroscopic law in the sense that the initial values of the variables a_j determine their later values.

The characterisation of a set of macroscopic variables conforming to such a law presents a serious problem. It is evident from elementary phenomenological considerations that the set depends not

*Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

only on the system under consideration but also on its thermodynamic phase. For example, in the case of an insulating crystal, the variables $\{a\}$ could be chosen to be energies of macroscopically small subvolumes each containing an enormously large number of atoms—these variables would then change in time according to the law of heat conduction. On the other hand, in the case of a fluid, additional macroscopic variables (mass currents, energies, energy currents of the subvolumes) would be needed in order to obtain a closed macroscopic law (hydrodynamics).

The principal tasks of our statistical-mechanical theory will be to answer the following questions:

(i) How is it possible that a set of macroscopic variables can evolve according to a closed causal law, that determines their values at time t (>0) from their initial values, despite the fact that these variables provide only an incomplete kinematical description of the system?

(ii) How is it possible that the macroscopic law constitutes an irreversible evolution of $\{a\}$, despite the fact that the microscopic equations of motion for the system, whether classical or quantal, are invariant under time reversals?

(iii) What are the special properties of a set $\{a\}$ that leads to a closed, time-irreversible law?

(iv) What are the forms of the functions $\varphi_j(a)$, in Eq. (I.1), which determine the explicit form of the macroscopic law for a specified system?

(v) What is the relationship of phenomenological quantities, such as transport coefficients, that occur in $\varphi_j(a)$, to microscopic properties of the system?

In connection with the paradoxes raised by questions (i), (ii), we shall show that the variables $\{a\}$ can evolve according to a closed, time-irreversible law only in an approximation which is, however, extremely good for suitably chosen macroscopic variables. To be more specific, we shall show that $\{a\}$ may evolve according to such a law only to zero order in a very small dimensionless parameter, Γ^{-1}, where Γ represents a certain characteristic ratio of macroscopic to microscopic quantities. Thus, as Γ is enormously large, the actual evolution of a suitably chosen set of variables $\{a\}$ is insensibly different from that described by the closed, time-irreversible law (I.1). Higher order corrections will be seen to correspond to fluctuations and to 'memory effects,' i.e., to effects whereby $\dot{a}$ depends not only on the instantaneous value of a but also on its values at earlier times.

Our formulation of macroscopic laws will be based on a treatment of the Liouville equation (classical or quantal) for the system, which is assumed to be initially prepared by measurement of a set of

macroscopic variables $\{a\}$. The treatment proceeds in four principal stages. Firstly, an exact, generalised master equation,[1),2)] governing the evolution of the macroscopic variables, is derived from the quantum-mechanical Liouville equation. This master equation takes the form of an integro-differential equation for a well-defined distribution function, P, governing the macroscopic variables:

$$\frac{dP_t}{dt} = \int_0^t dt' G(t') P_{t-t'} \qquad (I.2)$$

where G is an operator on P. The form of G depends on the microstructure of the system and on the variables $\{a\}$. It is important to note that this master equation is, of itself, too general to provide any answer to our above questions (i)-(iii). This may be seen from the fact that the equation is derived for an arbitrary set of variables, $\{a\}$, which may or may not conform to a closed causal law, and which might not even be macroscopic.

The second stage is to introduce a characterisation[3)] of the macroscopic variables in terms of general properties that these variables possess as a result of their many-particle structures. This characterisation will be expressed in terms of the very large dimensionless size parameter, Γ, referred to above. Specifically, it will be designed to evince the Γ-dependence of the variables pertinent to the master equation; in this connection it is important to realise that the kernel G contains microscopic as well as macroscopic variables. In view of the tremendous complexity of many-body problems, our formulation of the 'essential' characteristics of macroscopic and other variables rests inevitably on a number of assumptions, which are discussed fully in Section VI.

The third stage of the theory[3)] is to incorporate these characteristic properties of the variables into our formulation of the kernel G. In this way, we are able to obtain conditions on these variables for which the kernel G(t) decays in such a way that 'memory effects' are eliminated from the master equation, to lowest order in Γ^{-1}. Under these conditions, which depend crucially on the many-particle structures of the macroscopic variables, the master equation reduces to a Markoffian form

$$\frac{dP_t}{dt} = \overline{G} P_t; \qquad \overline{G} = \int_0^\infty dt\, G(t). \qquad (I.3)$$

In general, this Markoffian equation describes an irreversible approach to equilibrium.

The final stage of the theory is to analyse the Markoffian master equation[7)] and thereby to obtain conditions on $\overline{G}$, and thus on

negligible at all times—i.e., under which the dispersion is of the order of an appropriately low power of Γ. Under these conditions the master equation yields a closed, causal law of the form (I.1). The functions φ_j, which appear in this law, are expressed in terms of microscopic properties of the system.

II. Microstates and Macrostates

We shall be concerned with a description of an arbitrary system composed of a very large number of particles. It will be assumed that, in the situations considered, only a relatively few of the variables of the system are subjected to measurement. Consequently, the values of the complete set of compatible microscopic variables can be specified only in statistical terms. This means that we must describe the microscopic properties of the system in terms of mixed, rather than pure states. In other words, we must represent the microstate of a classical system by the Liouville distribution function for the full set of its coordinates and momenta; and that of a quantal system by its density matrix. We shall use a formalism, constructed by Emch[2] for quantal systems, in which the mixed states are represented by vectors in an appropriate Hilbert space—needless to say, this is not the Hilbert space of the pure states, but an associated space. I shall find it useful, for pedagogical reasons, to present not only the quantal formalism but also its classical analogue in which a mixed state is likewise represented by a vector in an appropriate Hilbert space. This will serve the purpose of demonstrating that the classical and quantal statistical theories possess the same mathematical structure—the essential difference between the two theories lies in their descriptions of pure states.

Classical Case.

Denote the full set of Cartesian coordinates and momenta for the particles of the system by x,p. The space of this set of continuous variables is the phase space, Σ, for the system. A pure microstate corresponds to a point in Σ, while a mixed one is represented by a single-valued function on Σ—namely, the Liouville distribution function f. It will be assumed that this function is real, non-negative, square-integrable over Σ so that

$$\int_{\Sigma} f\,dx\,dp = 1.$$

Accordingly the microstates, which we shall henceforth take to be mixed, correspond to elements f of a Hilbert space $\mathfrak{Q}$, defined as the

space of complex† square-integrable functions, g, on Σ, with inner products defined by

$$(g_1, g_2) = \int_\Sigma dx\, dp\, g_1^*(x,p) g_2(x,p). \tag{II.1}$$

We shall refer to $\mathfrak{L}$ as the Liouville space.

Corresponding to each state f, there is a time reversed state $\mathfrak{I}f$, the time-reversal operator $\mathfrak{I}$ being defined by

$$\mathfrak{I}g(x,p) = g(x,-p). \tag{II.2}$$

In order to define the macrostates of the system we introduce a set of macroscopic variables (observables)

$$A(x,p) \equiv (A_1(x,p), \ldots, A_m(x,p)).$$

These are single-valued functions, defined everywhere on Σ. In general they will represent extensive variables—but this property will not be invoked until a later stage of the theory (Section VI). It will be assumed that each A_j is either an odd or an even function of the momenta p, i.e.,

$$A_j(x,-p) = \tau_j A_j(x,p) \tag{II.3}$$

where each τ_j is ± 1. This assumption is satisfied for the macroscopic variables of usual interest, e.g., masses, energies, mass currents, energy currents in subvolumes of a system.

We idealise a macroscopic measurement of A_j, say, as one which establishes that the value of that variable lies in a well-defined interval, $(a_j - \frac{1}{2}\Delta_j, a_j + \frac{1}{2}\Delta_j)$, say (cf. van Kampen[5]). Evidently, Δ_j represents the experimental error in the measurement of A_j. In order to obtain a geometrical representation of the measurement, we divide the phase space Σ into cells $C^{(r)}$, each one of which is bounded by $2m$ hypersurfaces

$$A_j = a_j^{(r)} \pm \tfrac{1}{2}\Delta_j, \quad \text{for} \quad j = 1, \ldots, m;$$

the values of $a_j^{(r)}$, for given j, being spaced at intervals Δ_j. Hence $C^{(r)}$ corresponds to the set of intervals

---

† The Hilbert space needs to include complex functions on Σ in order to be large enough for a full description of our operations.

$$\left(a_j^{(r)} - \tfrac{1}{2}\Delta_j,\ a_j^{(r)} + \tfrac{1}{2}\Delta_j\right),\quad \text{for } j = 1,\ldots,m.$$

Consequently, a measurement of the variables $\{A\}$ may be idealised as one that determines which of the cells $C^{(r)}$ contains the phase point for the system

For notational convenience we shall henceforth denote the phase cells by $C(a)$ rather than $C^{(r)}$, where

$$a = (a_1,\ldots,a_m) \equiv (a_1^{(r)},\ldots,a_m^{(r)}). \tag{II.4}$$

The space of the discrete variables a will be referred to as the macrospace and denoted by M.

As a macromeasurement determines $C(a)$ precisely, it constitutes an exact measurement of a set of variables $\tilde{A} = (\tilde{A}_1,\ldots,\tilde{A}_m)$, defined so that A takes the value a throughout $C(a)$. Thus

$$\tilde{A}(x,p) = \sum_a a\, D(a;x,p) \tag{II.5}$$

where $D(a)$ is the characteristic function for $C(a)$; i.e., it is unity if the phase point lies in $C(a)$, and is otherwise zero. We shall henceforth refer to $\tilde{A}$, rather than A, as the macroscopic observables, since it is the former set of variables that are precisely measurable, at least in our idealisation. It may be seen that the value of $\tilde{A}$ at any phase point approximates to that of A, within an accuracy $\Delta = (\Delta_1,\ldots,\Delta_m)$.

It follows from our definition of the Liouville space and of $D(a)$ that this latter function (on Σ) is an element of $\mathfrak{L}$. It also follows from Eq. (II.3) and our definition of $D(a)$ that

$$D(a;x,-p) = D(a^T;x,p) \tag{II.6}$$

with

$$a^T = (\tau_1 a_1,\ldots,\tau_m a_m). \tag{II.7}$$

Hence, by Eq. (II.2),

$$\mathcal{J}D(a) = D(a^T). \tag{II.8}$$

The probability that $\tilde{A}$ takes the value a, when the microstate is f, is given by

$$P(a) = \int_{C(a)} f(x,p)dx\,dp \equiv (D(a), f). \tag{II.9}$$

This function P, on M, is therefore the probability distribution for the macroscopic observables. We shall henceforth represent the macrostate by this function. It follows then from Eqs. (II.8) and (II.9) that the macrostate which corresponds to the time-reversed microstate $\mathfrak{J}f$ is given by

$$(D(a), \mathfrak{J}f) \equiv (\mathfrak{J}D(a), f) \equiv (D(a^T), f) \equiv P(a^T). \tag{II.10}$$

Thus we shall henceforth refer to $P(a^T)$ as the time-reversed macrostate.

As macroscopic measurements identify phase cells rather than points, the properties of the system pertinent to these measurements are naturally expressed in terms of a coarse-graining operator (on $\mathfrak{g}$) whose application to g serves to replace the value of that function at a phase point by the average of the function over the cell containing the point. Thus the coarse-graining operator $\wp$ is defined by

$$\wp g(x,p) = \sum_a D(a;x,p) \int_{C(a)} g(x,p) \Big/ W(a) \tag{II.11}$$

where W(a) is the volume of C(a). This definition is equivalent to

$$\wp g = \sum_a D(a)\ (D(a), g) \Big/ W(a). \tag{II.12}$$

It is readily seen that $\wp$ is a projection operator, i.e.,

$$\wp = \wp^* = \wp^2. \tag{II.13}$$

It follows now from Eqs. (II.9) and (II.12) that the application of $\wp$ to f yields

$$\wp f = \sum_a D(a)\ P(a) \Big/ W(a). \tag{II.14}$$

Since, by the definition of D(a), the inner product (D(a), D(a')) is equal to $W(a)\delta_{aa'}$, it follows from Eq. (II.14) that

$$(D(a), \wp f) = P(a). \tag{II.15}$$

Consequently, by Eqs. (II.14), (II.15), P is in one-to-one

correspondence with $\mathcal{P}f$; and therefore the macrostate can be represented equivalently by P or $\mathcal{P}f$. Further, it follows from Eqs. (II.8), (II.10) and (II.14) that the application of $\mathcal{P}$ to the time-reversed microstate yields

$$\mathcal{P}\mathcal{J}f \equiv \mathcal{J}\mathcal{P}f \equiv \sum_a \frac{D(a)P(a^T)}{W(a)} \equiv \sum_a \frac{D(a^T)P(a)}{W(a)} \tag{II.16}$$

and hence

$$(D(a), \mathcal{P}\mathcal{J}f) \equiv (D(a), \mathcal{J}\mathcal{P}f) \equiv P(a^T). \tag{II.17}$$

Consequently the time-reversed macrostate $P(a^T)$ is in one-to-one correspondence with $\mathcal{J}\mathcal{P}f \equiv \mathcal{P}\mathcal{J}f$; and therefore the macrostate may be equivalently represented by $P(a^T)$ or $\mathcal{J}\mathcal{P}f = \mathcal{P}\mathcal{J}f$.

Quantal Case.

In quantum mechanics, the pure states of a system are represented by vectors, ψ, of a Hilbert space $\mathfrak{H}$. These vectors correspond to wave-functions, $\psi(x)$, of the configuration coordinates x and, if necessary, the spins, of the particles of the system.

The mixed states are represented by operators (density matrices) on $\mathfrak{H}$. These operators are Hermitian, non-negative and possess the property that

$$\mathrm{Tr}\, f^2 \leq \mathrm{Tr}\, f = 1. \tag{II.18}$$

Consequently, the states f may be represented as vectors in a second Hilbert space, $\mathfrak{L}$, whose elements are the operators g, which act on $\mathfrak{H}$ and for which $\mathrm{Tr}_{\mathfrak{H}}(g^*g)$ is finite. Inner products in $\mathfrak{L}$ are defined by

$$(g_1, g_2) = \mathrm{Tr}_{\mathfrak{H}}(g_1^* g_2). \tag{II.19}$$

We shall refer to $\mathfrak{L}$ as the Liouville space, as in the classical case. The formal equivalence between our descriptions of microstates f, for the classical and quantal cases, is now self-evident.

The time-reversed states may be formulated according to the Wigner[8] prescription. Thus as the elements of $\mathfrak{L}$ correspond to operators on $\mathfrak{H}$, the state f transforms under time-reversal to $\mathcal{J}f$, where

$$\mathcal{J}g = T g T^{-1} \tag{II.20}$$

and $T(=T^{-1})$ is the time-reversal operator (on $\mathfrak{H}$), defined by

$$T\psi(x) = \psi^*(x). \tag{II.21}$$

It is a simple matter to include spin in this prescription, if necessary.

The macrostates are defined analogously to the classical case. Thus, we start by introducing a set of macroscopic observables $\{A\}$, operators on $\mathfrak{H}$. We then idealise a macroscopic measurement as one that specifies the value of $\{A\}$ to limited accuracy, but which determines the value of an associated set $\{\tilde{A}\}$ both precisely and simultaneously. It follows from the basic principles of quantum theory that $\{\tilde{A}\}$ must form an intercommuting set of operators on $\mathfrak{H}$, even though the primitive set $\{A\}$ might not do so. As in the classical case, $\{\tilde{A}\}$ is so defined that a measurement of this set specifies $\{A\}$ to within $\Delta = (\Delta_1, \ldots, \Delta_m)$. This can be arranged (cf. von Neumann,[4] van Kampen[5]) provided that the quantities Δ exceed the Heisenberg uncertainties involved in simultaneous specifications of $\{A\}$.

Thus, as in the classical case, we represent the macroscopic observables $\{\tilde{A}\}$ as possessing simultaneous eigenvalues

$$\{a\} = (a_1, \ldots, a_m),$$

where the values taken by a_j, say, are spaced at intervals Δ_j. The subspace of $\mathfrak{H}$, spanned by the eigenvectors of $\{A\}$ with eigenvalue $\{a\}$, will be denoted by $C(a)$. This subspace corresponds to a phase cell of classical statistics, and will sometimes be called by the same name. Denoting the projection operator for $C(a)$ by $D(a)$, it follows that

$$\tilde{A} = \sum_a aD(a), \tag{II.22}$$

in precise analogy to the classical case.

The probability that the set $\{\tilde{A}\}$ takes the value $\{a\}$ is given by

$$P(a) = \mathrm{Tr}_{\mathfrak{H}}(D(a)f),$$

i.e.,

$$P(a) = (D(a), f) \tag{II.23}$$

as in the classical case. We again designate the macrostate by this function $P(a)$—also we refer to the space, M, of the variables $\{a\}$ as the macrospace.

Further properties of the quantal description follow by analogy with the classical case. Thus, if each of the macroscopic variables

A_j is either odd or even with respect to time-reversals. Then, by analogy with (II.10), we find that the time-reversed macrostate is given by

$$P(a^T) = (D(a), \mathcal{T}f), \tag{II.24}$$

provided that the phase cells are suitably chosen. Also, by analogy with Eq. (II.12), we define the coarse-graining operator $\mathcal{P}$ by

$$\mathcal{P}g = \sum_a D(a)\,(D(a), g)\Big/W(a), \tag{II.25}$$

with

$$W(a) = \mathrm{Tr}_{\mathfrak{D}}\,(D(a))\,. \tag{II.26}$$

Hence, again we have

$$\mathcal{P}f = \sum D(a)\,P(a)\Big/W(a) \tag{II.27}$$

and

$$(D(a), \mathcal{P}f) = P(a). \tag{II.28}$$

We also have again that the time-reversed macrostate can be represented by $P(a^T)$ or, equivalently, by

$$\mathcal{P}\mathcal{T}f = \mathcal{T}\mathcal{P}f. \tag{II.29}$$

III. The Irreversibility Problem and Initial States

We shall formally demonstrate that the principle of microscopic reversibility forbids the macrostate to evolve according to a closed law independently of the initial microstate.

We first note that the time-dependent microstate, f_t, evolves according to the Liouville equation

$$\frac{df_t}{dt} + i\mathfrak{L} f_t = 0 \tag{III.1}$$

where

$$\mathfrak{L} = \begin{matrix} i\{H, \}_{PB} & \text{(classical)} \\ [H, \]_- & \text{(quantal)} \end{matrix} \Big\} \tag{III.2}$$

Thus the evolution of the microstate from f_0 at $t=0$ to f_t at time t is given by

$$f_t = U_t f_o \tag{III.3}$$

where $U_t = e^{-i\mathcal{L}t}$ is the unitary operator† generated by $\mathcal{L}$.

The essential content of the principle of microscopic reversibility can be stated as follows: If the microstate evolves from f_o to f_t in time t, then a state which is initially $f_o' = \mathcal{I} f_t$ will evolve to $f_t' = \mathcal{I} f_o$ in the same time. Thus, corresponding to Eq. (III.3), we have

$$\mathcal{I} f_o = U_t \mathcal{I} f_t. \tag{III.4}$$

This equation can readily be derived from Eq. (III.3) for the microscopic reversibility does apply.

We shall now show that the time-dependent macrostate, which will be represented by $\mathcal{P} f_t$, cannot evolve according to a closed, non-reversible law

$$\mathcal{P} f_t = \tilde{U}_t \mathcal{P} f_o \tag{III.5}$$

where the operator $\tilde{U}_t$ is independent of f_o. For if we assume the validity of such a law, for the moment, we see from Eqs. (III.3) and (III.5) that

$$\mathcal{P} U_t f_o = \tilde{U}_t \mathcal{P} f_o \tag{III.6}$$

for all f_o. Hence

$$\mathcal{P} U_t f \equiv \tilde{U}_t \mathcal{P} f \tag{III.7}$$

and thus, replacing f by $\mathcal{I} f_t$,

$$\mathcal{P} U_t \mathcal{I} f_t = \tilde{U}_t \mathcal{P} \mathcal{I} f_t. \tag{III.8}$$

---

† Strictly speaking, although U_t can be formulated as a bona fide operator on $\mathfrak{L}$, its generator $\mathcal{L}$ does not act on the whole Liouville space—at least it does not do so in the classical case. This can be seen from the fact that the application of $\mathcal{L}$ to certain elements of $\mathfrak{L}$, namely, square-integrable but discontinuous functions, leads to divergences. This is irrelevant to the theory of the present section, as this is formulated in terms of U_t. As will be shown in Section IV, $\mathcal{L}$ can be formulated as an operator on $\mathfrak{L}$ in the quantal case, provided that the system is insulated so that its energy is continued to a restricted range.

It follows from this equation and Eq. (III.4) that

$$\mathcal{P}\mathcal{J}f_o = \tilde{U}_t \mathcal{P}\mathcal{J}f_t, \tag{III.9}$$

i.e., by Eq. (II.29),

$$\mathcal{J}\mathcal{P}f_o = \tilde{U}_t \mathcal{J}\mathcal{P}f_o \tag{III.10}$$

which means simply that the macroscopic law given by Eq. (III.5) must be time-reversible.

Hence a statistical-mechanical theory of irreversibility must limit itself to a restricted class of initial microstates f_o. It must have, then, some prescription for formulating f_o. The prescription could be designed to correspond, for example, to situations where the system is prepared in a specified manner; and could thus purport to give a 'realistic' formulation of the initial microstate. One thing which emerges from our above analysis is that, if f_o is a 'realistic' initial state that leads to an irreversible evolution of the macrostate, then $\mathcal{J}f_{t_o}$ ($t_o > 0$) is not.

IV. Generalised Master Equation

We shall employ the projective method of Zwanzig[1] and Emch[2] to derive a so-called generalised master equation for the evolution of the macrostate, P_t, of a closed system that is initially prepared by measurement of $\tilde{A}$. This method constitutes an exact treatment of the microscopic equations of motion for the system, together with a statistical assumption concerning the initial state. It is important to note that the method is free from any further statistical assumptions (e.g., Boltzmann's stosszahlansatz) concerning the actual dynamics of the system—this is a distinct merit of the method since assumptions of the latter type can conflict with the microscopic dynamical laws.

Our formulation of the theory will henceforth be restricted, for the sake of definiteness, to quantal systems. It will be assumed that the system under consideration is thermally insulated, so that its states are confined to an energy shell $(E, E+\Delta E)$. The Hamiltonian governing the dynamics of the system can them be represented as a bounded operator according to the following prescription: —We first denote the full Hamiltonian for the system by H^o, and the Hilbert space of its eigenstates by $\mathfrak{H}^o$. We then define $\mathfrak{H}$ as the subspace of $\mathfrak{H}^o$ corresponding to the energy shell $(E, E+\Delta E)$. Thus, if Π is the projection operator from $\mathfrak{H}^o$ to $\mathfrak{H}$, then the dynamics of the system, when confined to this shell, will be governed by the truncated Hamiltonian $\Pi H^o \Pi = H$. This is a bounded operator on $\mathfrak{H}$, with eigenvalues in the range $(E, E+\Delta E)$. The boundedness of H is

important, since it ensures that the corresponding Liouville operator, $\mathfrak{L}$, is also bounded. This enables us to carry out our operations involving $\mathfrak{L}$ with full mathematical justification.

We note also that $\tilde{A}$ and, for that matter, other observables of the system, may likewise be represented by operators on $\mathfrak{H}$.

The system is assumed to be initially prepared by measurement of $\{\tilde{A}\}$. In order to formulate the initial microstate, f_o, we assume that the macromeasurement does not discriminate between the (pure) state vectors within a cell. Thus we assume equal probabilities and random phases for the vectors within each cell. Hence we obtain

$$f_o = \sum_a D(a) P_o(a) \Big/ W(a). \tag{IV.1}$$

This formulation of the initial state constitutes the basic statistical assumption of the theory. It should be emphasised that we are not claiming this to be the only possible form for f_o, since other initial preparations are quite conceivable—e.g., where some additional observables are also measured.

The microstate evolves according to the quantum-mechanical Liouville equation

$$\frac{df_t}{dt} + i\mathfrak{L} f_t = 0 \tag{IV.2}$$

where

$$\mathfrak{L} = [H, \ ]_- \tag{IV.3}$$

in units where $\hbar = 1$.

As mentioned above, the boundedness of H ensures that $\mathfrak{L}$ is a bounded operator on $\mathfrak{H}$, as is necessary for the operations that will be considered.

We now use the projective method to obtain the master equation for the macrostate. Thus we apply the operators $\mathcal{P}$, $(I-\mathcal{P})$ to the Liouville equation and thereby obtain a pair of coupled equations for $\mathcal{P} f_t$, the part of f_t representing the macrostate, and $(I-\mathcal{P})f_t$, the complementary part. Thus

$$\left(\frac{d}{dt} + i\mathcal{P}\mathfrak{L}\mathcal{P}\right)\mathcal{P} f_t + i\mathcal{P}\mathfrak{L}(I-\mathcal{P})f_t = 0 \tag{IV.4}$$

and

$$\left(\frac{d}{dt} + i(I-\mathcal{P})\mathfrak{L}(I-\mathcal{P})\right)(I-\mathcal{P})f_t + i(I-\mathcal{P})\mathfrak{L}(\mathcal{P} f_t) = 0 \tag{IV.5}$$

where we have used the projective property $\mathcal{P}^2 = \mathcal{P}$. Equation (IV.5) may be formally integrated to yield

$$(I-\mathcal{P})f_t = \mathcal{U}(t)(I-\mathcal{P})f_o - i\int_o^t dt'\ \mathcal{U}(t-t')(I-\mathcal{P})\mathcal{L}\mathcal{P}f_{t'} \qquad \text{(IV.6)}$$

where

$$\mathcal{U}(t) = \exp\left\{-i(I-\mathcal{P})\mathcal{L}(I-\mathcal{P})t\right\}. \qquad \text{(IV.7)}$$

We now insert the formula (IV.6) for $(I-\mathcal{P})f_t$ into Eq. (IV.4) and thereby obtain the following equation for $\mathcal{P}f_t$, the 'relevant' part of f_t:

$$\left(\frac{d}{dt} + i\mathcal{P}\mathcal{L}\mathcal{P}\right)\mathcal{P}f_t + \int_o^t dt'\,\mathcal{P}\mathcal{L}\,\mathcal{U}(t-t')(I-\mathcal{P})\mathcal{L}(\mathcal{P}f_{t'}) = -i\mathcal{P}\mathcal{L}\,\mathcal{U}(t)(I-\mathcal{P})f_o. \qquad \text{(IV.8)}$$

Moreover, it follows from our formulae (II.27), (IV.3) and (IV.1) for $\mathcal{P}$, $\mathcal{L}$ and f_o that both $\mathcal{P}\mathcal{L}\mathcal{P}$ and $(I-\mathcal{P})f_o$ vanish. Hence, Eq. (IV.8) reduces to

$$\frac{d}{dt}\mathcal{P}f_t + \int_o^t dt'\,\mathcal{P}\mathcal{L}\,\mathcal{U}(t-t')(I-\mathcal{P})\mathcal{L}\mathcal{P}f_{t'} = 0. \qquad \text{(IV.9)}$$

On expressing $\mathcal{P}f_t$ in terms of P_t, by means of Eq. (II.27), this last equation takes the form

$$\sum_a \frac{D(a)}{W(a)}\left\{\frac{d}{dt}P_t(a) - \sum_{a'}\int_o^t dt'\,G(a,a'|\,t-t')P_{t'}(a')\right\} = 0 \qquad \text{(IV.10)}$$

with

$$G(a,a'|\,t) = -\Big(D(a),\mathcal{L}\,\mathcal{U}(t)(I-\mathcal{P})\mathcal{L}D(a')\Big)\Big/W(a'). \qquad \text{(IV.11)}$$

The master equation for P_t is now obtained by taking the inner product of Eq. (IV.10) with $D(a)$:

$$\frac{d}{dt}P_t(a) = \sum_{a'}\int_o^t dt'\,G(a,a'|\,t-t')P_{t'}(a'); \qquad \text{(IV.12)}$$

or, as we shall sometimes write it,

$$\frac{dP_t}{dt} = \int_0^t dt'\, G(t-t')P_{t'} \equiv \int_0^t dt'\, G(t')P_{t-t'} \qquad \text{(IV.13)}$$

where G is the matrix $[G(a,a')]$, and corresponds to an operator acting on functions on M.

We may express the master equation (IV.12) in a slightly more convenient form by noting that, as the set of cells C(a) comprises the space $\mathfrak{D}$,

$$\sum_a D(a) = I;$$

and therefore, by Eq. (IV.11),

$$\begin{aligned}\sum_a G(a,a'|t) &= \Big(I,\ \mathfrak{L}\,U(t)(I-\mathcal{P})\mathfrak{L}D(a')\Big)\Big/W(a') \\ &= \Big(U(t)(I-\mathcal{P})\mathfrak{L}D(a'),\ \mathfrak{L}I\Big)^*\Big/W(a') \\ &= 0,\end{aligned}$$

since $\mathfrak{L}I = 0$. Hence the master equation (IV.12) may be written in the gain-loss form

$$\frac{dP_t(a)}{dt} = \sum_{a'} \int_0^t dt'\Big[G(a,a'|t-t')P_{t'}(a') - G(a',a|t-t')P_{t'}(a)\Big]. \qquad \text{(IV.14)}$$

Finally, we note that, by Eq. (IV.11), the properties of the kernel G depend not only on the microstructure of the system, as represented by the Liouville operator $\mathfrak{L}$, but also on the macroscopic observables and the construction of the phase cells, as represented by D(a) and $\mathcal{P}$. The central problem now is to delineate those properties of the macroscopic observables which can result in a kernel G that generates a Markoffian, causal law. Clearly, such properties have to be considered in relation to the microstructure of the system, as represented by H or $\mathfrak{L}$.

V. The Interaction Representation

In order to express the properties of G in terms of the microscopic properties of the system, we start by splitting H into two parts, H_o and V, where the matrix elements of H_o are all intracellular and those of V are intercellular. Thus

$$H = H_o + V \qquad \text{(V.1)}$$

with

$$H_o = \sum_a D(a)\, H\, D(a) \tag{V.2}$$

and

$$V = \sum_{a \neq a'} V(a,a') \tag{V.3}$$

where

$$V(a,a') = D(a)\, H\, D(a'). \tag{V.4}$$

As V is the part of H governing intercellular transitions, it represents the interaction that engenders changes in the macrostate.

Corresponding to Eq. (V.1), we split the Liouville operator into two parts

$$\mathfrak{L} = \mathfrak{L}_o + \mathfrak{S} \tag{V.5}$$

where

$$\mathfrak{L}_o = [H_o, \ ] \tag{V.6}$$

and

$$\mathfrak{S} = [V, \ ] \tag{V.7}$$

It follows from Eqs. (II.22), (II.27), (IV.3), (V.2) and (V.6) that $\mathfrak{L}_o$ possesses the properties

$$\mathfrak{L}_o D(a) = 0; \quad \mathfrak{L}_o \tilde{A} = 0 \tag{V.8}$$

$$(D(a), \mathfrak{L}_o g) \equiv (g, \mathfrak{L}_o D(a))^* = 0 \tag{V.9}$$

$$\mathcal{P}\mathfrak{L}_o = \mathfrak{L}_o \mathcal{P} = 0. \tag{V.10}$$

Consequently, by Eqs. (IV.7) and (V.10) the propagator $\mathfrak{U}(t)$, that occurs in the kernel of the master equation, may be written in the form

$$\mathfrak{U}(t) = \exp -i(\mathfrak{L}_o + \mathfrak{S}_1)t \tag{V.11}$$

with

$$\mathfrak{S}_1 = (I-\mathcal{P})\mathfrak{S}(I-\mathcal{P}). \tag{V.12}$$

Using a well-known formula,[9] we may expand this equation for $U(t)$ in the form

$$U(t) = U_0(t)\Big(1 + \sum_{n=1}^{\infty} (-i)^n \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{n-1}} dt_n \mathcal{S}_1(t_1) \cdots \mathcal{S}_1(t_n)\Big) \tag{V.13}$$

where

$$U_0(t) = e^{-i\mathcal{L}_0 t} \tag{V.14}$$

and

$$\mathcal{S}_1(t) = U_0(-t)\,\mathcal{S}_1\,U_0(t)\,.$$

i.e., by Eqs. (V.10),

$$\mathcal{S}_1(t) = (I-\mathcal{P})\,\mathcal{S}(t)\,(I-\mathcal{P}) \tag{V.15}$$

with

$$\mathcal{S}(t) = U_0(-t)\,\mathcal{S}\,U_0(t). \tag{V.16}$$

Hence, by Eqs. (V.13) and (V.15),

$$U(t) = U_0(t)\Big[1 + \sum_{n=1}^{\infty} (-i)^n \int_0^t dt_1 \cdots \int_0^{t_{n-1}} dt_n (I-\mathcal{P})\,\mathcal{S}(t_1)(I-\mathcal{P})\,\mathcal{S}(t_2)\cdots$$

$$(I-\mathcal{P})\,\mathcal{S}(t_n)(I-\mathcal{P})\Big]. \tag{V.17}$$

On substituting this formula into Eq. (IV.17) and using Eq. (V.9), we obtain the following formula for the kernel G:

$$G(a,a'|t) = G_0(a,a'|t) + \sum_{n=1}^{\infty} (-i)^n \int_0^t dt_1 \cdots \int_0^{t_{n-1}} dt_n G_n(t,t_1,\ldots,t_n) \tag{V.18}$$

where

$$G_0(a,a'|t) = -\Big(D(a),\,\mathcal{S}\,U_0(t)(I-\mathcal{P})\,\mathcal{S}D(a')\Big)\Big/W(a') \tag{V.19}$$

and

$$G_n(a,a'|t,t_1,\ldots,t_n) =$$

$$-\Big(D(a),\mathcal{S}\,\mathcal{U}_0(t)(I-\mathcal{P})\,\mathcal{S}(t_1)(I-\mathcal{P})\,\mathcal{S}(t_2)\ldots(I-\mathcal{P})\,\mathcal{S}(t_n)(I-\mathcal{P})\,\mathcal{S}D(a')\Big)\Big/W(a'). \tag{V.20}$$

We aim to express these components of G in terms of correlation functions of the type used extensively in many-body theory. For this purpose we shall now make a slight digression and introduce the concept of a reference system, S_0, whose Hamiltonian is H_0. Our reason for doing this is that we will be able to express the components G_n as time-correlations for fluctuations of Heisenberg operators of S_0 about equilibrium states of that system. This will be useful as these are precisely the kind of correlations studied in many-body theory; and so we hope to be able to draw on some results of general character that have been obtained from studies of such correlations.

We shall now express some relevant properties of Heisenberg operators for S_0, and of correlation functions for that system, in terms of the Liouville space formalism. Thus, by Eqs. (V.6) and (V.14), we may express Heisenberg operators for S_0 in the two equivalent forms

$$g(t) \equiv \mathcal{U}_0(-t)g \equiv e^{iH_0t}\,g\,e^{-iH_0t}. \tag{V.21}$$

Hence, by Eqs. (V.7), (V.14), (V.16) and (V.21), the interaction Liouville operator $\mathcal{S}(t)$ may be expressed in terms of the Heisenberg operator $V(t)$ (for S_0) by the relation

$$\mathcal{S}(t) = [V(t), \ ]. \tag{V.22}$$

It follows from Eqs. (V.14), (V.21) and (V.22) that

$$\big(g_1(t), g_2\big) \equiv \big(g_1, \mathcal{U}_0(t)g_2\big) \tag{V.23}$$

and

$$\big(g_1, \mathcal{S}(t)g_2\big) \equiv \big(\mathcal{S}(t)g_1, g_2\big). \tag{V.24}$$

In order to formulate correlations for equilibrium states in S_0 we note that, in view of Eqs. (V.8), the observables $\tilde{A}$ are constants of the motion and the states $D(a)/W(a)$ are equilibrium states for S_0, the latter states corresponding to microcanonical distributions for which $\tilde{A}$ takes the values a. Averages over these distributions will be denoted by

$$\overline{g}(a) \equiv \langle g \rangle_a \equiv (D(a), g)/W(a). \tag{V.25}$$

Likewise, correlation functions for these microcanonical states will be denoted by

$$\Phi_a(g_1; g_2) = \left\langle \left(g_1^* - \overline{g}_1^*(a)\right)\left(g_2 - \overline{g}_2(a)\right) \right\rangle_a . \tag{V.26}$$

It follows from Eqs. (II.25), (V.25) and (V.26) that these correlation functions are related to matrix elements of $(I-\mathcal{P})$, between elements of $\mathcal{Q}$, by the equation

$$\left(g_1, (I-\mathcal{P})g_2\right) \equiv \left((I-\mathcal{P})g_1, g_2\right) \equiv \sum_a W_a \Phi_a\left(g_1; g_2\right) \equiv \Phi\left(g_1; g_2\right), \text{ say.} \tag{V.27}$$

Thus the inner product Φ is determined by the correlations between fluctuations of g_1^*, g_2 about their mean values for the various equilibrium states, $D(a)/W(a)$, of S_o.

Returning now to the properties of the kernel, G, we see from Eqs. (V.24) and (V.27) that its components G_n may be expressed in terms of correlation functions for S_o as follows:

$$G_o(a, a' \mid t) = -\Phi(F(a,t); F(a'))/W(a') \tag{V.28}$$

and

$$G_n(a, a' \mid t, t_1, \cdots, t_n)$$

$$= -\Phi\left(F^{(r)}(a \mid t_r, t_{r-1}, \ldots, t_1, t); F^{(n-r)}(a' \mid t_{r+1}, \ldots, t_n, 0)\right) \Big/ W(a'), \tag{V.29}$$

where r may take any value from 0 to n,

$$F(a) = \mathcal{S}D(a) \equiv \sum_{a'} (V(a', a) - V(a, a')), \tag{V.30}$$

$$F(a,t) \equiv F^o(a,t) = \mathcal{U}_o(-t)F(a) \equiv \mathcal{S}(t)D(a) \tag{V.31}$$

is the corresponding Heisenberg operator, and

$$F^{(j)}\left(a \mid t_1', \ldots, t_{j+1}'\right) \equiv \mathcal{S}(t_1')(I-\mathcal{P})\mathcal{S}(t_2') \cdots (I-\mathcal{P})\mathcal{S}(t_j')(I-\mathcal{P})F(a, t_{j+1}') \tag{V.32}$$

for any set $(t_1', \ldots, t_j')$.

It is evident from Eqs. (V.30) and (V.32) that F(a) corresponds to a 'force' leading to transitions between C(a) and other cells and that $F^{(j)}$ corresponds to a 'higher order force.' Thus, Eqs. (V.28) and (V.29) serve to express the components of G in terms of time-correlations between such forces. Other equivalent expressions for these components are

$$G_o(a,a'|t) = -\langle F^{(1)}(a|o,t)\rangle^*_{a'}, \tag{V.33}$$

and

$$G_n(a,a'|t,t_1,\ldots,t_n) = -\langle F^{(n+1)}(a|0,t_n,\ldots,t_1,t)\rangle^*_{a'}.$$

It will be useful for us to express G_n in terms of correlations between simple products of Heisenberg operators, since many-body theory is generally formulated in terms of such correlations. Thus we note that, by Eqs. (II.25) and (V.4), $\mathcal{P}V(a,a'|t) = 0$ and thus that G_o may be expressed in the equivalent forms

$$G_o(a,a'|t) = 2\,\mathrm{Re}\langle V(a',a|t)V(a,a')\rangle_{a'} = 2\,\mathrm{Re}\,\Phi_{a'}\big(V(a,a'|t);V(a',a)\big). \tag{V.34}$$

In considering the form of G_n, for $n>1$, we note that, by Eq. (V.32), the 'force' $F^{(j)}(a|\tilde{t}_1,t_2,\ldots,t_{j+1})$ is a sum of contributions

$$\Lambda_1\mathcal{P}\Lambda_2\mathcal{P}\ldots\Lambda_k\mathcal{P}\Lambda_{k+1}F(a,\tilde{t}_{j+1}), \tag{V.35}$$

each Λ being a product of operators $\mathcal{S}(t)$. It follows from Eq. (II.25) that this expression (V.35) is a sum of terms

$$\mathcal{S}(\tilde{t}'_1)\ldots\mathcal{S}(\tilde{t}'_k)F(a'',\tilde{t}_{k+1}) \times \text{(c-number)}$$

where $\tilde{t}'_1,\ldots,\tilde{t}'_k$ belong to $(\tilde{t}_1,\ldots,\tilde{t}_{j+1})$. Hence, by Eqs. (V.3), (V.22) and (V.30), $F^{(j)}$ is a sum of terms, each being a product of a c-number and an operator $V_{K_1}(t'_1)\ldots V_{K_\ell}(t'_\ell)$; where each K_r denotes a pair of a's, say (a^r,a'^r), and $V_{K_r}(t)$ is the Heisenberg operator $V(a^r,a'^r|t)$. Thus, by Eq. (V.29), G_n is a sum of terms containing, as factors, correlation functions

$$\Phi\big(V_{K_1}(t'_1)\cdots V_{K_\ell}(t'_\ell);\, V_{K'_1}(t''_1)\cdots V_{K'_m}(t''_m)\big) \tag{V.36}$$

where the times t', t" belong to the ranges (t_r,t) and (o,t_{r+1}), respectively.

VI. Characterisation of Macroscopic Observables

Up to this point our formalism contains no reference to the fact that the observables $\tilde{A}$ are macroscopic—only in verbal statements have we asserted that they are so. We now aim to formulate the characteristic properties of these observables, and also of the interaction V, that stem specifically from the many-particle structure of the system and the macroscopicality of $\tilde{A}$. This will entail a number of assumptions, of general character, that are designed to correspond to properties of systems composed of many-particles. The assumptions are formulated so as to be self-consistent; and some of them are readily justified for particular models. However, it is clear that a deeper justification for the full set of assumptions, other than by empirical means, would have to rest on the study of models that are both fully tractable and non-trivial.

We start by assuming that the observables $\tilde{A}$ correspond to extensive variables. Thus, each of these variables may be expressed as a product of an intensive variable and a dimensionless size parameter: For example, if the extensive variable were the energy of a subvolume, v, of crystal, then the size parameter could be chosen as the number of atoms in v, and the intensive variable would then be the energy per atom of that subvolume. It will be assumed, for simplicity, that the size parameters for the observables $\tilde{A}$ can be suitably chosen to have the same value, Γ.

The introduction of this parameter, which is always enormously large by comparison with unity, will serve the crucial purpose of providing a natural ratio between the scales appropriate to macroscopic and microscopic descriptions of the system. It will be our aim, from now on, to classify the quantities relevant to the theory in terms of their Γ-dependence. In this way, we aim to obtain conditions under which the macroscopic variables follow a Markoffian, causal law, to lowest order in Γ^{-1}.

Our next step is to introduce a set of parameters $\alpha = (\alpha_1,\ldots,\alpha_m)$ whose magnitudes are of the order of characteristic (eigen) values of the intensive variables $\tilde{A}/\Gamma$: for example, if $\tilde{A}_j$ were the energy of a subvolume, v, of insulating crystal, then α_j could be chosen to be the Debye energy quantum. In general, the parameters α and $\Gamma\alpha$ provide 'natural' units for the scales respectively appropriate to the intensive variables $\tilde{A}/\Gamma$ and the extensive variables $\tilde{A}$.

Next, we introduce a dimensionless 'coarseness' parameter, Λ, defined as the ratio of the spacing, Δ_j, between eigenvalues of $\tilde{A}_j$, to the microscopic parameter α_j. For simplicity, it will be assumed that Λ takes the same value for all the observables $\tilde{A}$. Consequently, the eigenvalues of $\tilde{A}$ may be expressed as

$$a = (\Lambda n_1\alpha_1, \ldots, \Lambda n_m\alpha_m) \qquad \text{(VI.1)}$$

where the n_j's are integers, which we shall sometimes term the cellular quantum numbers. The parameter Λ is chosen so that its magnitude lies in an intermediate region between Γ and 1, i.e.,

$$\Gamma \gg \Lambda \gg 1. \tag{VI.2}$$

We shall later express Λ as a suitably chosen positive power of Γ.

We now have three different scales for the specification of $\tilde{A}$. These are the macroscopic, cellular and microscopic scales for which the respective units are $\Gamma\alpha$, $\Lambda\alpha$ and α.

It follows from Eq. (VI.1) that the eigenvalues of $\tilde{A}$, as specified in ratio to the macroscopic units $\Gamma\alpha$, are given by the intensive variables $x = (x_1, \ldots, x_m)$ where

$$x_j = \frac{a_j}{\Gamma\alpha_j} = \frac{n_j}{\Omega} \tag{VI.3}$$

and

$$\Omega = \Gamma/\Lambda. \tag{VI.4}$$

It is evident from Eqs. (VI.2)-(VI.4) that the values of x_j are spaced at extremely small intervals, Ω^{-1}, and therefore the intensive variables x are almost continuous.

The form of the interaction V, as defined by Eqs. (V.3) and (V.4), depends on the construction of the phase cells. Its strength, corresponding to the state $D(a)/W(a)$, may be represented by the dimensionless quantity

$$\lambda_a = \omega^{-1}\langle V^2\rangle_a^{\frac{1}{2}}, \tag{VI.5}$$

where ω is a characteristic energy for a microscopic quantum, e.g., a Debye quantum in the case of an insulating crystal. We now make an assumption which is absolutely essential if the system is to follow smooth macroscopic laws. This assumption is that the relative change, $\delta\lambda_a/\lambda_a$, in λ_a, due to a small charge, δx, in x, is of the order of δx, up to a factor dependent only on intensive variables. The significance of this assumption may be seen from the fact that its alternative would be that $\delta\lambda_a/\lambda_a$ were of the order $\Gamma\delta x$, say, or $\Lambda\delta_x$. This would mean that an infinitesimal change in the intensive variables for the macrostate ($|\delta x_j| \ll |$) could lead to a change by an enormous factor ($\delta\lambda_a/\lambda_a \gg 1$) in the strength of the interaction leading to macroscopic changes, which would clearly preclude the possibility of smooth† macroscopic laws.

† For example, if one of the macroscopic variables were the energy of a subvolume v of crystal, this alternative would mean that an

It follows from our assumption concerning $\delta\lambda_a/\lambda_a$ that

$$\lambda_a = \lambda\varphi(x) \qquad \text{(VI.6)}$$

where φ depends on the intensive variables x only, and λ is a constant, representing a strength parameter for V. In general, λ will depend on both Γ and Λ. We shall formulate its dependence on these quantities with the aid of further assumptions that will be made below.

Our next assumption is that the phase cells are so constructed that V possesses a typical property of interactions in many-particle systems, namely, that it contains only matrix elements between state vectors that differ from one another by only a few $(=0(\Gamma^0))$ particles, or quanta. This is another way of saying that V leads to macroscopic changes by elementary processes (e.g., collisions), each involving only a few quanta. Formally,‡ it means that V has only matrix elements between pure states $|\,\rangle$ and $Q|\,\rangle$, where Q is a product of a 'few' creation and annihilation operators, C^* and C, for the system—or, more generally, Q could be a sum of such products. Thus, one could have $Q=C_1^*C_2C_3^*C_4$, but not a product of, say, $\Gamma^{\frac{1}{2}}$ operators C^*, C.

It should be appreciated that this assumption concerning V is far from innocuous. For, although the primitive interactions between

---

infinitesimal change in the temperature (or specific energy) of v could lead to an enormous relative change in the rate of energy transfer between v and the rest of the crystal.

‡ This assumption can be easily justified in simple cases. Consider, for example, a system S, composed of parts S_1 and S_2, with Hamiltonians H_1 and H_2. Then the Hamiltonian for S is H_1+H_2+U, where U is the interaction between S_1 and S_2. Let the macroscopic observables be coarse-grained energies for S_1 and S_2. Then the projection operators, $D(a)$, for S_1 are direct products of those for the energy shells of S_1, S_2; i.e., $D(a)=D_1(E_1)\otimes D_2(E_2)$. It follows from our definition of V that

$$V = \sum_{a'\neq a} D(a)\, U\, D(a').$$

For typical couplings, U contains only matrix elements between states differing by a few quanta, e.g., $|\,\rangle$ and $C_{1j}^*C_{1k}C_{2\ell}^*C_{2m}$, where $C_{1j}^*C_{1k}$ are creation and annihilation operators for S_1; and $C_{2\ell}^*$ and C_{2m} are such operators for S_2. It follows from our formula expressing V in terms of U and $D(a)$ that V has also only matrix elements between states differing by a few quanta, the role of $D(a)$ being to provide an energy selection rule.

the particles of a system are generally two-body forces and therefore have matrix elements only between states differing by a few quanta, the same will not be true of V unless the cells are suitably constructed.

In view of the way the cells were constructed in relation to the 'fine-grained' observables A, we shall assume† that state vectors which differ from one another by a few quanta, in the sense specified above, belong either to the same cell or to neighbouring cells for which the values of the respective quantum numbers n_j differ by ~ 1; we take the symbol $\sim$ to signify equality up to a factor, or order Γ^0, dependent on intensive variables only. Thus, as the matrix elements of V are intercellular (by definition) and connect only states differing by few quanta (by assumption), it follows that V has only matrix elements between states in neighbouring cells ($\Delta n_j \sim 1$).

We now aim to obtain the dependence of λ, the strength of V, on the parameters Λ and Γ. Of course, if we were working with a tractable model, we would be in a position to evaluate λ_a directly from explicit formulae for V and D(a). As we are formulating the theory in general terms, however, we have no explicit formulae for V, D(a); and so we shall employ a kinetic argument, based on assumptions which lead to a self-consistent theory, in order to obtain the dependence of λ on Γ and Λ. Thus, as λ is the strength of the interaction governing the dynamics of the macrostate, we first relate λ to Λ and Γ by an elementary kinetic argument based on the assumption (among others) that the macrostate evolves according to a Markoffian, deterministic law. This leads to the conclusion that λ may be rendered extremely small, equal to a negative power of Γ, by constructing the phase cells to be sufficiently large, i.e., by choosing Λ large enough. We then show, in Sections VII and VIII, that for such small λ and a suitable choice of macroscopic variables, the generalised master equation yields a Markoffian, deterministic law. Thus the whole procedure constitutes a self-consistent theory of closed dynamical laws for suitable macro-observables.

In order to formulate our elementary kinetic argument, we introduce parameters representing characteristic times for changes at the microscopic, cellular and macroscopic levels; by a characteristic

† The essential reason why this should be justified is that $\tilde{A}_j$ provides a specification of the fine-grained observable A_j to within $\sim \Lambda\alpha_j$. Now the difference between the mean values of A_j, for states that differ by a few quanta from one another, will be $\sim \alpha_j$. Consequently, the difference between the values of $\tilde{A}_j$ for these states will lie within $\sim\Lambda\alpha_j$ of one another, as $\Lambda \gg 1$. Hence the states must lie either in the same cell or in neighbouring cells, for which $\Delta n_j \sim 1$.

time for a process, we mean one that represents its duration up to a factor dependent only on intensive variables. Thus we choose $\bar{\tau}_\mu$ to represent the duration of a typical microscopic process (e.g., a collision). This time, an intensive quantity, may be suitably chosen as the reciprocal of the energy quantum (= frequency, as $\hbar = 1$) ω, i.e.,

$$\bar{\tau}_\mu = \omega^{-1}. \tag{VI.7}$$

The parameter representing a characteristic time taken for an intercellular transition will be denoted by $\bar{\tau}_c$. For a Markoffian system in which such transitions are engendered by V, this time $\sim \omega/\langle V^2\rangle_a$ (cf. van Hove[10]). Hence, by Eqs. (V.5) and (V.6), we may choose

$$\bar{\tau}_c = (\omega\lambda^2)^{-1}. \tag{VI.8}$$

The parameter representing a characteristic time for a macroscopic change, $\sim \Gamma_\alpha$, in $\tilde{A}$ will be denoted by $\bar{\tau}_M$. In order to formulate an expression for $\bar{\tau}_M$, we invoke van Kampen's[5] precept that macroscopic observables are slowly-varying quantities. We formalise this by assuming that $\bar{\tau}_M$ is greater than the microscopic time $\bar{\tau}_\mu$ by a factor which is very large, specifically because of the macroscopicality of $\tilde{A}$, i.e., because of the largeness of Γ. Accordingly, we shall assume that $\bar{\tau}_M/\bar{\tau}_\mu$ is a function of Γ which takes large values when Γ is large. For simplicity, we shall state this function to be a positive power of Γ, as it may readily be shown† to be in usual cases of phenomenological laws. Thus we assume

$$\bar{\tau}_M = \bar{\tau}_\mu \Gamma^q = \omega^{-1}\Gamma^q; \quad \text{with} \quad q > 0. \tag{VI.9}$$

Our formulae for the time $\bar{\tau}$ enable us to obtain λ as a function of Γ and Λ. For it follows from our definition of $\bar{\tau}_M$ that the mean rate of change of the variables $\tilde{A}$ is $\sim \Gamma a/\bar{\tau}_M$. On the other hand, as V is assumed to cause transitions between neighbouring cells in time $\bar{\tau}_c$, this mean rate of change of $\tilde{A}$ is also $\sim \Lambda a/\bar{\tau}_c$. Hence,

---

† Consider, for example, the escape of gas, through a small hole, from a container of volume V. The mass, M, of gas in V will (presumably) change at a rate depending on an intensive variable, namely the density M/V, and the geometry of the hole. Thus, defining Γ as a characteristic number of molecules in V, $M^{-1}\,dM/dt = \Gamma^{-1}$ x an intensive variable. It follows that $\bar{\tau}_M$ is proportional to Γ.

$$\frac{\Gamma_a}{\bar{\tau}_M} \sim \frac{\Lambda_a}{\bar{\tau}_c} , \tag{VI.10}$$

i.e., by Eqs. (VI.8) and (VI.9),

$$\lambda^2 \sim \Gamma^{1-q}\Lambda^{-1} . \tag{VI.11}$$

The requirements of this relation and the inequalities (VI.2) may be fulfilled by choosing

$$\lambda = \Gamma^{-r}$$

$$\Lambda = \Gamma^{1+2r-q} \tag{VI.12}$$

and

$$\Omega = \Gamma^{q-2r}$$

with

$$2r+1 > q > 2r > 0 .$$

Thus, constructing the cells to be sufficiently large, we can render λ and Ω^{-1} extremely small, equal to negative powers of Γ. It should be noted, however, that further restrictions on the value of the index r may be imposed by the Markofficity conditions (cf. discussion at the end of Section VII).

Finally, we shall use our assumptions to formulate the Γ-dependence of the kernel G. We shall restrict our consideration to times $t \ll \bar{\tau}_p$, the Poincaré recurrence period. In view of the enormity of $\bar{\tau}_p$, this restriction is only a formal device, irrelevant to any experimentally observable process.

We see from Section V that G is a functional of V(t) and its components, $V(a,a'|t)$. The properties of these time-dependent interactions may be conveniently expressed in terms of their spectral functions. Thus we divide the full range of possible energy differences between eigenvalues of H_o into intervals $(\epsilon, \epsilon + \Delta\epsilon)$, chosen so that (cf. Reference 10)

$$\tau_p \gg (\Delta\epsilon)^{-1} \gg \bar{\tau}_M , \ t. \tag{VI.13}$$

The former inequality ensures that each interval contains an enormous number of eigenstates of H_o. We now resolve V into components, V_ϵ, with matrix elements between eigenstates of H_o whose energy differences lie in the respective intervals $(\epsilon, \epsilon + \Delta\epsilon)$. Thus, denoting eigenstates and eigenvalues of H_o by ψ_j, E_j, we have

$$V = \sum_{\epsilon} V_{\epsilon} \tag{VI.14}$$

where

$$\left(\psi_j, V_{\epsilon}\psi_k\right) = \left(\psi_j, V\psi_k\right)\chi_{\epsilon} \tag{VI.15}$$

where $\chi_{\epsilon} = 1$ or 0 according to whether or not $(E_1 - E_2)$ lies in $(\epsilon, \epsilon + \Delta\epsilon)$. The spectral function σ is then defined by

$$\sigma_a(\epsilon)\Delta\epsilon = \langle V_{\epsilon}^* V_{\epsilon}\rangle . \tag{VI.16}$$

By Eqs. (VI.14)-(VI.16), this function satisfies the sum rule

$$\sum_{\epsilon} \sigma_a(\epsilon)\Delta\epsilon = (\lambda\omega\varphi(x))^2 . \tag{VI.17}$$

Thus $(\sigma_a(\epsilon)/(\lambda\omega\varphi(x))^2)$ is a normalised function $\left(\sum_{\epsilon} \ldots \Delta\epsilon = 1\right)$,

representing the strength of the interaction V corresponding to the transition energy ϵ. Accordingly, we may formalise our assumption that V contains matrix elements only between states that differ from one another by a few quanta, i.e., by energies $\sim \omega$, by postulating that the various moments of this relative strength function are given by appropriate powers of ω, apart from Γ-independent numerical factors. Hence

$$\sigma_a(\epsilon) = \lambda^2 \omega\, \xi(f/\omega) \tag{VI.18}$$

where the function $\xi(s)$ is dimensionless and contains no Γ-dependence.

Hence, by Eqs. (V.21), (VI.13) and (VI.15),

$$\left(\psi_r, V_{\epsilon}(t)\psi_s\right) = \left(\psi_r, V_{\epsilon}\psi_s\right)e^{i\epsilon t} .$$

It follows from Eqs. (VI.14)-(VI.16) that the spectral function σ is related to the autocorrelation function for V(t) by the formula

$$\langle V(t)V\rangle_a = \sum_{\epsilon} \sigma_a(t)e^{i\epsilon t}\Delta\epsilon . \tag{VI.19}$$

Assuming σ to be a smooth function of ϵ, we can replace the sum by an integral in this equation. Thus

$$\langle V(t)V\rangle_a = \int d\varepsilon\, \sigma_a(\varepsilon) e^{i\varepsilon t}; \tag{VI.20}$$

and, consequently, by Eq. (VI.18)

$$\langle V(t)V\rangle_a = (\lambda\omega)^2 \eta(\omega t) \tag{VI.21}$$

where the function $\eta(\tau)$ is dimensionless and Γ-independent. It is important to note that the replacement of the sum, in Eq. (VI.19), by an integral, in Eq. (VI.20), opens up the possibility that the resultant function η decays to zero as $t\to\infty$. Clearly this possibility has arisen only through our formal elimination of Poincaré cycles. It will be assumed that these cycles are similarly eliminated from all correlation functions relevant to the theory.

In the formula (VI.21) for the autocorrelation function for $V(t)$, the time t occurs only in ratio to the microscopic time $\bar{\tau}_\mu = \omega^{-1}$. Consequently, it is natural to express $V(t)$ in terms of the time, τ, for a scale whose unit is ω^{-1}, i.e.,

$$\tau = \omega t. \tag{VI.22}$$

Likewise it is natural to express V in ratio to a characteristic value, $\lambda\omega$, of its root mean square. Thus we define a dimensionless reduced interaction $v(\tau)$ by

$$V(t) = \lambda\omega\, v(\tau). \tag{VI.23}$$

Correspondingly we define reduced forms of the interaction component $V(a,a'|t)$, the 'force' $F(a,t)$ and the Liouville interaction operator $\mathcal{S}(t)$ by

$$v(a,a'|\tau) = (\lambda\omega)^{-1} V(a,a'|t) \tag{VI.24}$$

$$f(a,\tau) = (\lambda\omega)^{-1} F(a,t) \tag{VI.25}$$

and

$$s(\tau) = (\lambda\omega)^{-1} \mathcal{S}(t). \tag{VI.26}$$

It follows now from Eqs. (VI.21) and (VI.24) that the autocorrelation function $\langle v(\tau)v\rangle_a$ is a Γ-independent function of τ, as it was obviously designed to be. This means that $v(\tau)$ simulates an intensive variable, at least in the context of its autocorrelation function. We shall assume that both this interaction and its derivative reduced operators $f(a,\tau)$ and $\ell(\tau)$ behave likewise in the context of the

various correlation functions in which they appear; the validity of this assumption will depend on the microstructure of V.

Inserting these assumptions into Eqs. (V.32) and (V.33), we see that the components of the kernel G may be expressed in the forms

$$G_o(a,a'|t) = (\lambda\omega)^2 g_o(a,a'|\tau) \tag{VI.27}$$

and

$$G_n(a,a'|t,t_1,\ldots,t_n) = (\lambda\omega)^{n+2} g_n(a,a'|\tau,\tau_1,\ldots,\tau_n), \tag{VI.28}$$

where the functions g_n are intensive† variables. Hence, by Eq. (V.18), the kernel G may be expressed in terms of the interaction parameter λ and these intensive variables by the equation

$$G(t) = (\lambda\omega)^2 g(\tau;\lambda) \tag{VI.29}$$

where

$$g(\tau;\lambda) = \sum_0^\infty (-i\lambda)^n g_n(\tau) \tag{VI.30}$$

and

$$g_n(\tau) = \int_0^\tau d\tau_1 \cdots \int_0^{\tau_{n-1}} d\tau_n g_n(\tau,\tau_1,\ldots,\tau_n), \quad \text{for } n>0. \tag{VI.31}$$

No confusion should result from our use of the symbol g_n to refer both to a function of τ and to a function of $\tau,\ldots,\tau_n$.

VII. Parametrised Master Equation and Markofficity Conditions

On expressing the kernel G in the parametrised form of Eq. (VI.29), the master equation (IV.12) becomes

$$\frac{dP_t}{dt} = \lambda^2\omega \int_0^{\omega t} g(\tau';\lambda) P_{t-\omega-1_\tau}\, d\tau'. \tag{VII.1}$$

† Here one should not be misled by the dependence of g_n on a,a' since, in this context, these symbols are merely labels for the cells C(a), C(a').

We now express the time-dependence of P on a temporal scale of which the unit is the cellular time $\bar{\tau}_c$; this is clearly a 'natural' scale for the evolution of P. Thus we define the time of this scale as

$$\sigma = t/\bar{\tau}_c , \tag{VII.2}$$

i.e., by Eqs. (VI.7) and (VI.8),

$$\sigma = \omega\lambda^2 t = \lambda^2\tau . \tag{VII.3}$$

Correspondingly we express P_t as a function of σ

$$P_t = p(\sigma). \tag{VII.4}$$

Hence by Eqs. (VII.1), (VII.3) and (VII.4), the master equation now takes the form

$$\frac{dp(\sigma)}{d\sigma} = \int_0^{\sigma/\lambda^2} d\tau'(\tau';\lambda)p(\sigma - \lambda^2\tau'). \tag{VII.5}$$

This is the parametrised master equation in which g and p are expressed in terms of their respective 'natural' time-scales. The fact that the units for these scales differ by an enormous factor, λ^{-2}, is crucial to the possibility of Markoffisation.

The problem concerning the reducibility of the master equation, in an appropriate sense, to Markoffian form is best studied by the method of Laplace transforms. Thus we define the transforms of p, g and g_n as

$$\bar{p}(y) = \int_0^\infty d\sigma\, e^{-y\sigma}\, p(\sigma) \tag{VII.6}$$

$$\bar{g}(z) = \int_0^\infty d\tau\, e^{-z\tau}\, g(\tau) \tag{VII.7}$$

and

$$\bar{g}_n(z) = \int_0^\infty d\tau\, e^{-z\tau}\, g_n(\tau), \tag{VII.8}$$

the transformation variables y, z corresponding to reciprocal times on the σ and τ scales, respectively. It follows from these definitions and Eq. (VI.30) that the transform of the parametrised master equation is given by

$$y\bar{p}(y) - p(o) = \bar{g}(\lambda^2 y;\lambda)\bar{p}(y) \qquad \text{(VII.9)}$$

with

$$\bar{g}(\lambda^2 y;\lambda) = \sum_{o}^{\infty} (-i\lambda)^n \bar{g}_n(\lambda^2 y). \qquad \text{(VII.10)}$$

We shall be concerned with the evolution of the system as 'time-smoothed' over an interval $\sim \lambda$, on the σ-scale. The duration of such an interval lies in an intermediate region between the characteristic microscopic and cellular times, $\bar{\tau}_\mu$ and $\bar{\tau}_c$, whose values on the τ-scale are λ^2 and 1, respectively. The time-smoothing over an interval $\sim \lambda$ may be achieved by treating the transformed master equation (VII.9) on the basis of an approximation which is valid, to lowest order in λ, when $y \leq \lambda^{-1}$. Thus we treat the variable $\lambda^2 y$, which appears in the kernel of the transformed master equation, as being of first order in λ. Hence, as the normalized functions $\bar{g}_n(z)$ contain no explicit Γ-dependence, it follows from Eq. (VII.10) that the kernel $\bar{g}(\lambda^2 y;\lambda)$ reduces, in zero order in λ, to the y-independent quantity $\bar{g}_o(0)$, provided that: (a) the functions $\bar{g}_n(z)$ tend to finite limits as $z \to 0$; and (b) the increment in $\bar{g}_n(z)$, due to a small change ($\leq \lambda$) in z from 0 to δz, is $0(\delta z)$. We shall discuss the microscopic significance of these conditions a little later, following Eq. (VII.17). For the moment we observe that, if the conditions are fulfilled, the kernel $\bar{g}(\lambda^2 y;\lambda)$ of the transformed master equation may be replaced by the y-independent quantity

$$k = \bar{g}_o(0) \equiv \int_o^\infty d\tau\, g_o(\tau). \qquad \text{(VII.11)}$$

In this case the transformed master equation (VII.9) reduces to

$$y\bar{p}(y) - p(0) = k\bar{p}(y) \qquad \text{(VII.12)}$$

which is simply the transform of the Markoffian master equation

$$\frac{dp(\sigma)}{d\sigma} = kp(\sigma), \qquad \text{(VII.13)}$$

i.e.,

$$\frac{dp(a,\sigma)}{d\sigma} = \sum_{a'} k(a,a')\,p(a',\sigma), \qquad \text{(VII.14)}$$

or equivalently (cf. procedure from Eq. (IV.12) to Eq. (IV.14)),

$$\frac{dp(a,\sigma)}{d\sigma} = \sum_{a'}\Big[\, k(a,a')p(a',\sigma) - k(a',a)p(a,\sigma)\,\Big], \qquad \text{(VII.15)}$$

the matrix $k(a',a)$ being defined in correspondence with Eq. (VII.11):

$$k(a',a) = \int_0^\infty d\tau\; g_0(a',a|\tau)\,, \qquad \text{(VII.16)}$$

i.e., by Eqs. (V.34), (VI.24) and (VI.27)

$$k(a',a) = 2\,\mathrm{Re}\int_0^\infty d\tau\,\langle v(a,a'|\tau)v(a',a)\rangle_a d\tau. \qquad \text{(VII.17)}$$

In general, the Markoffian master equation (VII.15) describes an irreversible law. One can even define a macroscopic entropy function which, under this law, will increase monotonically.[11]

We see, then, that (a) and (b) constitute sufficient conditions for Markofficity. In order to relate these conditions to microscopic correlations in the system, we first make a change of variables

$$\left.\begin{aligned} &g_n(\tau,\tau_1,\ldots,\tau_n) = \gamma_n(\tau',\tau_1',\ldots,\tau_n'), \quad \text{for} \quad n>0\\ &\text{with}\\ &\tau' = \tau-\tau_1,\ \tau_1' = \tau_1-\tau_2,\ \ldots,\ \tau_{n-1}' = \tau_{n-1}-\tau_n,\ \tau_n' = \tau_n;\\ &\text{and}\\ &g_0(\tau) = \gamma_0(\tau). \end{aligned}\right\} \qquad \text{(VII.18)}$$

Hence, by Eqs. (VII.8) and (VII.18),

$$\bar{g}_n(z) = \int_0^\infty d\tau\; e^{-z\tau}\int_0^\tau d\tau_1\cdots\int_0^{\tau_{n-1}} d\tau_n\,\gamma_n(\tau-\tau_1,\tau_1-\tau_2,\ldots,\tau_{n-1}-\tau_n,\tau_n). \qquad \text{(VII.19)}$$

This reduces, by the Faltung theorem, to

$$\bar{g}_n(z) = \bar{\gamma}_n(z,z,\ldots,z) \qquad ((n+1)\text{ variables } z) \qquad \text{(VII.20)}$$

where $\bar{\gamma}_n$ is the multiple Laplace transform, defined by

$$\bar{\gamma}_n(z,z_1,\ldots,z_n) = \int_0^\infty d\tau' \int_0^\infty d\tau_1' \cdots \int_0^\infty d\tau_n' \gamma_n(\tau',\tau_1',\ldots,\tau_n') \times \exp -(z\tau' + z_1\tau_1' + \cdots + z_n\tau_n'). \tag{VII.21}$$

It follows from Eq. (VII.20) that the above conditions (a) and (b) will be fulfilled if the following, slightly stronger ones, are satisfied: (a') the functions $\bar{\gamma}_n(z,z_1,\ldots,z_n)$ tend to finite limits as $z, z_1\ldots, z_n$ tend independently to zero; and (b') the increments in γ_n, corresponding to increments for 0 to $\delta z, \delta z_j$ ($\leq \lambda$) in z, z_j, are $0(\delta z, \delta z_j)$. Thus (a') and (b') constitute sufficient conditions for Markofficity. It is evident from Eq. (VII.21) that these conditions correspond to the requirement that the functions $\gamma_n(\tau',\tau_1',\ldots,\tau_n')$ decay to zero with increase of <u>each</u> of the variables $\tau',\tau_1',\ldots,\tau_n'$ in some characteristic microscopic time, i.e., in a time whose duration depends only on intensive parameters. By Eqs. (VI.27), (VI.28) and (VII.18) this decay condition for γ_n is equivalent to the one that the function $G_n(t,t_1,\ldots,t_n)$ decays to zero with each of the temporal differences $(t-t_1), (t_1-t_2),\ldots,(t_{n-1}-t_n), t_n$ in some characteristic microscopic time $\bar{\tau}_o$, whose ratio to ω^{-1} depends only on intensive variables. Hence, as the times $t,t_1,\ldots,t_n,0$ are arranged in descending order, it follows from Eqs. (V.28), (V.29) that G_n will decay in this way, and thus the system will become Markoffian, if the various 'forces' $F^{(r)}(t_r,\ldots,t)$ and $F^{(n-r)}(t_{t+1},\ldots,0)$ become mutually uncorrelated, in the microcanonical states $D(a)/W(a)$, whenever their respective temporal ranges (t_r,t) and $(0,t_{r+1})$ become separated by $\gtrsim \bar{\tau}_o$.

In order to interpret these decay conditions in terms of many-body theory we shall express them in terms of correlations between Heisenberg operators $V_k(t) \equiv V(a,a'|t)$ for the reference system S_o. Thus, as G_n is composed of contributions of the form (V.36), we see that the above decay conditions are fulfilled if the interaction components $V_k(t)$ simulate random forces, in the sense that products $V_{k_1}(t_1')\ldots V_{k_\ell}(t_\ell')$ and $V_{k_1'}(t_1'')\ldots V_{k_m'}(t_m'')$ become uncorrelated with one another, in states $D(a)/W(a)$, whenever their respective temporal ranges are separated by $\gtrsim \bar{\tau}_o$.

As regards the fulfillability of these conditions, we note the functions (V.36) represent equilibrium correlations between fluctuations of Heisenberg operators for S_o. Such correlations between Heisenberg operators, in equilibrium states of a system, have been studied extensively in many-body theory;[12] where it is found that,

in a large class of cases, the correlations do decay† in microscopic times, corresponding to 'lifetimes' of quasi-particles in the system. This supports our contention that, in principle, the functions G_n can decay in accordance with our conditions.

Finally, we recall that the form of V, and therefore that of G, depends on the construction of the phase cells. Our conclusion, then, is that the system will be Markoffian if the cells are so constructed as to render the interaction components $V(a,a'|t)$ both weak ($\lambda = \Gamma^{-r} \ll 1$) and random, in the sense specified above. This combination of requirements might impose restrictions on the coarseness parameter beyond those specified in (VI.12). For example, it might turn out that such restrictions arise through a dependence of the decay properties of G on the index, r, of the formulae (VI.12).

VIII. Phenomenological Law

We have seen that the procedure from the microscopic to the cellular level of description of the system has led, subject to specified conditions, to a statistical law of Markoffian form. The final stage of the theory will be to proceed from the cellular to the macroscopic level and thereby to obtain a phenomenological law, together with a description of fluctuations of the macroscopic variables. In this way we shall obtain a condition that these fluctuations remain sufficiently small for the phenomenological law to be regarded as deterministic.

The phenomenological law may be extracted from the Markoffian master equation by means of a method due originally to van Kampen.[5] This method, like the one we have used to Markoffise the master equation, depends on expressing all relevant variables in suitable units. In order to employ this method, we first express $k(a',a)$, the transition rate for a jump from $C(a)$ to $C(a')$, in terms of the value, a, of $\tilde{A}$ in the former cell, and the change, $(a'-a)$, in this value, due to the jump. Thus

$$k(a,a') \equiv \tilde{k}(a|\Delta a) \qquad \text{(VIII.1)}$$

with

$$\Delta a = a' - a.$$

On comparing this equation with Eq. (VII.17), we observe two things. Firstly, as the interaction components $v(a,a')$ engender transitions only between neighbouring cells, the function $\tilde{k}$ depends on Δa only through the change Δn ($= \Delta a/\Omega$) in the cellular quantum numbers n,

† We should add, though, that except in simple cases, the decay properties are obtained only by means of certain truncated perturbative expansions.

defined in Eq. (VI.1). Secondly, as microcanonical averages $\langle\,\rangle_a$ are equal to corresponding canonical averages, expressed in terms of intensive variables rather than a, the function $\tilde{k}$ depends on a only via the intensive variables x. Thus we may write

$$\tilde{k}(a|\Delta a) = w(x|\Delta n) \tag{VIII.2}$$

where the function w contains no explicit Γ-dependence. Expressing the distribution p in terms of x, rather than a, i.e.,

$$p(a,\sigma) = f(x,\sigma), \tag{VIII.3}$$

it follows from (VIII.1)-(VIII.3) that the master equation (VII.15) now takes the form

$$\frac{d}{d\sigma} f(x,\sigma) = \sum_{\Delta n} \left[w\left(x + \frac{\Delta n}{\Omega} \mid -\Delta n\right) f\left(x + \frac{\Delta n}{\Omega}, \sigma\right) - w(x|\Delta n) f(x,\sigma) \right].$$

This equation may be expressed in a more convenient form by noting that its right-hand side will be unchanged if Δn is replaced by $-\Delta n$ in the first product wf of the summand. Thus we rewrite the master equation as

$$\frac{d}{d\sigma} f(x,\sigma) = \sum_{\Delta n} \left[w\left(x - \frac{\Delta n}{\Omega} \mid \Delta n\right) f\left(x - \frac{\Delta n}{\Omega}, \sigma\right) - w(x|\Delta n) f(x,\sigma) \right]. \tag{VIII.4}$$

For simplicity we shall treat this master equation for the case of only a single macroscopic variable, other than the energy; i.e., for the case of only one intensive variable x. The result will not differ in any essential way from the one we have obtained elsewhere[3] for the general case of several variables. This latter result will be quoted at the end of this Section.

We shall treat x as a continuous variable in the master equation (VIII.4), since, as follows from Eq. (VI.3), the spacing between its eigenvalues is the infinitesimal quantity Ω^{-1}. Thus we expand the right-hand side of the master equation (VIII.4) in a Taylor series:

$$\frac{df(x,\sigma)}{d\sigma} = \sum_{r=1}^{\infty} \frac{(-1)^r \Omega^{-r}}{r!} \frac{\partial^r}{\partial x^r} \left(w_r(x) f(x,\sigma) \right) \tag{VIII.5}$$

where

$$w_r(x) = \sum_{\Delta n} w(x|\Delta n)(\Delta n)^r, \tag{VIII.6}$$

which is an intensive variable. We now transform from a cellular to a macroscopic time scale, for which the temporal unit is $\bar{\tau}_M$. The time on this latter scale is therefore

$$s = t/\bar{\tau}_M = \sigma/\Omega. \tag{VIII.7}$$

We also make a transformation

$$x = \bar{x}(s) + \Omega^{-\frac{1}{2}} y \tag{VIII.8}$$

where $\bar{x}(s)$ is the mean of x for the distribution $f(x,s)$, and y is designed to be a suitable reduced variable to represent the fluctuations of x about its mean. Thus, anticipating that the fluctuations of x have dispersion of order Ω^{-1}, we shall treat y as a variable of order Ω^0 in our treatment of the master equation; and then we shall obtain a self-consistency condition for the justification of this treatment—the condition will be that for which the analysis leads to the conclusion that the dispersion in y is indeed $0(\Omega^0)$.

On expressing the distribution function in terms of y and s,

$$f(x,\sigma) = \Omega^{\frac{1}{2}} F(y,s) \tag{VIII.9}$$

where the factor $\Omega^{\frac{1}{2}}$ ensures that $\int F\,dy = 1$. It follows from Eqs. (VIII.7)-(VIII.9) that the master equation (VIII.5) transforms to

$$\frac{\partial F}{\partial s} - \Omega^{\frac{1}{2}} \frac{d\bar{x}(s)}{ds} \frac{\partial F}{\partial y} = \Omega \sum_{r=1}^{\infty} (-1)^r \frac{\Omega^{-r}}{r!} \frac{\partial^r}{\partial y^r}\left[w_r(x(s) + \Omega^{-\frac{1}{2}} y)F \right]. \tag{VIII.10}$$

We now expand the sum on the right-hand side of this equation in powers of $\Omega^{-\frac{1}{2}}$. Thus

$$\frac{\partial F}{\partial s} - \Omega^{\frac{1}{2}} \left(\frac{d\bar{x}(s)}{ds} - w_1(\bar{x}(s)) \right) \frac{\partial F}{\partial y} =$$

$$\frac{1}{2} w_2(\bar{x}(s)) \frac{\partial^2 F}{\partial y^2} - \frac{\partial w_1(\bar{x}(s))}{\partial(\bar{x}(s))} \frac{\partial}{\partial y}(yF) + 0(\Omega^{-\frac{1}{2}}). \tag{VIII.11}$$

Equating the coefficient of $\Omega^{\frac{1}{2}}$, the leading power of Ω, to zero, we obtain the phenomenological law governing the time-development of $\bar{x}$:

$$\frac{d\bar{x}(s)}{ds} = w_1(\bar{x}(s)). \tag{VIII.12}$$

In view of this equation, we see that Eq. (VIII.11) reduces, in zero order in $\Omega^{-\frac{1}{2}}$, to the form

$$\frac{\partial F}{\partial s} = \tfrac{1}{2}\ell(s) - m(s)\frac{\partial}{\partial y}(yF) \tag{VIII.13}$$

where

$$\ell(s) = w_2(\bar{x}(s)) \tag{VIII.14}$$

and

$$m(s) = \frac{\partial w_1(\bar{x}(s))}{\partial(\bar{x}(s))} \tag{VIII.15}$$

are intensive variables, of order Ω^0. Equation (VIII.13) evidently governs the fluctuations of x about its mean, $\bar{x}(s)$. This equation will now be used to obtain the time-dependence of the dispersion of these fluctuations. Thus, denoting the dispersion of y by

$$\Delta(s) = \int dy\, y^2 F(y,s), \tag{VIII.16}$$

it follows from Eq. (VIII.13) that

$$\frac{d\Delta(s)}{ds} = \ell(s) + 2m(s)\Delta(s). \tag{VIII.17}$$

It will be assumed that the initial measurement of $\tilde{A}$ is sufficiently precise to ensure that Δ is zero, or at most of order Ω^0, at $s=0$. The question we seek to answer is: Does $\Delta(s)$ subsequently remain bounded to values of order Ω^0 or does it grow to greater values? If the answer is affirmative, then the fluctuations in x are always $0(\Omega^{-1})$, which means that the phenomenological law (VIII.13) is (almost) a deterministic one. It also means that our treatment of y as a variable of order Ω^0 is self-consistently justified. On the other hand, if the answer is negative, the macroscopic variables $\tilde{A}$ evolve according to a stochastic law, as in hydrodynamical turbulence.

In order to answer this question regarding $\Delta(s)$, we first note that, by Eqs. (VIII.6) and (VIII.12), $\ell(s)$ is the second moment of the transition probability w, and is therefore always positive. It follows from Eq. (VIII.17) that $\Delta(s)$ will increase indefinitely if m(s) is positive, but will remain bounded to values $<|\ell/2m|$ if m is negative. Consequently, the dispersion in y will be $0(\Omega^0)$ at all times if m is negative. Hence, our condition for macroscopic causality is that m(s) should be a negative quantity.

It is interesting to observe that this causality condition is exactly the same as the condition for stability of the phenomenological law given by Eq. (VIII.12). For it follows from this latter equation and Eq. (VIII.15) that a disturbance $\delta\bar{x}(x)$ in $\bar{x}(s)$ will evolve according to the law

$$\frac{d}{ds}\delta\bar{x}(s) = m(s)\delta\bar{x}(s). \qquad \text{(VIII.18)}$$

Consequently, the disturbance $\delta\bar{x}$ will be amplified or damped out in time according to whether m is positive or negative. Therefore, the condition for stability of the phenomenological law is simply that $m < 0$, which is the same as the macroscopic causality condition.

Finally, we remark that the form of the phenomenological law, and the condition for macroscopic causality, for the case of several variables x are given by the natural generalisation[3] of the results we have obtained for one variable. Thus, reverting to the notation we have used for a description of several variables, the phenomenological law corresponds to the m equations

$$\frac{d\bar{x}_j(s)}{ds} = w_1^j(\bar{x}(s)), \quad \text{for } j = 1,\ldots,m \qquad \text{(VIII.19)}$$

where

$$w_1^j(x) = \sum_{\Delta n} w(x|\Delta n)\Delta n_j \; ; \qquad \text{(VIII.20)}$$

while the condition for macroscopic causality is the same as that for the stability of this law. This means that the causality condition may be investigated by analysis of the stability of the phenomenological law, as given by the set of ordinary differential equations (VIII.19), without further reference to the more complicated partial differential equation from which it was extracted. It should be noted that Eq. (VIII.20), together with Eqs. (VIII.1), (VIII.2) and (VIII.17), serves to express the functions w_1^j, which determine the form of the phenomenological law, in terms of the microscopic properties of the system.

References

The theory of generalised master equations has been formulated in many ways. The form of this theory employed in the present course is due to:

1. R. Zwanzig, J. Chem. Phys. 33 (1960), 1338;
R. Zwanzig, Lectures in Theoretical Physics, Boulder, Vol. III, 1960, and by
2. G. G. Emch, Helv. Phys. Acta. 37 (1964), 532.

The characterisation of the macroscopic observables in the form presented here, and employed for the extraction of phenomenological laws from the generalised master equation, was proposed by the author:

3. G. L. Sewell, Physica 34 (1967), 493.

Earlier treatment of the properties of these observables are due to

4. J. von Neumann, Ch. V of Mathematical Foundations of Quantum Mechanics, Princeton University Press, 1955;
5. N. G. van Kampen, Physica 20 (1954), 603; and
6. G. Ludwig, in Ergodic Theories, pp. 57-132, Ed. P. Caldirola, Academic Press, 1961.

The method employed for extraction of a phenomenological law from a Markoffian master equation is due to

7. N. G. van Kampen, Can. J. Phys. 39 (1961), 551.
8. E. P. Wigner, Chapter 26 of Group Theory, Acad. Press, 1957.
9. M. L. Goldberger and E. M. Watson, Section 2.5 of Collision Theory, J. Willy and Sons, Inc., 1964.
10. L. Van Hove, Fundamental Problems in Statistical Mechanics, pp. 157-172, Ed. E. G. D. Cohen, North Holland, 1962.
11. N. G. van Kampen, Physica 25 (1959), 1294.

For a many-body theoretical treatment of correlation functions, see, for example,

12. V. L. Bonch-Bruerich and S. V. Tyablikov, The Green Function Method in Statistical Mechanics, North Holland, 1962.

A DIAGRAM METHOD FOR THE DENSITY MATRICES[†]

A. Isihara
Department of Physics
State University of New York
Buffalo, New York

Contents

* * *

I. Introduction

The object of these lectures is to discuss some quantum field theoretical methods for the density matrices or the distribution functions. Recent progress in the quantum field theoretical methods in statistical mechanics is reflected in the publication of many articles, review papers, monographs and the lectures at various summer schools such as Montroll's at Boulder in 1960 or de Dominicis' at les Houches in 1959.[1)] While most of them emphasize the evaluation of partition functions or the energies of many-body systems, it is well-known that distribution functions also play important roles in the evaluation of average quantities, spatial correlation and the intensity of scattered electromagnetic waves. In this connection, diagram methods for the distribution functions will be discussed; however, since there are many distribution functions, attention will be limited to equilibrium systems.

Roughly speaking, the investigation on the density matrices and the distribution functions has been developed in three stages. In the first ten-year period between 1932 and 1942 the pioneering

[†] Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

works by Wigner and Uhlenbeck on quantum systems and by Mayer on classical systems are to be mentioned.[2] In 1940 Husimi[3] published his well-known thesis on the formal structure of the density matrices and introduced the resolvent solution of the Bloch equation. Concerning a more specific theory, the work by London[4] published in 1943 may be cited. In this work London investigated the spatial correlation of particles in a Bose gas in connection with the Bose-Einstein condensation.

It seems that the most significant achievement in the next ten year period between 1942 and 1952 was on classical distribution functions. Needless to mention, we find in this period the important contributions by Bogolyubov, Kirkwood, Born and Green, Mayer and others.[5]

In the last stage, namely since about 1952, quantum field theoretical methods have been applied to statistical problems. However, it is surprising that most of the review articles and monographs did not give credit to the pioneering work by Goldberger and Adams[6] published in the Journal of Chemical Physics in 1952. I think that they were the first to introduce the Feynman propagator technique to the Bloch equation for the density matrix. Actually, this is the reason why the year 1952 has been designated as the beginning of the last stage. However, they did not apply their formalism to specific problems nor introduce diagrams, although it was clear that one could introduce Feynman diagrams in a reciprocal temperature and coordinate space.

Thus, it took a few more years until physicists started developing the new diagrammatic methods in statistical mechanics. It is very interesting to observe that significant and sudden development was made almost simultaneously in 1957 by many physicists in different countries, although preceded by important articles by Schwinger, Brueckner, Watson, Matsubara, Bogolyubov and other Russian colleagues.[7] For instance, there are the papers by Goldstone and Hubbard in the Proceedings of Royal Society, the articles by Gell-Mann and Brueckner, Brueckner and Sawada, Lee, Huang, Yang and Luttinger, Bardeen, Cooper and Schriefer, et cetera, in the Physical Review, those by Landau, Galitzkii and Migdal, Beliaev and others in the Soviet Physics, JETP, and so on.[8] As a consequence, very rapid development in the methods and techniques resulted. It might be mentioned that counting from the year 1957 this is the tenth anniversary for the new statistical mechanics.

We have now quite a few papers on many subjects. Thus, let us pick up a few important subjects which will be related to the present talks.

First, we have an electron gas. Its treatment is of prime importance and connected with the investigation of metals, plasmas and ionic solutions. There are important contributions on the evaluation

of the correlation energy or the equation of state expanded in linked clusters. Particular mention should be made of the fine work by Montroll and Ward[9] who further developed the Goldberger and Adams formalism to the evaluation of the correlation energy. At the same time, in a Belgian journal, Fujita, Isihara and Montroll developed a general linked cluster expansion theory of the distribution functions.[10] This theory is discussed in the next section, while in subsequent sections emphasis will be placed on its applications and further developments.[11]

Another important subject of interest is liquid helium. It is indeed very strange that the nature has provided us with twin ideal objects of investigation, liquids He^4 and He^3. These particles are physicists' particles: Their chemical structures are simple but their condensed phases are fascinating. They look similar, yet one obeys Bose statistics and the other Fermi statistics and differs from the other.

The investigation of superfluidity in liquid He^4 has made good strides since Landau introduced intuitively the well-known energy spectrum.[12] It was Bogolyubov who first derived the phonon spectrum from a more microscopic theory.[13] However, his treatment assumed that the interactions are weak. Thus, Lee, Huang and Yang[14] considered a hard-sphere Bose gas as an idealized model for liquid helium.

The treatment of a hard-sphere system is important not only for liquid helium but also for nuclear matter. Theoretically, it is significant since the usual perturbation expansion in potentials breaks. Therefore, modification of the standard many-body theory based on perturbation is required.

For this purpose, Lee, Huang and Yang introduced the pseudopotential method.[14] This method is based on introducing a new continuous potential which is equivalent to the hard-sphere potential if one uses it in the Schroedinger equation with suitable boundary conditions. Lee and Yang developed another method known as the binary kernel method.[14] In this method a two-body problem is solved first and is used in a many-body problem. In the sense that all the two-body interactions are taken into consideration, this binary kernel method is analogous to the t-matrix method developed by Brueckner.

The pseudopotential and the binary kernel methods are different. In practice, however, both require certain approximations. We shall apply these methods to the evaluation of the distribution functions and compare the subtle differences.

The development of the investigation on liquid He^3 has been remarkable. This liquid was obtained only twenty years ago by the devoted efforts of Los Alamos people overcoming various difficulties such as producing low enough temperatures and obtaining the isotope. There were good reasons for the development, of course. Since the

particles obey Fermi statistics, one might expect an independent proof of Sommerfeld's theory of the electronic specific heat, an independent confirmation of the essential role of statistics in the transition in liquid He^4. One could even be ambitious and look for a transition into a new phase due to formation of Cooper pairs.[15] It is not possible to spend more time here to discuss the details of the research on liquid He^3, but we shall later discuss the spatial distribution in a hard-sphere Fermi system and its difference from the case of a Bose gas.

Thus, we have at least three important systems of our concern: (1) an electron gas, (2) a hard-sphere Bose gas, and (3) a hard-sphere Fermi gas. For convenience sake, we shall use the units such that $\hbar = 1$ and $2m = 1$ in these talks.

II. The Propagator Formalism

We start with the density matrix for an N particle system:

$$\langle \vec{r}' | \rho | \vec{r} \rangle = \sum_n \psi_n^*(\vec{r}') e^{-\beta H} \psi_n(\vec{r}), \tag{II.1}$$

where $\vec{r}$ is for the N particle coordinates, $\beta = 1/kT$ and $\{\psi_n(\vec{r})\}$ is an orthonormal complete set of eigenfunctions. The density matrix is useful in many-body theories since the right hand side of this expression may be expressed in terms of any orthonormal set of functions and since the Hamiltonian is usually given. On the other hand, the partition function

$$Z_N = \sum_n e^{-\beta E_n} \tag{II.2}$$

requires to evaluate the energies E_n solving the Schroedinger equation. This step is, of course, not necessarily simple.

Once the density matrix is known, one obtains the partition function by taking the trace

$$\int \langle \vec{r} | \rho | \vec{r} \rangle d\vec{r} = Z_N. \tag{II.3}$$

The diagonal elements of the density matrix are the probabilities of finding particles in a given configuration $\vec{r}$ and $\beta = 1/kT$. Therefore, the distribution functions will be defined in terms of the diagonal elements.

We rewrite Eq. (II.1) as follows:

$$\langle \vec{r}' | \rho | \vec{r} \rangle = \sum_n \Psi_n^*(0, \vec{r}') \Psi_n(\beta, \vec{r}), \tag{II.4}$$

where $\Psi(\beta,\vec{r})$ is defined by

$$\Psi(\beta,\vec{r}) = e^{-\beta H}\psi(\vec{r}). \tag{II.5}$$

Formally, this new wave function defined in the coordinate-reciprocal temperature space satisfies the Bloch equation:

$$\frac{\partial\Psi(\beta,\vec{r})}{\partial\beta} = -H\Psi. \tag{II.6}$$

When we take the conjugate complex of this function indicated by the asterisk, we treat β like a complex quantity.

The form of the right hand sum of Eq. (II.4) reminds us of an integral equation:

$$\Psi_n(\beta,\vec{r}) = \int K(\vec{r}\,\beta;\vec{r}'\beta')\Psi_n(\beta',\vec{r}')d\vec{r}'. \tag{II.7}$$

Here, it is remarked that the β-integration is not performed in the right hand side. It looks a bit strange that one can get β'-independent function $\Psi(\beta,\vec{r})$ in the left hand side, but this is due to the artificial nature of the differential equation (II.6).

The kernel or the propagator function $K(\vec{r}\beta;\vec{r}'\beta')$ may be expanded in terms of the eigenfunctions of the integral equation:

$$K(\vec{r}\beta;\vec{r}'\beta') = \sum_n \Psi_n^*(\beta',\vec{r}')\Psi_n(\beta,\vec{r}). \tag{II.8}$$

Thus, we may establish the relation

$$\langle\vec{r}|\rho|\vec{r}\rangle = K(\vec{r}\beta;\vec{r}0). \tag{II.9}$$

This relation gives a new rule to evaluate the distribution functions: we evaluate the <u>total</u> propagation of particles from $(0,\vec{r})$ to $(\beta,\vec{r})$.

The propagator is supposed to satisfy the Green's function equation corresponding to Eq. (II.6):

$$\frac{\partial K(2,1)}{\partial\beta} + H\,K(2,1) = \delta(2-1), \tag{II.10}$$

where we have used 2 for the set of coordinates $(\beta,\vec{r})$. Thus, in case the total Hamiltonian is split into two parts:

$$H = H_o + H_1, \tag{II.11}$$

we may divide the propagator correspondingly into two parts:

$$K = K_o + K_1 \tag{II.12}$$

such that K_o satisfies

$$\frac{\partial K_o(2,1)}{\partial \beta} + H_o K_o = \delta(2-1). \tag{II.13}$$

Then we find from Eqs. (II.10), (II.11), (II.12) and (II.13) the following integral equation:

$$K(2,1) = K_o(2,1) - \int K_o(2,3) H_1(3) K(3,1) d3, \tag{II.14}$$

where the numbers represent $(\vec{r},\beta)$ coordinates as before, and the integration is over the entire volume with respect to $\vec{r}$ and from 0 to β concerning β. Thus, assuming that H_1 is small, one obtains a perturbation series:

$$K(2,1) = K_o(2,1) - \int K_o(2,3) H_1(3) K_o(3,1) d3 + \cdots . \tag{II.15}$$

This series determines the propagator. The distribution function of ℓ particles may be defined by

$$\rho_\ell^{(N)}(r^\ell) = \frac{N!}{(N-\ell)! Z_N} \int \cdots \int K(\vec{r}^N \beta; \vec{r}^N 0) d\vec{r}_{\ell+1} \cdots d\vec{r}_N, \tag{II.16}$$

where

$$\vec{r}^\ell = \vec{r}_1, \vec{r}_2, \cdots, \vec{r}_\ell .$$

The definition (II.15) corresponds to the normalization

$$\int \rho_\ell^{(N)}(\vec{r}^\ell) d\vec{r}^\ell = N(N-1)\cdots(N-\ell). \tag{II.17}$$

In many-body theories we find the grand ensemble definition:

$$\rho_\ell(\vec{r}^\ell) = \frac{1}{\Xi} \sum_{N=\ell}^{\infty} \rho_\ell^{(N)}(\vec{r}) Z_N z^N \tag{II.18}$$

more convenient. Here Ξ is the grand partition function:

$$\Xi = \sum_N z^N Z_N . \tag{II.19}$$

The normalization of ρ_ℓ is then the expectation value:

$$\int \rho_\ell (\vec{r}^\ell) d\vec{r}^\ell = \langle N(N-1)\cdots(N-\ell+1)\rangle . \tag{II.20}$$

III. Linked Cluster Expansion

The propagator must have the same symmetry property as the original wavefunction Ψ as one can see from the structure of Eq. (II.7). In particular, if Ψ_0 is given by the Slater determinant (or permanent), the corresponding propagator K_0 will be an N by N determinant.

Generally, if we expand a determinant of N rows and columns we get terms composed of N elements by cyclic permutation of particles. When ℓ particles form a closed cycle in such terms, we say that they form an ℓ-toron. We expect torons of all order in the expansion in accordance with splitting of an integer N into sub-groups:

$$N = \sum_\ell \ell\, m_\ell \tag{III.1}$$

where m_ℓ is the number of ℓ-torons.

It is remarked that K_0 in the determinant-form appears only once in the perturbation series (II.15). Although we have used the same notation, $K_0(3,1)$ in the second term, for instance, must be the diagonal term of the determinant. This is simply because the variable 2 is supposed to carry the symmetry property of the wavefunction.

For this reason, the right hand side members of Eq. (II.15) are obtained by connecting the torons from the first propagators in all possible ways. The resulting diagrams are, however, not necessarily connected because H_1 is generally a sum of pair-interactions. So, the right hand side may be regrouped in terms of the number of particles in connected graphs. We then observe that the summation over N in Eq. (II.18) removes the restriction which arises from a constant N. Thus, in the grand ensemble we are able to discuss the evaluation of the distribution functions only in terms of connected graphs, forgetting from which determinant they have been brought about.

Suppose now we take the steps in accordance with Eqs. (II.3) and (II.9). In the integration process all the particles are treated equally, and the result of integration must be constant. It is then natural to introduce the constants b_ℓ which arise from the terms

corresponding to connected graphs of ℓ particles. The b_ℓ might appear m_ℓ times. We expect

$$Z_N = \sum_m \prod_\ell \frac{b_\ell^{m_\ell}}{m_\ell!} , \tag{III.2}$$

where the factor m_ℓ in the denominator is due to the splitting of an integer N as in Eq. (III.1) and 1/N attached to the determinant K_0. Thus, the grand partition function of Eq. (II.19) is given by

$$\ln \Xi = \sum_\ell z^\ell b_\ell , \tag{III.3}$$

the so-called linked cluster expansion.

Let us now consider to determine the pair distribution function $\rho_2(\vec{r}_1, \vec{r}_2)$ in a similar cluster expansion form. $\rho_2(\vec{r}_1, \vec{r}_2)\, d\vec{r}_1 d\vec{r}_2$ is the probability of finding two particles $\vec{r}_1$ and $\vec{r}_2$ in the volume element $d\vec{r}_1 d\vec{r}_2$ at $(\vec{r}_1, \vec{r}_2)$. We are now going to distinguish the two particles from the rest of the particles. Correspondingly, all the connected graphs are classified as follows:

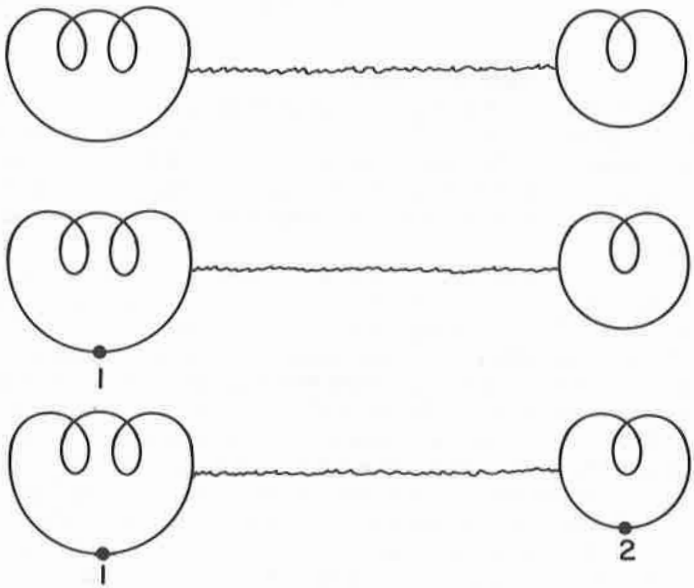


Figure 1. Three possible types of diagram.

1) Graphs which are independent of 1 and 2. These contribute terms such as b_ℓ appeared in Eq. (III.2), and altogether compensate Ξ in the denominator of Eq. (II.18).

2) Graphs which include either 1 or 2. These contribute terms such as $\ell b_\ell/V$ since all the integrands depend on the difference of coordinates and the integration over 1 or 2 must bring b_ℓ. The factor ℓ is necessary since there are ℓ possible positions for 1 (or 2) in the ℓ connected graphs. Altogether, the graphs contribute n^2 to the pair distribution function. This is easy to see because n is the probability to find a particle in a uniform random mixture and because from Eq. (III.3)

$$n = \frac{1}{V}\frac{\ln \Xi}{\ln z} = \frac{1}{V}\sum_\ell \ell b_\ell z^\ell . \tag{III.4}$$

3) Graphs which include both 1 and 2. The corresponding contribution will be of the form $b_\ell(\vec{r}_1,\vec{r}_2)$. Here $b_\ell(\vec{r}_1,\vec{r}_2)$ may be interpreted as the probability to find the particles 1 and 2 in connected graphs of ℓ particles. Since the probability of collecting ℓ particles in an interacting system may be given by z^ℓ, we arrive at $z^\ell b(\vec{r}_1,\vec{r}_2)$.

Consideration of these three contributions in the integrations of Eq. (II.16) gives

$$\rho_2(\vec{r}_1,\vec{r}_2) = n^2 + \sum_\ell z^\ell b_\ell(\vec{r}_1,\vec{r}_2), \tag{III.5}$$

where the summation should start with $\ell = 2$. Equation (III.5) is what we call the linked cluster expansion of the pair distribution function. From the normalization condition for ρ_2 we find

$$\int b_\ell(\vec{r}_1,\vec{r}_2)\,d\vec{r}_1\,d\vec{r}_2 = \ell(\ell-1)b_\ell . \tag{III.6}$$

By a similar consideration we can expand higher order distribution functions.

IV. Boltzmann Chains

Equation (II.8) suggests that the free particle propagator $K_o(\vec{r}\beta;\vec{r}'\beta')$ is given by

$$K_o(\vec{r}\beta;\vec{r}'\beta') = \frac{1}{(2\pi)^3}\int e^{-(\beta-\beta')p^2 - i\vec{p}\cdot(\vec{r}-\vec{r}')}d\vec{p}. \tag{IV.1}$$

We use this expression in the unit process illustrated in Figure 2.

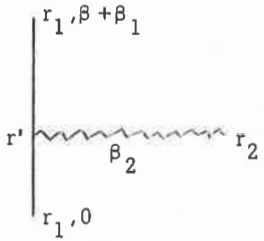


Figure 2. The propagation of G_1.

We find

$$F_1(\vec{r}_2\beta_2;\vec{r}_1\beta_1) = -z \int K_o(\vec{r}_1, \beta+\beta_1; \vec{r}'\beta_2)\phi(\vec{r}_2 - \vec{r}') \times K_o(\vec{r}'\beta_2; \vec{r}_1 0)d\vec{r}'\, d\vec{q}. \quad \text{(IV.2)}$$

After integrating over the intermediate coordinates $\vec{r}'$ we find

$$F_1(\vec{r}_2\beta_2;\vec{r}_1\beta_1) = \frac{-1}{(2\pi)^3} \int G_1(\vec{q};\beta_2 - \beta_1)\, u(\vec{q}) e^{i\vec{q}\cdot(\vec{r}_2-\vec{r}_1)}\, d\vec{q}. \quad \text{(IV.3)}$$

Here

$$G_1(\vec{q}, |\beta'' - \beta'|) = z\lambda^{-3} \exp(-\alpha(\beta-\alpha)q^2/\beta), \quad \text{(IV.3)}$$

where

$$\alpha = |\beta'' - \beta'|;$$

$$u(\vec{q}) = \int \phi(\vec{r}) e^{i\vec{q}\cdot\vec{r}}\, d\vec{r}, \quad \text{(IV.4)}$$

$\phi(r)$ being the potential.

It is to be remarked that the particle coordinates appear in Eq. (IV.3) only on the exponential factor. Thus, if we have another similar process represented by $F_1(\vec{r}_3\beta_2;\vec{r}_2\beta_2)$, following the process $F_1(\vec{r}_2\beta_2;\vec{r}_1\beta_1)$ and if the intermediate coordinates $\vec{r}_2$ are integrated, we find the total process represented by the function:

$$F_2(\vec{r}_3\beta_3;\vec{r}_1\beta_1) = \frac{1}{(2\pi)^3} \int G_1(q;\beta_3 - \beta_2)\, G_1(q;\beta_2 - \beta_1) e^{iq\cdot(\vec{r}_3-\vec{r}_1)} u(q)^2 d\vec{q}. \quad \text{(IV.5)}$$

Repeating the same processes and expanding $G_1(q;\beta_2-\beta_1)$ in a Fourier series:

$$G_1(\vec{q};\beta''-\beta') = \sum \lambda_j \psi_j(\beta')\psi_j^*(\beta'') \tag{IV.6}$$

$$\psi_j(\beta') = \beta^{-\frac{1}{2}} \exp 2\pi i j \beta''/\beta , \tag{IV.7}$$

one arrives at (after performing the β-integrations)

$$F_\ell(\vec{r}_2\beta_2;\vec{r}_1\beta_1) \sim \frac{1}{\beta(2\pi)^3} \sum_j \int dq[-u(q)]^{\ell-1}\lambda_j^\ell(q)$$

$$\times\ e^{i\vec{q}\cdot(\vec{r}_2-\vec{r}_1)} e^{2\pi i j(\beta_2-\beta_1)/\beta} . \tag{IV.8}$$

Multiplying this expression by z^ℓ and summing over all ℓ, we end up with

$$\rho_2(r) - n^2 = -\sum_j \frac{1}{\beta(2\pi)^3} \int \frac{u(q)\lambda_j^2}{1+u(q)\lambda_j} e^{i\vec{q}\cdot\vec{r}}\, d\vec{q} , \tag{IV.9}$$

where

$$\lambda_j = \frac{z}{\lambda^3} \int_0^\beta e^{-q^2\alpha(1-\alpha/\beta)} e^{2\pi i j\alpha/\beta}\, d\alpha . \tag{IV.10}$$

V. Electron Gas

Let us apply the formula (IV.9) to an electron gas at high temperatures. The Fourier transform of the potential is

$$u(q) = \frac{4\pi e^2}{q^2} . \tag{V.1}$$

The eigenvalues λ_j are given by

$$\lambda_j = \frac{z}{\lambda^3} \int_0^\beta \exp[-q^2\alpha(\beta-\alpha)/\beta - 2\pi i j\alpha/\beta]\, d\alpha . \tag{V.2}$$

At high temperatures one finds

$$\lambda_j = z\lambda^{-3}\beta\, \delta_{jo} . \tag{V.3}$$

Thus,

$$\rho_2(r) - n^2 = -\frac{z\kappa^2}{\lambda^3 (2\pi)^3} \int \frac{e^{i\vec{q}\cdot\vec{r}}}{q^2+\kappa^2} \, d\vec{q}$$

$$= -\frac{n\kappa^2 e^{-\kappa r}}{4\pi r}, \tag{V.4}$$

where the screening constant κ is

$$\kappa^2 = \frac{4\pi n e^2}{kT} . \tag{V.5}$$

It is remarked that Eq. (V.4) is due to one chain. Namely, the effect of having a linear array of interacting particles between the two particles 1 and 2 is to have the Debye screening. If the two particles are connected by many chains we would obtain in the same approximation the following result:

$$\rho_2(r) = n^2 \exp(-\Phi/kT), \tag{V.6}$$

where

$$\Phi = e^2 \frac{e^{-\kappa r}}{r} . \tag{V.7}$$

In other words, Eq. (V.4) is the first expansion term of the exponential function in Eq. (V.6).

As temperature is reduced it becomes necessary to take quantum exchange effects into consideration. The particles forming a toron may be considered to be bound by pseudo-statistical forces. Thus, one can show that a factor

$$(1 - e^{-r^2/2\beta}) \tag{V.8}$$

appears from the combination of the Boltzmann chain and free particle diagrams, λ being the de Broglie wavelength. Also, one can show that the first order exchange diagrams yield a contribution

$$\sim e^{-r^2/2\beta} \ln(r/\lambda) \tag{V.9}$$

for large distances.

Note that temperature dependences come in now. It becomes necessary in such cases to evaluate higher order eigenvalues. A

systematic expansion for fermion hard-sphere particles has been given by Isihara and Gupta.[16]

VI. Hard-Sphere Boltzmann Gas

In view of our interest in liquid helium, let us consider the lowest temperature limit. The eigenvalues λ_j for such a case are

$$\lambda_j = \frac{2n\beta}{\beta q^2 + (2\pi j)^2/\beta q^2} . \tag{VI.1}$$

Lee, Huang and Yang showed that the hard sphere potential may be replaced by the pseudopotential which is in the first order in the hard sphere potential

$$\phi(r) = 8\pi a\delta(r). \tag{VI.2}$$

This gives

$$u(q) = 8\pi a. \tag{VI.3}$$

The summation over j may be replaced by an integral:

$$\lambda_j \longrightarrow \frac{\lambda^2}{2\pi^2} \int dx. \tag{VI.4}$$

Thus, we end up with

$$\rho_2(r) - n^2 = -\frac{4an^2}{\pi^3} \int d\vec{q}\, e^{2i\vec{q}\cdot\vec{r}} \int dx\, q^4 (q^4+x^2)^{-1}(q^4+x^2+\gamma q^2)^{-1}$$

$$= -\frac{4an^2}{\pi^2\gamma} \int d\vec{q}\, e^{2i\vec{q}\cdot\vec{r}} \left[1 - \frac{1}{(1+\gamma/q^2)^{\frac{1}{2}}}\right], \tag{VI.5}$$

where

$$\gamma = 4\pi an. \tag{VI.6}$$

The right hand side integral has been evaluated in the following form:[11]

$$\rho_2(r) = n^2\left[1 - \frac{4a}{r} g(x)\right], \tag{VI.7}$$

where

$$g(x) = G_0(x) - \frac{G_1(x)}{x};$$

$$G_1(x) = I_0(x) - L_1(x); \qquad \text{(VI.8)}$$

$$x = 2^{\frac{1}{2}} r.$$

$I_1(x)$ are $L_1(x)$ are the modified Bessel and the Struve functions.

The result obtained in Eq. (VI.7) is a smoothly varying function of the reduced variable x. It is to be remarked that the result should not be used for $x \sim 2^{\frac{1}{2}}$ as it becomes necessary to take higher order terms into consideration. As a matter of fact, the pair distribution function vanishes at $r=a$. The result gives correct long distance values for the pair distribution function. To find the behavior, one may use the known asymptotic expressions of L and I functions, or adopt[17]

$$\Re \frac{2\pi}{ir} \int_0^\infty q\left[1 - q(q^2+\gamma)^{-\frac{1}{2}}\right] \exp 2iqr\, dq =$$

$$= \frac{2\pi\gamma}{r} \int_0^{\pi/2} \sin^2\theta \exp(-2\gamma^{\frac{1}{2}} \sin\theta) d\theta \qquad \text{(VI.9)}$$

$$\sim \frac{4\pi\gamma}{rx^3} .$$

Namely, we find a decay of the pair distribution function to n^2 in proportion to r^{-4}.

The large distance behavior of the pair distribution function is related to the phonon spectrum. To see this we use the Feynman-Bijl expression:

$$\epsilon_k = \frac{k^2}{S(k)} \qquad \text{(VI.10)}$$

where the structure factor $S(k)$ is obtained from

$$S(k) = 1 + n\int \left(\frac{\rho_2(r)}{n^2} - 1\right) e^{i\vec{k}\cdot\vec{r}}\, d\vec{r}. \qquad \text{(VI.11)}$$

One finds

$$S_k = k\left(k^2 + 16\pi an\right)^{-\frac{1}{2}} \qquad \text{(VI.12)}$$

which is the result Lee, Huang and Yang[14] obtained by a different method. Thus, for small momentum we have a phonon spectrum:

$$\epsilon_k = ck, \tag{VI.13}$$

where the phonon velocity c is

$$c = 4(\pi a n)^{\frac{1}{2}}. \tag{VI.14}$$

Note $c \to 0$ as $a \to 0$. It is also to be remarked that the phonon spectrum has been obtained for a Boltzmann hard-sphere gas.

VII. Hard-Sphere Bose Gas

When particles obey quantum statistics we must start with the zero order term, namely, the ideal gas term. The free toron diagrams may be labeled with 1 and 2. Generally, we may have s cycles between 1 and 2, and t cycles between 2 and 1. The corresponding cluster function $b_\ell^{o}(\vec{r}_1,\vec{r}_2)$ with $\ell = s+t$ is given by

$$b_\ell^{o}(\vec{r}_1,\vec{r}_2) = \frac{1}{(2\pi)^6}\int \exp\left[-s\beta p^2 + i\vec{p}\cdot(\vec{r}_2-\vec{r}_1)\right]d\vec{p}$$

$$\times \int \exp\left[-t\beta q^2 + i\vec{q}\cdot(\vec{r}_1-\vec{r}_2)\right]d\vec{q}. \tag{VII.1}$$

This yields the following contribution to the pair distribution function

$$\sum_\ell z^\ell b_\ell^{o}(\vec{r}_1,\vec{r}_2) = \frac{1}{(2\pi)^3}\int \frac{ze^{-\beta p^2}e^{-i\vec{p}\cdot(\vec{r}_2-\vec{r}_1)}}{1 - z\exp(-\beta p^2)}\,d\vec{p}, \tag{VII.2}$$

where z is to be determined by

$$n = \frac{1}{(2\pi)^3}\int \frac{ze^{-\beta p^2}}{1 - ze^{-\beta p^2}}\,d\vec{p}. \tag{VII.3}$$

However, when the condensation takes place we must separate out the contributions from the $\vec{p}=0$ in these expressions. Thus,

$$\sum_{\ell} z^{\ell} b_{\ell}^{o}(\vec{r}_1,\vec{r}_2) = \begin{cases} [g_{3/2}(z,r)]^2, & (T_o < T) \\ [n_o + g_{3/2}(1,r)]^2 - n_o^2, & (T_o > T) \end{cases} \tag{VII.4}$$

where n_o is the number density of the particles in the lowest state, and T_o is the degeneracy temperature:

$$T_o = \frac{4\pi}{k}\left[\frac{n}{g_{3/2}(1,0)}\right]^{2/3}. \tag{VII.5}$$

In these expressions, we have used

$$g_{3/2}(z,r) = \lambda^{-3}\sum z^{\ell}\ell^{-3/2}\exp(-\pi r^2/\ell\lambda^2). \tag{VII.6}$$

The contribution from the interaction graphs is also reduced into the evaluation of the eigenvalue λ_j. Separating again $p=0$ term, we find

$$\lambda_j(q) = (2n - n_o)\left[q^2 + (2\pi j/\beta q)^2\right]^{-1} \tag{VII.7}$$

for $T < T_o$.

Thus, at large distances, we end up with

$$\rho_2(r) = n^2\left[1 - \frac{4a}{r}\left(1 - \frac{n_o}{2n}\right)^2 g(x)\right], \tag{VII.8}$$

where $g(x)$ is given by Eq. (VI.8).

It is very interesting to observe that the effect of condensation appears in the lowering of the first peak near $r=a$. Due to this, the decay $\rho_2(r)\to n^2$ for $r\to\infty$ becomes slower than the case without condensation.[18] Figure 3 shows schematically the pair distribution function of a hard-sphere gas. $\rho_2(r)/n^2$ is close to 2 near $r=a$ and approaches 1 in proportion to r^{-4}. The dotted curve indicates the effect of condensation.

Because of condensation, the phonon spectrum is also changed. One can show that

$$c = 4(\pi a n)^{\frac{1}{2}}(1 - n_o/2n). \tag{VII.9}$$

Thus, the sound velocity becomes temperature dependent.

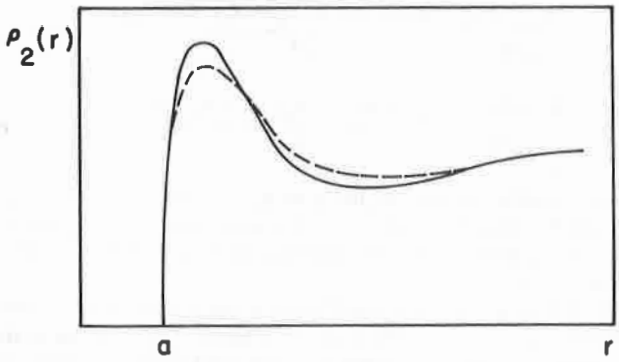


Figure 3. The Pair distribution function of a hard-sphere Bose gas.

VIII. Hard-Sphere Fermi Gas

The question of whether or not liquid He^3 shows superfluidity is extremely interesting. In 1964 Peshkov[19] reported an anormaly in specific heat, and early this year a very similar anormaly was found by Osgood and Goodkind.[20] However, according to the author's personal conversation with Dr. Goodkind and Dr. Atkins at the last New York meeting, the anormaly might have been due to experimental errors of unknown sources.

The author has recently developed a general theory for a fermion system of hard-sphere potentials. According to this, the pair distribution function is obtained by a Fourier transform of a function N(q):

$$\sum_{\ell} \frac{z^{\ell} b_{\ell}(|r_2 - r_1|)}{(\ell - 2)!} = \frac{1}{(2\pi)^3} \int N(q) e^{-2i\vec{q}\cdot\vec{r}} \, d\vec{q} . \qquad \text{(VIII.1)}$$

In view of the time, however, it is remarked here only that a general method to evaluate this function N(q) has been given. The results corresponding to short distances are as follows:

In view of the nuclear spin of He^3 let us introduce a spin eigenvalue J for this case. The non-interacting term yields for $T \to 0$:

$$\sum_{\ell} z^{\ell} b_{\ell}^{o}(r) = -\frac{(2J+1)}{4\pi 4_r 2} \left[\frac{k_o r \cos k_o r - \sin k_o r}{r^2} \right]^2 . \qquad \text{(VIII.2)}$$

The terms to first order in the interaction are approximately

$$\sum_{\ell} z^{\ell} b_{\ell}^{\ 1}(r) = \frac{2aJ(2J+1)k_o^3}{3\pi^5} \frac{\sin k_o r(\sin k_o r - k_o r \cos k_o r)}{r^4}$$

$$- \frac{4aJ(2J+1)}{4\pi^4} \frac{(\sin k_o r - k_o r \cos k_o r)^2}{r^7} . \qquad \text{(VIII.3)}$$

These results are characterized by the Fermi momentum k_o and correspond to high degeneracy. We see that the pair distribution function decays slowly to n^2 oscillating around a curve which is decreasing roughly as r^{-3}.

Since the pair distribution function depends on the density, one must carefully approach the lowest temperature. If we reduce the density first and then the temperature, we will have a dilute system. In this case it is not difficult to recover the phonon spectrum which we discussed before. The above results correspond to the reverse process where we reduce the temperature first.

IX. Concluding Remarks

We have discussed that chain diagrams are important to find many-body effects on the spatial correlations of particles. In conclusion, it is remarked that for degenerate fermions or bosons considerations only of simple chain diagrams are not good enough. As indicated in Section V, we need to take into consideration of more complicated diagrams arising from exchanges of particles.[18] Since, however, the time has come, I simply conclude these talks mentioning that such a consideration has been made and a very general formula has been obtained. According to this, the pair distribution function is expressed by essentially two eigenvalues which come from conjugate diagrams. These conjugate diagrams cancel each other's contribution at $r=0$ for fermion cases.[16] For bosons their contributions are added together.

Figure 4 illustrates typical one effective interaction diagrams. Note that (A_s) is a separated diagram and does not appear in $b_{\ell}(\vec{r}_1,\vec{r}_2)$. The diagrams (A), (A^*) and (A_s^*) contribute to the pair distribution function through the eigenvalues:

$$\lambda_j(q,r,z) = \frac{1}{(2\pi)^3} \int f(p)[1+f(\vec{p}+\vec{q})]\, e^{\alpha\{p^2-(\vec{p}+\vec{q})^2\}} e^{-i\vec{r}\cdot\vec{p}} e^{2\pi i j\alpha/\beta}\, d\alpha\, d\vec{p}$$

$$\mu_j(q,r,z) = \frac{1}{(2\pi)^3} \int f(p)f(\vec{p}+\vec{q})[1+f(\vec{p}+\vec{q})]\, e^{\alpha\{p^2-(\vec{p}+\vec{q})^2\}} e^{-i\vec{r}\cdot\vec{p}} \qquad \text{(IX.1)}$$

$$\times\ e^{2\pi i j\alpha/\beta} d\alpha\, d\vec{p}$$

for bosons. $f(p)$ is the Bose distribution.

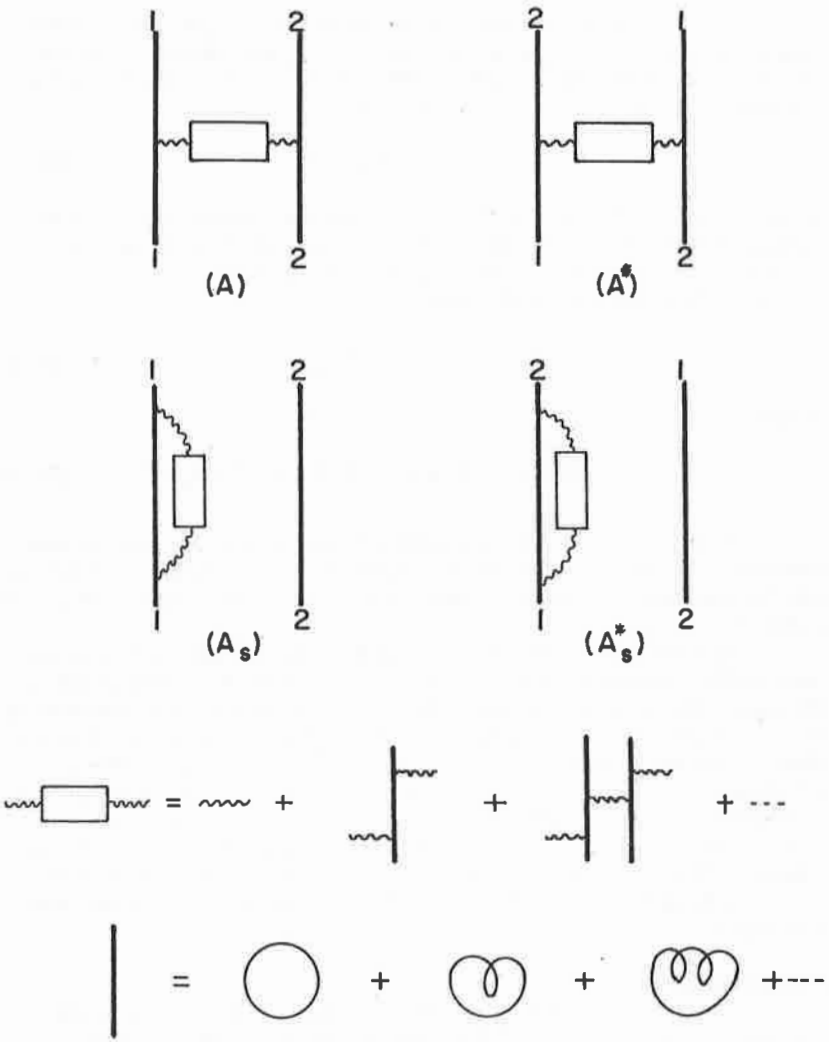


Figure 4. One Effective Line Diagrams. (A) and (A^*) are conjugate to each other. The conjugate diagram (A_s) to (A_s^*) is unconnected and involved in the n^2 term in the pair distribution function. A wavy line with a box represents an effective interaction which is a series of interactions. An effective propagation illustrated by a bold line is composed of many torons.

We have used for hard-sphere systems the simplified pseudopotential. This is just for simplicity's sake, and caution is necessary in using this approximation. On the other hand, one could use the binary kernel method. It is defined by

$$B = -\phi \exp(-\beta H_2), \tag{IX.2}$$

where H_2 is the Hamiltonian for two particles. Because the exponential factor B can be finite, even the potential ϕ is divergent. Actually, the binary kernel B assumes the following expression in which ϕ does not appear explicitly:

$$B = \frac{\partial U_2}{\partial \beta} + H_2^{o} U_2, \tag{IX.3}$$

where

$$U_2 = \exp(-\beta H_2) - \exp\left(-\beta H_1(r_1)\right) \exp\left(-\beta H_1(r_2)\right). \tag{IX.4}$$

Thus, use of the binary kernel is convenient for hard-sphere systems. Moreover, at short distances the pair distribution function may be approximated by U_2. Then, one finds a self-consistent result $\rho_2(a) = 0$.

Nevertheless, the structure of B in higher order terms in the hard-sphere potential is complcated. The situation is analogous to the exact pseudopotential when this is compared with the approximate one. Indeed, we can show that the first order binary kernel is equivalent to the first-order pseudopotential. Use of the first order expressions is convenient, but it must be remembered that they are not divergent at the hard-sphere boundary at $r = a$. Thus, use of these may be considered to be equivalent to assuming a weak repulsive potential. Therefore, it is not surprising that the results agree with what Bogolyubov obtained based on the assumption of weak repulsive potentials.

Acknowledgement

The author acknowledges the hospitality of the Department of Physics at the University of Colorado where he spent part of the summer of 1967.

References

1. E. W. Montroll, Lectures in Theoretical Physics, edited by W. E. Brittin et al., Vol. III (Interscience Publishers, 1960). C. De Dominicis, Cours de l'Ecole de Physique Théorique des Houches (Dunod, 1962). See also the lecture notes by

Brueckner, Hugenholtz and Huang in the 1958 Summer School at Grenoble.
2. E. P. Wigner, Phys. Rev. 40, 749 (1932). G. E. Uhlenbeck and E. Beth, Physica 3, 729 (1936). J. E. Mayer, Thermodynamics of Gases, Encyclopedia of Physics, edited by Flugge (J. Springer, 1958).
3. K. Husimi, Proc. Phys. Math. Soc. Japan, Ser. 3 22, 264 (1940).
4. F. London, J. Chem. Phys. 11, 203 (1943).
5. See the chapter written by Mayer in the Encyclopedia in Physics, Reference 2.
6. M. L. Goldberger and E. N. Adams, J. Chem. Phys. 20, 240 (1952).
7. K. M. Watson, Phys. Rev. 103, 489 (1956), T. Matsubara, Prog. Theo. Phys. 14, 351 (1955). N. N. Bogolyubov, J. Phys. U.S.S.R. 11, 23 (1947). The Green's Function Method in Statistical Mechanics, V. L. Bonch Bruevich and S. V. Tyablikov, English translation by D. ter Haar (North Holland, 1960). A. A. Abrikozov, L. P. Gorkov and I. E. Dzyaloshinski, Methods of Quantum Field Theory in Statistical Physics, English translation by R. A. Silverman (Prentice-Hall, 1963). L. Kadanoff and G. Baym, Quantum Statistical Mechanics (Benjamin, 1962).
8. Well-organized references on these articles can be found in P. Noziér, Le Probleme a N Corps (Dunod, 1963).
9. E. W. Montroll and J. C. Ward, Phys. Fluids 1, 55 (1958).
10. S. Fujita, A. Isihara and E. W. Montroll, Bull. Classe Sci. Acad. Roy. Belg. 44, 1018 (1958).
11. A. Isihara and D. D. H. Yee, Phys. Rev. 154, 150 (1964); Physica 30, 2123 (1964); C. J. Nisteruk and A. Isihara, Phys. Rev. 154, 150 1967).
12. L. Landau, J. Phys. U.S.S.R. 11, 91 (1947).
13. N. N. Bogolyubov, J. Phys. U.S.S.R. 11, 23 (1947).
14. K. Huang and C. N. Yang, Phys. Rev. 105, 767 (1957); K. Huang, C. N. Yang and J. M. Luttinger, Phys. Rev. 105, 776 (1957); T. D. Lee, K. Huang and C. N. Yang, Phys. Rev. 106, 1135 (1957). See also, K. A. Brueckner and K. Sawada, Phys. Rev. 106, 1117, 1128 (1957). T. D. Lee and C. N. Yang, Phys. Rev. 113, 1165 (1959).
15. K. A. Brueckner and T. Soda, P. W. Anderson and P. Morrel, Phys. Rev. 118, 1442 (1960). V. J. Emery and A. M. Sessler, Phys. Rev. 119, 43 (1960). L. Pitaevskii, J.E.T.P. 37, 1794 (1959) [Soviet Physics 10, 1267 (1960)]. L. P. Gor'kov and V. M. Galitskii, J.E.T.P. 40, 1124 (1961) [Soviet Physics 13, 792 (1963)]. P. Morel and P. Noziér, Phys. Rev. 126, 1909 (1962).
16. A. Isihara and A. Gupta, Phys. Rev. (to be published).

17. A. Isihara and Donald W. Jepsen, Phys. Rev. 158, 112 (1967).
18. A. Isihara, Proc. Internat. Conf. on Statis. Mech. Copenhagen (1966).
19. V. P. Peshkov, J.E.T.P. 46, 1510 (1964) [Soviet Physics 19, 1023 (1964)].
20. E. B. Osgood and J. M. Goodkind, Phys. Rev. Letters 18, 894 (1967).

QUANTIZATION OF HIGHER SPIN FIELDS[†]

Y. Takahashi
Dublin Institute for Advanced Studies
Dublin, Ireland

Notes compiled by
Amarjit H. Singh
University of Georgia
Athens, Georgia

Contents

* * *

[†]Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

I. Introduction

Let us briefly indicate our approach and the material to be covered.

Our ultimate aim is to evolve a unified approach for quantizing fields of arbitrary spin. We shall find that such a unified approach is possible for a large class of fields which satisfy a certain set of conditions (to be explicitly stated in due course). It turns out that most fields of physical interest actually do fall in this class. Not included are the Maxwell field and massless fields of arbitrary spin——these must be treated separately. Subsequently whenever we speak of fields in general, it is to be understood that the Maxwell field and massless fields of arbitrary spin are excluded from the discussion. A peculiar case is that of the neutrino: two-component neutrinos are excluded from the aforementioned class while four-component neutrinos are included.

If one uses the ordinary Hamiltonian formalism for quantization, one obtains equal-time commutation relations which are not covariant in form. This procedure for quantization is quite tedious and one must consider each case separately. In other words, the canonical formalism does not lend itself easily for evolving a unified approach to quantization. So we try to develop an alternative approach.

In our quantization procedure the most fundamental role is played by the equation

$$-i\partial_\lambda \phi(x) = [\phi(x), P_\lambda], \tag{I.1}$$

which is the covariant generalization of the Heisenberg equation of motion

$$i\frac{\partial \mathcal{O}(t)}{\partial t} = [\mathcal{O}(t), H], \tag{I.2}$$

where $\mathcal{O}(t)$ is any dynamical variable. We regard (I.2) as the most fundamental relation and the reason for doing so is that one can derive from it Bohr's frequency-energy relation which has a firm experimental basis. This is in sharp contrast to the purely canonical approach. We shall take the following interpretation of (I.2): Suppose the time evolution of $\mathcal{O}(t)$ is known. Then we can determine both H and the commutation relations. However, there are too many unknowns in this equation so that our determinations of H and the C.R.'s may not be unique. In other words, there may exist a number of possible quantizations. The unknowns can be eliminated by imposing physically meaningful requirements on the solutions. Use of (I.1) then leads to commutation relations which are covariant in form, and the P_λ determined turns out to be unique as we shall prove.

In our approach the knowledge of an explicit Lagrangian is not required. However we do assume that a Lagrangian must exist even though its explicit form may not be known. This is necessary for the conservation laws to follow from the invariance properties of the field equations.

First we study the classical field theory for relativistic free fields from a unified point of view. In Section II we review the wave equations for the various spin fields. In each case the field equation and the associated subsidiary conditions (if any) can be cast into the general form

$$\Lambda_{\alpha\beta}(\partial)\phi_{\beta}(x) = 0, \tag{I.3}$$

where $\alpha,\beta = 1,2,\ldots,n(s)$. Here $n(s)$ is a positive integer depending on the spin. Even though we cannot prove the preceding statement in general it is possible to verify it for all the known fields. Common properties of the various fields are discussed next. In the following two sections we develop the tools to be used later. In Section III we introduce the "first identity" which is essentially the configuration space version of the generalized Ward identity. Conservation laws are studied making use of the first identity but without reference to any explicit Lagrangian. The Klein-Gordon divisor and the second identity are introduced and discussed in Section IV. A discussion of how to construct the Klein-Gordon divisor is also given. Making use of these tools a unified treatment of the normalization and closure conditions for the c-number solutions of the general wave equation (for arbitrary spin) is presented in Section V.

Quantization is carried through in Section VI using the procedure already indicated above. In this connection the raising and lowering operators are introduced and the connection between spin and statistics is also discussed briefly. Simple examples are given and the uniqueness of P_{μ} and $M_{\mu\nu}$ is proved.

In Appendix A we have set down the notation and conventions used in this course. Appendix B discusses how one can construct explicit wave functions for arbitrary spin fields.

II. Wave Equations for Relativistic Free Fields

In this section we shall review the wave equations for various spin fields. In our discussion we shall often digress to point out correspondences between different spin fields and features common to various fields in order to provide clues as well as motivation for evolving a unified approach to arbitrary spin fields.

We shall see that for each known case the field equation and the associated subsidiary conditions, if any, can be cast into the form

$$\Lambda_{\alpha\beta}(\partial)\phi_\beta(x) = 0, \tag{II.1}$$

where α, β run from 1 to n(s), and n(s) is a positive integer which depends on the spin of the field. We shall also sort out other properties which the various fields share in common.

For a fuller discussion of the various wave equations than presented here we refer the reader to the books by Corson[1] and Umezawa.[2]

A. The Klein-Gordon Field.

The Klein-Gordon (to be abbreviated as K-G hereafter) equation is written

$$(\Box - m^2)\phi(x) = 0, \tag{II.2a}$$

where m is a real constant. We write the K-G equation for the adjoint field $\phi^\dagger(x)$ as

$$\phi^\dagger(x)(\overleftarrow{\Box} - m^2) \equiv (\Box - m^2)\phi^\dagger(x) = 0. \tag{II.2b}$$

The four-current defined by

$$j_\mu(x) = i\phi^\dagger(x)\left(\partial_\mu - \overleftarrow{\partial}_\mu\right)\phi(x) \tag{II.3}$$

is conserved, since

$$\begin{aligned}
\partial_\mu j_\mu &= i\phi^\dagger\left(\partial_\mu + \overleftarrow{\partial}_\mu\right)\left(\partial_\mu - \overleftarrow{\partial}_\mu\right)\phi \\
&= i\phi^\dagger\left(\Box - \overleftarrow{\Box}\right)\phi \\
&\equiv i\phi^\dagger\left[\left(\Box - m^2\right) - \left(\overleftarrow{\Box} - m^2\right)\right]\phi \\
&= 0,
\end{aligned} \tag{II.4}$$

by virtue of (II.2).

Let us present here some of the standard formalism in order to familiarize the reader with our notation. We introduce two auxiliary fields $\phi^{(+)}$ and $\phi^{(-)}$ via

$$\phi^{(\pm)}(x) = \frac{1}{2\pi i}\int_{C_+} \frac{d\tau}{\tau}\,\phi(x \mp n\tau), \tag{II.5}$$

where τ is a parameter and n is a unit time-like vector with positive time component, i.e.,

$$n_\mu n_\mu = -1, \qquad n_o > 1. \tag{II.6}$$

The contour C_+ runs along the real axis in the τ-plane from $\tau = -\infty$ to $\tau = +\infty$ and is indented below the origin to avoid the singularity. We can show that $\phi^{(+)}(x)$ contains positive frequencies only while $\phi^{(-)}$ contains negative frequencies only: Since $\phi(x)$ is a solution of the K-G equation, we can write

$$\phi(x) = \int d^4k\, e^{ikx}\, \delta(k^2+m^2)\, \hat{\phi}(k). \tag{II.7}$$

Consider $\phi^{(+)}(x)$ first. If we substitute (II.7) in the definition (II.5) for $\phi^{(+)}(x)$, we get

$$\phi^{(+)}(x) = \int d^4k\, \delta(k^2+m^2)\, \hat{\phi}(k) e^{ikx} \frac{1}{2\pi i} \int_{C_+} \frac{d\tau}{\tau}\, e^{-i(kn)\tau}$$

$$= \int d^4k\, \delta(k^2+m^2)\, \theta(-kn) e^{ikx}\, \hat{\phi}(k), \tag{II.8}$$

where we have used the well-known formula

$$\frac{1}{2\pi i} \int_{C_+} \frac{d\tau}{\tau}\, e^{ix\tau} = \theta(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases} \tag{II.9}$$

Now we can write

$$-kn = k_o n_o \left(1 - \frac{\underline{k}\cdot\underline{n}}{k_o n_o}\right),$$

and for time-like k, the factor

$$\left(1 - \frac{\underline{k}\cdot\underline{n}}{k_o n_o}\right)$$

is positive in virtue of (II.6), so that

$$\theta(-kn)\delta(k^2+m^2) = \theta(k_o n_o)\delta(k^2+m^2) = \theta(k_o)\delta(k^2+m^2). \tag{II.10}$$

Using (II.10) in (II.8) we get

$$\phi^{(+)}(x) = \int d^4k\, e^{ikx}\, \theta(k_o)\, \delta(k^2+m^2)\, \hat{\phi}(k). \tag{II.11}$$

It is now clear that $\phi^{(+)}(x)$ contains positive frequencies only. In a similar fashion one can show

$$\phi^{(-)}(x) = \int d^4k\, e^{ikx}\, \theta(-k_o)\, \delta(k^2+m^2)\, \hat{\phi}(k). \tag{II.12}$$

It may be remarked that the separation into positive and negative frequencies is consistent with relativistic invariance. From (II.7), (II.11) and (II.12) it follows that

$$\phi(x) = \phi^{(+)}(x) + \phi^{(-)}(x). \tag{II.13}$$

The familiar invariant singular function $\Delta(x)$ is defined by

$$\Delta(x) = \frac{-i}{(2\pi)^3} \int d^4k\, e^{ikx}\, \epsilon(k_o)\, \delta(k^2+m^2). \tag{II.14}$$

It satisfies the K-G equation,

$$(\Box - m^2)\Delta(x) = 0, \tag{II.15a}$$

is Lorentz invariant under all proper Lorentz transformations and has the properties:

$$\Delta(x) = 0 \quad \text{for} \quad x^2 > 0 \quad \text{(i.e., space-like x)}, \tag{II.15b}$$

$$\Delta(-x) = -\Delta(x), \tag{II.15c}$$

and

$$\delta(x_o)\, \frac{\partial}{\partial x_o}\, \Delta(x) = -\delta^4(x). \tag{II.15d}$$

We can separate $\Delta(x)$ into positive and negative frequency parts by writing

$$\Delta(x) = \Delta^{(+)}(x) + \Delta^{(-)}(x), \tag{II.16a}$$

where

$$\Delta^{(+)}(x) = \frac{-i}{(2\pi)^3} \int d^4k\, e^{ikx}\, \theta(k_o)\, \delta(k^2+m^2), \tag{II.16b}$$

$$\Delta^{(-)}(x) = \frac{+i}{(2\pi)^3}\int d^4k\, e^{ikx}\,\theta(-k_o)\,\delta(k^2+m^2). \qquad \text{(II.16c)}$$

Now we can prove

$$\phi(x) = \int_\sigma d\sigma_\lambda(x')\,\Delta(x-x')\left(\partial_\lambda' - \overleftarrow{\partial}_\lambda'\right)\phi(x'), \qquad \text{(II.17)}$$

where σ is an arbitrary space-like surface. It is readily seen that the right-hand side of (II.17) is independent of the particular choice of σ, since if we take its functional derivative at the point x' and make use of (A.2) of the Appendix A, we get

$$\begin{aligned}
&\frac{\delta}{\delta\sigma(x')}\int d\sigma_\lambda(x')\,\Delta(x-x')\left(\partial_\lambda' - \overleftarrow{\partial}_\lambda'\right)\phi(x') \\
&\quad = \partial_\lambda'\left[\Delta(x-x')\left(\partial_\lambda' - \overleftarrow{\partial}_\lambda'\right)\phi(x')\right] \\
&\quad = \Delta(x-x')\left(\partial_\lambda' + \overleftarrow{\partial}_\lambda'\right)\left(\partial_\lambda' - \overleftarrow{\partial}_\lambda'\right)\phi(x') \\
&\quad = \Delta(x-x')\left(\Box' - \overleftarrow{\Box}'\right)\phi(x') \\
&\quad = \Delta(x-x')\left[\left(\Box' - m^2\right) - \left(\overleftarrow{\Box}' - m^2\right)\right]\phi(x') \\
&\quad = 0, \qquad \text{(II.18)}
\end{aligned}$$

in virtue of (II.2a) and (II.15a). So we choose the space-like surface $x_o' = \text{const.} = x_o$, then

$$d\sigma_\lambda(x') = (0,0,0,-id^3x').$$

We then have

$$\int_\sigma d\sigma_\lambda(x')\Delta(x-x')\left(\partial'_\lambda - \overleftarrow{\partial}'_\lambda\right)\phi(x')$$

$$= \int_{x'_0=x_0} d\sigma_4(x')\Delta(x-x')\left(\partial'_4 - \overleftarrow{\partial}'_4\right)\phi(x')$$

$$= -\int_{x'_0=x_0} d^3x'\Delta(x-x')\left(\frac{\partial}{\partial x'_0} - \frac{\overleftarrow{\partial}}{\partial x'_0}\right)\phi(x')$$

$$= -\int d^3x'\left[\Delta(x-x')\right]_{x'_0=x_0}\left[\frac{\partial}{\partial x'_0}\phi(x')\right]_{x'_0=x_0}$$

$$+ \int d^3x'\left[\frac{\partial}{\partial x'_0}\Delta(x-x')\right]_{x'_0=x_0}\phi(\underset{\sim}{x}', x_4)$$

$$= 0 + \int d^3x'\left[+\delta^3(\underset{\sim}{x}-\underset{\sim}{x}')\right]\phi(\underset{\sim}{x}', x_4)$$

$$= \phi(x, x_4) \equiv \phi(x),$$

where we have made use of (II.15b,c and d).

Next we want to study the transformation property of the K-G field under infinitesimal Lorentz transformations. Recall that under an infinitesimal Lorentz transformation

$$x_\mu \to x'_\mu = a_{\mu\nu}x_\nu, \tag{II.19a}$$

where $a_{\mu\nu}$ has the form

$$a_{\mu\nu} = \delta_{\mu\nu} + \epsilon_{\mu\nu}, \tag{II.19b}$$

and it satisfies

$$a_{\mu\lambda}a_{\nu\lambda} = a_{\lambda\mu}a_{\lambda\nu} = \delta_{\mu\nu}, \tag{II.19c}$$

hence $\epsilon_{\mu\nu}$ is antisymmetric,

$$\epsilon_{\mu\nu} + \epsilon_{\nu\mu} = 0, \tag{II.19d}$$

in virtue of (II.19c) as we neglect bilinear and higher terms in $\epsilon_{\mu\nu}$.

Note that the inverse of (II.19a) is

$$x_\nu = x'_\mu a_{\mu\nu}. \tag{II.19e}$$

Now

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial x_\nu}{\partial x'_\mu}\frac{\partial}{\partial x_\nu}$$

$$= a_{\mu\nu}\frac{\partial}{\partial x_\nu}, \tag{II.20a}$$

so that

$$\Box' = \frac{\partial}{\partial x'_\mu}\frac{\partial}{\partial x'_\mu} = a_{\mu\nu}\frac{\partial}{\partial x_\nu}\, a_{\mu\lambda}\frac{\partial}{\partial x_\lambda}$$

$$= \left(a_{\mu\nu}a_{\mu\lambda}\right)\frac{\partial}{\partial x_\nu}\frac{\partial}{\partial x_\lambda}$$

$$= \delta_{\nu\lambda}\frac{\partial}{\partial x_\lambda}\frac{\partial}{\partial x_\lambda} = \Box. \tag{II.20b}$$

In view of (II.20b), invariance of the field equation under infinitesimal Lorentz transformation follows from the fact that the field $\phi(x)$ transforms according to

$$\phi(x) \to \phi'(x') = \phi(x). \tag{II.21}$$

B. The Dirac Field.

The Dirac field $\psi(x)$ satisfies the field equation

$$[\gamma\partial + m]\psi(x) = 0. \tag{II.22a}$$

In our notation the γ-matrices are all hermitian and they satisfy the commutation relation

$$\{\gamma_\mu, \gamma_\nu\} \equiv \gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\delta_{\mu\nu}. \tag{II.23}$$

The adjoint field $\bar{\psi}(x)$ is defined as

$$\bar{\psi}(x) = \psi^\dagger(x)\gamma_4, \tag{II.24}$$

and its field equation is

$$\bar{\psi}(x)\left[-\gamma\overleftarrow{\partial} + m\right] = 0. \qquad \text{(II.22b)}$$

The four-current defined by

$$j_\mu(x) = i\bar{\psi}(x)\gamma_\mu\psi(x) \qquad \text{(II.25)}$$

is conserved, since

$$i\partial_\mu(\bar{\psi}\gamma_\mu\psi) = i\bar{\psi}(\gamma\partial + \gamma\overleftarrow{\partial})\psi$$
$$= i\bar{\psi}(-m + m)\psi = 0,$$

using (II.22).

If we write

$$\Lambda(\partial) = -(\gamma\partial + m), \qquad \text{(II.26a)}$$

then (II.22a) does indeed have the form (II.1). We want our fields to satisfy the K-G equation irrespective of their spins. This is certainly true for $\psi(x)$: If we define

$$d(\partial) = -(\gamma\partial - m), \qquad \text{(II.26b)}$$

then we have

$$d(\partial)\Lambda(\partial) = \Lambda(\partial)d(\partial) = \Box - m^2, \qquad \text{(II.27)}$$

so that $\psi(x)$ will automatically satisfy the K-G equation if it satisfies (II.22a). As we shall see later, it is possible to find $\Lambda(\partial)$ and $d(\partial)$ satisfying (II.1) and (II.27) for most known cases. This is obviously true for the K-G field.

Recall that for the Dirac case the charge conjugate field

$$\psi_c(x) = C\psi^*(x) \qquad \text{(II.28)}$$

also satisfies (II.22a). The charge conjugation matrix C has the property

$$C^{-1}\gamma_\mu C = g_{\mu\nu}\gamma_\nu^t, \qquad \text{(II.29)}$$

or equivalently,

$$C^{-1}\gamma_i C = \gamma_i^t$$
$$C^{-1}\gamma_4 C = -\gamma_4^t. \qquad \text{(II.29')}$$

One can then show that

$$\left[\gamma_4(\gamma\partial + m)\right]^t = -C^{-1}\gamma_4(\gamma\partial + m)C. \tag{II.30}$$

Later we shall introduce charge conjugation matrix for arbitrary spin fields which satisfies a condition which is a generalization of (II.30).

Next, by defining in the usual way

$$S(x) = (\gamma\partial - m)\Delta(x), \tag{II.31}$$

we can prove

$$\psi(x) = \int_\sigma d\sigma_\lambda(x')S(x-x')\gamma_\lambda\psi(x') \tag{II.32}$$

in the same fashion as we proved (II.17). Comparing (II.32) with (II.17) for the Dirac case, we note the correspondence:

$$\left\{\begin{array}{ll} \left(\partial'_\lambda - \overleftarrow{\partial}'_\lambda\right) & \text{is now replaced by } \gamma_\lambda, \\ \Delta(x-x') & \text{is replaced by } S(x-x'). \end{array}\right\} \tag{II.33}$$

We shall see that this correspondence leads to a natural generalization for arbitrary spin fields.

Under an infinitesimal Lorentz transformation the Dirac field transforms according to

$$\psi_\alpha(x) \to \psi'_\alpha(x') = L_{\alpha\beta}\psi_\alpha(x), \tag{II.34a}$$

where L is a 4×4 matrix whose most general form may be written as

$$L_{\alpha\beta} = \left(1 + \frac{i}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)_{\alpha\beta}, \tag{II.34b}$$

since we neglect bilinear and higher terms in $\epsilon_{\mu\nu}$. Clearly $S_{\mu\nu}$ is an anti-symmetric tensor whose each element is a 4×4 matrix. On actually doing the algebra, we find

$$S_{\mu\nu} = \frac{1}{4i}\left(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu\right). \tag{II.34c}$$

Let us give a brief outline of the derivation of (II.34c). Upon transformation the field equation (II.22a) becomes (in matrix notation)

$$(\gamma\partial' + m)\psi'(x') = 0.$$

We write this as

$$\left(\gamma_\mu \frac{\partial}{\partial x'_\mu} + m\right) L\psi(x) = 0.$$

Using

$$\frac{\partial}{\partial x'_\mu} = \frac{\partial x_\nu}{\partial x'_\mu} \frac{\partial}{\partial x_\nu}$$

$$= a_{\mu\nu} \frac{\partial}{\partial x_\nu},$$

and multiplying from the left by L^{-1}, we get

$$\left(L^{-1}\gamma_\mu L a_{\mu\nu} \frac{\partial}{\partial x_\nu} + m\right)\psi(x) = 0.$$

If the field equation is invariant under infinitesimal Lorentz transformations, then we must have

$$L^{-1}\gamma_\mu L a_{\mu\nu} = \gamma_\nu. \tag{II.35}$$

If we multiply both sides of (II.35) by $a_{\lambda\nu}$ and use (II.19c), we get

$$L^{-1}\gamma_\lambda L = a_{\lambda\nu}\gamma_\nu. \tag{II.36}$$

Upon substituting the explicit forms (II.34b) and (II.19b) for L and $a_{\lambda\nu}$, respectively, we find that (II.36) reduces to the following condition for $S_{\mu\nu}$:

$$-i\left[S_{\mu\nu}, \gamma_\lambda\right] = \delta_{\mu\lambda}\gamma_\nu - \delta_{\nu\lambda}\gamma_\mu. \tag{II.37}$$

The explicit form (II.34c) for $S_{\mu\nu}$ can then easily be derived from (II.37).

If we write

$$S_1 = S_{23}, \quad S_2 = S_{31}, \quad S_3 = S_{12}, \tag{II.38}$$

then from (II.34c) and (II.23) we have

$$S_1 = \frac{1}{2i}\gamma_2\gamma_3, \quad S_2 = \frac{1}{2i}\gamma_3\gamma_1, \quad S_3 = \frac{1}{2i}\gamma_1\gamma_2, \tag{II.39}$$

so that

$$S_1^2 + S_2^2 + S_3^2 = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} = \frac{1}{2}\left(\frac{1}{2} + 1\right). \qquad \text{(II.40)}$$

The factor $\frac{1}{2}(\frac{1}{2}+1)$ is of the familiar form $\ell(\ell+1)$ for the spin-vector squared from quantum mechanics.

It should be clear that the general way of writing the transformation property of an arbitrary spin field $\phi_\alpha(x)$, $\alpha = 1,2,\ldots n(s)$, under an infinitesimal Lorentz transformation, is the same as in (II.34a) and (II.34b) where α,β now run from 1 to n(s). The $S_{\mu\nu}$'s, which are now $n(s) \times n(s)$ matrices, can be found in each case—at least in principle. With the definitions (II.38), one can compute $(S_1^2 + S_2^2 + S_3^2)$ and expect to obtain $\ell(\ell+1)$. This seems to be a reasonable way for constructing higher spin fields.

As a parenthetical remark, we note here that $S_{\mu\nu} = 0$ for the K-G field as one would expect. (See (II.21).)

C. The Vector Field.

Consider the vector field $U_\sigma(x)$, $\sigma = 1,\ldots,4$. Each component of this field satisfies the K-G equation

$$(\Box - m^2)U_\sigma(x) = 0 \qquad (\sigma = 1,2,3,4). \qquad \text{(II.41)}$$

Since $U_\sigma(x)$ is a four-vector, under infinitesimal Lorentz transformations it must transform in the same way as x_μ, i.e.,

$$U_\sigma(x) \rightarrow U'_\sigma(x') = a_{\sigma\rho}U_\rho(x)$$

$$= \left(\delta_{\sigma\rho} + \epsilon_{\sigma\rho}\right)U_\rho(x). \qquad \text{(II.42a)}$$

On the other hand, according to (II.34a,b) we should be able to write it as

$$U_\sigma(x) \rightarrow U'_\sigma(x') = \left[\delta_{\sigma\rho} + \frac{i}{2}\epsilon_{\mu\nu}(S_{\mu\nu})_{\sigma\rho}\right]U_\rho(x). \qquad \text{(II.42b)}$$

Comparing these two forms, we find at once

$$(S_{\mu\nu})_{\sigma\rho} = -i\left[\delta_{\mu\sigma}\delta_{\nu\rho} - \delta_{\mu\rho}\delta_{\nu\sigma}\right]. \qquad \text{(II.43)}$$

According to (II.38), we then have for the vector field:

$$S_1 = S_{23} = i\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$S_2 = S_{31} = i\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$S_3 = S_{12} = i\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

So one gets

$$S_1^2+S_2^2+S_3^2 = \begin{bmatrix} 2 & & & \\ & 2 & & 0 \\ & & 2 & \\ & 0 & & 0 \end{bmatrix},$$

where the factors 2 come from $1(1+1)$, while 0 comes from $0(0+1)$, i.e., we have spin 1 and spin 0 associated with $U_\sigma(x)$. We can get rid of spin 0 by imposing the auxiliary condition $\partial_\sigma U_\sigma(x) = 0$. Then the equations

$$\begin{aligned} (\Box - m^2)U_\sigma(x) &= 0, \\ \partial_\sigma U_\sigma(x) &= 0, \end{aligned} \tag{II.44}$$

describe a pure spin 1 field. The two equations above can be combined as

$$\left[(\Box - m^2)\delta_{\sigma\rho} - \partial_\sigma\partial_\rho\right]U_\rho(x) = 0. \tag{II.45}$$

The spin 1 field $U_\sigma(x)$ satisfying (II.44) or (II.45) is called the Proca field.

One easily shows that (II.45) is equivalent to (II.44): If we apply ∂_σ to (II.45), we get

$$(\Box - m^2)\partial_\rho U_\rho - \Box\partial_\rho U_\rho = 0,$$

or,

$$-m^2\partial_\rho U_\rho = 0.$$

If $m \neq 0$, we obtain $\partial_\rho U_\rho = 0$ which is the second of equations (II.44). Substituting this in (II.45) we obtain the other equation in (II.44). We note that (II.45) is of the form (II.1). By studying the invariance of (II.45) under infinitesimal Lorentz transformations we can confirm (II.42) and (II.43).

To summarize, we have learned that a spinor index carries spin $\frac{1}{2}$ while a tensor index carries spin 1 and spin 0. Now we are ready to construct fields of higher spin.

D. The Bargmann-Wigner Fields.

Suppose we want to construct a field with spin s. Consider the field $\psi_{\alpha_1\alpha_2\ldots\alpha_{2s}}(x)$ which has 2s spinor indices. In general, this will contain spin s and lower spins. This is so because we know from quantum mechanics that vector addition of spin $\frac{1}{2}$ with spin $\frac{1}{2}$ yields spin 1 and spin 0:

$$\mathfrak{D}_{\frac{1}{2}} \otimes \mathfrak{D}_{\frac{1}{2}} = \mathfrak{D}_0 \oplus \mathfrak{D}_1.$$

If we add another spin $\frac{1}{2}$, we get

$$\mathfrak{D}_0 \otimes \mathfrak{D}_{\frac{1}{2}} = \mathfrak{D}_{\frac{1}{2}},$$

$$\mathfrak{D}_1 \otimes \mathfrak{D}_{\frac{1}{2}} = \mathfrak{D}_{\frac{1}{2}} \oplus \mathfrak{D}_{\frac{3}{2}}.$$

Thus it is clear that $\psi_{\alpha_1\alpha_2\ldots\alpha_{2s}}(x)$ will have spin s and lower spins. If we construct this field in such a way that it is totally symmetric in all its spinor indices, then this field will carry spin s only since total symmetry with respect to the spinor indices means that all the 2s spin $\frac{1}{2}$ vectors add in parallel. Also the Dirac equation with respect to any spinor index must be satisfied. In other words, the field $\psi_{\alpha_1\alpha_2\ldots\alpha_{2s}}(x)$ satisfying

$$\left.\begin{aligned} &(\gamma\partial + m)_{\alpha_1\alpha_1'}\psi_{\alpha_1'\alpha_2\ldots\alpha_{2s}}(x) = 0; \\ &\alpha_i = 1,2,3,4; \\ &\psi_{\alpha_1\alpha_2\ldots\alpha_{2s}}(x): \text{ totally symmetric} \\ &\text{in } \alpha_1,\alpha_2,\ldots\alpha_{2s} \end{aligned}\right\}, \qquad \text{(II.46)}$$

carries spin s. The field equations (II.46) are called the Bargmann-Wigner equations.

This way we can construct any spin field, integer or half-integer.

The above construction of the Bargmann-Wigner field equations implies that the Bargmann-Wigner field transforms according to

$$\psi_{\alpha_1\ldots\alpha_{2s}}(x) \to \psi'_{\alpha_1\ldots\alpha_{2s}}(x') = \left(1+\frac{i}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)_{\alpha_1\ldots\alpha_{2s};\beta_1\ldots\beta_{2s}}\psi_{\beta_1\ldots\beta_{2s}}(x), \tag{II.47a}$$

with

$$\begin{aligned}(S_{\mu\nu})_{\alpha_1\ldots\alpha_{2s};\beta_1\ldots\beta_{2s}} &= \left(S_{\mu\nu}^{(\frac{1}{2})}\right)_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\cdots\delta_{\alpha_{2s}\beta_{2s}} \\ &+ \delta_{\alpha_1\beta_1}\left(S_{\mu\nu}^{(\frac{1}{2})}\right)_{\alpha_2\beta_2}\delta_{\alpha_3\beta_3}\cdots\delta_{\alpha_{2s}\beta_{2s}} + \cdots \\ &+ \delta_{\alpha_1\beta_1}\cdots\delta_{\alpha_{2s-1}\beta_{2s-1}}\left(S_{\mu\nu}^{(\frac{1}{2})}\right)_{\alpha_{2s}\beta_{2s}},\end{aligned} \tag{II.47b}$$

where $S_{\mu\nu}^{(\frac{1}{2})}$ is the matrix appropriate to the Dirac field:

$$S_{\mu\nu}^{(\frac{1}{2})} = \frac{1}{4i}\left(\gamma_\mu\gamma_\nu - \gamma_\nu\gamma_\mu\right). \tag{II.47c}$$

Explicit form of the Lagrangian density for the Bargmann-Wigner fields is known for $s \leq 3/2$ but not for higher spins. For details we refer the reader to the paper by Kamefuchi and Takahashi.[3]

E. The Pauli-Fierz Fields.

If we want to construct a higher spin field of integral spin s, then we can do so by using tensor indices exclusively; i.e., we consider $U_{\mu_1\ldots\mu_s}(x)$. However, since a tensor index carries spin 1 as well as spin 0, in general an object like $U_{\mu_1\ldots\mu_s}(x)$ will have spin s and lower spins associated with it. One must impose supplementary conditions to eliminate all other spins except the spin s. The field $U_{\mu_1\ldots\mu_s}(x)$ satisfying

$$
\begin{aligned}
&(\Box - m^2) U_{\mu_1 \ldots \mu_s}(x) = 0, \\
&\qquad U_{\mu_1 \ldots \mu_s}(x) \ : \text{totally symmetric in indices}, \\
&U_{\mu_1 \mu_1 \mu_3 \ldots \mu_s}(x) = 0 \qquad\qquad (II.48) \\
&\partial_{\mu_1} U_{\mu_1 \mu_2 \ldots \mu_s}(x) = 0, \\
&\qquad \mu_i = 1, \ldots, 4,
\end{aligned}
$$

is called the Pauli-Fierz field and it is a pure spin s field. Let us check this for $s=2$. The appropriate field then is $U_{\mu\nu}(x)$ which, according to (II.48), satisfies

$$
\begin{aligned}
&(\Box - m^2) U_{\mu\nu}(x) = 0, \\
&\qquad U_{\mu\nu}(x) = U_{\nu\mu}(x), && \text{(6 conditions)} \\
&\qquad U_{\mu\mu}(x) = 0, && \text{(1 condition)} \\
&\qquad \partial_\mu U_{\mu\nu}(x) = 0. && \text{(4 conditions).}
\end{aligned}
$$

These eleven conditions on $U_{\mu\nu}(x)$ leave only five out of its sixteen components independent—which is exactly $(2s+1)$ independent components corresponding to $s=2$. To see this in more detail, recall that each tensor index carries spin 1 and spin 0. So we get:

$$
\begin{array}{lll}
(\mu) & (\nu) & \\
\mathfrak{D}_0 \otimes \mathfrak{D}_0 = \mathfrak{D}_0 & & \\
\mathfrak{D}_0 \otimes \mathfrak{D}_1 = \mathfrak{D}_1 & & \text{These are eliminated by} \\
\mathfrak{D}_1 \otimes \mathfrak{D}_0 = \mathfrak{D}_1 & & \text{supplementary conditions.} \\
\mathfrak{D}_1 \otimes \mathfrak{D}_1 = \mathfrak{D}_0 \oplus \mathfrak{D}_1 & & \\
\qquad\qquad \oplus \mathfrak{D}_2 & & \\
\qquad\qquad \longrightarrow & & \text{This is what we get.}
\end{array}
$$

Under infinitesimal Lorentz transformations the Pauli-Fierz fields transform according to

$$U_{\sigma_1 \ldots \sigma_s}(x) \to U'_{\sigma_1 \ldots \sigma_s}(x') = \left(1 + \frac{i}{2} \epsilon_{\mu\nu} S_{\mu\nu}\right)_{\sigma_1 \ldots \sigma_s ; \rho_1 \ldots \rho_s} U_{\rho_1 \ldots \rho_s}(x), \tag{II.49a}$$

with

$$(S_{\mu\nu})_{\sigma_1 \ldots \sigma_s ; \rho_1 \ldots \rho_s} = \left(S^{(1)}_{\mu\nu}\right)_{\sigma_1 \rho_1} \delta_{\sigma_2 \rho_2} \ldots \delta_{\sigma_s \rho_s}$$

$$+ \delta_{\sigma_1 \rho_1} \left(S^{(1)}_{\mu\nu}\right)_{\sigma_2 \rho_2} \delta_{\sigma_3 \rho_3} \ldots \delta_{\sigma_s \rho_s} + \ldots +$$

$$+ \delta_{\sigma_1 \rho_1} \ldots \delta_{\sigma_{s-1} \rho_{s-1}} \left(S^{(1)}_{\mu\nu}\right)_{\sigma_s \rho_s} . \tag{II.49b}$$

where the matrix $S^{(1)}_{\mu\nu}$ is the one we discovered for the Proca field, namely,

$$\left(S^{(1)}_{\mu\nu}\right)_{\sigma\rho} = -i\left[\delta_{\mu\sigma}\delta_{\nu\rho} - \delta_{\mu\rho}\delta_{\nu\sigma}\right]. \tag{II.49c}$$

Explicit form of the Lagrangian density for the Pauli-Fierz fields is known only for $s \leq 2$. For details, see the article by Bhargava and Watanabe.[4]

F. The Rarita-Schwinger Fields (cf. II.H).

Now we discuss a method for constructing half-integral spin fields due to Rarita and Schwinger. The idea is simple: one uses $(s-\frac{1}{2})$ tensor indices (here s is a half-integer) and one spinor index. One needs to symmetrize with respect to the tensor indices and impose supplementary conditions to get rid of the unwanted lower spins, of course. The field must satisfy the Dirac equation with respect to the spinor index. For the sake of concreteness, consider $s = 3/2$ first. The appropriate Rarita-Schwinger field then is $\psi_{\alpha\mu}(x)$, where α is a spinor index and μ is a tensor index. We take this field to satisfy

$$(\gamma\partial + m)_{\alpha\beta} \psi_{\beta\mu}(x) = 0,$$

$$\alpha, \beta, \mu = 1, \ldots, 4;$$

$$\partial_\mu \psi_{\alpha\mu}(x) = 0, \qquad \text{(II.50)}$$

$$(\gamma_\mu)_{\alpha\beta}\psi_{\beta\mu}(x) = 0,$$

which are the Rarita-Schwinger equations for spin 3/2 field.

It is quite straightforward to generalize (II.50) to half-integral spins greater than 3/2. The field $\psi_{\alpha\mu_1\mu_2\ldots\mu_{s-\frac{1}{2}}}(x)$ satisfying the Rarita-Schwinger equations

$$\begin{aligned}
&(\gamma\partial + m)_{\alpha\beta}\psi_{\beta\mu_1\ldots\mu_{s-\frac{1}{2}}}(x) = 0,\\
&\qquad \alpha,\beta,\mu_i = 1,\ldots,4;\\
&\qquad \psi_{\alpha\mu_1\ldots\mu_{s-\frac{1}{2}}}(x) : \text{totally symmetric in } \mu_1\ldots\mu_{s-\frac{1}{2}},\\
&(\gamma_{\mu_1})_{\alpha\beta}\psi_{\beta\mu_1\mu_2\ldots\mu_{s-\frac{1}{2}}}(x) = 0\\
&\partial_{\mu_1}\psi_{\alpha\mu_1\mu_2\ldots\mu_{s-\frac{1}{2}}}(x) = 0,\\
&\psi_{\alpha\mu_1\mu_1\mu_3\ldots\mu_{s-\frac{1}{2}}}(x) = 0
\end{aligned} \qquad \text{(II.51)}$$

carries spin s.

Under infinitesimal Lorentz transformations the Rarita-Schwinger fields transform according to

$$\psi_{\alpha\sigma_1\ldots\sigma_{s-\frac{1}{2}}}(x) \to \psi'_{\alpha\sigma_1\ldots\sigma_{s-\frac{1}{2}}}(x')$$

$$= \left(1 + \frac{1}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)_{\alpha\sigma_1\ldots\sigma_{s-\frac{1}{2}};\beta\rho_1\ldots\rho_{s-\frac{1}{2}}} \psi_{\beta\rho_1\ldots\rho_{s-\frac{1}{2}}}(x), \qquad \text{(II.52a)}$$

with

$$(S_{\mu\nu})_{\alpha\sigma_1\ldots\sigma_{s-\frac{1}{2}};\beta\rho_1\ldots\rho_{s-\frac{1}{2}}} = \left(S_{\mu\nu}^{(\frac{1}{2})}\right)_{\alpha\beta}\delta_{\sigma_1\rho_1}\cdots\delta_{\sigma_{s-\frac{1}{2}}\rho_{s-\frac{1}{2}}}$$

$$+ \delta_{\alpha\beta}\left(S_{\mu\nu}^{(1)}\right)_{\sigma_1\rho_1}\delta_{\sigma_2\rho_2}\cdots\delta_{\sigma_{s-\frac{1}{2}}\rho_{s-\frac{1}{2}}} + \cdots$$

$$+ \delta_{\alpha\beta}\delta_{\sigma_1\rho_1}\cdots\left(S_{\mu\nu}^{(1)}\right)_{\sigma_{s-\frac{1}{2}}\rho_{s-\frac{1}{2}}}, \qquad \text{(II.52b)}$$

where $S_{\mu\nu}^{(\frac{1}{2})}$ and $S_{\mu\nu}^{(1)}$ are the same as in (II.47c) and (II.49c).

Explicit form of the Lagrangian density for the Rarita-Schwinger fields is known for $s \leq 5/2$. For details see the paper by Kawakami and Kamefuchi.[5)]

G. The Duffin-Kemmer and Harish-Chandra Fields.

By the Duffin-Kemmer field we mean a 16-component field $\psi_\xi(x)$ satisfying the equation

$$\left[\beta_\mu\partial_\mu + m\right]_{\xi\eta}\psi_\eta(x) = 0, \qquad \text{(II.53a)}$$

where β_μ's ($\mu = 1,\ldots,4$) are 16×16 matrices satisfying

$$\beta_\mu\beta_\lambda\beta_\nu + \beta_\nu\beta_\lambda\beta_\mu = \delta_{\mu\lambda}\beta_\nu + \delta_{\nu\lambda}\beta_\mu . \qquad \text{(II.53b)}$$

If we write

$$Q \equiv \beta_\mu\partial_\mu, \qquad \text{(II.54a)}$$

by applying $\partial_\mu\partial_\nu\partial_\lambda$ to (II.53b) we find

$$Q^3 = \Box Q. \qquad \text{(II.54b)}$$

Suppressing the indices ξ, η, we can rewrite (II.53a) as

$$(Q + m)\psi(x) = 0. \qquad \text{(II.53'a)}$$

If we multiply this equation from the left by Q^2 and make use of (II.54b) we get

$$(\Box Q + mQ^2)\psi(x) = 0.$$

In view of (II.53'a) we can replace Q by -m in the last equation to get

$$-m(\Box - m^2)\psi(x) = 0,$$

or

$$(\Box - m^2)\psi(x) = 0,$$

if $m \neq 0$. We have shown that each component of the Duffin-Kemmer field does indeed satisfy the K-G equation—a necessary requirement for any relativistic free field.

By studying the algebra of the β-matrices, i.e., Eq. (II.53b), one can show that the β-matrices are reducible into the form

$$\beta_\mu = \left(\begin{array}{|c|c|c|} \hline 1 \times 1 & 0 & 0 \\ \hline 0 & 5 \times 5 & 0 \\ \hline 0 & 0 & 10 \times 10 \\ \hline \end{array} \right) \qquad \text{(II.55)}$$

The 1×1 representation is a trivial one since we get $\beta_\mu = 0$, which leads to $\psi = 0$. The 5×5 representation corresponds to the spin 0 case. Historically, Kemmer arrived at the equation (II.53) with 5×5 β-matrices in his attempt to find a first-order matrix differential equation for the scalar field. The five-component field $\psi(x)$ in this case really has only one independent component, namely the K-G field $\varphi(x)$; the other four are derivatives of $\varphi(x)$: $\partial_1\varphi, \ldots, \partial_4\varphi$. Lastly, the 10×10 representation corresponds to the spin 1 case. In other words, if one tries to find a first-order matrix-differential equation for the Proca field, one arrives at (II.53) with 10×10 β-matrices. For details, see, for example, the book by Umezawa.[2)]

Under infinitesimal Lorentz transformations the Duffin-Kemmer field transforms according to

$$\psi_\xi(x) \to \psi'_\xi(x') = \left(1 + \frac{i}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)_{\xi\eta}\psi_\eta(x), \qquad \text{(II.56a)}$$

where

$$(S_{\mu\nu})_{\xi\eta} = i\left(\beta_\mu\beta_\nu + \beta_\nu\beta_\mu\right)_{\xi\eta}, \qquad \text{(II.56b)}$$

$$\begin{pmatrix} \mu, \nu = 1, \ldots, 4. \\ \xi, \eta = 1, \ldots, 16. \end{pmatrix}$$

Harish-Chandra has generalized the Duffin-Kemmer equations (II.53) to the field with maximum spin s as follows:

$$\left[\beta_\lambda \partial_\lambda + m\right]_{\xi\eta} \psi_\eta(x) = 0, \qquad \text{(II.57a)}$$

where β-matrices satisfy the relation

$$\sum_{(p)} \beta_{\mu_1}\beta_{\mu_2}\ldots\beta_{\mu_{2s-1}}\left[\beta_{\mu_{2s}}\beta_{\mu_{2s+1}} - \delta_{\mu_{2s}\,\mu_{2s+1}}\right] = 0. \qquad \text{(II.57b)}$$

Here λ and μ_i run from 1 to 4, while ξ and η run from 1 to n, where n is a positive integer which depends upon the spin. In (II.57b) sum over all permutations of the indices μ_i is to be taken.

If we write $Q = \beta\partial$, we find from (II.57b) the relation

$$Q^{2s+1} = Q^{2s-1}\Box.$$

Then we can prove easily that each component of ψ_η in (II.57) satisfies the K-G equation, just as in the Duffin-Kemmer case.

The equations (II.57) reduce to the Duffin-Kemmer equations if we set $s=1$, $n=16$. (We actually end up with Proca as well as K-G fields.) The Dirac field equation is obtained by setting $s=\frac{1}{2}$, $n=4$. In principle, all higher spin fields can be obtained.

The form of the matrix $(S_{\mu\nu})$ which specifies the transformation property of the Harish-Chandra field under infinitesimal Lorentz transformations is known for spins up to 3/2 only. For $s=0$ and 1, it is given by (II.56b); for $s=\frac{1}{2}$ it is given by (II.47c), while for $s=3/2$ it is known but is very complicated. The form (II.56b) does not work for higher spins.

The Harish-Chandra equation follows from the simple Lagrangian density

$$\mathcal{L}(x) = -\bar{\psi}(\beta\partial + m)\psi.$$

However, things have not been worked out explicitly for $s > 3/2$. For further details of spin 3/2 field we refer the reader to the papers of K. K. Gupta[6] and S. N. Gupta.[7]

As a final remark, Bhabha has suggested writing the field equation for arbitrary spin in the form (II.57a) and further demanding the matrix $(S_{\mu\nu})$ to have the form (II.56b). For $s > 1$ one gets the result that the individual components of $\psi(x)$ do not satisfy the K-G equation. One has to then resort to a multiple mass formulation. However, we shall not pursue this any further here.

H. Some Remarks on Higher Spin Field Formulations and the Quantization Problem.

We have seen that the Klein-Gordon, Dirac, Proca and Harish-Chandra field equations are easily written in the form (II.1). We would like to be able to do the same for the Bargmann-Wigner, Pauli-Fierz and Rarita-Schwinger equations. However, we cannot prove this in general. Therefore, we shall be content here to show that we can arrive at the form (II.1) for a specific case. For simplicity, let us take the Rarita-Schwinger field for spin 3/2. Suppressing the spinor indices, the appropriate equations are

$$(\gamma\partial + m)\psi_\mu(x) = 0, \tag{II.58a}$$

$$\gamma_\mu\psi_\mu(x) = 0, \tag{II.58b}$$

$$\partial_\mu\psi_\mu(x) = 0. \tag{II.58c}$$

We remark here that, in fact, only two of the above equations are independent. Using the first two we can derive the third. To prove this, we multiply the second equation by $(\gamma\partial)$ and then use the commutation relations for the γ-matrices and the first two equations. In detail:

$$\begin{aligned} 0 &= (\gamma\partial)\gamma_\mu\psi_\mu, \\ &= -\gamma_\mu(\gamma\partial)\psi_\mu + 2\partial_\mu\psi_\mu, \\ &= +m\gamma_\mu\psi_\mu + 2\partial_\mu\psi_\mu, \qquad \text{(using (II.58a))} \\ &= 2\partial_\mu\psi_\mu, \qquad \text{(using (II.58b))} \end{aligned}$$

which proves our assertion.

Next, it can be easily shown that the equations (II.58) can be put into the form

$$\Lambda_{\sigma\rho}(\partial)\psi_\rho(x) = 0, \tag{II.59a}$$

where

$$\Lambda_{\sigma\rho}(\partial) = -(\gamma\partial + m)\delta_{\sigma\rho} + \frac{1}{3}\left[\gamma_\sigma\partial_\rho + \gamma_\rho\partial_\sigma\right] - \frac{1}{3}\left[\gamma_\sigma(\gamma\partial - m)\gamma_\rho\right]. \tag{II.59b}$$

Lest the reader be misled into believing that in (II.59a) we already

have the desired form (II.1), let us rewrite it with the spinor indices restored:

$$\left[\Lambda_{\sigma\rho}(\partial)\right]_{\alpha\beta}\psi_{\beta\rho}(x) = 0. \qquad \text{(II.59'a)}$$

We indicate briefly how one shows that (II.59) is equivalent to (II.58). If we apply ∂_σ to (II.59a) we obtain

$$\left[2(\gamma\partial)\partial_\rho - m(\gamma\partial)\gamma_\rho + 3m\partial_\rho\right]\psi_\rho(x) = 0,$$

while multiplying (II.59a) by γ_σ gives

$$\left[-2\partial_\rho + m\gamma_\rho\right]\psi_\rho(x) = 0.$$

If we assume $m \neq 0$, these two equations imply

$$\gamma_\rho\psi_\rho = 0, \qquad \partial_\rho\psi_\rho = 0,$$

which are just (II.58b) and (II.58c). Substituting these in (II.59a) we obtain (II.58a). Therefore, (II.59) and (II.58) are indeed equivalent.

(It seems that the $m \neq 0$ condition plays a rather important role always. This is the reason that we have to exclude the Maxwell field from our discussion.)

In order to bring the Rarita-Schwinger equations into the Harish-Chandra form, define:

$$\Gamma_{\mu,\sigma\rho} = -\gamma_\mu\delta_{\sigma\rho} + \frac{1}{3}\left(\gamma_\rho\delta_{\sigma\mu} + \gamma_\sigma\delta_{\rho\mu}\right) - \frac{1}{3}\gamma_\sigma\gamma_\mu\gamma_\rho, \qquad \text{(II.60a)}$$

$$\beta_{\sigma\rho} = -\left(\delta_{\sigma\rho} - \tfrac{1}{3}\gamma_\sigma\gamma_\rho\right). \qquad \text{(II.60b)}$$

With these definitions, (II.59) may be rewritten as

$$\left(\Gamma_{\mu,\sigma\rho}\partial_\mu + m\beta_{\sigma\rho}\right)\psi_\rho(x) = 0. \qquad \text{(II.60c)}$$

One can find a set of matrices $\beta^{-1}_{\sigma\rho}$ having the property:

$$\beta^{-1}_{\tau\sigma}\beta_{\sigma\rho} = \delta_{\tau\rho}.$$

If we multiply (II.60c) by $\beta^{-1}_{\tau\sigma}$ and define

$$\beta^{-1}_{\tau\sigma}\Gamma_{\mu,\sigma\rho} = \beta_{\mu,\tau\rho},$$

the equation (II.60c) becomes

$$\left(\beta_{\mu,\tau\rho}\partial_\mu + m\delta_{\tau\rho}\right)_{\alpha\beta}\psi_{\beta\rho}(x) = 0, \qquad \text{(II.61)}$$

where we have restored the hitherto suppressed spinor indices. Now we combine the two indices β,ρ into a single index $\eta=(\beta\rho)=1,\ldots,16$; similarly, we write $\xi=(\alpha\tau)=1,\ldots,16$. This enables us to write (II.61) in the Harish-Chandra form

$$\left(\beta_\mu\partial_\mu + m\right)_{\xi\eta}\psi_\eta(x) = 0.$$

It can be shown that the β's arrived at this way turn out to have the correct properties appropriate to spin 3/2 Harish-Chandra equation. So we have succeeded in reducing the Rarita-Schwinger equation for spin 3/2 into the form (II.1).

We have studied various higher spin field formulations. One may ask if the various formulations for a given spin are equivalent. It would seem that the answer is yes, though it is not possible to prove it in general. Anyway, we have shown the equivalence between the Rarita-Schwinger and the Harish-Chandra equations for spin 3/2 above. One can also construct the Bargmann-Wigner field $\psi_{\alpha\beta\gamma}(x)$ for spin 3/2 from the Rarita-Schwinger field:

$$\psi_{\alpha\beta\gamma} = -\frac{1}{m}\left[(\gamma\partial - m)\gamma_\mu\gamma_4 C\right]_{\alpha\beta}\psi_{\gamma\mu}(x), \qquad \text{(II.62)}$$

where C is the charge conjugation matrix:

$$C^{-1}\gamma_\mu C = g_{\mu\nu}\gamma_\nu^t.$$

The field $\psi_{\alpha\beta\gamma}(x)$ as constructed above can be shown to satisfy the Bargmann-Wigner equation and to possess the correct symmetry property with respect to α,β,γ.[3]

A few words on the quantization problem are in order now. The usual method for quantization is based on the Hamiltonian formalism. Since our aim is to evolve a unified approach to quantization, the Hamiltonian formalism is not very helpful for the following reasons:

(1) Straightforward application of the Hamiltonian formalism works only for the Klein-Gordon field.

As an example, consider the Duffin-Kemmer equation:

$$(\beta\partial + m)\psi(x) = 0, \qquad \text{(II.53a)}$$

$$\beta_\mu\beta_\lambda\beta_\nu + \beta_\nu\beta_\lambda\beta_\mu = \delta_{\mu\lambda}\beta_\nu + \delta_{\nu\lambda}\beta_\mu. \qquad \text{(II.53b)}$$

Setting $\mu=\lambda=\nu=4$ in (II.53b) we get

$$\beta_4^3 = \beta_4,$$

which implies the condition

$$\beta_4\left(\beta_4^2-1\right) = 0. \tag{II.63a}$$

Next, if we set $\lambda=\nu=4$, $\mu=i$, we get

$$\beta_i\beta_4^2 + \beta_4^2\beta_i = \beta_i$$

which can be written as

$$\left\{\beta_i, 2\beta_4^2 - 1\right\} = 0. \tag{II.63b}$$

The adjoint field $\bar{\psi}(x)$ is defined by

$$\bar{\psi}(x) = \psi^\dagger(x)\eta, \tag{II.64a}$$

where

$$\eta = 2\beta_4^2 - 1. \tag{II.64b}$$

The Lagrangian density is written as

$$\mathcal{L} = -\bar{\psi}(x)(\beta\partial + m)\psi(x). \tag{II.65}$$

Then the canonically conjugate field $\pi(x)$ is

$$\pi(x) = \frac{\partial\mathcal{L}}{\partial\dot{\psi}(x)} = i\bar{\psi}(x)\beta_4,$$

so that we get the canonical commutation relation

$$\left[\psi(x), i\bar{\psi}(y)\beta_4\right]_{x_0=y_0} \equiv \left[\psi(x), \pi(y)\right]_{x_0=y_0} = i\delta^3(x-y).$$

If we multiply this by (β_4^2-1) and use (II.63a), we get

$$0 = \left(\beta_4^2-1\right)\delta^3(x-y),$$

which is a contradiction! (Note that $\beta_4^2 \neq 1$ in view of (II.63a).)

The straightforward application of the Hamiltonian formalism failed in this case as not all the components of $\psi(x)$ are canonically independent. So one must separate the canonically independent components before applying the Hamiltonian formalism. Schwinger[8] has done this. However, the procedure is quite complicated and what we finally obtain is:

(2) Equal time commutators only.

From the relativistic invariance viewpoint it is desirable to be able to write the commutators for arbitrary times.

We do not wish to imply that the Hamiltonian approach to quantization is no good. However, one must be careful in its application which often turns out to be quite tedious. Besides that, it yields only equal time commutation relations and does not seem to lend itself easily to treating fields of arbitrary spin in a unified manner. Because of these reasons we try to find a quantization procedure without using the canonical formalism. We have already discussed the basis of our approach in the Introduction. We proceed to discuss the common properties of the fields for which our approach works.

I. Common Properties of Field Equations.

We have already made it plausible that the free field equations and the supplementary conditions, if any, can be brought into the form

$$\Lambda_{\alpha\beta}(\partial)\phi_{\beta}(x) = 0, \tag{II.1}$$

where α, β run from 1 to $n(S)$ which is a positive integer depending on the spin associated with the field. From now on we shall assume that the form (II.1) is true for fields of interest to us—we shall point out any exceptions as we go along. In matrix notation (II.1) is simply written as

$$\Lambda(\partial)\phi(x) = 0. \tag{II.1'}$$

Further, we assume $\Lambda(\partial)$ to satisfy the following conditions:

(A) There exists a non-singular matrix η such that

$$[\eta\Lambda(\partial)]^{\dagger} = \eta\Lambda(-\partial).$$

Making use of this assumption we can construct the Lagrangian

$$\mathfrak{L} = -\overline{\phi}(x)\Lambda(\partial)\phi(x),$$

with

$$\bar{\phi}(x) = \phi^{\dagger}(x)\eta .$$

This Lagrangian leads to the field equation

$$\eta \Lambda(\partial)\psi(x) = 0 ,$$

which is the same as

$$\Lambda(\partial)\psi(x) = 0 ,$$

since η is non-singular. Of course, other Lagrangians may be constructed which lead to the same equation of motion. However, this Lagrangian and the condition (A) imply the desirable property

$$\mathfrak{L} = \mathfrak{L}^{\dagger} + 4\text{-divergence}.$$

Recall that most of the familiar field equations do have these properties. For example, for the Dirac field we have

$$\Lambda(\partial) = -(\gamma\partial + m)$$

$$\eta = \gamma_4 ,$$

$$\mathfrak{L} = -\bar{\psi}(\gamma\partial + m)\psi ,$$

$$\mathfrak{L} = \mathfrak{L}^{\dagger} - \partial_{\mu}(\bar{\psi}\gamma_{\mu}\psi).$$

The assumption that η exists is equivalent to the assumption that a Lagrangian exists.

(B) $\Lambda(\partial)$ is of the form

$$\Lambda_{\alpha\beta}(\partial) = (\Lambda_0)_{\alpha\beta} + (\Lambda_{\mu})_{\alpha\beta}\partial_{\mu} + \left(\Lambda_{\mu_1\mu_2}\right)_{\alpha\beta}\partial_{\mu_1}\partial_{\mu_2} + \cdots ,$$

where

$$\left(\Lambda_{\mu_1 \ldots \mu_{\ell}}\right)_{\alpha\beta}$$

is symmetric in all pairs of indices $\mu_1 \ldots \mu_{\ell}$, and independent of x. We assume there is only a finite number of terms on the right hand side.

Obviously this form is true for all cases that we know.

(C) From $\Lambda(\partial)\phi(x) = 0$, we should be able to arrive at the Klein-Gordon equation, $(\Box - m^2)\phi(x) = 0$, by a finite number of operations. We can state this condition in an alternate form:

The Klein-Gordon divisor $d(\partial)$ exists such that

$$\Lambda(\partial)d(\partial) = d(\partial)\Lambda(\partial) = \Box - m^2.$$

The K-G divisor satisfying the above condition must be derivable by a finite number of operations.

For the Dirac case we have

$$\Lambda(\partial) = -(\gamma\partial + m),$$
$$d(\partial) = -(\gamma\partial - m).$$

For the Duffin-Kemmer equation:

$$\Lambda(\partial) = -(\beta\partial + m),$$

$$d(\partial) = -\left[\frac{1}{m}(\Box - m^2) + \beta\partial - \frac{1}{m}(\beta\partial)^2\right].$$

The Maxwell field does not satisfy this condition.

(D) $[\Lambda(\partial), \ell_{\mu\nu}] = 0$, where

$$(\ell_{\mu\nu})_{\alpha\beta} = \left(x_\mu\partial_\nu - x_\nu\partial_\mu\right)\delta_{\alpha\beta} + i(S_{\mu\nu})_{\alpha\beta}.$$

This is just a statement of the Lorentz invariance requirement.

Proof:

The field equation $\Lambda(\partial)\phi(x) = 0$, under infinitesimal Lorentz transformation, becomes

$$\Lambda(\partial')\phi'(x') = 0,$$

which is the same as

$$\Lambda(\partial)\phi'(x) = 0.$$

Here

$$\phi'(x') = \left(1 + \frac{i}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)\phi(x).$$

With $x' = x + \delta x$, and using Taylor's expansion

$$\phi'(x) = \left(1 + \frac{i}{2}\epsilon_{\mu\nu}S_{\mu\nu}\right)\phi(x - \delta x)$$

$$= \left[1 + \frac{1}{2}\left(x_\mu\partial_\nu - x_\nu\partial_\mu + iS_{\mu\nu}\right)\epsilon_{\mu\nu}\right]\phi(x),$$

where we have neglected terms of second order and higher in $\epsilon_{\mu\nu}$. In terms of $\ell_{\mu\nu}$:

$$\phi'(x) = \left(1 + \frac{1}{2}\epsilon_{\mu\nu}\ell_{\mu\nu}\right)\phi(x).$$

Going back to the transformed field equation

$$0 = \left(1 - \frac{1}{2}\ell_{\mu\nu}\epsilon_{\mu\nu}\right)\Lambda(\partial)\left(1 + \frac{1}{2}\ell_{\mu\nu}\epsilon_{\mu\nu}\right)\phi(x)$$

$$= \Lambda(\partial)\phi(x) + \frac{1}{2}\epsilon_{\mu\nu}\left[\Lambda(\partial), \ell_{\mu\nu}\right]\phi(x)$$

$$= \frac{1}{2}\epsilon_{\mu\nu}\left[\Lambda(\partial), \ell_{\mu\nu}\right]\phi(x),$$

so that

$$\left[\Lambda(\partial), \ell_{\mu\nu}\right] = 0,$$

as asserted.

(E) There exists a unitary, symmetric matrix C such that

$$[\eta\Lambda(\partial)]^t = \rho C^{-1}\eta\Lambda(-\partial)C,$$

where

$$\rho = \begin{cases} +1 & \text{for fields with integer spin,} \\ -1 & \text{for fields with half-integer spin.} \end{cases}$$

In the Dirac case C is simply the charge conjugation matrix. The significance of this condition will become apparent later. The two-component neutrino does not satisfy this condition and is therefore excluded from our discussion.

If any of these conditions (A)-(E) is not satisfied by a particular field, then our quantization procedure is not directly applicable to that field.

III. Conservation Laws and the First Identity

In this section we shall discuss the derivation of conservation laws directly from the field equations without reference to any Lagrangian.[9] As we have emphasized earlier, we do have to assume the existence of a Lagrangian from which the field equations are derivable. This is necessary for the conservation laws to arise from the invariance properties of the field equations. The explicit form of the Lagrangian need not be known and, indeed, we never use the

Lagrangian. In our approach the assumption of the existence of a Lagrangian is expressed by the condition (A) of the last section which says that a non-singular matrix η must exist which has the property

$$[\eta\Lambda(\partial)]^{\dagger} = \eta\Lambda(-\partial). \tag{III.1}$$

First, we shall present a configuration space version of the generalized Ward identity which is a relation between the propagator and the vertex. Then we shall make use of this identity in deriving the conservation laws.

Recall that $\Lambda(\partial)$ has the form [condition (B)]:

$$\Lambda(\partial) = \Lambda_0 + \Lambda_\mu\partial_\mu + \Lambda_{\mu\nu}\partial_\mu\partial_\nu + \cdots, \tag{III.2a}$$

then

$$\Lambda(-\overleftarrow{\partial}) = \Lambda_0 - \Lambda_\mu\overleftarrow{\partial}_\mu + \Lambda_{\mu\nu}\overleftarrow{\partial}_\mu\overleftarrow{\partial}_\nu - \cdots. \tag{III.2b}$$

If we define

$$\Gamma_\mu(\partial, -\overleftarrow{\partial}) = \Lambda_\mu + \Lambda_{\mu\nu}(\partial_\nu - \overleftarrow{\partial}_\nu) + \Lambda_{\mu\nu\lambda}\left(\partial_\nu\partial_\lambda - \partial_\nu\overleftarrow{\partial}_\lambda + \overleftarrow{\partial}_\nu\overleftarrow{\partial}_\lambda\right) + \cdots, \tag{III.3}$$

then from (III.2) and (III.3) it easily follows that

$$\left(\partial_\mu + \overleftarrow{\partial}_\mu\right)\Gamma_\mu(\partial, -\overleftarrow{\partial}) = \Lambda(\partial) - \Lambda(-\overleftarrow{\partial}). \tag{III.4}$$

We shall call (III.4) the "first identity" and it will play an important role subsequently. For the Dirac field $\Gamma_\mu = -\gamma_\mu$ and the momentum space form of (III.4) becomes

$$i(p-q)_\mu\gamma_\mu = S_F^{-1}(p) - S_F^{-1}(q),$$

which is the generalized Ward identity. So our first identity is the configuration space version of the generalized Ward identity for arbitrary spin case. This identity basically follows from condition (B).

We list some properties of $\Gamma_\mu(\partial, -\overleftarrow{\partial})$:

$$\Gamma_\mu(\partial, -\overleftarrow{\partial}) = \Gamma_\mu(-\overleftarrow{\partial}, \partial), \tag{III.5}$$

$$\left[\eta\Gamma_\mu(\partial, -\overleftarrow{\partial})\right]^{\dagger} = -g_{\mu\nu}\eta\Gamma_\nu(-\partial, \overleftarrow{\partial}), \tag{III.6}$$

$$C^{-1}\eta\Gamma_\mu(\partial, -\overleftarrow{\partial})C = \rho\left[\eta\Gamma_\mu(-\partial, \overleftarrow{\partial})\right]^t. \tag{III.7}$$

The relation (III.5) follows directly from (III.4); (III.6) follows from the condition (A) while (III.7) follows from the condition (E).

Recall that our free field $\phi(x)$ satisfies the field equation

$$\Lambda(\partial)\phi(x) = 0, \tag{III.8a}$$

while the adjoint field $\bar{\phi}(x) = \phi^\dagger(x)\eta$ satisfies

$$\bar{\phi}(x)\Lambda(-\overleftarrow{\partial}) = 0. \tag{III.8b}$$

Suppose that $F[x]$ and $G[x]$ are functionals of $\phi, \bar{\phi}$ and their derivatives. Then

$$\partial_\mu[G\Gamma_\mu F] \equiv G(\partial_\mu + \overleftarrow{\partial}_\mu)\Gamma_\mu F$$

$$= G\Lambda(\partial)F - G\Lambda(-\overleftarrow{\partial})F.$$

If $G=\bar{\phi}$, $F=\phi$, then because of (III.8)

$$\partial_\mu[\bar{\phi}\Gamma_\mu\phi] = 0,$$

which says that $\bar{\phi}\Gamma_\mu\phi$ is a conserved current. We shall consider two broad types of situations in which we get conservation laws:

(I) If $G[x]\Lambda(\partial)F[x] - G[x]\Lambda(-\overleftarrow{\partial})F[x] = 0$, then $J_\mu(x) = G[x]\Gamma_\mu(\partial, -\overleftarrow{\partial})F[x]$ is the conserved current.

(II) If $G[x]\Lambda(\partial)F[x] - G[x]\Lambda(-\overleftarrow{\partial})F[x] = \partial_\mu K_\mu[x]$, then $J_\mu(x) = G[x]\Gamma_\mu(\partial, -\overleftarrow{\partial})F[x] - K_\mu[x]$ is the conserved current.

The conserved quantity is written in covariant form as

$$Q = \int d\sigma_\mu(x)\, J_\mu(x), \tag{III.9}$$

where σ is a space-like surface. Clearly

$$\frac{\delta}{\delta\sigma} Q = \partial_\mu J_\mu = 0.$$

Then choose $d\sigma_\lambda = (0,0,0,-id^3x)$ to get

$$Q = -i\int d^3x\, J_4(x). \tag{III.9'}$$

Now we come to the question as to how one constructs F, G, J_μ, etc. Suppose that the field equation

$$\Lambda(\partial)\phi(x) = 0$$

is invariant under some transformation

$$\phi(x) \to \phi'(x'),$$
$$x \to x',$$

so that

$$\Lambda(\partial')\phi'(x') = 0,$$

which implies

$$\Lambda(\partial)\phi'(x) = 0.$$

Since $\phi(x)$ and $\phi'(x)$ both satisfy the free field equation, so does the quantity $\delta\phi(x) = \phi'(x) - \phi(x)$. If we take $G = \bar{\phi}$, $F = \delta\phi$, we have the conserved current $J_\mu = \bar{\phi}\Gamma_\mu\delta\phi$. This construction is, of course, not unique. There are many other possibilities.

Now we shall consider several examples of conservation laws. First we take the situations of type (I) where

$$G\Lambda(\partial)F - G\Lambda(-\overleftarrow{\partial})F = 0. \qquad \text{(III.10)}$$

There are two possibilities within this type. The first one is the case when

$$\Lambda(\partial)F[x] = 0 \qquad \text{and} \qquad G[x]\Lambda(-\overleftarrow{\partial}) = 0, \qquad \text{(III.11)}$$

i.e., the functionals F and G satisfy the free field equations. The second one is the case when F and G do not satisfy the free field equations but the weaker condition (III.10) holds. Of the examples that follow, the first five satisfy the stronger condition (III.11) while the sixth example satisfies the weaker condition (III.10). The seventh and the last example is of the type (II) where

$$G\Lambda(\partial)F - G\Lambda(-\overleftarrow{\partial})F = \partial_\mu K_\mu. \qquad \text{(III.12)}$$

Example (1): Phase transformation.

Suppose the field equation is invariant under infinitesimal phase transformation

$$\phi(x) \to \phi'(x') = (1 - i\alpha)\phi(x),$$

$$\bar{\phi}(x) \to \bar{\phi}'(x') = (1 + i\alpha)\bar{\phi}(x).$$

If we choose

$$G \sim \delta\bar{\phi} = i\alpha\bar{\phi}, \qquad F \sim \delta\phi = -i\alpha\phi,$$

then both (III.10) and (III.11) are satisfied. We get the conserved current

$$J_\mu(x) = -i\bar{\phi}(x)\Gamma_\mu(\partial, -\overleftarrow{\partial})\phi(x).$$

Example (2): Space-time translation.

Suppose the field equation is invariant under infinitesimal coordinate translation,

$$x_\mu \to x'_\mu = x_\mu + \varepsilon_\mu,$$

$$\phi(x) \to \phi'(x') = \phi(x),$$

$$\bar{\phi}(x) \to \bar{\phi}'(x') = \phi(x),$$

$$\begin{aligned} \delta\phi(x) &= \phi'(x) - \phi(x) \\ &= \phi(x-\varepsilon) - \phi(x) \\ &= -\varepsilon_\mu \partial_\mu \phi(x), \end{aligned}$$

$$\delta\bar{\phi}(x) = -\varepsilon_\mu \partial_\mu \bar{\phi}(x) \equiv -\varepsilon_\mu \bar{\phi}(x)\overleftarrow{\partial}_\mu.$$

If we take the combination proportional to

$$\bar{\phi}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\delta\phi(x) - \delta\bar{\phi}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\phi(x),$$

which satisfies (III.10) and (III.11), and write it as

$$t_{\lambda\mu} = -\frac{\varepsilon}{2}\bar{\phi}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})(\partial_\mu - \overleftarrow{\partial}_\mu)\phi(x),$$

where

$$\varepsilon = \begin{cases} 1 & \text{for non-Hermitian fields } (\phi \neq \bar{\phi}) \\ \frac{1}{2} & \text{for Hermitian fields } (\phi = \bar{\phi}) \end{cases},$$

we can show that $t_{\lambda\mu}$ as defined above is Hermitian and is the usual energy-momentum tensor to within a divergence term. The continuity equation

$$\partial_\lambda t_{\lambda\mu}(x) = 0$$

expresses the conservation of energy-momentum vector

$$P_\mu = \int d\sigma_\lambda(x) t_{\lambda\mu}(x)$$

$$= -\frac{\epsilon}{2}\int d\sigma_\lambda(x)\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})(\partial_\mu - \overleftarrow{\partial}_\mu)\phi(x).$$

The factor $-\epsilon/2$ can be understood only after quantization has been carried through.

Example (3): Infinitesimal Lorentz transformation.

The invariance of the field equation under infinitesimal Lorentz transformation gives

$$\phi(x) \to \phi'(x') = \left(1 + \frac{i}{2} S_{\mu\nu}\epsilon_{\mu\nu}\right)\phi(x),$$

$$\bar{\phi}(x) \to \bar{\phi}'(x') = \bar{\phi}(x)\left(1 - \frac{i}{2} S_{\mu\nu}\epsilon_{\mu\nu}\right),$$

if $\bar{\phi}\phi$ is a scalar. One finds

$$\delta\phi(x) = \phi'(x) - \phi(x)$$

$$= \tfrac{1}{2}\epsilon_{\mu\nu}\ell_{\mu\nu}\phi(x),$$

$$\delta\bar{\phi}(x) = \bar{\phi}'(x) - \bar{\phi}(x)$$

$$= -\tfrac{1}{2}\epsilon_{\mu\nu}\bar{\phi}(x)\overleftarrow{\ell}_{\mu\nu},$$

where

$$\ell_{\mu\nu} = x_\mu\partial_\nu - x_\nu\partial_\mu + iS_{\mu\nu},$$

$$\overleftarrow{\ell}_{\mu\nu} = -x_\mu\overleftarrow{\partial}_\nu + x_\nu\overleftarrow{\partial}_\mu + iS_{\mu\nu}.$$

From these results one can write down the conserved angular momentum density as

$$M_{\lambda\mu\nu}(x) = -\tfrac{1}{2}\epsilon\,\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})\ell_{\mu\nu}\phi(x)$$

$$-\tfrac{1}{2}\epsilon\,\bar{\phi}(x)\overleftarrow{\ell}_{\mu\nu}\Gamma_\lambda(\partial,-\overleftarrow{\partial})\phi(x),$$

$$\partial_\lambda M_{\lambda\mu\nu}(x) = 0.$$

$$M_{\mu\nu} = \int d\sigma_\lambda(x) M_{\lambda\mu\nu}(x) \quad \text{is conserved.}$$

The factor ϵ is the same as in the previous example.

Example (4): Frequency-dependent phase transformation.

The positive and negative frequency parts of a field satisfy the same wave equation:

$$\Lambda(\partial)\phi^{(\pm)}(x) = 0,$$

$$\overline{\phi^{(\pm)}}(x)\Lambda(-\overleftarrow{\partial}) = 0.$$

Invariance under frequency-dependent infinitesimal phase transformation implies

$$\phi^{(\pm)}(x) \to \phi^{(\pm)'}(x) = (1 \mp i\alpha)\phi^{(\pm)}(x),$$

or

$$\delta\phi^{(\pm)}(x) = \mp i\alpha\phi^{(\pm)}(x). \qquad (\alpha : \text{real}).$$

So we can take, for example,

$$F \sim \mp i\alpha\phi^{(\pm)}, \quad G \sim \overline{\phi^{(\pm)}}.$$

The currents

$$j_\mu^{(\pm)}(x) = \mp i\overline{\phi^{(\pm)}}(x)\Gamma_\mu(\partial, -\overleftarrow{\partial})\phi^{(\pm)}(x)$$

are conserved, that is, $\partial_\mu j_\mu^{(\pm)} = 0$. $\int d\sigma_\mu(x) j_\mu^{(\pm)}(x)$ are conserved quantities.

Example (5): Space reflection.

If the field equation is invariant under space reflection,

$$x_\mu \to x'_\mu = -x^*_\mu,$$

then the field transforms according to

$$\phi(x) \to \phi^P(x') = P\phi(x),$$

where P is some appropriate matrix. We have

$$0 = \Lambda(\partial')\phi^P(x')$$

$$= \Lambda(\partial)\phi^P(x).$$

Let us take $F=\phi^P$, $G=\bar{\phi}$. Then the relevant "conserved current" is

$$j_\mu(x) = \bar{\phi}(x)\Gamma_\mu(\partial, -\overleftarrow{\partial})\phi^P(x)$$

$$= \bar{\phi}(x)\Gamma_\mu(\partial, -\overleftarrow{\partial})P\phi(-x^*).$$

Then the quantity

$$S = \int d\sigma_\mu(x) j_\mu(x)$$

is conserved.

Recall that in the Dirac case

$$\Gamma_\mu = -\gamma_\mu,$$

$$P = i\gamma_4,$$

$$\psi^P(x') = i\gamma_4\psi(x),$$

so that in this case

$$S = -i\int d\sigma_\mu(x)\bar{\psi}(x)\gamma_\mu\gamma_4\psi(-x^*).$$

If we define

$$\mathcal{P} = e^{i(\pi/2)S},$$

we can show that for the quantized theory

$$\mathcal{P}\psi(x)\mathcal{P}^{-1} = i\gamma_4\psi(-x^*) = \psi^P(x).$$

It is a peculiar feature of our approach that we have a conserved current for a discrete symmetry such as space reflection. We can also write a conserved current associated with charge conjugation. On the other hand, we can also find many other "conserved currents" by our method which have no physical significance.

The following two examples concern conservation laws for interacting fields:

Example (6): Dirac field interacting with an electromagnetic field.
The field equations are

$$-\Lambda(\partial)\psi = (\gamma\partial + m)\psi = ie\gamma_\mu A_\mu \psi ,$$

$$-\bar{\psi}\Lambda(-\overleftarrow{\partial}) = \bar{\psi}(-\gamma\overleftarrow{\partial} + m) = ie\bar{\psi}\gamma_\mu A_\mu .$$

If we choose $F=\psi$ and $G=\bar{\psi}$, then (III.11) obviously does not hold. However, the weaker condition (III.10) does hold:

$$-G\Lambda(\partial)F + G\Lambda(-\overleftarrow{\partial})F = \bar{\psi}(\gamma\partial + m)\psi - \bar{\psi}(-\gamma\overleftarrow{\partial} + m)\psi$$

$$= ie\bar{\psi}\gamma_\mu\psi A_\mu - ie\bar{\psi}\gamma_\mu\psi A_\mu = 0 .$$

So the conserved current is

$$j_\mu = -ieG\Gamma_\mu(\partial, -\overleftarrow{\partial})F = ie\bar{\psi}\gamma_\mu\psi ,$$

as we know quite well.

Example (7): Charged scalar field interacting with an electromagnetic field.
The field equations are:

$$(\Box - m^2)\phi = 2ieA_\mu\partial_\mu\phi + e^2A_\mu A_\mu\phi ,$$

$$\phi^\dagger(\overleftarrow{\Box} - m^2) = -2ie\phi^\dagger\overleftarrow{\partial}_\mu A_\mu + e^2\phi^\dagger A_\mu A_\mu ,$$

$$\partial_\mu A_\mu = 0 .$$

Here

$$\Lambda = (\Box - m^2),$$

$$\Gamma_\mu = \partial_\mu - \overleftarrow{\partial}_\mu .$$

If we choose $G=\phi^\dagger$ and $F=\phi$, then both (III.11) and (III.10) are not satisfied but (III.12) is:

$$G\Lambda(\partial)F - G\Lambda(-\overleftarrow{\partial})F = \phi^\dagger(\Box - m^2)\phi - \phi^\dagger(\overleftarrow{\Box} - m^2)\phi$$

$$= 2ie\phi^\dagger(\partial_\mu + \overleftarrow{\partial}_\mu)\phi \cdot A_\mu = \partial_\mu[2ie\phi^\dagger\phi A_\mu],$$

so that

$$K_\mu = 2ie\phi^\dagger A_\mu \phi .$$

Therefore, if we write

$$j_\mu = ie\left[G\Gamma_\mu(\partial, -\overleftarrow{\partial})F - K_\mu\right]$$

$$= ie\left[\phi^\dagger(\partial_\mu - \overleftarrow{\partial}_\mu)\phi - 2ie\phi^\dagger A_\mu\phi\right],$$

we have the well-known conserved current for this case.

Since the interaction of a matter field with the electromagnetic field is introduced by the replacements

$$\partial_\mu \longrightarrow \partial_\mu - ieA_\mu$$

$$\overleftarrow{\partial}_\mu \longrightarrow \overleftarrow{\partial}_\mu + ieA_\mu$$

in the free field equations, the first identity can be generalized to the case of matter fields interacting with the electromagnetic field:

$$(\partial_\mu + \overleftarrow{\partial}_\mu)\Gamma_\mu(\partial - ieA, -\overleftarrow{\partial} - ieA) = \Lambda(\partial - ieA) - \Lambda(-\overleftarrow{\partial} - ieA).$$

This is a further generalization of the generalized Ward identity in the configuration space.

IV. The Klein-Gordon Divisor and the Second Identity

The Klein-Gordon divisor, $d(\partial)$, is defined by

$$\Lambda(\partial)d(\partial) = d(\partial)\Lambda(\partial) = \Box - m^2. \qquad \text{(IV.1)}$$

We shall call the relation (IV.1) the second identity.

From the conditions (A) and (E) [see Section II] on $\Lambda(\partial)$ and the second identity, we obtain

$$\left[d(\partial)\eta^{-1}\right]^\dagger = d(-\partial)\eta^{-1}, \qquad \text{(IV.2)}$$

$$\left[d(\partial)\eta^{-1}\right]^t = \rho C^{-1} d(-\partial)\eta^{-1} C. \tag{IV.3}$$

In our formalism, when considering particular cases, knowledge of the explicit form of $d(\partial)$ is required. So we would like to discuss methods of obtaining $d(\partial)$ when $\Lambda(\partial)$ is known:

The most naïve way is to assume

$$d(\partial) = d_o + d_\mu \partial_\mu + d_{\mu\nu}\partial_\mu\partial_\nu + \cdots, \tag{IV.4}$$

and substitute this in the second identity and determine $d_o, d_\mu, d_{\mu\nu}, \ldots$ by equating coefficients on both sides for each particular derivative term. As an example, consider the Dirac case:

$$\Lambda(\partial) = -(\gamma\partial + m), \tag{IV.5}$$

so we write $d(\partial)$ as

$$d(\partial) = a + b\gamma\partial + c\Box + \cdots .$$

Then

$$(\Box - m^2) = \Lambda(\partial)d(\partial)$$

$$= -am - (a+mb)\gamma\partial - (b+cm)\Box + \cdots$$

so that

$$a = m, \quad b = -1, \quad c = 0, \quad d = 0 \cdots,$$

and we get

$$d(\partial) = -(\gamma\partial - m). \tag{IV.6}$$

As another example, consider the Proca field (spin 1):

$$\Lambda_{\sigma\rho}(\partial) = (\Box - m^2)\delta_{\sigma\rho} - \partial_\sigma\partial_\rho, \tag{IV.7}$$

and we write

$$d_{\rho\lambda}(\partial) = a\delta_{\rho\lambda} + b\delta_{\rho\lambda}\Box + c\partial_\rho\partial_\lambda + \cdots .$$

Then

$$(\Box - m^2)\delta_{\sigma\lambda} = \Lambda_{\sigma\rho}(\partial)d_{\rho\lambda}(\partial)$$

$$= -m^2 a\delta_{\sigma\lambda} + a(\Box\delta_{\sigma\lambda} - \partial_\sigma\partial_\lambda)$$

$$-m^2(b\delta_{\sigma\lambda}\Box + c\partial_\sigma\partial_\lambda) + \cdots ,$$

on equating coefficients

$$-m^2 a = -m^2 \longrightarrow a = 1$$

$$a - bm^2 = 1 \longrightarrow b = 0$$

$$-a - cm^2 = 0 \longrightarrow c = -\frac{1}{m^2} .$$

We get

$$d_{\rho\lambda}(\partial) = \delta_{\rho\lambda} - \frac{1}{m^2}\partial_\rho\partial_\lambda . \tag{IV.8}$$

The naïve method discussed above, even though of general applicability, can get very tedious in many cases. On the other hand, it is possible to invent tricks for individual cases to make the derivation of $d(\partial)$ easier. We consider a few typical cases.

First consider the Harish-Chandra case for which

$$\Lambda(\partial) = -(\beta_\lambda\partial_\lambda + m) = -(Q_1 + m),$$
$$Q_1 = \beta_\lambda\partial_\lambda . \tag{IV.9}$$

Recall that from the relation (II.57b) for the β-matrices, it follows that (s = maximum spin)

$$(Q_1)^{2s+1} = \Box(Q_1)^{2s-1}, \tag{IV.10'}$$

or, equivalently,

$$\left(\Box - Q_1^2\right)Q_1^{2s-1} = 0. \tag{IV.10}$$

Now

$$
\begin{aligned}
d(\partial) &= \left(\Box - m^2\right)\Lambda^{-1}(\partial) \\
&= -\left(\Box - Q_1^2 + Q_1^2 - m^2\right)\left(Q_1 + m\right)^{-1} \\
&= -\left(\Box - Q_1^2\right)\left(Q_1 + m\right)^{-1} - \left(Q_1^2 - m^2\right)\left(Q_1 + m\right)^{-1} \\
&= -\left(\Box - Q_1^2\right)\left(Q_1 + m\right)^{-1} - \left(Q_1 - m\right)\left(Q_1 + m\right)\left(Q_1 + m\right)^{-1} \\
&= -\left(Q_1 - m\right) - \frac{1}{m}\left(\Box - Q_1^2\right)\left(1 + \frac{Q_1}{m}\right)^{-1} \\
&= m - Q_1 - \frac{1}{m}\left(\Box - Q_1^2\right)\left[1 - \frac{Q_1}{m} + \frac{Q_1^2}{m^2} + \cdots \right. \\
&\qquad \left. + (-1)^{\ell}\left(\frac{Q_1}{m}\right)^{\ell} + \cdots\right] \\
&= m - Q_1 - \frac{1}{m}\left(\Box - Q_1^2\right)\left[1 - \frac{Q_1}{m} + \left(\frac{Q_1}{m}\right)^2 + \cdots \right. \\
&\qquad \left. + (-1)^{2s-2}\left(\frac{Q_1}{m}\right)^{2s-2}\right], \qquad \text{(IV.11)}
\end{aligned}
$$

since

$$\left(\Box - Q_1^2\right)\left(\frac{Q_1}{m}\right)^{\ell} = 0$$

for $\ell \geq 2s-1$ because of (IV.10). Note that the highest order derivative in the expression (IV.11) for $d(\partial)$ is ∂^{2s}.

We apply (IV.11) to the Duffin-Kemmer case, for which

$$Q_1^3 = \Box Q_1, \qquad \text{(IV.12)}$$

that is to say $s = 1$. We obtain

$$d(\partial) = m - Q_1 - \frac{1}{m}\left(\Box - Q_1^2\right). \qquad \text{(IV.13)}$$

To check (IV.13), we have

$$\Lambda(\partial)d(\partial) = -\left(m + Q_1\right)\left[m - Q_1 - \frac{1}{m}\left(\Box - Q_1^2\right)\right]$$

$$= -m^2 + Q_1^2 + \frac{1}{m}\left(m + Q_1\right)\left(\Box - Q_1^2\right)$$

$$= -m^2 + Q_1^2 + \left(\Box - Q_1^2\right) + \frac{1}{m} Q_1\left(\Box - Q_1^2\right)$$

$$= \Box - m^2.$$

Consider the case where $\Lambda(\partial)$ takes the form

$$\Lambda(\partial) = \Lambda_o + \Lambda_{\mu\nu}\partial_\mu\partial_\nu, \tag{IV.14}$$

$$= \Lambda_o + Q_2, \tag{IV.14a}$$

$$Q_2 = \Lambda_{\mu\nu}\partial_\mu\partial_\nu. \tag{IV.14b}$$

We have the identity

$$\Box - m^2 = \left(\Lambda_o + Q_2\right) - \left(m^2 + \Lambda_o\right) + \left(\Box - Q_2\right). \tag{IV.15}$$

Now we can write

$$d(\partial) = \left(\Box - m^2\right)\Lambda^{-1}(\partial)$$

$$= 1 - \left(m^2 + \Lambda_o\right)\left(\Lambda_o + Q_2\right)^{-1} + \left(\Box - Q_2\right)\left(\Lambda_o + Q_2\right)^{-1}$$

$$= 1 - \left(m^2 + \Lambda_o\right)\Lambda_o^{-1}\left[1 - \Lambda_o^{-1}Q_2 + \cdots + (-1)^{L+1}\left(\Lambda_o^{-1}Q_2\right)^{L+1}\right]$$

$$+\left(\Box - Q_2\right)\Lambda_o^{-1}\left[1 - \Lambda_o^{-1}Q_2 + \cdots + (-1)^{L}\left(\Lambda_o^{-1}Q_2\right)^{L}\right], \tag{IV.16}$$

assuming that the highest derivative appearing in $d(\partial)$ is of order $2(L+1)$, i.e.,

$$d(\partial) \sim \partial^{2(L+1)}.$$

In our previous example we had found that

$$d(\partial) \sim \partial^{2s}, \tag{IV.17}$$

so we conjecture now

$$L = s - 1. \tag{IV.18}$$

The relation (IV.18) can at best be regarded as an empirical law and it seems to hold when $\Lambda(\partial)$ has the form (IV.14). Consider, for instance, the Proca (spin 1) case:

$$\Lambda_{\sigma\rho}(\partial) = (\Box - m^2)\delta_{\sigma\rho} - \partial_\sigma\partial_\rho, \tag{IV.19}$$

which implies

$$(Q_2)_{\sigma\rho} = \Box\,\delta_{\sigma\rho} - \partial_\sigma\partial_\rho, \tag{IV.19a}$$

$$(\Lambda_o)_{\sigma\rho} = -m^2\delta_{\sigma\rho}. \tag{IV.19b}$$

Since $s=1$, we set $L=0$ in (IV.16) to obtain

$$d_{\rho\lambda}(\partial) = \delta_{\rho\lambda} - \left[(m^2+\Lambda_o)\Lambda_o^{-1}(1 - \Lambda_o^{-1}Q_2)\right]_{\rho\lambda} + \left[(\Box - Q_2)\Lambda_o^{-1}\right]_{\rho\lambda}. \tag{IV.20}$$

The second term in the right hand side of (IV.20) vanishes because $\Lambda_o + m^2 = 0$ from (IV.19b). Also we have

$$\Lambda_o^{-1} = -\frac{1}{m^2},$$

so the third term can be simplified as

$$\left[(\Box - Q_2)\Lambda_o^{-1}\right]_{\rho\lambda} = -\frac{1}{m^2}(\Box - Q_2)_{\rho\lambda}$$

$$= -\frac{1}{m^2}\partial_\rho\partial_\lambda,$$

where in the last step we made use of (IV.19a). We get

$$d_{\rho\lambda}(\partial) = \delta_{\rho\lambda} - \frac{1}{m^2}\partial_\rho\partial_\lambda, \tag{IV.21}$$

a result which we obtained in (IV.8) using a different method.

Finally, if $\Lambda(\partial)$ has the form

$$\Lambda(\partial) = \Lambda_o + \Lambda_\mu \partial_\mu + \Lambda_{\mu\nu}\partial_\mu\partial_\nu, \tag{IV.22}$$

we are not able to find a general formula for $d(\partial)$ and also the empirical rule $d(\partial)\sim\partial^{2s}$ does not hold.

Now we discuss a few important properties of the K-G divisor:

(a) If

$$\Lambda(\partial) = -(\beta\partial + m) \equiv -(Q_1 + m) \tag{IV.23a}$$

then we have

$$d(\partial)d(\partial)\Delta^{(\ )}(x-y) = 2m\, d(\partial)\Delta^{(\ )}(x-y), \tag{IV.23b}$$

on the mass shell. Here $\Delta^{(\ )}(x-y)$ is any solution of the K-G equation.

Proof:

$$-(Q_1 + m)d(\partial)\Delta^{(\ )}(x-y) = (\Box - m^2)\Delta^{(\ )}(x-y)$$

$$= 0,$$

which implies

$$Q_1 d(\partial)\Delta^{(\ )}(x-y) = -m\, d(\partial)\Delta^{(\ )}(x-y). \tag{IV.24}$$

Now making use of (IV.11) and (IV.24), we have

$$d(\partial)d(\partial)\Delta^{(\ )}(x-y)$$

$$= \left[m - Q_1 + \frac{1}{m}(\Box - Q_1^2)\left\{1 - \frac{Q_1}{m} + \cdots + (-1)^{2s-2}\left(\frac{Q_1}{m}\right)^{2s-2}\right\}\right] \cdot d(\partial)\Delta^{(\ )}(x-y)$$

$$= \left[m + m + \frac{1}{m}(\Box - m^2)\{2s-1\}\right] d(\partial)\Delta^{(\ )}(x-y)$$

$$= \left[2m + \frac{(2s-1)}{m}(\Box - m^2)\right] d(\partial)\Delta^{(\ )}(x-y)$$

$$= 2m\, d(\partial)\Delta^{(\ )}(x-y),$$

since $\Delta^{(\ )}(x-y)$ satisfies the K-G equation.

(b) If

$$\Lambda(\partial) = \Lambda_o + Q_2, \quad Q_2 = \Lambda_{\mu\nu}\partial_\mu\partial_\nu \tag{IV.25a}$$

then

$$d(\partial)d(\partial)\Delta^{(\ )}(x-y) = -m^2\Lambda_o^{-1}\, d(\partial)\Delta^{(\ )}(x-y), \tag{IV.25b}$$

on the mass shell. Here again $\Delta^{(\ )}(x-y)$ is any solution of the Klein-Gordon equation.
Proof of (IV.25) is similar to that of (IV.23), and we leave it to the reader.

(c) Because of the second identity

$$\Lambda(\partial)d(\partial) = \Box - m^2,$$

it is very simple to construct the Green function corresponding to any field equation.
Suppose $\Delta_G(x-y)$ is the Green function appropriate to the K-G equation:

$$(\Box - m^2)\Delta_G(x-y) = \delta^4(x-y). \tag{IV.26}$$

We define, for the case under consideration,

$$G_{\alpha\beta}(x-y) = d_{\alpha\beta}(\partial)\Delta_G(x-y). \tag{IV.27}$$

Then this $G(x-y)$ is indeed the appropriate Green function because

$$\Lambda(\partial)\, G(x-y) = \delta^4(x-y).$$

V. Normalization and Closure Conditions for the Wave Functions

In this section we want to discuss the c-number solutions of

$$\Lambda(\partial)u(x) = 0. \tag{V.1}$$

We have already learned how to calculate $\Gamma_\mu(\partial, -\overleftarrow{\partial})$ and $d(\partial)$ when $\Lambda(\partial)$ is given. Making use of these quantities we shall obtain normalization and closure conditions for wave functions corresponding to arbitrary spin in a compact, universal form.

Since the equation (V.1) is homogeneous, normalization is at our disposal. As the wave functions satisfy the Klein-Gordon equation also, decomposition into positive and negative frequencies is Lorentz invariant. We shall denote the positive frequency wave functions by $u(x)$ and the negative frequency wave functions by $v(x)$.

Let us discuss the positive frequency wave functions first. For the sake of simplicity we take the momentum representation and write our wave functions as $u_k^{(r)}(x)$, with

$$u_k^{(r)}(x) = u^{(r)}(k) f_k(x), \tag{V.2a}$$

where

$$f_k(x) = \frac{1}{(2\pi)^{3/2}} e^{i\underset{\sim}{k}\cdot\underset{\sim}{x} - i\omega(k)t}, \tag{V.2b}$$

$$\omega(k) = +\sqrt{\underset{\sim}{k}^2 + m^2}\,. \tag{V.2c}$$

The superscript r is the spin orientation or helicity. We have

$$hu^{(r)}(k) = h^{(r)} u^{(r)}(k) \tag{V.3}$$

where h is the helicity operator

$$h = (\underset{\sim}{k}\cdot\underset{\sim}{S})/|k|\,. \tag{V.4}$$

Here $\underset{\sim}{S}$ is the spin operator with [recall (II.38)]

$$S_1 = S_{23}, \quad S_2 = S_{31}, \quad S_3 = S_{12}, \tag{V.5}$$

so that (V.4) may be written as

$$h = \frac{1}{2|\underset{\sim}{k}|}\, \epsilon_{ijk} k_i S_{jk}. \tag{V.6}$$

From (V.1) and (V.2) it follows that

$$\Lambda(ik) u^{(r)}(k) = 0. \tag{V.7}$$

We claim that

$$[\Lambda(ik), h] = 0. \tag{V.8}$$

Proof: From the Lorentz invariance requirement we have

$$[\Lambda(\partial), \ell_{\mu\nu}] = 0,$$

which, in momentum space, becomes

$$\left[\Lambda(ik),\ k_\mu \frac{\partial}{\partial k_\nu} - k_\nu \frac{\partial}{\partial k_\mu} + iS_{\mu\nu}\right] = 0. \tag{V.9}$$

Let us set $\mu = j$, $\nu = k$:

$$\left[\Lambda(ik),\left(k_j \frac{\partial}{\partial k_k} - k_k \frac{\partial}{\partial k_j}\right) + iS_{jk}\right] = 0. \tag{V.9'}$$

Note that

$$\epsilon_{ijk} k_i \left(k_j \frac{\partial}{\partial k_k} - k_k \frac{\partial}{\partial k_j}\right) = 0 \tag{V.10}$$

because of the anti-symmetry property of ϵ_{ijk}. Therefore, multiplying (V.9') by

$$\frac{-i}{2|k|} \epsilon_{ijk} k_i ,$$

and making use of (V.10) and (V.6), we obtain

$$[\Lambda(ik), h] = 0,$$

as we had claimed.

We define the adjoint wave function by

$$\bar{u}_k^{(r)}(x) = u_k^{(r)\dagger}(x)\eta. \tag{V.11}$$

Now consider the quantity

$$N_{k'k}^{(r'r)} = -i \int d\sigma_\lambda(x)\ \bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x). \tag{V.12}$$

First we show that it vanishes for $k' \neq k$ or $r' \neq r$. From (V.2b) we have

$$-i\partial_\mu u_k^{(r)}(x) = k_\mu u_k^{(r)}(x), \tag{V.13a}$$

$$-i\partial_\mu \bar{u}_{k'}^{(r')}(x) = k'_\mu \bar{u}_k^{(r')}(x). \tag{V.13b}$$

If we multiply (V.13a) by

$$\bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})$$

from the left and (V.13b) by

$$\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x)$$

from the right and add the two relations, we obtain

$$\left(k'_\mu - k_\mu\right) \bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x)$$

$$= i\partial_\mu\left[\bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x)\right]. \qquad (V.14)$$

From the last relation we easily deduce

$$\left(k'_\mu - k_\mu\right) N_{k'k}^{(r'r)} = \int d\sigma_\lambda(x)\partial_\mu\left[\bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x)\right]. \qquad (V.15)$$

In Appendix A we have shown that

$$\int d\sigma_\lambda(x)\partial_\mu f(x) = \int d\sigma_\mu(x)\partial_\lambda f(x). \qquad (V.16)$$

In proving (V.16) one has to assume that $f(x) \to 0$ for $|\underset{\sim}{x}| \to \infty$. This is not true for the term in square brackets in (V.15) with our plane-wave wave functions. However, in a more careful treatment one would use wave packets rather than plane waves so that the condition of vanishing at large distances would be satisfied. Therefore, the use of (V.16) in (V.15) is quite justified. We obtain

$$\left(k'_\mu - k_\mu\right) N_{k'k}^{(r'r)} = \int d\sigma_\mu(x)\partial_\lambda\left[\bar{u}_{k'}^{(r')}(x)\Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u_k^{(r)}(x)\right] = 0,$$

where we make use of the first identity in the last step. Therefore, we have the result

$$N_{k'k}^{(r'r)} = 0 \qquad \text{if} \qquad k' \neq k. \qquad (V.17)$$

Next we want to show that $N_{k'k}^{(r'r)}$ vanishes also when $r' \neq r$. First, note that

$$\int d\sigma_\lambda(x) f_{k'}^*(x)\partial_\mu f_k(x) = -\int d\sigma_\lambda(x) f_{k'}^*(x)\overleftarrow{\partial}_\mu f_k(x), \qquad (V.18)$$

which means that in (V.12) we can replace $-\overleftarrow{\partial}$ by ∂ or vice versa:

$$N_{k'k}^{(r'r)} \equiv -i \int d\sigma_\lambda \, \bar{u}_{k'}^{(r')}(x) \Gamma_\lambda(\partial, -\overleftarrow{\partial}) u_k^{(r)}(x) \tag{V.12}$$

$$= -i \int d\sigma_\lambda \, \bar{u}_{k'}^{(r')}(x) R_\lambda(\partial) u_k^{(r)}(x) \tag{V.19a}$$

$$= -i \int d\sigma_\lambda \, \bar{u}_{k'}^{(r')}(x) R_\lambda(-\overleftarrow{\partial}) u_k^{(r)}(x), \tag{V.19b}$$

where

$$R_\lambda(\partial) = \Gamma_\lambda(\partial,\partial) = \left.\frac{\partial \Lambda(ik)}{\partial ik_\lambda}\right|_{ik=\partial} . \tag{V.19c}$$

In order to study the dependence on r,r' we set $\underset{\sim}{k}=\underset{\sim}{k}'=0$ and drop the subscript k on the wave functions for convenience. We know

$$S_3 u^{(r)}(x) = r u^{(r)}(x),$$

$$(S_1+iS_2)u^{(r)}(x) = g(r,s)u^{(r+1)}(x),$$

$$(S_1-iS_2)u^{(r)}(x) = g(r-1,s)u^{(r-1)}(x),$$

$$\bar{u}^{(r)}(x)(S_1+iS_2) = g(r-1,s)\,\bar{u}^{(r-1)}(x),$$

$$\bar{u}^{(r)}(x)(S_1-iS_2) = g(r,s)\,\bar{u}^{(r+1)}(x),$$

$$\bar{u}^{(r)}(x)S_3 = r\,\bar{u}^{(r)}(x),$$

where

$$g(r,s) = \sqrt{s(s+1) - r(r+1)} ,$$

and r takes the values $-s, -s+1, \cdots, s-1, s$. In a general coordinate system we have from Lorentz invariance

$$[\Lambda(ik), \ell_{ij}] = 0.$$

If we apply $\partial/\partial ik_4$ to this relation, we get

$$[R_4(ik), \ell_{ij}] = 0.$$

which, in the frame where $\underset{\sim}{k}=0$, takes the form

$$[R_4(ik), S_k] = 0.$$

Choose $k=3$ and sandwich this relation between $\bar{u}^{(r')}(x)$ and $u^{(r)}(x)$:

$$\begin{aligned} 0 &= \bar{u}^{(r')}(x)R_4(ik)S_3u^{(r)}(x) - \bar{u}^{(r')}(x)S_3R_4(ik)u^{(r)}(x) \\ &= (r-r')\,\bar{u}^{(r')}(x)R_4(ik)\,u^{(r)}(x) \\ &= (r-r')\,\bar{u}^{(r')}(x)R_4(\partial)\,u^{(r)}(x). \end{aligned} \tag{V.20}$$

That $N_{k'k}^{(r'r)}$ is independent of σ can be easily shown by applying $\delta/\delta\sigma(x)$ to (V.12) and making use of the first identity. Choosing the x_0 = constant surface,

$$d\sigma_\lambda(x) = (0,0,0,-id^3x),$$

we have

$$\begin{aligned} N_{k'k}^{(r'r)} &= -\int d^3x\, \bar{u}_{k'}^{(r')}(x)R_4(\partial)\,u_k^{(r)}(x) \\ &= 0 \qquad \text{if} \quad r' \neq r, \end{aligned} \tag{V.21}$$

from (V.20).

Next, we show that $N_{k'k}^{(r'r)}$ is independent of r. From

$$(S_1+iS_2)\,u^{(r)}(x) = g(r,s)\,u^{(r+1)}(x).$$

We have

$$\begin{aligned} g(r,s)\,\bar{u}^{(r+1)}(x)R_\lambda(\partial)\,u^{(r+1)}(x) &= \bar{u}^{(r+1)}(x)R_\lambda(\partial)(S_1+iS_2)\,u^{(r)}(x) \\ &= \bar{u}^{(r+1)}(x)(S_1+iS_2)R_\lambda(\partial)\,u^{(r)}(x) \\ &= g(r,s)\,\bar{u}^{(r)}(x)R_\lambda(\partial)\,u^{(r)}(x). \end{aligned}$$

Hence, we obtain

$$N_{k',k}^{(r+1,r+1)} = N_{k',k}^{(r,r)}.$$

Finally we have to show that $N_{kk}^{(rr)}$ is a real number. From Eq. (V.12):

$$\left(N^{(r'r)}_{k'k}\right)^* = +i \int d\sigma_\lambda^*(x)\, u^{(r)}_k(x)^\dagger \left[\eta \Gamma_\lambda(\overleftarrow{\partial}, -\partial)\right]^\dagger u^{(r')}_{k'}(x),$$

where we have replaced $\Gamma_\lambda(\partial, -\overleftarrow{\partial})$ by $\Gamma_\lambda(\overleftarrow{\partial}, -\partial)$ because the wave function originally standing to the right (left) now stands to the left (right). Recalling that

$$d\sigma_\lambda^*(x) = g_{\lambda\nu} d\sigma_\nu(x),$$

$$\left[\eta \Gamma_\lambda(\overleftarrow{\partial}, -\partial)\right]^\dagger = -g_{\lambda\mu} \eta \Gamma_\mu(-\overleftarrow{\partial}, +\partial),$$

$$\Gamma_\mu(-\overleftarrow{\partial}, \partial) = \Gamma_\mu(\partial, -\overleftarrow{\partial}),$$

we get immediately

$$\left[N^{(r'r)}_{k'k}\right]^* = N^{rr'}_{kk'}, \tag{V.22}$$

which implies

$$N^{(rr)}_{kk} \text{ is real.} \tag{V.23}$$

In view of (V.17), (V.21) and (V.23) we can write the normalization condition for the positive frequency wave functions as

$$-i \int d\sigma_\lambda(x)\, \bar{u}^{(r')}_{k'}(x) \Gamma_\lambda(\partial, -\overleftarrow{\partial})\, u^{(r)}_k(x) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'). \tag{V.24}$$

Other equivalent forms are

$$-i \int d\sigma_\lambda(x)\, \bar{u}^{(r')}_{k'}(x) R_\lambda(\partial)\, u^{(r)}_k(x) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'), \tag{V.24'}$$

and

$$-i \int d\sigma_\lambda(x)\, \bar{u}^{(r')}_{k'}(x) R_\lambda(-\overleftarrow{\partial})\, u^{(r)}_k(x) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'), \tag{V.24''}$$

where $R_\lambda(\partial)$ is given by (V.19c).

We apply (V.24) to the K-G and Dirac cases. For the K-G case,

$$(\Box - m^2)\, u(x) = 0,$$

$$\Gamma_\lambda(\partial, -\overleftarrow{\partial}) = \partial_\lambda - \overleftarrow{\partial}_\lambda,$$

and there is no spin. We get

$$-i \int d\sigma_\lambda(x)\, \bar{u}_{k'}(x)(\partial_\lambda - \overleftarrow{\partial}_\lambda) u_k(x) = \delta(\underset{\sim}{k} - \underset{\sim}{k'}),$$

which is the familiar orthonormality relation for the K-G wave functions.

For the Dirac case

$$-(\gamma\partial + m) u(x) = 0,$$

i.e.,

$$\Lambda(\partial) = -(\gamma\partial + m),$$

then

$$\Gamma_\mu = -\gamma_\mu,$$

so we obtain

$$i \int d\sigma_\lambda(x)\, \bar{u}_{k'}^{(r')}(x)\, \gamma_\mu u_k^{(r)}(x) = \delta_{rr'}\delta(\underset{\sim}{k} - \underset{\sim}{k'}).$$

To examine this more closely, choose the x_o = constant surface to get

$$\int d^3x\, \bar{u}_{k'}^{(r')}(x)\, \gamma_4 u_k^{(r)}(x) = \delta_{rr'}\delta(\underset{\sim}{k} - \underset{\sim}{k'})$$

or

$$\int d^3x\, u_{k'}^{(r')\dagger}(x)\, u_k^{(r)}(x) = \delta_{rr'}\delta(\underset{\sim}{k} - \underset{\sim}{k'}),$$

which is correct.

We should point out here that the normalization condition fixes the sign of $\Lambda(\partial)$, and therefore that of Γ_μ and $d(\partial)$, which the homogeneous wave equation does not. For example, if we write for the Dirac case

$$\Lambda(\partial) = \gamma\partial + m,$$

we get the wrong result

$$-\int d^3x\, u^\dagger(x)\, u(x) = \delta_{rr'}\delta(\underset{\sim}{k} - \underset{\sim}{k'}).$$

Now we prove the closure condition

$$\sum_r \int d^3k\, u_k^{(r)}(x)\, \bar{u}_k^{(r)}(y) = id(\partial)\, \Delta^{(+)}(x - y). \qquad (V.25)$$

Since the quantity $id(\partial)\Delta^{(+)}(x-y)$ has only positive frequencies, we can expand it in terms of $u_k^{(r)}(x)$'s:

$$id(\partial)\Delta^{(+)}(x-y) = \sum_{r'} \int d^3k' u_{k'}^{(r')}(x) C_{k'}^{(r')}(y). \qquad (V.26)$$

If we apply

$$-i \int d\sigma_\lambda(x) \bar{u}_k^{(r)}(x) \Gamma_\lambda(\partial, -\overleftarrow{\partial})$$

to (V.26) and make use of (V.24), we get

$$C_k^{(r)}(y) = \int d\sigma_\lambda(x) \bar{u}_k^{(r)}(x) \Gamma_\lambda(\partial, -\overleftarrow{\partial}) d(\partial) \Delta^{(+)}(x-y). \qquad (V.27)$$

One easily shows that this relation is independent of σ by applying $\delta/\delta\sigma(x)$. One makes use of the first identity and the equations of motion:

$$\bar{u}_k^{(r)}(x) \Lambda(-\overleftarrow{\partial}) = 0,$$

$$(\Box - m^2)\Delta^{(+)}(x-y) \equiv \Lambda(\partial) d(\partial) \Delta^{(+)}(x-y) = 0.$$

We shall evaluate (V.27) at the surface $x_o = y_o$. First replace $\Gamma_\lambda(\partial, -\overleftarrow{\partial})$ by $R_\lambda(\partial)$ as we did in (V.19a). We need to calculate $R_\lambda(\partial)d(\partial)$. Starting with

$$\Lambda(\partial) d(\partial) = \Box - m^2,$$

we get

$$\Lambda(ik) d(ik) = -(k^2 + m^2). \qquad (V.28)$$

Applying $\partial/\partial ik_\lambda$ to (V.28),

$$R_\lambda(ik) d(ik) + \Lambda(ik) \frac{\partial d(ik)}{\partial ik_\lambda} = 2ik_\lambda,$$

or

$$R_\lambda(\partial) d(\partial) = \left[2ik_\lambda - \Lambda(ik) \frac{\partial d(ik)}{\partial ik_\lambda} \right]_{ik=\partial}. \qquad (V.29)$$

We can now evaluate $C_k^{(r)}(y)$ explicitly.

$$C_k^{(r)}(y) = \int_{x_o=y_o} d\sigma_\lambda(x)\, \bar{u}_k^{(r)}(x)\, R_\lambda(\partial) d(\partial)\, \Delta^{(+)}(x-y)$$

$$= \int_{x_o=y_o} d\sigma_\lambda(x)\, \bar{u}_k^{(r)}(x) \left[2ik_\lambda - \Lambda(ik) \frac{\partial d(ik)}{\partial ik_\lambda} \right]_{ik_\lambda = \partial_\lambda = -\overleftarrow{\partial}_\lambda} \Delta^{(+)}(x-y).$$

But $\bar{u}_k^{(r)}(x)\, \Lambda(ik) = 0$, since $\bar{u}_k^{(r)}(x)\, \Lambda(-\overleftarrow{\partial}) = 0$. Therefore,

$$C_k^{(r)}(y) = \int_{x_o=y_o} d\sigma_4(x)\, \bar{u}_k^{(r)}(x)\, 2\partial_4\, \Delta^{(+)}(x-y)$$

$$= -2 \int d^3x\, \bar{u}_k^{(r)}(x) \left[\frac{\partial}{\partial x_o} \Delta^{(+)}(x-y) \right]_{x_o=y_o}$$

$$= \bar{u}_k^{(r)}(y), \qquad \text{(V.30)}$$

where we have made use of the well-known result

$$\frac{\partial}{\partial x_o} \Delta^{(+)}(x-y) \Big|_{x_o=y_o} = -\tfrac{1}{2}\, \delta(\underset{\sim}{x}-\underset{\sim}{y}). \qquad \text{(V.31)}$$

Substituting (V.30) in (V.26) we get the closure condition (V.25).

The negative frequency solutions are given by

$$v_k^{(r)}(x) = C\, u_k^{(r)*}(x), \qquad \text{(V.32a)}$$

with the inverse transformation

$$C\, v_k^{(r)*}(x) = u_k(x). \qquad \text{(V.32b)}$$

Using the transformation properties of the various quantities with respect to charge conjugation, one can easily write the normalization and closure conditions for the negative frequency solutions.

From the relation

$$C^{-1} \eta \Gamma_\mu(\partial, -\overleftarrow{\partial}) C = \rho\, g_{\mu\nu} \left[\eta \Gamma_\nu(\partial, -\overleftarrow{\partial}) \right]^*, \qquad \text{(V.33)}$$

the relations (V.32) and (V.24), we get the normalization condition

$$-i\int d\sigma_\lambda(x)\,\bar{v}_{k'}^{(r')}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})\,v_k^{(r)}(x) = -\rho\,\delta_{rr'}\delta(k-k'), \tag{V.34}$$

where $\Gamma_\lambda(\partial,-\overleftarrow{\partial})$ may be replaced as before. Recall

$$\rho = \begin{cases} +1 & \text{for integer spin} \\ -1 & \text{for half-integer spin.} \end{cases}$$

The closure condition turns out to be

$$\sum_r \int d^3k\, v_k^{(r)}(x)\,\bar{v}_k^{(r)}(y) = -i\rho\, d(\partial)\,\Delta^{(-)}(x-y), \tag{V.35}$$

as the reader can readily prove. We can combine (V.25) and (V.35) to read

$$\sum_r \int d^3k\left\{u_k^{(r)}(x)\,\bar{u}_k^{(r)}(y) - \rho\, v_k^{(r)}(x)\,\bar{v}_k^{(r)}(y)\right\} = i\,d(\partial)\,\Delta(x-y). \tag{V.36}$$

VI. Quantization

In this section we shall present a quantization procedure which does not make use of the canonical formalism.

First, we discuss the raising and lowering operators. Next, we carry out the quantization procedure and in this connection we establish the relation between spin and statistics. Simple applications are discussed. Finally, we discuss the uniqueness of P_μ and $M_{\mu\nu}$.

A. Raising and Lowering Operators.

Consider operators $a, a^\dagger$ satisfying

$$[a,a^\dagger] \equiv aa^\dagger - a^\dagger a = 1. \tag{VI.1}$$

For this case everyone knows that the operator

$$N = a^\dagger a \tag{VI.2}$$

has the eigenvalues $0,1,2,\ldots$ and is therefore called the number operator. We shall briefly go through the proof as it is instructive. This is just a mathematical exercise and no physics is involved.

The operator N is Hermitian, $N = N^\dagger$, so its eigenvalues are real. Suppose ψ is an eigenstate of N with the eigenvalue λ, i.e.,

$$N\psi = \lambda\psi,$$

where λ is real. Then

$$(\psi, N\psi) = \lambda(\psi, \psi),$$

and, again

$$(\psi, N\psi) = (\psi, a^\dagger a\psi) = (a\psi, a\psi),$$

so that

$$(a\psi, a\psi) = \lambda(\psi, \psi).$$

Since the norm of a state is positive definite, it follows that λ is non-negative. The statement

$$N = N^\dagger \geq 0 \tag{VI.3}$$

summarizes the fact that N is Hermitian and its eigenvalues are non-negative real numbers. We have yet to show that the eigenvalues are integers. We work in a representation where N is diagonal and denote by $|n\rangle$ the eigenstate corresponding to the eigenvalue n, thus

$$N|n\rangle = n|n\rangle. \tag{VI.4}$$

Consider the state $a|n\rangle$, then

$$\begin{aligned} Na|n\rangle &= a^\dagger aa|n\rangle \\ &= (aa^\dagger - 1)a|n\rangle \\ &= aN|n\rangle - a|n\rangle = (n-1)a|n\rangle, \end{aligned} \tag{VI.5}$$

that is to say that $a|n\rangle$ is also an eigenstate of N with the eigenvalue $(n-1)$. By repeating this process, we obtain

$$Na^\ell|n\rangle = (n-\ell)a^\ell|n\rangle. \tag{VI.6}$$

According to (VI.3), the eigenvalue $(n-\ell)$ cannot be negative; therefore, we must have

$$a^\ell|n\rangle = 0 \quad \text{for} \quad \ell > n. \tag{VI.7}$$

In other words, if ℓ_0 is the integer satisfying

$$n - 1 < \ell_0 \leq n, \tag{VI.8a}$$

then

$$a^{\ell_o}|n\rangle \neq 0, \tag{VI.8b}$$

but

$$a^{\ell_o+1}|n\rangle = 0. \tag{VI.8c}$$

Applying $a^\dagger$ to the last relation, we have

$$0 = a^\dagger a^{\ell_o+1}|n\rangle = Na^{\ell_o}|n\rangle = (n-\ell_o)a^{\ell_o}|n\rangle,$$

which leads to

$$n = \ell_o, \tag{VI.8d}$$

in view of (VI.8b). Therefore, n is restricted to integral values only, including zero. Thus the eigenstates are $|0\rangle, |1\rangle, \ldots, |n\rangle, \ldots$. The state $|0\rangle$ satisfies

$$a|0\rangle = 0, \tag{VI.9a}$$

$$N|0\rangle = 0. \tag{VI.9b}$$

Using the commutation relation (VI.1), one easily establishes the following:

$$Na^\dagger|0\rangle = a^\dagger|0\rangle, \tag{VI.10a}$$

$$N(a^\dagger)^n|0\rangle = n(a^\dagger)^n|0\rangle \tag{VI.10b}$$

$$[a,N] = a, \tag{VI.11a}$$

$$[a^\dagger,N] = -a^\dagger. \tag{VI.11b}$$

If we sandwich the relations (VI.11) between states $\langle n'|$ and $|n\rangle$, we obtain

$$\langle n'|a|n\rangle(n-1-n') = 0,$$

$$\langle n'|a^\dagger|n\rangle(n+1-n') = 0,$$

so that $\langle n'|a|n\rangle$ survives only when $n'=n-1$, and $\langle n'|a^\dagger|n\rangle$ survives only when $n'=n+1$. If we write

$$a|n\rangle = \sqrt{n}\,|n-1\rangle\,, \tag{VI.12a}$$

$$a^\dagger|n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \tag{VI.12b}$$

it can easily be checked that these relations are consistent with all the equations above. In an obvious matrix notation we can write the lowering and raising operators as

$$a = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{2} & 0 & 0 & & \\ 0 & 0 & 0 & \sqrt{3} & 0 & & \\ 0 & 0 & 0 & 0 & & & \\ 0 & & & & & & \\ \vdots & & & & & & \end{bmatrix}, \tag{VI.13a}$$

and

$$a^\dagger = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & . & \\ 0 & \sqrt{2} & 0 & 0 & . & \\ 0 & 0 & \sqrt{3} & 0 & . & \\ 0 & 0 & 0 & . & . & \\ 0 & & & & & \\ . & & & & & \\ . & & & & & \end{bmatrix}. \tag{VI.13b}$$

In a similar fashion we can handle the operators $a, a^\dagger$ satisfying

$$\{a, a^\dagger\} \equiv aa^\dagger + a^\dagger a = 1, \tag{VI.14a}$$

$$\{a, a\} \equiv aa + aa = 0. \tag{VI.14b}$$

In this case we find

$$N = N^\dagger = a^\dagger a = 0 \text{ or } 1. \tag{VI.15}$$

The raising and lowering operators characterized by (VI.1) or (VI.14) are not the only such operators. However, we shall not discuss the other possibilities here.

Now we are ready to quantize.

B. Quantization.

We shall regard the Heisenberg equation of motion,

$$i\hbar \frac{\partial F(t)}{\partial t} = \left[F(t), H \right], \tag{VI.16}$$

as the most fundamental equation in quantized field theory. Here $F(t)$ is any dynamical variable.

Let us explain why we consider (VI.16) so fundamental. Since $F^{\dagger}(t)$ is also a dynamical variable, it too satisfies (VI.16), i.e.,

$$i \frac{\partial F^{\dagger}(t)}{\partial t} = \left[F^{\dagger}(t), H \right],$$

where we have set $\hbar = 1$. If we take the Hermitian conjugate of (VI.16), we obtain

$$i \frac{\partial F^{\dagger}(t)}{\partial t} = \left[F^{\dagger}(t), H^{\dagger} \right].$$

From these two equations we get

$$\left[F^{\dagger}(t), H - H^{\dagger} \right] = 0,$$

which implies

$$H - H^{\dagger} = \text{c-number}.$$

This means that H is essentially Hermitian. Then we can diagonalize H so that

$$H|E\rangle = E|E\rangle, \tag{VI.17}$$

and the eigenvalue E is real. If we sandwich the equation (VI.16) between $\langle E'|$ and $|E\rangle$, we get the relation

$$i \frac{\partial}{\partial t} \langle E'|F(t)|E\rangle = (E - E')\langle E'|F(t)|E\rangle. \tag{VI.18}$$

If we introduce the Fourier transform

$$\langle E'|F(t)|E\rangle = \int d\omega\, e^{-i\omega t} \langle E'|\hat{F}(\omega)|E\rangle, \tag{VI.19}$$

then (VI.18) yields

$$\omega = E - E', \tag{VI.20}$$

since F is arbitrary. This is just Bohr's frequency-energy relation (with $\hbar = 1$), if we can interpret E as energy. Since Bohr's frequency-energy relation has a firm experimental basis, it is eminently reasonable to regard (VI.16) as a most fundamental equation.

There are a number of possible interpretations of the equation

$$i\frac{\partial F(t)}{\partial t} = \left[F(t), H\right]. \tag{VI.16}$$

For example:

(i) The time-evolution of F(t) and H are known, then the above equation will determine the commutation relations of F(t).

(ii) The commutation relations and H are known, then the above equation will determine the time-evolution of F(t).

However, the interpretation that we shall use differs from the above two and we state it as

(iii) From the knowledge of the time-evolution of F(t), the commutation relations as well as H can be determined.

According to this interpretation there are too many unknowns in the equation. As a consequence, there exist a number of possible quantizations. For example, one can include parastatistics and hyperquantization within the framework of this approach. We shall not discuss parastatistics and hyperquantization in this course, however. Suffice it to say that this approach is much more flexible than the canonical formalism.

The relativistic generalization of (VI.16) is

$$-i\partial_\mu F(x) = [F(x), P_\mu], \tag{VI.21}$$

as a particular case of which we have

$$-i\partial_\mu \phi(x) = [\phi(x), P_\mu]. \tag{VI.22}$$

Since ∂_μ is a four-vector, P_μ must be a four-vector too. This is not a trivial statement in field theory as we shall see later. So, we try to find the commutation relations and P_μ when

$$\Lambda(\partial)\phi(x) = 0 \tag{VI.23}$$

is given.

As we have mentioned earlier, when we take the interpretation (iii) for the equation (VI.16), there are too many unknowns. In order to arrive at a physically meaningful theory, we restrict ourselves to those solutions only which obey the following conditions:

1) P_i and $H = -iP_4$ are Hermitian.
2) H is non-negative (for it to be the energy).
3) Only bosons and fermions exist in nature. (This excludes the parastatistical particles from our consideration.)
4) P_μ is a four-vector.
5) All physical quantities at finite distance exterior to the null-cone are commutative. (This is just microcausality.)

We expand $\phi(x)$ in terms of c-number wave functions as

$$\phi(x) = \sum_r \int d^3k \left\{ a^{(r)}(k) u_k^{(r)}(x) + b^{(r)\dagger}(k) v_k^{(r)}(x) \right\}, \quad \text{(VI.24)}$$

where, at this stage, $a^{(r)}(k)$ and $b^{(r)\dagger}(k)$ are just the relevant expansion coefficients. Then

$$-i\partial_\mu \phi(x) = \sum_r \int d^3k \, k_\mu \left\{ a^{(r)}(k) u_k^{(r)}(x) - b^{(r)\dagger}(k) v_k^{(r)}(x) \right\}, \quad \text{(VI.25)}$$

due to

$$-i\partial_\mu u_k^{(r)}(x) = k_\mu u_k^{(r)}(x),$$

$$-i\partial_\mu v_k^{(r)}(x) = -k_\mu v_k^{(r)}(x).$$

If we substitute (VI.24) and (VI.25) in (VI.22), we get

$$\left[a^{(r)}(k), P_\mu \right] = k_\mu a^{(r)}(k), \quad \text{(VI.26a)}$$

$$\left[b^{(r)\dagger}(k), P_\mu \right] = -k_\mu b^{(r)\dagger}(k). \quad \text{(VI.26b)}$$

If we make the guess

$$P_\mu = \sum_r \int d^3k \, k_\mu \left\{ a^{(r)\dagger}(k) a^{(r)}(k) + b^{(r)\dagger}(k) b^{(r)}(k) \right\}, \quad \text{(VI.27)}$$

which is consistent with conditions 1) and 2), then the possibilities are

$$a^{(r)}(k) a^{(r')\dagger}(k') - \rho' a^{(r')\dagger}(k') a^{(r)}(k) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'), \quad \text{(VI.28a)}$$

$$b^{(r)}(k) b^{(r')\dagger}(k') - \rho' b^{(r')\dagger}(k') b^{(r)}(k) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'), \quad \text{(VI.28b)}$$

where ρ' may be either +1 or -1. There are other possibilities but they are excluded by the condition 3). (The reader interested in these other possibilities is referred to the paper by Kamefuchi and Takahashi.[10]) Here ρ' specifies the statistics:

$$\rho' = \begin{cases} 1 & \text{for bosons} \\ -1 & \text{for fermions.} \end{cases} \tag{VI.28c}$$

Since P_μ has the eigenvalues

$$P_\mu = k_\mu \begin{cases} 0,1,2,\ldots & \text{for } \rho' = 1 \\ 0,1 & \text{for } \rho' = -1 \end{cases} \tag{VI.29}$$

for each mode, k_μ is interpreted as the energy-momentum of individual particles. In particular, we have $k_0 = \sqrt{\boldsymbol{k}^2 + m^2}$ so that m is indeed the mass and the relativistic relation between energy and momentum is satisfied. The origin of this relationship can be traced to

$$(\Box - m^2)\phi(x) = 0.$$

Next we want to study the commutation relation of $\phi(x)$. First note that

$$\begin{aligned} \bar{\phi}(y) &\equiv \phi^\dagger(y)\eta \\ &= \sum_r \int d^3k \left\{ a^{(r)\dagger}(k)\, \bar{u}_k^{(r)}(y) + b^{(r)}(k)\bar{v}_k^{(r)}(y) \right\} . \end{aligned} \tag{VI.30}$$

Then from (VI.26), (VI.30) and (VI.28) one shows

$$\begin{aligned} &\phi(x)\bar{\phi}(y) - \rho'\bar{\phi}(y)\phi(x) \\ &= \sum_r \int d^3k \left\{ u_k^{(r)}(x)\bar{u}_k^{(r)}(y) - \rho' v_k^{(r)}(x)\bar{v}_k^{(r)}(y) \right\} . \end{aligned} \tag{VI.31}$$

Now we make use of the closure conditions (V.25) and (V.35) to rewrite this as

$$\begin{aligned} &\phi(x)\bar{\phi}(y) - \rho'\bar{\phi}(y)\phi(x) \\ &\quad = id(\partial)\Delta^{(+)}(x-y) + \rho\rho' id(\partial)\Delta^{(-)}(x-y) \\ &\quad = id(\partial)\left\{ \Delta^{(+)}(x-y) + \rho\rho'\Delta^{(-)}(x-y) \right\} , \end{aligned} \tag{VI.32}$$

where ρ specifies the spin:

$$\rho = \begin{cases} 1 & \text{when spin is an integer} \\ -1 & \text{when spin is a half-integer.} \end{cases} \tag{VI.33}$$

In order for the condition 5) to be satisfied, the right-hand side of (VI.32) must be zero for (x-y) space-like. For this to be true, we must have $\rho\rho'=+1$, i.e.,

$$\rho' = \rho, \tag{VI.34}$$

which is the relation between spin and statistics. Then we obtain the commutation relation

$$\phi(x)\bar{\phi}(y) - \rho'\bar{\phi}(y)\phi(x) = id(\partial)\Delta(x-y). \tag{VI.35}$$

We assume $\rho'=\rho$ from now on.

We want to see how P_μ looks in configuration space. Consider the quantity

$$-i\int d\sigma_\lambda(x)\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})(\partial_\mu-\overleftarrow{\partial}_\mu)\phi(x).$$

If we make use of the expansions (VI.24) and (VI.30) and the normalization conditions for the c-number wave functions, we find

$$\begin{aligned} &-i\int d\sigma_\lambda(x)\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})(\partial_\mu-\overleftarrow{\partial}_\mu)\phi(x) \\ &\quad = 2i\sum_r\int d^3k\, k_\mu\left\{a^{(r)\dagger}(k)a^{(r)}(k) + \rho\, b^{(r)}(k)b^{(r)\dagger}(k)\right\} \\ &\quad = 2i\,P_\mu + \text{C-number}; \end{aligned}$$

therefore, we can write

$$\begin{aligned} P_\mu = &-\frac{1}{2}\int d\sigma_\lambda\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})(\partial_\mu-\overleftarrow{\partial}_\mu)\phi(x) \\ &+\frac{1}{2}\int d\sigma_\lambda\langle 0|\bar{\phi}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})(\partial_\mu-\overleftarrow{\partial}_\mu)\phi(x)|0\rangle. \end{aligned} \tag{VI.36}$$

Again, by making use of the expansions for $\phi(x)$ and $\bar{\phi}(x)$ and the normalization conditions for the C-number wave functions, one can easily verify the following expressions for the lowering and raising operators:

$$a^{(r)}(k) = -i\int d\sigma_\lambda(x)\,\bar{u}_k^{(r)}(x)\Gamma_\lambda(\partial,-\overleftarrow{\partial})\,\phi(x), \tag{VI.37a}$$

$$a^{(r)\dagger}(k) = -i\int d\sigma_\lambda(x)\,\bar{\phi}(x)\,\Gamma_\lambda(\partial,-\overleftarrow{\partial})\,u_k^{(r)}(x), \tag{VI.37b}$$

$$b^{(r)}(k) = i\rho\int d\sigma_\lambda(x)\,\bar{\phi}(x)\,\Gamma_\lambda(\partial,-\overleftarrow{\partial})v_k^{(r)}(x), \tag{VI.37c}$$

$$b^{(r)\dagger}(k) = i\rho\int d\sigma_\lambda(x)\,\bar{v}_k^{(r)}(x)\,\Gamma_\lambda(\partial,-\overleftarrow{\partial})\,\phi(x). \tag{VI.37d}$$

If we substitute the above expressions for $a^{(r)}(k)$ and $b^{(r)\dagger}(k)$ back into

$$\phi(x) = \sum_r \int d^3k\left\{u_k^{(r)}(x)a^{(r)}(k) + v_k^{(r)}(x)\,b^{(r)\dagger}(k)\right\},$$

and make use of the closure conditions, we obtain

$$\phi(x) = \int d\sigma_\lambda(x')\,d(\partial)\,\Delta(x-x')\,\Gamma_\lambda(\partial',-\overleftarrow{\partial'})\,\phi(x'), \tag{VI.38}$$

and a similar relation for $\bar{\phi}(x)$. Note that for the Klein-Gordon case,

$$\left\{\begin{array}{l} d(\partial) = 1 \\ \Gamma_\lambda(\partial,-\overleftarrow{\partial}) = \partial_\lambda - \overleftarrow{\partial}_\lambda \end{array}\right\},$$

(VI.38) reduces to (II.17), and for the Dirac case,

$$\left\{\begin{array}{l} d(\partial) = -(\gamma\partial - m) \\ \Gamma_\lambda(\partial,-\overleftarrow{\partial}) = -\gamma_\lambda \end{array}\right\},$$

it reduces to (II.32), as one should expect.

Now we indicate a few applications briefly. One easily shows that

$$\langle 0|\phi_\alpha(x)\bar{\phi}_\beta(y)|0\rangle = \sum_r \int d^3k\, u_{k\alpha}^{(r)}(x)\,\bar{u}_{k\beta}^{(r)}(y)$$

$$= id_{\alpha\beta}(\partial)\,\Delta^{(+)}(x-y) \tag{VI.39a}$$

and, similarly,

$$\langle 0|\bar{\phi}_\beta(y)\phi_\alpha(x)|0\rangle = -i\rho\, d_{\alpha\beta}(\partial)\Delta^{(-)}(x-y). \tag{VI.39b}$$

From these

$$\langle 0|\phi_\alpha(x)\bar{\phi}_\beta(y) + \rho\,\bar{\phi}_\beta(y)\phi_\alpha(x)|0\rangle = d_{\alpha\beta}(\partial)\Delta^{(1)}(x-y), \tag{VI.40a}$$

where

$$\Delta^{(1)}(x) = i\left\{\Delta^{(+)}(x) - \Delta^{(-)}(x)\right\}. \tag{VI.40b}$$

Next

$$\begin{aligned}
&\langle 0|T\left[\phi_\alpha(x), \bar{\phi}_\beta(y)\right]|0\rangle \\
&\equiv \theta(x_o-y_o)\langle 0|\phi_\alpha(x)\bar{\phi}_\beta(y)|0\rangle + \rho\,\theta(y_o-x_o)\langle 0|\bar{\phi}_\beta(y)\phi_\alpha(x)|0\rangle \\
&= i\theta(x_o-y_o)\, d_{\alpha\beta}(\partial)\Delta^{(+)}(x-y) - i\theta(y_o-x_o)\, d_{\alpha\beta}(\partial)\Delta^{(-)}(x-y) \\
&= i d_{\alpha\beta}(\partial)\Delta^{(+)}(x-y)\,\theta(x_o-y_o) - i d_{\alpha\beta}(\partial)\Delta^{(-)}(x-y)\,\theta(y_o-x_o) \\
&\quad + i\left[\theta(x_o-y_o),\, d_{\alpha\beta}(\partial)\right]\Delta^{(+)}(x-y) - i\left[\theta(y_o-x_o),\, d_{\alpha\beta}(\partial)\right]\Delta^{(-)}(x-y) \\
&= i d_{\alpha\beta}(\partial)\Delta_c(x-y) + N_{\alpha\beta}(x-y),
\end{aligned} \tag{VI.41a}$$

where

$$\Delta_c(x) = \theta(x_o)\Delta^{(+)}(x) - \theta(-x_o)\Delta^{(-)}(x), \tag{VI.41b}$$

and

$$\begin{aligned}
N_{\alpha\beta}(x) &= i\left[\theta(x_o), d_{\alpha\beta}(\partial)\right]\Delta^{(+)}(x) - i\left[\theta(-x_o), d_{\alpha\beta}(\partial)\right]\Delta^{(-)}(x) \\
&= \frac{i}{2}\left[\varepsilon(x_o),\, d_{\alpha\beta}(\partial)\right]\Delta(x).
\end{aligned} \tag{VI.41c}$$

It can be shown that[11]

$$N_{\alpha\beta}(x) \propto \delta^4(x). \tag{VI.41c'}$$

Recall that the K-G divisor for the Duffin-Kemmer case is

$$d(\partial) = -\left[\frac{1}{m}(\Box - m^2) + Q_1 - \frac{1}{m}Q_1^2\right].$$

It turns out that, in general, one has

$$d(\partial) = d'(\partial) + d''(\partial)(\Box - m^2). \tag{VI.42}$$

Then the commutation relation may be written

$$\begin{aligned}\left[\phi(x), \bar{\phi}(y)\right]_{\pm} &= id(\partial)\Delta(x-y) \\ &= id'(\partial)\Delta(x-y),\end{aligned}$$

that is to say, $d''(\partial)$ does not contribute to the commutation relation. However, the Green function is

$$\begin{aligned}\langle 0|T\left[\phi(x)\bar{\phi}(y)\right]|0\rangle &= id(\partial)\Delta_c(x-y) + N(x-y) \\ &= id'(\partial)\Delta_c(x-y) + N'(x-y),\end{aligned}$$

where

$$N'(x) = \frac{i}{2}\left[\epsilon(x_o),\, d'(\partial)\right]\Delta(x-y).$$

In the Yang-Feldman formalism the interaction Hamiltonian has the form

$$\mathcal{H}_{int} = \mathcal{L}_{int} + (\text{term depending upon the normal } n_\mu).$$

In the S-matrix the normal-dependent term of the interaction Hamiltonian cancels N but not N'. In other words, we can neglect the normal-dependent term and N at the same time. In the Feynman rule we have to put

$$id(\partial)\Delta_c(x-y)$$

and not

$$id'(\partial)\Delta_c(x-y),$$

because, of the two, only the former is the appropriate Green function. We refer the reader to the works of Katayama[11] and Umezawa[2] for the details.

Now we shall prove the uniqueness of P_μ and $M_{\mu\nu}$ when the commutation relation is given.

C. Uniqueness of P_μ and $M_{\mu\nu}$.

Since the commutation relation

$$\left[\phi(x), \bar{\phi}(y)\right]_\pm = id(\partial)\Delta(x-y)$$

is relativistically invariant, i.e.,

$$\left[\phi'(x'), \bar{\phi}'(y')\right]_\pm = id(\partial')\Delta(x'-y'),$$

it implies that

$$\left[\phi'(x), \bar{\phi}'(y)\right]_\pm = id(\partial)\Delta(x-y). \tag{VI.43}$$

Now we can write

$$\phi'(x) = G_L^{-1}\phi(x)G_L, \tag{VI.44}$$

where for an infinitesimal Lorentz transformation we have

$$G_L = 1 + \frac{i}{2} M_{\mu\nu}\epsilon_{\mu\nu} - iP_\mu\epsilon_\mu, \tag{VI.45a}$$

$$G_L^{-1} = 1 - \frac{i}{2} M_{\mu\nu}\epsilon_{\mu\nu} + iP_\mu\epsilon_\mu . \tag{VI.45b}$$

From (VI.44) and (VI.45) we get

$$\phi'(x) = \phi(x) + \frac{i}{2}\left[\phi(x), M_{\mu\nu}\right]\epsilon_{\mu\nu} - i\left[\phi(x), P_\mu\right]\epsilon_\mu . \tag{VI.46}$$

On the other hand,

$$\phi'(x') = \left[1 + \frac{i}{2} S_{\mu\nu}\epsilon_{\mu\nu}\right]\phi(x), \tag{VI.47a}$$

where

$$x'_\mu = x_\mu + \epsilon_{\mu\nu}x_\nu + \epsilon_\mu . \tag{VI.47b}$$

This yields

$$\phi'(x) = \phi(x) + \frac{1}{2}\ell_{\mu\nu}\phi(x)\epsilon_{\mu\nu} - \partial_\mu\phi(x)\epsilon_\mu . \tag{VI.48}$$

Comparison of (VI.46) and (VI.48) gives

$$-i\partial_\mu \phi(x) = \left[\phi(x), P_\mu\right], \tag{VI.49}$$

and

$$-i\ell_{\mu\nu}\phi(x) = \left[\phi(x), M_{\mu\nu}\right]. \tag{VI.50}$$

From the integrability condition of Lorentz transformations, we have

$$\left[P_\mu, P_\nu\right] = 0, \tag{VI.51}$$

$$\left[P_\lambda, M_{\mu\nu}\right] = i\left(\delta_{\lambda\nu}P_\mu - \delta_{\lambda\mu}P_\nu\right) \tag{VI.52}$$

and

$$\left[M_{\lambda\kappa}, M_{\mu\nu}\right] = i\left(\delta_{\lambda\nu}M_{\mu\kappa} + \delta_{\lambda\mu}M_{\kappa\nu} + \delta_{\kappa\nu}M_{\lambda\mu} + \delta_{\kappa\mu}M_{\nu\lambda}\right). \tag{VI.53}$$

Now we show that P_μ and $M_{\mu\nu}$ are unique when the commutation relation is given. The proof of the theorem is based on the assumption that $\phi(x)$ forms an irreducible ring, i.e., any quantity that commutes with $\phi(x)$ and $\overline{\phi}(x)$ at all times is a C-number.

If there are two P_μ's, let us call them $P_\mu^{(1)}$ and $P_\mu^{(2)}$, then

$$-i\partial_\mu \phi(x) = \left[\phi(x), P_\mu^{(1)}\right],$$

$$-i\partial_\mu \phi(x) = \left[\phi(x), P_\mu^{(2)}\right],$$

so that

$$0 = \left[\phi(x), P_\mu^{(1)} - P_\mu^{(2)}\right],$$

which implies that

$$P_\mu^{(1)} - P_\mu^{(2)} = C_\mu, \tag{VI.54}$$

where C_μ is a C-number.

Similarly, if there are two $M_{\mu\nu}$'s, they also differ by a C-number

$$M_{\mu\nu}^{(1)} - M_{\mu\nu}^{(2)} = C_{\mu\nu}. \tag{VI.55}$$

Now

$$\left[M_{\lambda\kappa}^{(1)}, M_{\mu\nu}^{(1)}\right] - \left[M_{\lambda\kappa}^{(2)}, M_{\mu\nu}^{(2)}\right]$$

$$= i\left(\delta_{\lambda\nu} C_{\mu\kappa} + \delta_{\lambda\mu} C_{\kappa\nu} + \delta_{\kappa\nu} C_{\lambda\mu} + \delta_{\kappa\mu} C_{\nu\lambda}\right), \qquad \text{(VI.56)}$$

and, again,

$$\left[M_{\lambda\kappa}^{(1)}, M_{\mu\nu}^{(1)}\right] - \left[M_{\lambda\kappa}^{(2)}, M_{\mu\nu}^{(2)}\right]$$

$$\equiv \left[M_{\lambda\kappa}^{(1)} - M_{\lambda\kappa}^{(2)}, M_{\mu\nu}^{(1)}\right] + \left[M_{\lambda\kappa}^{(2)}, M_{\mu\nu}^{(1)} - M_{\mu\nu}^{(2)}\right]$$

$$= \left[C_{\lambda\kappa}, M_{\mu\nu}^{(1)}\right] + \left[M_{\lambda\kappa}^{(2)}, C_{\mu\nu}\right]$$

$$= 0. \qquad \text{(VI.57)}$$

From (VI.56) and (VI.57) one concludes that

$$C_{\mu\nu} = 0, \qquad \text{(VI.58)}$$

so that $M_{\mu\nu}$ is unique. Next,

$$\left[P_\lambda^{(1)}, M_{\mu\nu}\right] = i\left(\delta_{\nu\lambda} P_\mu^{(1)} - \delta_{\lambda\mu} P_\nu^{(1)}\right)$$

and

$$\left[P_\lambda^{(2)}, M_{\mu\nu}\right] = i\left(\delta_{\nu\lambda} P_\mu^{(2)} - \delta_{\lambda\mu} P_\nu^{(2)}\right).$$

Subtracting the second relation from the first, we get

$$0 = \left[C_\mu, M_{\mu\nu}\right] = i\left(\delta_{\nu\lambda} C_\mu - \delta_{\lambda\mu} C_\nu\right),$$

which implies

$$C_\mu = 0, \qquad \text{(VI.59)}$$

so that P_μ is also unique.

Appendix A. Notation and Conventions

In this course we use the so-called natural units in which $\hbar = c = 1$. Greek letters, when employed as tensor indices, run from 1 to 4, while Latin letters similarly employed signify space components only and run from 1 to 3. In our convention the fourth component of a four-vector is pure imaginary, thus

$$x_\mu = (\underset{\sim}{x}, ix_o), \quad \text{i.e.,} \quad x_4 = ix_o,$$

where x_o is real. Unless explicitly stated otherwise, sum over repeated indices is to be understood, e.g.,

$$P_i Q_i = \underset{\sim}{P}\cdot\underset{\sim}{Q},$$

$$P_\mu Q_\mu = \underset{\sim}{P}\cdot\underset{\sim}{Q} + P_4 Q_4 = \underset{\sim}{P}\cdot\underset{\sim}{Q} - P_o Q_o.$$

As typical examples of volume elements, we have

$$d\underset{\sim}{x} = dx_1 dx_2 dx_3, \quad d^4x = dx_o d\underset{\sim}{x}, \quad d^4p = dp_o d\underset{\sim}{p}, \quad \text{etc.}$$

The symbol ∂_μ stands for the differential operator

$$\partial_\mu = \frac{\partial}{\partial x_\mu} = \left(\underset{\sim}{\nabla}, -1\frac{\partial}{\partial x_o}\right),$$

and it acts on functions standing to its right, while $\overleftarrow{\partial}_\mu$ is a similar operator which acts on functions to its left. Thus

$$f(x)\overleftarrow{\partial}_\mu g(x) = \left[\partial_\mu f(x)\right] g(x),$$

$$f(x)\left(\partial_\mu + \overleftarrow{\partial}_\mu\right) g(x) = \partial_\mu\left[f(x)\, g(x)\right],$$

for arbitrary functions $f(x)$ and $g(x)$. The four-dimensional Laplace operator, called the D'Alembertian, is written

$$\Box = \partial_\mu\partial_\mu = \nabla^2 - \frac{\partial^2}{\partial x_o^2}.$$

The symbol $g_{\mu\nu}$, whenever it appears, stands for the 4×4 matrix:

$$g_{\mu\nu} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

A space-like three-dimensional hyper-surface in Minkowski space (for brevity we simply call it a space-like surface) is denoted by σ, or by $\sigma(x)$ when we wish to emphasize that this surface passes through the specific point x. We denote by $d\sigma_\mu(x)$ the four-vector differential surface area at the point x and define it to be

$$d\sigma_\mu(x) = \left(dx_o dx_2 dx_3,\ dx_o dx_1 dx_3,\ dx_o dx_1 dx_2,\ -i dx_1 dx_2 dx_3\right)$$

$$= \frac{-i}{3!} \left|\epsilon_{\mu\nu\lambda\rho}\right| dx_\nu dx_\lambda dx_\rho,$$

where the Levi-Civita symbol $\epsilon_{\mu\nu\lambda\rho}$ has the property

$$\epsilon_{\mu\nu\lambda\rho} = \begin{cases} 1 & \text{if } (\mu\nu\lambda\rho) \text{ is an even permutation of (1234),} \\ -1 & \text{if } (\mu\nu\lambda\rho) \text{ is an odd permutation of (1234),} \\ 0 & \text{otherwise.} \end{cases}$$

A convenient way of remembering the explicit form of $d\sigma_\mu(x)$ is to think of it as being d^4x/dx_μ. For example,

$$d\sigma_2(x) = \frac{d^4x}{dx_2} = \frac{dx_o dx_1 dx_2 dx_3}{dx_2} = dx_o dx_1 dx_3.$$

The symbol n_μ is used to denote a time-like unit vector with a positive time-component:

$$n_\mu n_\mu = -1, \qquad n_o > 1.$$

It is convenient to introduce here the notion of a functional derivative and give a couple of useful relations involving the same. We denote by $F[\sigma]$ some functional of a space-like surface σ. Let σ' be another space-like surface which overlaps with σ everywhere except in the infinitesimal vicinity of the point x. Let $\delta V(x)$ be the volume enclosed between the surfaces σ and σ'. The functional derivative of $F[\sigma]$ at the point x is then defined by

$$\frac{\delta F[\sigma]}{\delta \sigma(x)} = \lim_{\delta V(x) \to 0} \frac{F[\sigma'] - F[\sigma]}{\delta V(x)}. \tag{A.1}$$

If $F[\sigma]$ is given by

$$F[\sigma] = \int_{\sigma} d\sigma_\lambda(x)\, f_\lambda(x), \tag{A.2a}$$

then, by making use of Gauss' theorem, one can easily show that

$$\frac{\delta F[\sigma]}{\delta \sigma(x)} = \partial_\lambda f_\lambda(x). \tag{A.2b}$$

Two other useful relations are

$$\frac{\delta}{\delta\sigma(x)} \int d\sigma_\lambda(x)\, g(x) = \partial_\lambda g(x), \tag{A.3}$$

and

$$\int d\sigma_\lambda(x)\partial_\mu f(x) = \int d\sigma_\mu(x)\partial_\lambda f(x). \tag{A.4}$$

Proof of (A.4):

$$\frac{\delta}{\delta\sigma(x)}\Big[\int d\sigma_\lambda(x)\partial_\mu f(x) - \int d\sigma_\mu \partial_\lambda f(x)\Big] = \partial_\lambda\partial_\mu f(x) - \partial_\mu\partial_\lambda f(x)$$

$$= 0,$$

so that (A.4) is independent of the choice of the space-like surface σ. Choose t = constant surface:

$$d\sigma_\lambda(x) = (o,o,o,-id^3x).$$

Now we consider all possible choices of λ and μ.

$\lambda = i,\ \mu = j$:

$$\int d\sigma_\lambda \partial_\mu f - \int d\sigma_\mu \partial_\lambda f = 0 - 0 = 0.$$

$\lambda = i,\ \mu = 4$:

$$\int d\sigma_\lambda \partial_\mu f - \int d\sigma_\mu \partial_\lambda f$$

$$= 0 \quad + i \int d^3x \partial_i f$$

$$= 0,$$

since f is assumed to vanish at very large distances. Similarly,

$$\lambda = 4, \quad \mu = j:$$

$$\int d\sigma_\lambda \partial_\mu f - \int d\sigma_\mu \partial_\lambda f$$

$$= -i \int d^3x \, \partial_j f - 0$$

$$= 0,$$

and finally for $\mu = \lambda = 4$ both sides of (A.4) are equal. This proves (A.4) for arbitrary μ, ν.

Our matrix notation is as follows:

$\dagger$: Hermitian conjugation,

$*$: complex conjugation,

t : transposition.

As we shall have occasion to do contour integrals in the complex plane, let us recall the well-known Cauchy formula

$$\frac{1}{2\pi i} \oint_C \frac{f(z)}{z - a} \, dz = f(a), \tag{A.5}$$

where the symbols have their usual meanings.

Appendix B. Explicit Construction of Wave Functions

In this appendix we shall discuss how to construct the wave functions for a particular case. Recall that $u_k^{(r)}(x)$ satisfies

$$\Lambda(\partial) \, u_k^{(r)}(x) = 0, \tag{B.1}$$

where r takes the values $1, \ldots, 2s+1$. The normalization condition is

$$-i \int d\sigma_\lambda(x) \, \bar{u}_k^{(r)}(x) \Gamma_\lambda(\partial, -\overleftarrow{\partial}) \, u_{k'}^{(r')}(x) = \delta_{rr'} \delta(\underset{\sim}{k} - \underset{\sim}{k}'), \tag{B.2}$$

and the closure condition is

$$\sum_r \int d^3k\, u_k^{(r)}(x)\, \bar{u}_k^{(r)}(y) = id(\partial)\Delta^{(+)}(x-y). \tag{B.3}$$

Once $u(x)$ has been found, then making use of

$$v_k^{(r)}(x) = C u_k^{(r)*}(x), \tag{B.4a}$$

$$\bar{u}_k^{(r)}(x) = u_k^{(r)\dagger}(x)\eta \tag{B.4b}$$

$$\bar{v}_k^{(r)}(x) = v_k^{(r)\dagger}(x)\eta \tag{B.4c}$$

one can obtain v, $\bar{u}$ and $\bar{v}$.

Let us put

$$u_k^{(r)}(x) = u^{(r)}(k) f_k(x), \tag{B.5a}$$

$$f_k(x) = \frac{1}{(2\pi)^{3/2}} e^{ikx}. \tag{B.5b}$$

The normalization condition now becomes

$$\bar{u}^{(r)}(k)\Gamma_4(ik, ik)u^{(r')}(k) = -\delta_{rr'}, \tag{B.6}$$

while the closure condition takes the form

$$\sum_r u^{(r)}(k)\bar{u}^{(r)}(k) = \frac{d(ik)}{2\omega(k)}, \tag{B.7}$$

where k is now on the mass-shell, $k^2+m^2=0$, and the equations of motion are

$$\Lambda(ik)u^{(r)}(k) = 0, \tag{B.8a}$$

$$\bar{u}^{(r)}(k)\Lambda(ik) = 0. \tag{B.8b}$$

The equations (B.8) suggest writing

$$u^{(r)}(k) = d(ik)\xi^{(r)}(k), \tag{B.9a}$$

$$\bar{u}^{(r)}(k) = \bar{\xi}^{(r)}(k)d(ik), \tag{B.9b}$$

with

$$\bar{\xi}^{(r)}(k) = \xi^{(r)\dagger}(k)\eta, \tag{B.9c}$$

which follows from the relation

$$\left[d(ik)\eta^{-1}\right]^{\dagger} = d(ik)\eta^{-1}.$$

It should be clear that the wave functions $u^{(r)}(k)$ and $\bar{u}^{(r)}(k)$ as defined in (B.9) satisfy the equations (B.8) because of the mass-shell relation

$$\Lambda(ik)\,d(ik) = d(ik)\Lambda(ik) = -(k^2+m^2) = 0. \tag{B.10}$$

Since

$$[\Lambda(ik),h] = 0, \tag{B.11}$$

we have

$$[d(ik),h] = 0. \tag{B.12}$$

Therefore, if

$$h\xi^{(r)}(k) = h^{(r)}\xi^{(r)}(k), \tag{B.13}$$

then we have

$$hu_k^{(r)}(x) = h^{(r)}u_k^{(r)}(x). \tag{B.14}$$

In other words, in constructing the functions $\xi^{(r)}(k)$ one should make sure that they satisfy (B.13) so that (B.14) may hold true.

If we multiply (B.6) by $u^{(r)}(k)$ from the left and by $\bar{u}^{(r')}(k)$ from the right and sum over r and r', we get

$$d(ik)\Gamma_4(ik,ik)\,d(ik) = -2\omega(k)\,d(ik), \tag{B.15}$$

where use has been made of the closure condition (B.7). With the help of (B.15) we can now write the normalization condition in terms of $\xi^{(r)}(k)$ and $\bar{\xi}^{(r)}(k)$:

$$
\begin{aligned}
-\delta_{rr'} &= \bar{u}^{(r)}(k)\,\Gamma_4(ik,ik)\,u^{(r')}(k) \\
&= \bar{\xi}^{(r)}(k)\,d(ik)\,\Gamma_4(ik,ik)\,d(ik)\,\xi^{(r')}(k) \\
&= -2\omega(k)\,\bar{\xi}^{(r)}(k)\,d(ik)\,\xi^{(r')}(k),
\end{aligned}
$$

which we rewrite as

$$
\sqrt{2\omega(k)}\,\bar{\xi}^{(r)}(k)\,d(ik)\,\xi^{(r')}(k)\sqrt{2\omega(k)} = \delta_{rr'}. \tag{B.16}
$$

The closure condition reads

$$
\begin{aligned}
d(ik) &= 2\omega(k)\sum_r u^{(r)}(k)\,\bar{u}^{(r)}(k) \\
&= d(ik)\sqrt{2\omega(k)}\sum_r \xi^{(r)}(k)\,\bar{\xi}^{(r)}(k)\sqrt{2\omega(k)}\,d(ik).
\end{aligned} \tag{B.17}
$$

Recall that on the mass-shell $d(ik)$ is a projection operator

$$
d(ik)\,d(ik) = a\,d(ik), \tag{B.18}
$$

where a is a constant independent of k. Referring back to (IV.23) and (IV.25) we see that this statement is certainly true for the cases we studied. We take it to be true in general. Then (B.18) implies that a similarity transformation exists such that

$$
S^{-1}_{\alpha\sigma}(k)\,d_{\sigma\rho}(ik)S_{\rho\beta}(k) = \delta_{\alpha\beta}a^{(\beta)}, \tag{B.19a}
$$

where

$$
a^{(\beta)} = \begin{cases} a & \text{for } \beta = 1,2,\ldots,2S+1; \\ 0 & \text{for } \beta > 2S+1. \end{cases} \tag{B.19b}
$$

If we define a matrix $\hat{I}$ such that

$$
\hat{I}_{\alpha\beta} = \begin{cases} 1 & \text{if } \alpha = \beta = 1,2,\ldots,2S+1; \\ 0 & \text{otherwise,} \end{cases}
$$

we can rewrite (B.19) as

$$
S^{-1}(k)\,d(ik)\,S(k) = a\hat{I}, \tag{B.20}
$$

or

$$d(ik) = a\, S(k)\, \hat{I}\, S^{-1}(k), \tag{B.21a}$$

so that

$$d_{\alpha\beta}(ik) = a \sum_{r=1}^{2s+1} S_{\alpha r}(k)\, S^{-1}_{r\beta}(k). \tag{B.21b}$$

Again, if we multiply (B.21a) by $d(ik)$ from the right as well as the left and make use of (B.18), we get

$$a^2 d(ik) = a\, d(ik)\, S(k)\, \hat{I}\, S^{-1}(k)\, d(ik),$$

or

$$d(ik) = d(ik)\, \frac{S(k)}{\sqrt{a}}\, \hat{I}\, \frac{S^{-1}(k)}{\sqrt{a}}\, d(ik). \tag{B.22}$$

Now by comparing (B.20) to (B.16), and (B.22) to (B.17), we see that we can take

$$\xi^{(r)}_{\alpha}(k) = \frac{S_{\alpha r}(k)}{\sqrt{2a\omega(k)}}, \tag{B.23a}$$

$$\bar{\xi}^{(r)}_{\alpha}(k) = \frac{S^{-1}_{r\alpha}(k)}{\sqrt{2a\omega(k)}}. \tag{B.23b}$$

Since we have $\bar{\xi} = \xi^{\dagger}\eta$, the relation

$$\sum_{\alpha} S^{\dagger}_{r\alpha}\eta_{\alpha\beta} = S^{-1}_{r\beta}, \tag{B.24}$$

or, equivalently,

$$\hat{I}\, S^{\dagger}\eta = \hat{I}\, S^{-1}, \tag{B.24'}$$

must be satisfied. So it must be checked as an intermediate step. In view of (B.13) we also need

$$h_{\alpha\beta}S_{\beta r} = h^{(r)}S_{\alpha r}. \tag{B.25}$$

Because of the $(2s+1)$-fold degeneracy we can always rearrange S so that it satisfies (B.25).

Making use of (B.20) we can write $u^{(r)}(k)$ and $\bar{u}^{(r)}(k)$ in terms of S and d as follows:

$$u_\alpha^{(r)}(k) = d_{\alpha\beta}(ik)\,\frac{S_{\beta r}(k)}{\sqrt{2a\,\omega(k)}}$$

$$= a\,\frac{S_{\alpha r}(k)}{\sqrt{2a\,\omega(k)}}$$

$$= \sqrt{\frac{a}{2\omega(k)}}\,S_{\alpha r}(k)\ , \qquad \text{(B.26a)}$$

and

$$\bar{u}_\alpha^{(r)}(k) = \frac{S_{r\beta}^{-1}(k)}{\sqrt{2a\,\omega(k)}}\,d_{\beta\alpha}(ik)$$

$$= a\cdot\frac{S_{r\alpha}^{-1}(k)}{\sqrt{2a\,\omega(k)}}$$

$$= \sqrt{\frac{a}{2\omega(k)}}\,S_{r\alpha}^{-1}(k)\,. \qquad \text{(B.26b)}$$

As a concrete example of the application of our method as prescribed above, consider the Dirac case. Here (we use Pauli representation)

$$\Lambda(ik) = -(i\gamma k + m)\,, \qquad \text{(B.27a)}$$

$$d(ik) = -(i\gamma k - m)\,. \qquad \text{(B.27b)}$$

One finds

$$d(ik)\,d(ik) = 2m\,d(ik)\,, \qquad \text{(B.28a)}$$

on the mass-shell. So we have

$$a = 2m\,. \qquad \text{(B.28b)}$$

Recall that the boost

$$L(k) = \sqrt{\frac{\omega(k)+m}{2m}} - i\gamma_i\gamma_4 k_i\,\frac{1}{\sqrt{2m(\omega(k)+m)}}\,, \qquad \text{(B.29)}$$

has the property

$$L^{-1}(k)\, d(ik)\, L(k) = \lim_{k\to 0} d(ik) = m(1+\gamma_4)$$

$$= 2m\left(\frac{1+\gamma_4}{2}\right). \qquad \text{(B.30)}$$

In the Pauli representation

$$\hat{I} = \begin{pmatrix} 1 & & \bigcirc & \\ & 1 & & \\ \bigcirc & & 0 & \\ & & & 0 \end{pmatrix} = \frac{1+\gamma_4}{2}. \qquad \text{(B.31)}$$

Keeping in mind (B.28b) and (B.31), if we compare (B.30) and (B.20) it would seem that we can identify $L(k)$ with $S(k)$. Since

$$\eta = \gamma_4 \qquad \text{(B.32)}$$

and

$$L^{-1}(k) = \sqrt{\frac{\omega(k)+m}{2m}} + i\gamma_i\gamma_4 k_i \frac{1}{\sqrt{2m(\omega(k)+m)}}, \qquad \text{(B.33a)}$$

$$L^{\dagger}(k) = \sqrt{\frac{\omega(k)+m}{2m}} - i\gamma_i\gamma_4 k_i \frac{1}{\sqrt{2m(\omega(k)+m)}}, \qquad \text{(B.33b)}$$

it is clear that the condition (B.24'), namely,

$$\hat{I}\, L^{\dagger}\eta = \hat{I}\, L^{-1}, \qquad \text{(B.34)}$$

is satisfied. So $L(k)=S(k)$ is the correct choice.

According to (B.26) we have, therefore,

$$u_\alpha^{(r)}(k) = \sqrt{\frac{m}{\omega(k)}}\, L_{\alpha r}(k)$$

$$= \sqrt{\frac{m}{\omega(k)}} \cdot \frac{1}{\sqrt{2m(\omega(k)+m)}} \left[\omega + m - i\gamma_i\gamma_4 k_i\right]_{\alpha r}$$

$$= \frac{1}{\sqrt{2\omega(k)(\omega(k)+m)}} \left[\gamma_4\omega - i\gamma_i k_i + m\right]_{\alpha r}$$

$$= \frac{1}{\sqrt{2\omega(k)(\omega(k)+m)}} \left[-i\gamma k + m\right]_{\alpha r},$$

where

$$\alpha = 1, \ldots, 4,$$

$$r = 1, 2.$$

Note that above we replaced $\omega - i\gamma_i\gamma_4 k_i$ by $\gamma_4\omega - i\gamma_i k_i$ to make our expression covariant. Since we restrict r to the values 1 and 2, this is justified because

$$(\gamma_4)_{\alpha r} = \delta_{\alpha r} \quad \text{if} \quad r = 1, 2.$$

Similarly, one finds

$$\bar{u}^{(r)}_{\alpha}(k) = \frac{1}{\sqrt{2\omega(k)(\omega(k)+m)}}\left[-i\gamma k + m\right]_{r\alpha} .$$

Then

$$v^{(r)}_{\alpha}(k) = C_{\alpha\beta}\, u^{(r)*}_{\beta}(k)$$

$$= \frac{1}{\sqrt{2\omega(k)(\omega(k)+m)}}\left[(i\gamma k + m)C\right]_{\alpha r} ,$$

$$\bar{v}^{(r)}_{\alpha}(k) = -\frac{1}{\sqrt{2\omega(k)(\omega(k)+m)}}\left[C^{-1}(i\gamma k + m)\right]_{r\alpha} .$$

We refer the reader to the paper by Takahashi[12] for more examples.

References

1. E. M. Corson, Introduction to tensors, spinors and relativistic equations (Blackie & Son, London-Glasgow, 1953).
2. H. Umezawa, Quantum Field Theory (North-Holland Publishing Co., Amsterdam, 1956).
3. S. Kamefuchi and Y. Takahashi, Nuovo Cimento 44, 1(1966).
4. S. C. Bhargava and H. Watanabe, Nucl. Phys. 87, 273(1966).
5. A. Kawakami and S. Kamefuchi, Nuovo Cimento A48, 239 (1967).
6. K. K. Gupta, Proc. Ind. Acad. Sci. 35, 255(1952).
7. S. N. Gupta, Phys. Rev. 95, 1334(1954).
8. J. Schwinger, Phys. Rev. 82, 914(1951).
9. D. Lurie, Y. Takahashi and H. Umezawa, J. Math. Phys. 7, 1478 (1966).
10. S. Kamefuchi and Y. Takahashi, Nucl. Phys. 36, 177 (1962).
11. Y. Katayama, Progr. Theoret. Phys. 10, 31(1953).
12. Y. Takahashi, Nucl. Phys. 87, 481(1966).

FUNCTIONAL INTEGRALS IN QUANTUM FIELD THEORY AND RELATED TOPICS[†]

Jan Tarski[‡]
Columbia University
New York, New York

Contents

[†]Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

[‡]Present address: Laboratoire de Physique Théorique, Faculté des Sciences, Université de Paris, Paris, France.

I. Introduction

A. Preliminaries.

These notes are from a series of lectures which were devoted to functional methods in physics. We emphasize two functional integrals, namely, the Wiener Integral and the commutative Hilbert space integral. We also include functional derivatives and other examples of functional integrals.

The applications emphasize quantum field theory. We can offer two reasons for this choice. The first is the interest of the author in such applications. The second is the fact that quantum field theory provides a rich variety of situations where functional methods can be applied.

We may note in this connection that we listed several applications in the table of contents. However, a number of other short applications are included as well.

Let us make a few brief historical remarks. Significant early work on functionals was done by Volterra beginning in the 1880's. The Wiener integral dates back to 1923. The commutative Hilbert space integral was constructed in the 1950's by Friedrichs and Shapiro and by Segal. As we shall see, these two integrals are related in a simple way.

We mention the following as some of the more prominent contributions to the development of functional methods in quantum field theory: (1) Feynman's path and history integrals, (2) Schwinger's external field and variational techniques, (3) Segal's investigations of free fields and of canonical systems, and (4) Symansik's analysis of generating functionals and his more recent work on functional techniques in Euclidean field theory.

The essential prerequisite for these notes, from the physical side, is some acquaintance with quantum field theory. On the mathematical side, we presuppose general background rather than specific knowledge. (Some references to measure theory are made, but they are generally not crucial for following the arguments.)

As to the rigor, the material covers a variety of topics, and the discussion is rigorous in some places but heuristic in others. (For some of the topics, like the time-ordered generating functional, a rigorous discussion is as yet impossible.) Moreover, we emphasize manipulations and applications, while long and detailed proofs are avoided.

We made no attempt to be complete or consistent in citing references. In a number of cases we gave only one reference, which is neither the original work nor a comprehensive recent review. However, much of the material can now be considered classical, and we felt that this justified in part the aforementioned neglect.

The author gratefully acknowledges the help of Dr. John Lowenstein in the preparation of these notes. The research represented by these notes was carried out at Columbia University (with support from the Atomic Energy Commission) and also during the Summer Institute at the University of Colorado, and at the Université de Paris. The author would like to express his appreciation for the hospitality which he received.

B. Generalities Concerning Functionals. Notation.[1]

We are all familiar with the real- and complex-valued functions on the Euclidean n-space R^n:

$$y = f(x) \qquad x \in R^n, \quad y \in R^1 \text{ or } C^1.$$

Let us generalize this by letting x range over a linear space $\mathfrak{L}$ of functions, or over some related space or set (e.g., a space of distributions, or an abstract linear space). Then f will be a functional.

We will usually take for $\mathfrak{L}$ a linear space of functions defined on a region of R^n. The usual notations for a functional will be

$$F(x) \text{ or } F(\cdot) \text{ or } F, \ f\{\eta\}, \text{ etc., where } x, \eta \in \mathfrak{L}.$$

Let us now give a few examples of functionals.

(a) $$F_1\{\eta\} = \int d^n u\, w(u)\eta(u),$$

where the functions w and η must be such that the product is integrable (e.g., in the sense of Riemann or Lebesque).

We will discuss the restrictions like these (on w and η) more fully in Section II. For the present, such questions will be largely ignored.

(b) $$F_2\{\eta\} = \exp\left[-\frac{1}{2}\int d^n u\, d^n v\, K(u,v)\eta(u)\eta(v)\right],$$

(c) $$F_3\{\eta\} = \eta(v_o),$$

(d) $$F_4\{\eta\} = E_o\{\eta\},$$

where $E_o\{\eta\}$ is the ground state energy of a hydrogen atom in the external potential η.

For investigating various properties of a function, a basic tool is the derivative. For functionals, the analogous concept is that of a functional derivative. Let us approximate the functional F_1 above by

a discrete sum,

$$F_1\{\eta\} \approx \overline{F}_1\{\eta\} = \sum w(u_i)\eta(u_i)(\Delta u_i).$$

Then

$$\frac{1}{\Delta u_i}\frac{\partial \overline{F}_1}{\partial[\eta(u_i)]} = w(u_i),$$

and for the functional F_1, we set

$$\delta F_1/\delta\eta(u) = w(u).$$

More generally, one can define

$$\frac{\delta F\{\eta\}}{\delta\eta(u)} = \frac{d}{d\epsilon} F\{\eta + \epsilon\delta_u\}\Big|_{\epsilon=0}, \tag{I.1}$$

where $\delta_u(v)=\delta(u-v)$, the Dirac delta function. The usual rules for differentiation remain valid for functional derivatives, and we find, e.g.,

$$\delta F_2/\delta\eta(u) = -\Big[\int d^n v\, K(u,v)\eta(v)\Big]F_2,$$

if K is symmetric. Next we find

$$\delta\eta(v)/\delta\eta(u) = \delta(u-v). \tag{I.2}$$

We can also set

$$\frac{\delta}{\delta\eta(u)}\left[\frac{\partial\eta}{\partial u^i}(v)\right] = \frac{\partial}{\partial u^i}\delta(u-v),$$

provided the contribution from the endpoints (in integration by parts) vanishes.

A few words about the notation may be appropriate. We use a variety of different notations. However, the following points are generally followed.

(i) The points in R^n are denoted by u, v, etc., sometimes by $\vec{u}$, etc. In particular, for a relativistic four-vector we write $u=(u^o,\vec{u})$. We use x, y, also η, χ, f, etc., to denote functions, vectors in a Hilbert space, and the like.

(ii) Function spaces are denoted by capital script letters.

(iii) The coordinates in R^n and in Hilbert spaces are denoted by superscripts, e.g.,

$$u = (u^1, \ldots, u^n), \quad x = (x^1, x^2, \ldots).$$

Subscripts normally indicate distinct points or vectors.

(iv) A scalar quantum field under consideration is denoted by φ, and a free field by $\varphi^{(o)}$. The creation part of $\varphi^{(o)}$ is $\varphi^{(o,+)}$, and the annihilation part is $\varphi^{(o,-)}$.

(v) The metric is time-favored.

C. Simple Applications.

(1) (To quantum field theory.[2]) The canonical commutation relations of quantum mechanics,

$$\left[p^j, q^k\right] = i^{-1}\delta^{jk},$$

provide a heuristic basis for the field-theoretic relations. For a scalar field, the latter may be expressed as

$$\left[\dot{\varphi}(t,\vec{u}), \varphi(t,\vec{v})\right] = i^{-1}\delta(\vec{u}-\vec{v}). \tag{I.3}$$

This relation may be realized by taking a suitable space of functionals with φ a multiplicative operator and $\dot{\varphi}$ a functional derivative, cf. (I.2). Explicitly, we pick a fixed time t, and let

$$\left[\varphi(\vec{u})F\right]\{\eta\} = \eta(\vec{u})F\{\eta\}, \quad \left[\dot{\varphi}(\vec{v})F\right]\{\eta\} = i^{-1}\,\delta F\{\eta\}/\delta\eta(\vec{v}). \tag{I.4}$$

With this convention, however, one has the problem of identifying $\dot{\varphi}$ as $\partial_t\varphi$.

(2) (To classical particle mechanics.[3]) The equations of motion of a particle moving under the influence of a potential V are determined by Hamilton's principle,

$$\delta S\{\vec{x}\}/\delta x^j(t) = 0, \quad j = 1,2,3. \tag{I.5a}$$

Here the $x^j(t)$ define a hypothetical path of the particle, with the initial and the final points $\vec{x}(t_o)$, $\vec{x}(t_1)$ not subject to variation. Further, S is the action

$$S\{\vec{x}\} = \int_{t_o}^{t_1} dt\Big(T\{\vec{x}(t)\} - V\{\vec{x}(t)\}\Big). \tag{I.5b}$$

(3) (To the calculus of variations.[4]) The last example is typical of the calculus of variations, where we usually want to minimize or maximize a certain functional. For example, we may want to find a closed curve of given length and greatest area. For this problem, it is convenient to take the arclength s as the independent variable. The curve must clearly be symmetric with respect to, say, the u-axis. (Otherwise, the area could be increased by symmetrization.) Hence $\frac{1}{2}A = \int du\, y$ or

$$\tfrac{1}{2}A\{y\} = \int_0^{\ell} ds \frac{du}{ds}\, y = \int_0^{\ell} ds \left[1 - \left(\frac{dy}{ds}\right)^2\right]^{\frac{1}{2}} y.$$

For the maximum, we set

$$\delta A\{y\}/\delta y(\sigma) = 0, \quad y(0) = y(\ell) = 0.$$

This yields an ordinary differential equation whose solution describes a semicircle.

We should like to point out that manipulations such as in the foregoing can easily lead to anomalies. We give two examples.

First, consider the functional

$$J\{y\} = \int_0^1 du\, y^2(u),$$

where $y(0) = 0$, $y(1) = 1$, and y is a real continuous function. It is clear that J can be made arbitrarily small (and positive), but it can never take the value zero as long as y remains continuous.

For the second example, consider the operator

$$\delta^2/\delta[\eta(v)]^2 \equiv D. \tag{I.6}$$

This corresponds to the frequently encountered term $-\dot{\varphi}^2(v)$ in the representation (I.4). Let us consider the following two functionals:

$$J_1\{\eta\} = \int d^n u\, \rho(u)\eta^4(u), \tag{I.7a}$$

$$J_2\{\eta\} = \int d^n u_1 \ldots d^n u_4 w(u_1, \ldots, u_4)\eta(u_1)\ldots\eta(u_4). \tag{I.7b}$$

Then

$$DJ_1\{\eta\} = 12\rho(v)\eta^2(v)\delta(v-v).$$

The quantity $\delta(0)$ is undefined (even though it may sometimes be given the value infinity).

On the other hand, it may be unnecessarily restrictive to say that the operator D is meaningless. We have in fact, if w is symmetric,

$$DJ_2\{\eta\} = 12\int d^n u_1 d^n u_2 w(u_1,u_2,v,v)\eta(u_1)\eta(u_2),$$

an unambiguous expression for continuous integrable functions.

II. Several Kinds of Functional Derivatives

A. Examples of Function Spaces.

We will now develop a mathematically precise formulation of the foregoing notions. We start with a review of some commonly employed function spaces.

(1) A Banach space $\mathfrak{B}$ is defined[5] by the following properties valid for all $x,y\in\mathfrak{B}$ and $\alpha,\beta\in R^1$ or C^1 (depending on whether a real or a complex space is desired):

(i) $\mathfrak{B}$ is a linear vector space: $\alpha x+\beta y\in\mathfrak{B}$, etc.

(ii) There is a norm function, $\|x\|\in R^1$, such that

(ii-1) $\|x\|\geq 0$ and $\|x\|=0 \Leftrightarrow x=0$,

(ii-2) $\|\alpha x\| = |\alpha|\,\|x\|$,

(ii-3) $\|x+y\|\leq \|x\| + \|y\|$.

(iii) Completeness: every Cauchy sequence in $\mathfrak{B}$ has a limit in $\mathfrak{B}$. (Only strong limits, e.g., $\|x_n-x\|\to 0$, interest us.)

A Hilbert space can be defined as a Banach space, equipped with a Hermitian form $\langle x,y\rangle$ such that $\langle x,x\rangle = \|x\|^2$. We use physicists' convention, $\langle x,\alpha y\rangle = \alpha\langle x,y\rangle = \langle \alpha^* x,y\rangle$. We confine ourselves to separable Hilbert spaces, i.e., having a countable basis.

(2) Some other examples of Banach spaces are as follows.

(2a) The space of functions $x(u)$, $u\in S\subseteq R^n$, such that

$$X \equiv \int_S d^n u\, w(u)|x(u)|^p < \infty, \qquad \text{(II.1a)}$$

where $w\geq 0$. We set

$$\|x\| = X^{1/p}. \qquad \text{(II.1b)}$$

This space is variously denoted: $\mathfrak{L}^p(S,w)$, L_p, etc.

(2b) C—the space of bounded continuous functions on R^n, with the norm $\|x\| = \sup_u|x(u)|$.

(2c) C_o—the space of continuous functions on R^n, with compact support (equivalently, with bounded support). The norm is

$$\|x\| = \sup_u |x(u)| + \mu(\operatorname{supp} x),$$

where the last term indicates the measure of the support of x.

(2d) $\mathcal{K}^\mu$—the space of Hölder continuous functions over an interval $I \subseteq R^1$, with the index satisfying $0<\mu<1$. Here

$$\|x\| = \sup_u |x(u)| + \sup_{u \neq u'} \frac{|x(u) - x(u')|}{|u-u'|^\mu}. \tag{II.2}$$

Properties (ii-3) and (iii) can be verified.[6]

(3) We mention two basic spaces of test functions.[7]

(3a) $\mathcal{D}$—infinitely differentiable functions on R^n with compact support.

(3b) $\mathcal{S}$—infinitely differentiable functions on R^n which, together with all their derivatives, approach zero at infinity faster than the inverse of any polynomial. The following is the classical example of an element of $\mathcal{D}$ and of $\mathcal{S}$:

$$x(u) = \exp\left[-1/u^2 - 1/(1-u)^2\right] \quad \text{for} \quad 0<u<1,$$

$$= 0 \quad \text{otherwise}. \tag{II.3}$$

(4) If V is a linear vector space, the space of linear functionals f defined on elements $x \in V$ is the dual space V'. We give two examples.

(4a) For a Hilbert space $\mathcal{H}$ one can identify $\mathcal{H}$ and $\mathcal{H}'$; i.e., given $f \in \mathcal{H}'$, $\exists\, y \in \mathcal{H}$ such that[5]

$$f(x) = \langle y,x \rangle \quad \text{for all} \quad x \in \mathcal{H}. \tag{II.4}$$

(4b) $\mathcal{S}'$ is the space of tempered distributions, and $\mathcal{D}' \supset \mathcal{S}'$.

B. Functional Derivatives.

Suppose that F is a functional on a function space $\mathcal{K}$ such that for $\xi \in \mathcal{K}_0 \subseteq \mathcal{K}$ and for all $\eta \in \mathcal{K}$,

$$F\{\xi + \varepsilon\eta\} - F\{\xi\} = \varepsilon F_1\{\xi,\eta\} + 0(\varepsilon),$$

where F_1 is linear in η. Then for the given ξ,

$$F_1\{\xi,\cdot\} \in \mathcal{K}'. \tag{II.5}$$

We may write symbolically,

$$F_1\{\xi,\eta\} = \int d^n u\, h_1\{\xi;u\}\eta(u). \tag{II.6a}$$

We call h_1 the Volterra derivative of F. This derivative corresponds to the heuristic one of Section I, and we write

$$h_1\{\xi;u\} = \delta F\{\xi\}/\delta\xi(u). \tag{II.6b}$$

One may define similarly higher derivatives, e.g.,

$$h_2\{\xi;u,v\} = \delta^2 F\{\xi\}/\delta\xi(u)\delta\xi(v).$$

Note that the h_j need not be functions in the usual sense. For example, if $\mathcal{H} = \mathcal{S}$, then the h_j define elements of $\mathcal{S}'$ (cf. the examples in Section I).

It is often convenient to work with abstract Banach spaces without referring to any specific spaces of function. One then considers, quite generally, mappings from an open subset of a Banach space into another Banach space,

$$F: \mathcal{R}_1 \to \mathcal{B}_2 \quad \text{where} \quad \mathcal{R}_1 \subseteq \mathcal{B}_1. \tag{II.7}$$

The case $\mathcal{B}_2 = C^1$ gives functionals such as considered previously.

Since $\mathcal{R}_1$ is open,

$$x \in \mathcal{R}_1 \text{ and } h \in \mathcal{B}_1 \Rightarrow x + \epsilon h \in \mathcal{R}_1,$$

for all ϵ, real and sufficiently near zero. We construct, if possible,

$$\delta_h F(x) = (d/d\epsilon)F(x+\epsilon h)\Big|_{\epsilon=0}. \tag{II.8a}$$

If we now keep x fixed, then we can define a map $A: \mathcal{B}_1 \to \mathcal{B}_2$ as follows,

$$Ah = \delta_h F(x). \tag{II.8b}$$

If A is a bounded linear operator, defined for all $h \in \mathcal{B}_1$, then we say that F is Fréchet differentiable at x, and $\delta_h F$ is called the Fréchet differential.[8)-10)] Note that $\delta_h F$ is a map, in general nonlinear, from some $\mathcal{R}_1' \subseteq \mathcal{R}_1$ into $\mathcal{B}_2$.

If $\mathcal{B}_1$ is given as a function space, then we may relate the Fréchet and the Volterra derivatives as follows:

$$\delta_h F(x) = \int d^n u\, h(u) \frac{\delta F(x)}{\delta x(u)}. \tag{II.9}$$

The Volterra derivative, of course, is not restricted to the case $F \in C^1$.

If the operator A is unbounded, or not defined on all of $\mathfrak{B}_1$, then one is led to the notion of the Gâteaux differential. We shall not be concerned with such cases.

C. The Implicit Function Theorem.

The concept of the Fréchet differential enables us to extend various standard theorems of classical analysis to Banach spaces. One of the most basic of these is the implicit function theorem. We can now give it the following formulation.[9),10)]

Theorem. Let $\mathfrak{B}_X$, $\mathfrak{B}_Y$, and $\mathfrak{B}_Z$ be Banach spaces, let $\mathfrak{O}$ be an open subset of $\mathfrak{B}_X \times \mathfrak{B}_Y$ containing (x_0, y_0), and let $f: \mathfrak{O} \to \mathfrak{B}_Z$ be a map such that

(i) $f(x_0, y_0) = 0$,

(ii) f is continuously Fréchet-differentiable in $\mathfrak{O}$,

(iii) for all $(x,y) \in \mathfrak{O}$, the differential

$$\delta_{2,h} f(x,y) = (d/d\epsilon)\, f(x, y+\epsilon h)\Big|_{\epsilon=0} \qquad \text{(II.10)}$$

has a bounded inverse as a function of h; i.e., as a map $\mathfrak{B}_Y \to \mathfrak{B}_Z$.

Then there exists a map from $\mathfrak{O}_X \subseteq \mathfrak{B}_X$ to $\mathfrak{B}_Y$, to be denoted by $y(x)$, which is the unique solution of $f(x,y) = 0$ in a neighborhood of (x_0, y_0). In particular, $y(x_0) = y_0$. The map y can be extended until we reach the boundary of $\mathfrak{O}$, and is continuously Fréchet differentiable.

A function F (as in (II.7)) is continuously Fréchet differentiable if for every h (in $\mathfrak{B}_1$), $\delta_h F$ is a continuous function of x.

This theorem is of interest primarily for nonlinear functions. However, it is also instructive to see how it includes the standard result for Fredholm equations as a special case. Consider

$$f\{a,b\} \equiv b(u) + \int dv\, K(u,v) b(v) - a(u) = 0, \qquad \text{(II.11)}$$

where a, b, and f are elements of a certain Banach space, and we seek to find b in terms of a. We have

$$\delta_{2,h} f\{a,b\} = (d/d\epsilon)(b + \epsilon h + Kb + \epsilon Kh - a) = (I+K)h.$$

The implicit function theorem tells us that, if I+K has a bounded inverse, then the equation can be solved for arbitrary a. On the other hand, one knows that if I+K does not have a bounded inverse, then the equation cannot be solved for some a.

This example illustrates the need for condition (iii) in the statement of the theorem. In the finite-dimensional case, this condition can be replaced by that of non-vanishing of the Jacobian.

D. Application to Dispersion Theory.[9]

Let A(s) be a partial wave amplitude for a two-particle scattering process, s being the center-of-mass energy. We can express A in terms of the phase shift $\delta^{(o)}(s)$ and an elasticity parameter $\eta(s)$ $(0 < \eta \leq 1)$ as follows:

$$\text{Re } A \equiv x = \tfrac{1}{2}\eta \sin 2\delta^{(o)}, \tag{II.12a}$$

$$\text{Im } A \equiv y = \tfrac{1}{2}(1 - \eta \cos 2\delta^{(o)}). \tag{II.12b}$$

We assume that A satisfies a once-subtracted dispersion relation, that there are no bound states, and that the contributions from the left-hand cut can be neglected. Let s_o be the physical threshold, let

$$\delta^{(o)}(s_o) = x(s_o) = 0. \tag{II.13a}$$

Furthermore, let the functions approach definite limiting values at infinity,

$$\delta^{(o)}(\infty) = 0, \quad x(\infty) = 0, \quad \eta(\infty) \neq 0. \tag{II.13b}$$

The condition $\delta^{(o)}(\infty) = 0$ is somewhat special; normally one would allow $\delta^{(o)}(\infty) = \frac{1}{2} n\pi$, where n is an integer. We will comment on this later. In order to avoid centrifugal factors $k^{2\ell}$, we suppose the scattering to be s-wave.

We remove the kinematic singularities, as usual, by separating a factor $\rho(s)$ such that

$$\rho(s) \sim (s - s_o)^{\frac{1}{2}} \quad \text{near} \quad s = s_o, \quad \rho(\infty) \neq 0. \tag{II.14}$$

For example, for π-π scattering, $\rho = [(s - s_o)/4s]^{\frac{1}{2}}$.

A convenient way of taking the conditions at infinity into account is by means of the transformation

$$s \longrightarrow v(s) = (s - s_o)/s, \tag{II.15a}$$

so that $v(s_o) = 0$, $v(\infty) = 1$. The factors in dispersion relations are transformed according to

$$ds'\left[s/s'(s - s')\right] = dv'(v' - v)^{-1}. \tag{II.15b}$$

For brevity, we shall use the notation $x(v)$ for $x(s(v))$, etc.

The once-subtracted dispersion relation now is

$$x(v) - \alpha\rho(v) + \frac{\rho(v)}{\pi} \int_0^1 \frac{dv'\, y(v')}{\rho(v')(v'-v)} \equiv F\{\delta^{(o)}\} = 0, \qquad \text{(II.16)}$$

where α is a subtraction constant. We may ask if the perturbed equation

$$F\{\delta\}(v) - g(v) \equiv F_1\{g,\delta\}(v) = 0, \qquad \text{(II.17)}$$

where g is a given function, has a solution $\delta(v)$, and whether this solution is unique.

To formulate such a question properly, we have to require all the given functions to belong to specified spaces, and to seek the solution δ also within a definite space. For our problem, a suitable space is $\mathcal{H}^\mu$, that of real Hölder continuous functions on $[0,1]$ with an exponent $\mu \leq \frac{1}{2}$. (Note that ρ could not be incorporated if $\mu > \frac{1}{2}$, since $\rho \sim v^{\frac{1}{2}}$ near $v=0$.) For the perturbation g and the solution δ we require a stronger condition, namely that $g, \delta \in \mathcal{B}$ where

$$\mathcal{B} = \left\{ f \in \mathcal{H}^\mu : f(0) = f(1) = 0 \right\}. \qquad \text{(II.18)}$$

Clearly, $\mathcal{B}$ is a Banach space if it is equipped with the same norm as $\mathcal{H}^\mu$, i.e., (II.2).

In analogy with the example of the Fredholm equation, we compute

$$\left(\delta_{2,\Delta} F_1\{g,\delta\}\right)(v) = \left[\eta(v)\cos 2\delta(v)\right]\Delta(v) - \frac{\rho(v)}{\pi} P \int_0^1 \frac{dv'[\eta(v')\sin 2\delta(v')]}{\rho(v')(v'-v)} \Delta(v') \equiv (L\Delta)(v'). \qquad \text{(II.19)}$$

The question now is whether or not the operator L has a bounded inverse. To investigate this question, we consider the equation $L\Delta = h$ where h is arbitrary in $\mathcal{B}$. (We also restrict the solution Δ to $\mathcal{B}$.) We arrive this way at a singular integral equation of a kind for which there is an extensive theory available.[6]

The conclusion is:[9] There is a unique Δ for a given h, and the map $h \to \Delta$ is bounded. Consequently, by the implicit function theorem, there is a unique solution δ to $F\{\delta\} - g = 0$, which depends continuously on g and reduces to $\delta^{(o)}$ if $g = 0$.

If we had assumed $\delta^{(o)}(s=\infty) = \frac{1}{2} n\pi$, not necessarily zero, then two complications would develop. First, we could not utilize a

space like $\mathcal{B}$ so easily. Second, the equation $L\Delta = h$ would have an n-parameter family of solutions if $n > 0$, and if $n < 0$, then $(-n)$ supplementary conditions on h would be required. The implicit function theorem of Section II.C would no longer apply. We refer to the cited work of Lovelace[9] for further discussion.

III. Generating Functionals

A. Concept of a Generating Functional.

Now that we have the Volterra derivative, it is natural to consider the following expansion of a functional:

$$F\{J_o+zJ\} = F\{J_o\} + \sum_{k=1}^{\infty} \frac{z^k}{k!} \int d^n u_1 \dots d^n u_k f_k(u_1, \dots, u_k) J(u_1) \dots J(u_k). \tag{III.1}$$

The functions f_k will be assumed symmetric. This is a power series in z, and we call such an expansion a Volterra series. We see that

$$f_k(u_1, \dots, u_k) = \left[\delta^k / \delta J(u_1) \dots \delta J(u_k)\right] F\{J_o+J\}\Big|_{J=0}. \tag{III.2}$$

The functional $F\{J\}$ consequently contains the information contained in the infinite sequence of functions

$$f_o = F\{J_o\}, \quad f_1(u_1), \quad f_2(u_1, u_2), \dots, \tag{III.3}$$

and is called a generating functional of the sequence. Conversely, given a sequence of symmetric functions, we can construct the functional.

We assumed implicitly in the foregoing that the series (III.1) converges for z, J, and J_o suitably restricted. However, such series may be useful even when they do not converge.

A special case of generating functionals is the generating functions, which are familiar from the study of the special functions of mathematical physics. In fact, if $J(u) = \delta(u-t)$ $(u, t \in R^1)$, then (III.1) reduces to

$$F(t,z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} g_k(t). \tag{III.4}$$

Usefulness of generating functions may be illustrated with a simple example. Suppose that F depends only on the quantity[11] $z^2 - 2tz$. Then

$$z\partial F/\partial z = (t-z)\partial F/\partial t, \tag{III.5}$$

and this equation implies a prototype recursion relation,

$$kg_k(t) = tg_k'(t) - kg_{k-1}'(t). \tag{III.6}$$

This argument applies in particular to the Legendre and to the Hermite polynomials,

$$(z^2 - 2tz + 1)^{-\frac{1}{2}} = \sum (z^k/k!)[k!\, P_k(t)], \tag{III.7a}$$

$$\exp(-z^2 + 2tz) = \sum (z^k/k!)\, H_k(t). \tag{III.7b}$$

Both of these expansions have non-zero radii of convergence.

B. Digression on Vacuum Expectation Values.

In quantum field theory we encounter various sequences of functions (strictly speaking, of distributions). In particular, one introduces[12] (i) Wightman functions, i.e., the vacuum expectation values of products of field functions $\varphi(u_j)$,

$$W_n(u_1,\dots,u_n) = \langle \varphi(u_1)\dots\varphi(u_n)\rangle_o, \tag{III.8}$$

and (ii) time-ordered functions,

$$\tau_n(u_1,\dots,u_n) = \left\langle \left(\varphi(u_1)\dots\varphi(u_n)\right)_+ \right\rangle_o, \tag{III.9}$$

where

$$\left(\varphi(u_1)\dots\varphi(u_n)\right)_+ = \varphi(u_{i_1})\dots\varphi(u_{i_n}), \quad u^o_{i_1} \geq u^o_{i_2} \geq \dots \geq u^o_{i_n}.$$

Both the W_n and the τ_n are usually assumed to be distributions in $\mathcal{S}'$.

Among other important functions are the retarded functions r_n, of which we note a special case,

$$r_2(u,v) = i\theta(u-v)\left\langle [\varphi(u),\varphi(v)] \right\rangle_o. \tag{III.10}$$

[Here $\theta(u) = 1$ if $u^o > 0$, $= 0$ if $u^o \leq 0$.] We shall not consider the question as to whether or not the existence of, for example, the τ_n can be derived from other assumptions.

For the free scalar field of mass m, all the $W_n^{(o)}$ and the $\tau_n^{(o)}$ can be expressed in the following way. We recall,

$$W_2^{(o)}(1,2) \equiv W_2^{(o)}(u_1,u_2) = \langle \varphi^{(o)}(u_1)\varphi^{(o)}(u_2)\rangle$$

$$= \frac{1}{(2\pi)^3}\int \frac{d^3\vec{k}}{2k^o}\, e^{-ik(u_1-u_2)}. \qquad \text{(III.11a)}$$

Then

$$\tau_2^{(o)}(1,2) \equiv \tau_2^{(o)}(u_1,u_2) = \theta(u_1-u_2)W_2^{(o)}(1,2) + \theta(u_2-u_1)W_2^{(o)}(2,1), \qquad \text{(III.11b)}$$

$$W_{2k+1}^{(o)} = 0, \qquad \tau_{2k+1}^{(o)} = 0, \qquad \text{(III.11c)}$$

$$\tau_{2n}^{(o)}(1,\ldots,2n) = \sum \tau_2^{(o)}(i_1,i_2)\ldots\tau_2^{(o)}(i_{2n-1},i_{2n}), \qquad \text{(III.11d)}$$

where the sum extends over all partitions of $\{1,\ldots,2n\}$ into pairs. For the $W_{2n}^{(o)}$, we replace the $\tau_2^{(o)}$ by $W_2^{(o)}$ in (III.11d).

Equation (III.11d) gives in particular, if we write $\tau_2^{(o)}(1,2) \equiv (1,2)$,

$$\tau_4^{(o)}(1,\ldots,4) = (1,2)(3,4) + (1,3)(2,4) + (1,4)(2,3).$$

This function corresponds to the trivial scattering matrix,

$$\langle p,q|S|p',q'\rangle = \delta(p-p')\delta(q-q') + \delta(p-q')\delta(q-p').$$

For interacting fields, we are usually interested in what remains when this trivial effect is subtracted off. We therefore define recursively the truncated functions, as follows (as usual, we assume $W_1=\tau_1=0$):

$$\tau_2^T(u_1,u_2) = \tau_2(u_1,u_2), \qquad \text{(III.12a)}$$

$$\tau_n^T(u_1,\ldots,u_n) = \tau_n(u_1,\ldots,u_n)$$

$$- \sum \tau_{m_1}^T(u_1,\ldots,u_{i_{m_1}})\ldots\tau_{m_k}^T(u_{i_N},\ldots,u_n), \qquad \text{(III.12b)}$$

where $N = n - m_k + 1$, and where the sum is over all partitions of $\{1,\ldots,n\}$.

The truncated Wightman functions W_n^T are defined in an analogous way. For the free field functions we have, in particular,

$$W_n^{(o)T} = 0 \quad \text{and} \quad \tau_n^{(o)T} = 0, \quad \text{if} \quad n > 2. \tag{III.13}$$

For the functions W_n^T, the following cluster property can be proved.[13] Take n points at equal times, $u_1^o = \ldots = u_n^o$, let v be purely space-like, $v = (0, \vec{v})$, $\vec{v} \neq 0$, let $\lambda \in R^1$, and consider

$$W_n^T\big(u_1, \ldots, u_k, u_{k+1} + \lambda v, \ldots, u_n + \lambda v\big). \tag{III.14a}$$

Then as $\lambda \to \infty$, the arguments of W_n^T separate into two clusters, and asymptotically

$$W_n^T \lesssim e^{-\lambda |\vec{v}| m}. \tag{III.14b}$$

Here m is the lowest mass of the theory, assumed non-zero. It also follows that the corresponding function W_n factorizes in the limit $\lambda \to \infty$,

$$W_n \longrightarrow W_k(u_1, \ldots, u_k)\, W_{n-k}(u_{k+1}, \ldots, u_n). \tag{III.15}$$

(We made use of translational invariance.)

C. The Time-Ordered Generating Functional.

Of the vacuum expectation values introduced in the foregoing, the Wightman functions are not symmetric, but the time-ordered functions are. Hence the latter can be used to construct a generating functional. We should emphasize, however, that in view of our meager knowledge of interacting fields, the material of this part (III.C) is strictly heuristic.

The generating functional T can be given in closed form,[14]

$$T\{J\} = \left\langle \left(\exp i \int d^4u\, \varphi(u) J(u) \right)_+ \right\rangle_o . \tag{III.16}$$

The factor i was supplied to improve the chance of convergence and to make the operator unitary (i.e., heuristically). We shall see below [Eq. (III.26) and also Section XII] that J has a natural interpretation as an external c-number source. The time-ordered exponential is defined as a suitable limit of the product

$$\exp\Big(i \int_{u^o \le t_1} d^4u \varphi J\Big) \exp\Big(i \int_{t_1 \le u^o \le t_2} d^4u \varphi J\Big) \ldots \quad \exp\Big(i \int_{t_n \le u^o} d^4u\, \varphi J\Big) . \tag{III.17}$$

It follows that

$$\frac{\delta}{i\delta J(v)} T\{J\} = \left\langle \left(e^{i\int d^4 u \varphi J} \varphi(v) \right)_+ \right\rangle_o . \qquad \text{(III.18)}$$

Indeed, one may argue that if, for example, $t_1 < v^o < t_2$, then the differentiation will bring down $\varphi(v)$ just before or just after the second factor in (III.17) to a good approximation. Then, in the limit, one should get the expression in (III.18). We conclude also that

$$\left[\delta^2 / i^2 \, \delta J(v_1) \delta J(v_2) \right] T\{J\} \Big|_{J=0} = \tau_2(v_1, v_2), \qquad \text{(III.19)}$$

and similarly for the other functions τ_n.

One may also construct the generating functional $T_o\{J\}$ for the truncated time-ordered functions. It is a matter of easy combinatorics to show that

$$T\{J\} = \exp T_o\{J\} . \qquad \text{(III.20)}$$

We next want to construct a functional differential equation for T. Suppose that the field φ satisfies

$$(\Box - \mu^2)\varphi = \lambda(\varphi^3)_{ren} . \qquad \text{(III.21)}$$

For the renormalized interaction term we can take the following form (which has been established in perturbation theory[15]):

$$(\varphi^3)_{ren}(u) = \lim_{\xi \to o} \left[:\varphi(u-\xi)\varphi(u)\varphi(u-\xi): - F(\xi)\varphi(u) \right] \Big/ \left[1 + G(\xi) \right] . \qquad \text{(III.22)}$$

Here ξ is to be space-like, the functions F and $(1+G)^{-1}$ may be singular as $\xi \to 0$, and

$$:ABC: = ABC - A\langle BC \rangle_o - B\langle AC \rangle_o - C\langle AB \rangle_o . \qquad \text{(III.23)}$$

However, the detailed structure of $(\varphi^3)_{ren}$ is of no consequence for us.

We also assume the canonical commutation relations

$$\left[\dot{\varphi}(t, \vec{u}), \varphi(t, \vec{v}) \right] = i^{-1} \delta(\vec{u} - \vec{v}) . \qquad \text{(III.24)}$$

In general, one has $[\dot{\varphi},\varphi] = i^{-1}Z_3^{-1}\delta$, where Z_3 may be zero. It is important for us that $Z_3 > 0$, and then it can be made unity by a scale transformation of the field. (In perturbation theory, $Z_3 > 0$ for the φ^4-coupling in two-dimensions and in three-dimensions of space-time.)

Now, from (III.18) it may appear that we will obtain an identity by replacing $\varphi(v) \to \delta/i\delta J(v)$ in (III.21) and by letting each member act on T. Actually, differentiation with respect to time gives rise to an extra term, as the following argument shows[16] (here $t = v^0$):

$$\frac{\partial}{\partial t}\left(e^{i\int d^4u\varphi J}\varphi(v)\right)_+ = \frac{\partial}{\partial t}\left(e^{i\int_t^\infty d^4u\varphi J}\right)_+ \varphi(v)\left(e^{i\int_{-\infty}^t d^4u\varphi J}\right)_+$$

$$= \left(e^{i\int_t^\infty d^4u\varphi J}\right)_+ \left\{\dot{\varphi}(v) + i\int_{u^0=t} d^3\vec{u}\left[\varphi(v),\varphi(u)\right]J(u)\right\}\left(e^{i\int_{-\infty}^t d^4u\varphi J}\right)_+$$

$$= \left(e^{i\int d^4u\varphi J}\dot{\varphi}(v)\right)_+ . \qquad \text{(III.25)}$$

On the other hand, when we apply $\partial/\partial t$ again, the commutator $[\varphi,\varphi] = 0$ will be replaced by $[\dot{\varphi},\varphi] = i^{-1}\delta$. This will give a term proportional to $J(v)T$. One then obtains the equation

$$(\Box - \mu^2)\frac{1}{i}\frac{\delta}{\delta J(v)}T\{J\} = \left[\lambda\left(\frac{\delta^3}{i^3\delta J^3(v)}\right)_{\text{ren}} - J(v)\right]T\{J\} . \qquad \text{(III.26)}$$

We shall return to this equation in Section XII.C.

D. Molecular Distribution Functions.

We consider a system of N identical classical particles, which is described by the phase-space probability distribution function, called the molecular distribution function,[17]

$$D^N\left(\vec{q}_1,\vec{p}_1;\ldots;\vec{q}_N,\vec{p}_N\right) \quad \text{or} \quad D^N\left(r_1,\ldots,r_N\right), \qquad \text{(III.27)}$$

where $r_j = (\vec{q}_j,\vec{p}_j)$. This function sill be assumed symmetric, and of course it depends on the various parameters of the system such as volume and temperature.

In the typical cases of interest, D^N depends on a very large number of variables. More useful functions are the reduced distribution functions f_s (or D_s^N, below), which describe the correlations of a small number of particles (like molecules). To define these, we first integrate D^N $(N > s)$ over the phase space of all but s molecules,

$$D_s^N(r_1,\ldots,r_s) = V^s \int d^6r_{s+1}\, d^6r_N\, D^N(r_1,\ldots,r_N). \quad \text{(III.28)}$$

Note the following dependence on the volume V,

$$\int d^6r_{s+1}\ldots d^6r_N \sim V^{N-s}, \quad D^N \sim V^{-N}, \quad D_s^N \sim \text{constant}. \quad \text{(III.29)}$$

It is now reasonable to let $N\to\infty$, with the intensive quantities like the density and the temperature kept constant:

$$f_s = \lim_{N\to\infty} D_s^N. \quad \text{(III.30)}$$

The functions f_n have some analogies with the vacuum expectation values. In particular, the f_n are Euclidean invariant (provided there are no external fields and the interaction is Euclidean invariant). The f_n are also expected to have the cluster property: If $r_j' = r_j + \lambda(\vec{q}\,,0)$, where $\vec{q}$ is fixed and non-zero, then

$$f_s(r_1,\ldots,r_k, r'_{k+1},\ldots,r'_s) \to f_k(r_1,\ldots,r_k) \times f_{s-k}(r_{k+1},\ldots,r_s), \quad \text{(III.31)}$$

in the limit $\lambda\to\infty$, i.e., for infinite spatial separation of the two groups of points, in analogy with (III.14) and (III.15). The validity of (III.31) of course depends on the behavior of the potential at large separations.

We may note that we confine ourselves here to a heuristic discussion, even though various rigorous statements can be made.[18]

Let us introduce the generating functional for the f_s,

$$F\{\chi\} = 1 + \sum_{s=1}^{\infty} \int d^6r_1\ldots d^6r_s f_s(r_1,\ldots,r_n)\chi(r_1)\ldots\chi(r_s). \quad \text{(III.32)}$$

As far as we know, there is no direct interpretation for F, in contrast to the case of the functional T of Section III.C. We comment below on possible limitations for the argument χ.

The functional F can be utilized in various ways. For example, if the system is not in equilibrium, then the f_n are time-dependent. This time dependence can be described by a single equation for F, which we will give presently.

We start with Liouville's equation,

$$\frac{\partial D_N}{\partial t} = \sum_{\substack{\alpha=1,2,3 \\ k=1,\ldots,N}} \left(\frac{\partial H_N}{\partial q_k^\alpha} \frac{\partial D_N}{\partial p_k^\alpha} - \frac{\partial H_N}{\partial p_k^\alpha} \frac{\partial D_N}{\partial q_k^\alpha} \right) . \tag{III.33}$$

For the Hamiltonian we assume

$$H_N = \sum \vec{p}_j^{\,2}/2m + \sum_{i<j} \Phi \left(|\vec{q}_i - \vec{q}_j| \right). \tag{III.34}$$

The procedure is now straightforward, even though the equations are somewhat lengthy. We integrate both members of (III.33) as in (III.28). This yields an equation relating D_s^N and D_{s+1}^N, which is valid for all N, and hence also in the limit. The resulting hierarchy of equations is then seen to be equivalent to

$$\frac{\partial F\{\chi\}}{\partial t} = \int d^6r \left[\frac{\vec{p}^{\,2}}{2m} ; \frac{\delta F}{\delta \chi(r)} \right] \chi(r) + \frac{1}{2} \int d^6r d^6r' \left\{ \chi(r)\chi(r') + \right.$$

$$\left. + \sigma\chi(r) + \sigma\chi(r') \right\} \left[\Phi(|\vec{q} - \vec{q}'|) ; \frac{\delta^2 F}{\delta\chi(r)\delta\chi(r')} \right], \tag{III.35}$$

where σ is the density (i.e., N/V) and the brackets are Poisson brackets.

Let us now suppose that the system is time-independent and that it has a Maxwellian velocity distribution for each N and in the limit. Then the foregoing equation can be reduced to the following for $F\{\zeta\}$, ζ being a function of three-vectors $\vec{q}$: 19),20)

$$\frac{\partial}{\partial q^\alpha} \frac{\delta F\{\zeta\}}{\delta \zeta(\vec{q})} + \beta \int d^3\vec{q}' \, \frac{\partial \Phi(|\vec{q} - \vec{q}'|)}{\partial q^\alpha} \frac{\delta^2 F}{\delta\zeta(\vec{q})\delta\zeta(\vec{q}')} \left[\zeta(\vec{q}') + \sigma \right] = 0. \tag{III.36}$$

The constant $\beta = 1/kT$ corresponds to the assumed Maxwellian distribution.

At this point it seems natural to make the substitution

$$\zeta(\vec{q}) \longrightarrow \bar{\zeta}(\vec{q}) = \zeta(\vec{q}) + \sigma . \tag{III.37}$$

However, we see from (III.31) that the functions f do not vanish at infinity. Consequently the argument function χ in (III.35) must be restricted (for example, to compact support) and a similar restriction must be made on ζ. (In the field-theoretic case, Section III.C, if

the functions τ_n are assumed to be in $\mathcal{S}'$, then the source J must be in $\mathcal{S}$.)

It follows that the substitution $\zeta \to \overline{\zeta}$ cannot be effected in (III.36). However, we can introduce the analogue to the truncated vacuum expectation values. The desired functions have the generating functional G which satisfies, cf. (III.20) and set $F\{0\}=1$,

$$F\{\zeta\} = \exp G\{\zeta\}. \tag{III.38}$$

The transformation $F \to G$ is nonlinear, and the equation for G can be obtained from (III.36) by replacing

$$\frac{\delta F}{\delta \zeta(\vec{q})} \to \frac{\delta G}{\delta \zeta(\vec{q})}, \quad \frac{\delta^2 F}{\delta \zeta(\vec{q})\,\delta \zeta(\vec{q}')} \to \frac{\delta^2 G}{\delta \zeta(\vec{q})\,\delta \zeta(\vec{q}')} + \frac{\delta G}{\delta \zeta(\vec{q})}\,\frac{\delta G}{\delta \zeta(\vec{q}')}. \tag{III.39}$$

In the resulting equation for G, the substitution $\zeta \to \overline{\zeta}$ is admissible for potentials which vanish at infinity sufficiently rapidly. This substitution can be used, for example, for finding a solution in closed form, see Section VIII.A.

IV. The Wiener Integral

The present section, except for Part D, is based largely on the review article of Gelfand and Yaglom.[21]

A. Basic Notions.

One interpretation of the Wiener integral is as an expression for the average value of a quantity for a particle or particles undergoing Brownian motion. For the one-dimensional case, let $\psi(t,u)$ be the probability density that such a particle at the time t will be found at the point u. This function satisfies the diffusion equation

$$\partial\psi/\partial t = D\partial^2\psi/\partial u^2, \tag{IV.1}$$

where D is a constant. We will take $D=\frac{1}{2}$ for the remainder of Section IV. (But we note that one also often takes $D=\frac{1}{4}$.)

If the particle is known to have been at $u=0$ at time $t=0$, then we have the initial condition

$$\psi(0,u) = \delta(u), \tag{IV.2}$$

and the solution

$$\psi(t,u) = (2\pi t)^{-\frac{1}{2}} e^{-u^2/2t}. \tag{IV.3}$$

Similarly, the probability density that the particle, initially at $u=0$, passes through the points

$$u_1,\ldots,u_m \quad \text{at times} \quad 0<\tau_1<\ldots<\tau_m,$$

respectively, is

$$\psi(\tau_1,u_1;\ldots;\tau_m,u_m)=\left[(2\pi)^m\tau_1(\tau_2-\tau_1)\ldots(\tau_m-\tau_{m-1})\right]^{-\frac{1}{2}} \times \exp\left\{-\frac{1}{2}\left[\frac{u_1^2}{\tau_1}+\frac{(u_2-u_1)^2}{\tau_2-\tau_1}+\ldots+\frac{(u_m-u_{m-1})^2}{\tau_m-\tau_{m-1}}\right]\right\}. \tag{IV.4}$$

Let us now consider the probability that the coordinates $u_i=x(\tau_i)$ of such a particle satisfy

$$a_i\le x(\tau_i)\le b_i \quad \text{for} \quad i=1,2,\ldots,m. \tag{IV.5}$$

This will clearly be given by

$$\int_{a_1}^{b_1}du_1\ldots\int_{a_m}^{b_m}du_m\,\psi(\tau_1,u_1;\ldots;\tau_m,u_m). \tag{IV.6}$$

This integral is also the expectation value of the following functional defined on the space of particle paths:

$$\Phi\{x\}=\begin{cases}1 & \text{if (IV.5) is fulfilled,}\\ 0 & \text{otherwise.}\end{cases} \tag{IV.7}$$

This suggests that the expectation of more general functionals $F\{x\}$ can be found as follows. Suppose that $x(\tau)$ is continuous for $0\le\tau\le t$, and $x(0)=0$. We then select $m-1$ interior points,

$$0<\tau_1<\ldots<\tau_{m-1}<\tau_m=t,$$

and consider, in place of $x(\tau)$, the polygonal function $\hat{x}(\tau)$ determined by the vertices

$$\hat{x}(\tau_i)=x(\tau_i)\equiv u_i \quad \text{for} \quad i=1,\ldots,m.$$

We then approximate $F\{x\}$ by

$$F\{\hat{x}\} \equiv F_n(u_1, \ldots, u_m),$$

and form the expectation value

$$\int du_1 \ldots du_m \psi(\tau_1, u_1; \ldots; \tau_m, u_m) F_n(u_1, \ldots, u_m). \quad \text{(IV.8)}$$

Finally, we pass to the limit

$$m \to \infty, \qquad \max(\tau_{k+1} - \tau_k) \to 0. \quad \text{(IV.9)}$$

If this limit exists and is independent of the choice of partitions, we call it (provisionally) the Wiener integral of F, and denote it by

$$\int d\mu_W(x) F\{x\} \quad \text{or} \quad \int \mathfrak{D}_W(x) e^{\frac{i}{\hbar}\langle \dot{x}, \dot{x}\rangle} F\{x\} \quad \text{(IV.10)}$$

(or $\int d\mu_W F$, etc.). The first notation reflects the fact that the Wiener integral is a measure-theoretic integral. Consequently, the standard theorems, like the Fubini theorem and the dominated convergence theorem, remain valid.

The second notation shows the Wiener integral as the (commutative) integral over the Hilbert space of real functions $x(\tau)$ on $[0,t]$, satisfying $x(0) = 0$ and

$$\langle \dot{x}, \dot{x}\rangle = \int_0^t d\tau\, \dot{x}^2(\tau) < \infty. \quad \text{(IV.11)}$$

We may note in this connection that the exponent of $\psi(\tau_1, u_1 \ldots u_m)$ gives

$$\lim_{\max(\Delta_k\tau)\to 0} \left[-\frac{1}{2} \sum \left(\frac{\Delta_k x}{\Delta_k \tau}\right)^2 \Delta_k \tau \right] = -\tfrac{1}{2}\langle \dot{x}, \dot{x}\rangle. \quad \text{(IV.12)}$$

We also note that two functions such that $x_1 - x_2 = \text{constant}$ are equivalent with respect to the norm (IV.11), and hence the condition $x(0) = 0$ removes an inherent ambiguity.

The measure-theoretic aspect and the Hilbert space aspect of the Wiener integral will be discussed more fully later on. In particular, we shall see some drawbacks to the present definition in Section IX.C.

B. Two Examples.

We will now describe the evaluation of the Wiener integral for two kinds of functionals. We will assume a special choice of the subintervals $[\tau_k, \tau_{k+1}]$ and, in the second example, a variant to the

polygonal approximation. It will follow from the discussion in Sections VI, VII and IX that functionals of the two kinds are in fact integrable, and that a variety of approximating sequences may be employed for evaluating the integrals.

First, let

$$f\{x\} = x(t_1)\dots x(t_k). \tag{IV.13}$$

Let the points t_a be among the points τ_α. Then the approximating expression (IV.8) becomes exact and can be evaluated in closed form. If k is odd, then one encounters integrals like

$$\int_{-\infty}^{\infty} du\, e^{-u^2/s_u^{2j+1}}$$

as factors, and the result is zero. For $k=2\varkappa$, one obtains

$$\int d\mu_W(x)x(t_1)\dots x(t_{2\varkappa}) = \sum b(t_{i_1},t_{i_2})\dots b(t_{i_{2\varkappa-1}},t_{i_{2\varkappa}}), \tag{IV.14a}$$

where the sum is over all partitions of $\{1,\dots,2\varkappa\}$ into pairs, and

$$b(t_i,t_j) = \min(t_i,t_j) = \int d\mu_W(x)\,x(t_i)\,x(t_j). \tag{IV.14b}$$

For the second example, let

$$h\{x\} = \exp\Big[-\frac{1}{2}\int_0^t d\tau\, p(\tau)x^2(\tau)\Big], \tag{IV.15}$$

where p is continuous and ≥ 0. We let $\Delta\tau = t/m$, $\tau_j = j\Delta\tau$, also

$$p_j = p(\tau_j), \quad u_j = x(\tau_j),$$

and we approximate h by

$$\exp\Big(-\frac{1}{2}\,\Delta\tau\sum p_j u_j^2\Big).$$

The approximating integral (IV.8) now has the form

$$C\int du_1\dots du_n e^{-\sum a_{jk}u_j u_k} = C\pi^{\frac{1}{2}n}\Big[\det(a_{jk})\Big]^{-\frac{1}{2}}, \tag{IV.16}$$

where

$$(a_{jk}) = \frac{1}{2\Delta\tau}\begin{pmatrix} a_{11} & -1 & & & \\ -1 & a_{22} & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ & & & -1 & a_{mm} \end{pmatrix} \tag{IV.17}$$

$$a_{mm} = 1 - (\Delta\tau)^2 p_m, \quad a_{ii} = 2 - (\Delta\tau)^2 p_i$$

for $i < m$. The evaluation (IV.16) may be obtained by diagonalizing the quadratic form with the help of an orthogonal transformation.

We next set

$$D(\tau_k) = \det(a_{ij})_{i,j=k+1,\ldots,m}.$$

The expansion of a determinant in terms of its minors yields a recursion relation for D, and in the limit $m \to \infty$ one gets

$$d^2D(\tau)/d\tau^2 - p(\tau)D(\tau) = 0. \tag{IV.18}$$

The Wiener integral of h may then be expressed in terms of a solution of this equation. In conclusion we should say that the details of this particular solution are not especially important, but Gaussian functionals like (IV.15) occur constantly.

C. Two Generalizations.

One obvious way of generalizing the Wiener integral is to take the underlying space to be n-dimensional. Then, in place of the basic relation (IV.3) we have

$$\psi_n(t,\vec{u}) = (2\pi t)^{-\frac{1}{2}n} e^{-\vec{u}^2/2t}. \tag{IV.19}$$

Similarly, in Eq. (IV.8) we replace du_k by $d^n\vec{u}_k$, and the functionals depend on the vector-valued functions $\vec{x}(\tau)$. Except for other obvious modifications, the construction remains unchanged.

Note that if the functional factorizes,

$$F\{\vec{x}\} = \prod F_k\{x^k\},$$

then so does the Wiener integral,

$$\int d\mu_W(\vec{x})F\{\vec{x}\} = \prod \int d\mu_W(x^k)F_k\{x^k\}. \tag{IV.20}$$

Another generalization depends on specifying the final end-point as well as the initial one. Thus for n dimensions we may require

$$\vec{x}(t_1) = \vec{v}_1, \quad \vec{x}(t_2) = \vec{v}_2, \text{ where } t_1 < t_2, \tag{IV.21}$$

and we allow, in general, $t_1 \neq 0$ and $\vec{v}_1 \neq \vec{0}$. The resulting integral is called the conditional Wiener integral. In order to give an explicit definition, we can take the approximating integral (IV.8) and make the obvious modifications implied by (IV.21)—in particular, we integrate over $d^n\vec{u}_1 \ldots d^n\vec{u}_{m-1}$ but not over $d^n\vec{u}_m$.

We will indicate the conditional Wiener integral subject to (IV.21) by $C(\vec{v}_1, \vec{v}_2)$, with the dependence on the t_j usually suppressed. We have, for example, the evaluations

$$\int_{C(\vec{v}_1, \vec{v}_2)} d\mu_W(\vec{x}) = \left[2\pi(t_2-t_1)\right]^{-\frac{1}{2}n} e^{-(v_2-v_1)^2/2(t_2-t_1)} \equiv J, \tag{IV.22}$$

$$\int_{C(\vec{v}_1, \vec{v}_2)} d\mu_W(x)\vec{x}(t) = \left[(t-t_1)\vec{v}_2 + (t_2-t)\vec{v}_1\right](t_2-t_1)^{-1}J, \tag{IV.23}$$

in place of 1 and $\vec{v}_1$, respectively, for the unrestricted integral. In (IV.23) we assume $t_1 \leq t \leq t_2$. The two limiting cases give $\vec{v}_1 J$ and $\vec{v}_2 J$, respectively, as also follows directly from the contraints $\vec{x}(t_j) = \vec{v}_j$.

The following relation is obvious (for integrable functionals),

$$\int d^n\vec{v}_2 \int_{C(\vec{v}_1, \vec{v}_2)} d\mu_W(x)F\{x\} = \int_{\vec{x}(t_1)=\vec{v}_1} d\mu_W(\vec{x})F\{\vec{x}\}, \tag{IV.24}$$

and so is the composition law, for $t_1 < t < t_2$,

$$\int_{C(\vec{v}_1, \vec{v}_2)} d\mu_W F = \int d^n\vec{u} \int_{\substack{\vec{x}(t_1)=\vec{v}_1 \\ \vec{x}(t)=\vec{u}}} d\mu_W \int_{\substack{\vec{x}(t)=\vec{u} \\ \vec{x}(t_2)=\vec{v}_2}} d\mu_W F. \tag{IV.25}$$

In particular, if F depends only on the single value $\vec{x}(t)$, then the right hand side becomes

$$\int d^n\vec{u}\, F(\vec{u}) \int_{(\ldots)} d\mu_W \int_{(\ldots)} d\mu_W, \tag{IV.26}$$

a generalization of (IV.23). The two conditional integrals are as in (IV.22).

D. The Diffusion Equation with a Potential.

Suppose that $Q(t,\vec{u})$ is the solution to the equation

$$\partial Q/\partial t = \tfrac{1}{2}\Delta Q - V(\vec{u})Q, \tag{IV.27a}$$

with the initial condition

$$Q(0,\vec{u}) = \delta(\vec{u}). \tag{IV.27b}$$

Conditions on V will be specified presently. Now, it is a remarkable and basic fact that for $t>0$, Q can be expressed as a conditional Wiener integral:

$$Q(t,\vec{u}) = \int_{\substack{\vec{x}(0)=\vec{0} \\ \vec{x}(t)=\vec{u}}} d\mu_W(\vec{x}) \exp\Big[-\int_0^t d\tau\, V\{\vec{x}(\tau)\}\Big]. \tag{IV.28}$$

This formula is known as the Feynman-Kac formula.

A simple proof can be given along the following lines.[22] Let us assume that $0 \le V(\vec{u}) \le M$. Then we can expand the exponential and justify term-by-term integration. The result is also valid if only $0 \le V < \infty$, but then additional arguments must be provided. We also assume for simplicity that V is continuous. One can, however, allow V to depend explicitly on time, without additional complications.

We now convert the differential equation into an integral equation

$$Q(t,\vec{u}) = \psi_n(t,\vec{u}) - \int_0^t d\tau \int d^n\vec{\xi}\, \psi_n(t-\tau,\vec{u}-\vec{\xi})\, V(\vec{\xi}) Q(\tau,\vec{\xi}), \tag{IV.29}$$

where ψ_n is the solution for $V=0$ (see (IV.19)). We take $Q=\psi_n$ as the zero-order approximation, and iterate. For the term linear in V we find

$$-\int_0^t d\tau \int d^n\vec{\xi}\, V(\vec{\xi}) \Big[4\pi^2(t-\tau)\tau\Big]^{-\frac{1}{2}n} e^{-(\vec{u}-\vec{\xi})^2/2(t-\tau)} e^{-\vec{\xi}^2/2\tau}. \tag{IV.30a}$$

On the other hand, the expansion of the Wiener integral gives, for the linear term,

$$-\int_0^t d\tau \int_{C(\vec{0},\vec{u})} d\mu_W(x) V\{x(\tau)\} =$$

$$= -\int_0^t d\tau \int d^n\vec{\xi}\, V(\vec{\xi}) \int_{C(\vec{0},\vec{\xi})} d\mu_W \int_{C(\vec{\xi},\vec{u})} d\mu_W, \tag{IV.30b}$$

where we used Fubini's theorem and the discussion of Section IV.C. If we evaluate the conditional integrals as in (IV.22), we then obtain agreement with (IV.30a).

The higher-order terms can also be shown to agree. Convergence is no problem of $0 \le V \le M$, as we already stated. Hence the series for Q satisfies the integral equation (IV.29), and also the differential equation (IV.27a) with the initial condition. The proof of the F.-K. formula is complete.

The foregoing proof presupposed the measure-theoretic basis of the Wiener integral (for example, we used Fubini's theorem. Note also that the foregoing proof gives no clue as to how the formula might have been discovered. In Section XI we shall describe another way of arriving at this formula.

An immediate consequence of the F.-K. formula is

$$Q(t,\vec{u}) > 0 \quad \text{for} \quad t > 0,$$

which is not obvious from an inspection of the differential equation. Strict inequality follows from $Q \ge e^{-Mt}j$ (cf. (IV.22)) or, if $M \to \infty$, from measure-theoretic considerations in Section IX.

For future reference we note that the Feynman-Kac formula can also be written as:

$$\langle \vec{u} | e^{-tH} | \vec{v} \rangle = \int_{\substack{\vec{x}(0)=\vec{u} \\ \vec{x}(t)=\vec{v}}} d\mu_W(\vec{x}) \exp\Big(-\int_0^t d\tau V\Big). \tag{IV.31}$$

E. The Partition Function.

We derive here, without a detailed justification of the arguments, a representation for the partition function which will be needed in Section V.

Let us consider an N-particle system which is governed by the Hamiltonian $H = -(2m)^{-1}\Delta + V$ with a discrete spectrum,

$$H\chi_j = E_j\chi_j, \quad j = 0, 1, 2, \ldots.$$

Then the partition function of statistical mechanics is

$$Z = \sum_j e^{-\beta E_j}, \quad \text{where} \quad \beta = 1/kT, \tag{IV.32}$$

and may be written, if the eigenfunctions are normalized,

$$Z = \int d^{3N}\vec{u} \sum_j \chi_j^*(\vec{u})\chi_j(\vec{u})e^{-\beta E_j}.$$

Next we observe that

$$Q_{\vec{v}}(\beta,\vec{u}) = \sum_j \chi_j^*(\vec{v})\chi_j(\vec{u})e^{-\beta E_j} \tag{IV.33}$$

is the solution to

$$\partial Q/\partial\beta = \left[(2m)^{-1}\Delta - V\right]Q, \tag{IV.34a}$$

which satisfies the "initial" condition

$$Q(0,\vec{u}) = \delta(\vec{u}-\vec{v}). \tag{IV.34b}$$

By applying the Feynman-Kac formula (IV.28) we obtain

$$Z = \int d^{3N}\vec{u} \int_{C(\vec{u},\vec{u})} d\mu_W(x) \exp\left[-\int_0^\beta d\tau V\{x(\tau)\}\right], \tag{IV.35}$$

which is the desired representation. (Here we use units where $m=1$.)

Note that for a free particle, $V=0$, the Wiener integral is positive and independent of $\vec{u}$. Consequently, we get $Z=\infty$, as we should, for an infinite volume.

V. Inequalities and Their Applications.

A. Preliminaries. Convexity.

There are, relatively speaking, few functionals whose Wiener integrals can be evaluated in closed form. The best-known of such functionals are the polynomials, the Gaussians, and the linear exponentials, and we saw examples of the first two types in Section IV. However, the integrals which occur in the study of interacting systems with an infinite number of degrees of freedom ordinarily cannot be handled so easily.

The Wiener integral can be valuable also in such cases, as it may clarify and simplify some derivations. Moreover, significant conclusions can be obtained with the help of inequalities, as we shall see in the remainder of Section V.

We note, first of all, that the standard inequalities for integrals over finite-dimensional spaces carry over to the limiting case of the Wiener integral. This follows from the fact that $a_n \leq b_n$ for all n implies

$$\lim a_n \leq \lim b_n.$$

To give an example, the Minkowski (triangle) inequality for finite-dimensional integrals implies that

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p,$$

where

$$\|f\|_p = \left\{ \int d\mu_W(x) |f\{x\}|^p \right\}^{1/p}.$$

We now turn to convexity. This notion gives rise to many powerful inequalities. We recall that a real-valued function $\psi(u)$ on an interval I of the real line is convex if for all $u, v \in I$,

$$\psi\left(\tfrac{1}{2}(u+v)\right) \leq \tfrac{1}{2}\psi(u) + \tfrac{1}{2}\psi(v). \tag{V.1}$$

One can show that a convex function is necessarily continuous.[23] However, it need not be differentiable.

If ψ is in fact differentiable, then the inequality (V.1) is equivalent to each of the following (in the sense of validity for all $u, v \in I$):

$$\psi(u) - \psi(v) \geq (u-v)\psi'(v), \tag{V.2a}$$

$$\psi'(u) \geq \psi'(v) \quad \text{if} \quad u \geq v. \tag{V.2b}$$

If ψ is twice differentiable, (V.1) is also equivalent to

$$\psi''(v) \geq 0. \tag{V.3}$$

In particular, the real exponential function is convex on the whole line since

$$(d^2/du^2)e^{au} = a^2 e^{au} \geq 0.$$

The exponential function will be basic for what follows, and we will take R^1 as the interval I.

B. Two Important Inequalities.[24]

Suppose that ψ is a compound function, $\psi = \psi(x(u))$, where $u \in R^n$ and $x(u) \in R^1$. We define averages with respect to a non-negative weight function ρ:

$$\overline{\psi(x)} = \int d^n u\, \rho(u)\, \psi(x(u)) \Big/ \int d^n u\, \rho(u), \tag{V.4}$$

similarly for $\bar{x}$, etc. Now, if we replace in (V.2a) v by $\bar{x}$ and u by $x(u)$, and average both sides, then the right hand side gives zero, and we obtain (for a convex function ψ)

$$\overline{\psi(x)} \geq \psi(\bar{x}). \tag{V.5}$$

For the particular case $\psi(x) = e^x$, $x(u) = \log f(u)$ (where $f \geq 0$; the limiting case $f = 0$ causes no difficulties), (V.5) yields

$$\bar{f} \geq \exp(\overline{\log f}). \tag{V.6}$$

We call this the inequality of generalized arithmetic and geometric means. If $u \geq 0$, $f(u) = u$, $\bar{f} = \frac{1}{2}(u_1 + u_2)$, which corresponds to taking $\rho = \frac{1}{2}\delta(u-u_1) + \frac{1}{2}\delta(u-u_2)$, then (V.6) reduces to the more familiar result

$$\tfrac{1}{2}(u_1 + u_2) \geq (u_1 u_2)^{\frac{1}{2}}. \tag{V.7}$$

Suppose now that f is a function of two variables u and v. We will denote averaging with respect to u and v by a bar and a tilde, respectively. From (V.6) we obtain by averaging over v, and by replacing $f \to \tilde{f}$, respectively,

$$\tilde{\bar{f}} \geq \left[\exp(\overline{\log f})\right]^{\sim}, \quad \bar{\tilde{f}} \geq \exp(\overline{\log \tilde{f}}). \tag{V.8a,b}$$

By Fubini's theorem, $\tilde{\bar{f}} = \bar{\tilde{f}}$. It is significant that the two smaller members can also be related. One has

$$\exp(\overline{\log \tilde{f}}) \geq \left[\exp(\overline{\log f})\right]^{\sim}. \tag{V.9}$$

This is called the generalized Hölder inequality.

To prove this inequality, replace $f \to f/\tilde{f}$ in (V.6) and use $\overline{g+h} = \bar{g} + \bar{h}$:

$$\overline{f/\tilde{f}} \geq \exp(\overline{\log f}) \exp(\overline{-\log \tilde{f}}).$$

Now transpose the last factor to the left hand side, apply tilde-averaging, and use $\tilde{cg} = c\tilde{g}$, where c is constant. Then the left hand side becomes

$$\left[(\overline{f/\tilde{f}})\exp(\overline{\log \tilde{f}})\right]^{\sim} = (f/\tilde{f})^{-\sim}\exp(\overline{\log \tilde{f}}),$$

and (V.9) follows from

$$(f/\tilde{f})^{-\sim} = (f/\tilde{f})^{\sim -} = (\tilde{f}/\tilde{f})^{-} = 1.$$

We can easily obtain the usual Hölder inequality as a special case of the generalized one. Let u take the values 1,2; let

$$f(1,v) = |g(v)|^{p}, \quad f(2,v) = |h(v)|^{q},$$

where

$$p^{-1} + q^{-1} = 1 \quad (p,q \geq 1). \tag{V.10a}$$

For any function $F(u,v)$, let

$$\overline{F}(v) = p^{-1}F(1,v) + q^{-1}F(2,v),$$

$$\tilde{F}(u) = \int d^{n}v\,\rho(v)\,F(u,v) \Big/ \int d^{n}v\,\rho(v),$$

where $\rho \geq 0$. Now evaluate both sides of (V.9),

$$\exp(\overline{\log \tilde{f}}) = \tilde{f}^{1/p}(1)\tilde{f}^{1/q}(2) = \|g\|_{p}\|h\|_{q},$$

$$\left[\exp(\overline{\log f})\right]^{\sim} = \left[f^{1/p}(1,v)\,f^{1/q}(2,v)\right]^{\sim} = \|gh\|_{1}.$$

Thus (V.9) implies the standard result,

$$\|gh\|_{1} \leq \|g\|_{p}\|h\|_{q}. \tag{V.10b}$$

This last inequality can be easily put into a slightly more general form. If $p^{-1}+q^{-1}=r^{-1}$, where $p,q,r \geq 1$, then

$$\|gh\|_{r} \leq \|g\|_{p}\|h\|_{q}. \tag{V.11}$$

C. The Polaron Problem.[25]

An electron moving slowly in an ionic crystal (like NaCl) polarizes its own immediate neighborhood in the crystal. The system, electron and its polarization cloud, behaves like a single particle, and is called the polaron. Its effective mass is, of course, larger than that of an electron. The calculation of this mass, as suggested by Feynman, is an interesting application of the Wiener integral.

The real problem is, of course, extremely complicated, but it becomes manageable if we make the following approximations:

(i) All modes of vibration in the crystal have the same frequency ω.

(ii) The electron interacts with the crystal as if the latter were a continuous dielectric.

We refer to the cited literature for a fuller discussion of the physical ideas and of the manipulations.

Now, let $\vec{x}$ be the coordinate vector of the electron, and let $q_{\vec{k}}$ be the amplitude for the mode with the wave number $\vec{k}$. We set $\hbar = m = \omega = 1$ (then $c \neq 1$, but this is irrelevant here). The effective Lagrangian becomes

$$L = \tfrac{1}{2}\dot{\vec{x}}^2 + \tfrac{1}{2}\sum_{\vec{k}}\left(\dot{q}_{\vec{k}}^{\,2} - q_{\vec{k}}^{\,2}\right) + \alpha' \sum_{\vec{k}} q_{\vec{k}}\, e^{i\vec{k}\vec{x}} \Big/ |\vec{k}| \,. \tag{V.12}$$

The constant α' depends on the volume and on the dielectric properties of the crystal, while α'' below depends on the dielectric properties only.

We assume a given electron path $\vec{x}(t)$. Then each amplitude $q_{\vec{k}}$ satisfies the equation of a forced harmonic oscillator, and can be found explicitly in terms of the path $\vec{x}(t)$.

Let us specify that the oscillators be in the ground state initially and finally ($t = t_1, t_2$ respectively, and the ground state depends on the path $\vec{x}$). Then the resulting action of the electron, i.e., the time integral of T-V, is

$$S = \tfrac{1}{2}\int_{t_1}^{t_2} d\tau\, \dot{\vec{x}}^2(\tau) + \alpha'' i \int\int_{t_1}^{t_2} d\tau\, d\sigma\, |\vec{x}(\tau) - \vec{x}(\sigma)|^{-1} e^{-i|\tau-\sigma|}, \tag{V.13}$$

where we took the limit of infinite volume. We see an effective potential which is nonlocal in time. This is not surprising, since S describes only a part of an intricately coupled system.

We saw in Section IV.E that for an ordinary potential system, we obtain the partition function from the quantum-mechanical wave function which corresponds to imaginary time, i.e., we replace $t \to -it$. For the polaron problem this means that the effective interaction in (V.13) has to be modified, since it involves two times τ and σ rather than τ alone.

One can show that the partition function for this problem becomes

$$Z = \int_V d^3\vec{v} \int_{C(\vec{v},\vec{v})} d\mu_W(\vec{x}) e^{-F\{x\}}, \tag{V.14a}$$

$$F\{x\} = \alpha'' \int\int_0^\beta d\tau\, d\sigma\, |\vec{x}(\tau) - \vec{x}(\sigma)|^{-1} e^{-|\tau-\sigma|}. \tag{V.14b}$$

(Here V is the volume, see below.) One may also write Z as

$$Z = \int_V d^3\vec{v} \int_{C(\vec{v},\vec{v})} \mathcal{D}_W(\vec{x}) e^{-S_{eff}\{\vec{x}\}}, \quad S_{eff} = \tfrac{1}{2}\int_0^\beta d\tau\, \dot{\vec{x}}^2 + F. \tag{V.15}$$

Now,

$$Z = \sum_j e^{-\beta E_j},$$

and for large β only the ground state is significant, i.e.,

$$Z \sim e^{-\beta E_0} \quad \text{as} \quad \beta \to \infty. \tag{V.16}$$

If we could evaluate the integral (V.14a), our problem would be solved. However, we have to content ourselves with obtaining some estimates. (See Section V.D.)

We may point out that the restriction to oscillators in the ground state is now justified by (V.16).

We indicated a finite volume for the integral over $d^3\vec{v}$. In fact, the integrand is translation-invariant so that $Z \propto V$ and $Z \to \infty$ with volume. For quantities like E_0, the dependence on volume will of course disappear. One may, however, accuse us of inconsistency in taking the infinite volume limit in (V.13). But here the limit is finite, and we may regard the limiting value as an approximation to an already oversimplified problem.

D. Estimates for Z.

We first construct an action for which the integral (V.15) can be evaluated,[25] for example,

$$S' = \tfrac{1}{2}\int d\tau\, \dot{\vec{x}}^2 + \tfrac{1}{2}\xi \int\int d\tau\, d\sigma\, |\vec{x}(\tau) - \vec{x}(\sigma)|^2 e^{-\eta|\tau-\sigma|}, \tag{V.17}$$

where the parameters ξ and η can be varied so as to yield the best value for the ground state energy. We will now use $e^{-S'}$ as the basic weight for computing averages. For translation-invariant

functionals, as in (V.14b) and in (V.17), we can take the integral over $\vec{v}$ for a unit volume, and then

$$\langle G\rangle = \frac{1}{Z'}\int_{C(\vec{v},\vec{v})} \mathfrak{D}_W(\vec{x}) e^{-S'\{\vec{x}\}} G\{\vec{x}\}, \tag{V.18a}$$

where

$$Z' = \int_{C(\vec{v},\vec{v})} \mathfrak{D}_W(\vec{x}) e^{-S'\{\vec{x}\}}. \tag{V.18b}$$

Now, $e^{-S} = e^{S'-S} e^{-S'}$ [where we wrote S in place of S_{eff} of (V.15)] and it follows that $Z = Z'\langle e^{S'-S}\rangle$. The inequality (V.7) now yields

$$Z \geq Z' e^{\langle S'-S\rangle}. \tag{V.19}$$

If E_o' is the ground state energy corresponding to S', then we obtain finally, in view of (V.16),

$$E_o \leq E_o' - \langle S'-S\rangle/\beta. \tag{V.20}$$

The expectation $\langle S\rangle$ can be evaluated by employing a Fourier transformation and thus we get an upper limit for E_o.

Suppose now that we are dealing with a system governed by a Hamiltonian $H = H_o + V_o + \Delta V$, which can be approximated by the soluble Hamiltonian $H_o + V_o$. Then the actions S and S' are single integrals, and not double integrals as in the polaron problem. In this case (V.19) can be transformed to[24]

$$Z \geq Z' e^{-\beta\overline{\Delta V}} \tag{V.21}$$

where

$$Z = \mathrm{Tr}\, e^{-\beta(H_o+V_o+\Delta V)}, \quad Z' = \mathrm{Tr}\, e^{-\beta(H_o+V_o)}, \tag{V.22}$$

$$\overline{\Delta V} = \mathrm{Tr}\left(e^{-\beta(H_o+V_o)}\Delta V\right)\Big/ Z'. \tag{V.23}$$

To establish (V.21), note that by interchanging integrations and by using (IV.25), (IV.26) and (IV.31), we obtain

$$\langle S'-S\rangle = -\frac{1}{Z'}\int_0^\beta d\sigma \int d^3\vec{v} \int_{C(\vec{v},\vec{v})} d\mu_W(\vec{x}) \exp\Big[-\int_0^\beta d\tau\, V_0\{x(\tau)\}\Big]\Delta V\{x(\sigma)\}$$

$$= -\frac{1}{Z'}\int_0^\beta d\sigma \int d^3\vec{v} \int d^3\vec{u}\, \langle\vec{v}|\, e^{-\sigma(H_0+V_0)}\,|\vec{u}\rangle$$

$$\langle\vec{u}|\, e^{-(\beta-\sigma)(H_0+V_0)}\,|\vec{v}\rangle\, \Delta V(\vec{u})$$

$$= -\frac{1}{Z'}\beta \int d^3\vec{u}\, \langle\vec{u}|\, e^{-\beta(H_0+V_0)}\,|\vec{u}\rangle\, \Delta V(\vec{u}) = -\beta\overline{\Delta V}. \qquad (V.24)$$

An upper bound for Z can be obtained just as easily. We use the inequality (V.8a) as follows:

$$Z = (\exp\widehat{\log f})^{\sim} \le (\hat{f})^{\sim} = (\tilde{f})^{\wedge}$$

with

$$f\{x,\sigma\} = \exp\Big[-\beta\,\Delta V\{x(\sigma)\}\Big],$$

$$\hat{g} = \frac{1}{\beta}\int_0^\beta d\sigma\, g, \quad \tilde{g} = \int d^3\vec{v} \int_{C(\vec{v},\vec{v})} d\mu_W(\vec{x}) \exp\Big(-\int_0^\beta d\tau\, V_0\Big) g.$$

We proceed as before and obtain

$$Z \le \mathrm{Tr}\Big(e^{-\beta(H_0+V_0)}\, e^{-\beta\Delta V}\Big).$$

We can express the two bounds as follows:

$$Z' e^{-\beta\overline{\Delta V}} \le Z \le Z' \overline{e^{-\beta\Delta V}}. \qquad (V.25)$$

Note that if we had used the generalized Hölder inequality (V.9) in place of (V.8a), a sharper upper bound would have resulted.

VI. The Commutative Hilbert Space Integral

In this section we introduce another approach to functional integration. Here the integral is, in effect, over an abstract Hilbert space $\mathcal{H}$, and the realization of $\mathcal{H}$ as a function space is irrelevant. However, this integral is closely related to the Wiener integral. The present approach is based on the work of Friedrichs and Shapiro.[26)27)]

See also Karpacz lectures of this author.[28] Finally, we note a comprehensive recent review of Segal.[29]

A. Cylinder Functionals.

Let $\mathcal{K}$ be a real Hilbert space. A functional F over $\mathcal{K}$ is called a tame or cylinder functional if (i) there exists a subspace E of dimension $n<\infty$ such that

$$F(x) = F(P_E x), \tag{VI.1}$$

where P_E is the orthogonal projection onto E, and if (ii) the function F is measurable with respect to the Lebesque measure on E; i.e., $dx^1 \ldots dx^n$, where the x^j are the Cartesian coordinates. Under these conditions we say that F has a base in E.

We now introduce the orthogonal invariant measure in E,

$$d\mu_E(x) = (2\pi\sigma)^{-\frac{1}{2}n} dx^1 \ldots dx^n \exp\Big[-\sum (x^j)^2/2\sigma\Big]. \tag{VI.2}$$

Here σ is an arbitrary positive parameter, called variance, which will usually be suppressed in the notation. This choice of measure can be justified in several ways. For example, it is not difficult to show that it is a consequence of (i) finiteness and normalization, $\mu(R^n)=1$, (ii) orthogonal invariance, and (iii) factorizability,

$$d\mu(x) = dF_1(x^1) \ldots dF_n(x^n), \tag{VI.3}$$

which corresponds to probabilistic independence.

Let $\mathcal{C}$ be the set of all cylinder functionals on $\mathcal{K}$. If $F\in\mathcal{C}$ has a base in E, and is in L_1 with respect to $d\mu_E$, we set

$$I_E(F) = \int d\mu_E(x) F(P_E x). \tag{VI.4a}$$

If $\hat{E}\supset E$, then F also has a base in $\hat{E}$, and integration over each redundant coordinate gives unity, in view of the factor $(2\pi\sigma)^{-\frac{1}{2}}$. Consequently, $I_{\hat{E}}(F)=I_E(F)$.

Now, if F has a base in E and in E', it follows that

$$I_E(F) = I_{E+E'}(F) = I_{E'}(F).$$

Thus we may define an integral over $\mathcal{K}$ by

$$I_{\mathcal{K}}(F) = I_E(F), \tag{VI.4b}$$

and this definition does not depend on the choice of the base E.

We also define, for $\alpha \geq 1$,

$$\mathcal{C}_\alpha = \left\{ F \in \mathcal{C} : I_{\mathcal{K}}\left(|F|^\alpha \right) < \infty \right\}. \tag{VI.5a}$$

These spaces may be equipped with the respective norms,

$$\| F \|_\alpha = \left[I_{\mathcal{K}}\left(|F|^\alpha \right) \right]^{1/\alpha}, \tag{VI.5b}$$

and may be completed in the usual way.

B. Integration of Invariant Functionals.

The next problem is to extend the foregoing integral to more general functionals. One way to go about it is the following. Let $\{y_j\}$ be an orthonormal basis and let P_n be the projection onto the linear span of $\{y_1, \ldots, y_n\}$. A system like $\{P_n\}$ will be called a basic system.

For a given functional F, we define $F_n^P(x) = F(P_n x)$. The F_n^P are cylinder functionals, and they might form a Cauchy sequence in the space $\mathcal{C}_\alpha$ (with respect to the α-norm). If so, we could try to identify F with an element of the completion of $\mathcal{C}_\alpha$. The difficulty is that another basic system $\{Q_m\}$ might perhaps give us another element of the completion of $\mathcal{C}_\alpha$.

We are thus led to the following definition. A functional F is α-invariant if for any two basic systems $\{P_n\}$ and $\{Q_m\}$,

$$\text{(i)} \quad F_n^P, F_m^Q \in \mathcal{C}_\alpha \text{ for all } n, m \text{ sufficiently large}, \tag{VI.6a}$$

and

$$\text{(ii)} \quad \lim_{n,m \to \infty} \| F_n^P - F_m^Q \|_\alpha = 0. \tag{VI.6b}$$

The space of all α-invariant functionals will be denoted by $\mathfrak{L}_\alpha$. It follows easily from Hölder's inequality (V.11) and the fact that $\|1\|_\beta = 1$ for all $\beta \geq 1$, that

$$F \in \mathfrak{L}_\alpha \Rightarrow F \in \mathfrak{L}_\gamma \text{ for all } \gamma, \quad 1 \leq \gamma \leq \alpha. \tag{VI.7}$$

In particular, if $F \in \mathfrak{L}_\alpha$ then $F \in \mathfrak{L}_1$, and we may define the basis-independent integral $I_{\mathcal{K}}(F)$, which will also be denoted by

$$I_{\mathcal{K}}^\sigma(F), \qquad \int \mathfrak{D}(x) e^{-\langle x,x \rangle / 2\sigma} F(x), \quad \text{etc.} \tag{VI.8}$$

Invariance of the integrand is sufficient to establish independence of the integral of the basis used in the approximations, and

for various other desirable properties. But for some purposes it is convenient to introduce a more restrictive notion. We say that F is completely α-invariant if conditions (i) and (ii) above are satisfied for arbitrary systems of projections $\{P_n\}$, $\{Q_m\}$ such that $P_n, Q_m \to 1$ as $n, m \to \infty$.

Theorem. Every functional in C_α is completely α-invariant.

We do not prove this theorem, but illustrate it by an example. Let $\alpha = 2$, $b \in \mathcal{H}$, and $F(\eta) = \langle b, \eta \rangle$. For any projection P,

$$F(P\eta) = \langle b, P\eta \rangle = \langle Pb, \eta \rangle .$$

The cylinder functionals are now easily integrated,

$$I_{\mathcal{H}}\left(\left(F_n^P - F_m^Q\right)^2\right) = I_{\mathcal{H}}\left(\langle P_n b - Q_m b, \cdot \rangle^2\right)$$

$$= \sigma \| P_n b - Q_m b \|^2 \longrightarrow 0,$$

and we also find $I_{\mathcal{H}}(F^2) = \sigma \| b \|^2$. We note a slight generalization,

$$F_j(\eta) = \langle b_j, \eta \rangle \Rightarrow I_{\mathcal{H}}^{\sigma}(F_1 F_2) = \sigma \langle b_1, b_2 \rangle . \tag{VI.9}$$

One can also give an example of a non-invariant functional. Let $\mathcal{H}$ be the space of square-summable sequences; let $x = (x^1, x^2, \ldots)$, and let

$$G(x) = \sum c_{ij} x^i x^j . \tag{VI.10a}$$

If $g^{ij}(x) = x^i x^j$, then $I_{\mathcal{H}}(g^{ij}) = \sigma \delta^{ij}$ by (VI.9), and it follows that

$$I_{\mathcal{H}}(G) = \sigma \sum c_{ii} = \sigma \, \mathrm{Tr}(c_{ij}) . \tag{VI.10b}$$

But the trace is not, in general, basis-independent. Consider the two following operators which are related by an orthogonal transformation, and are in fact Hilbert-Schmidt:

$$\begin{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & & & \\ & \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix} & & \\ & & \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{1}{3} & 0 \end{pmatrix} & \\ & & & \ddots \end{pmatrix} \cdot \begin{pmatrix} \begin{pmatrix} 1 & \\ & -1 \end{pmatrix} & & & \\ & \begin{pmatrix} \frac{1}{2} & \\ & -\frac{1}{2} \end{pmatrix} & & \\ & & \begin{pmatrix} \frac{1}{3} & \\ & -\frac{1}{3} \end{pmatrix} & \\ & & & \ddots \end{pmatrix} \cdot \tag{VI.11}$$

The trace of the first vanishes, and that of the second is indeterminate.

C. Connection with the Wiener Integral.

We mentioned earlier that the Wiener integral for, say, the interval $[0,t]$ with $x(0)=0$, is in fact the integral over the Hilbert space of functions $x(\tau)$ which satisfy

$$x(0) = 0 \quad \text{and} \quad \langle \dot{x}, \dot{x} \rangle \equiv \int_0^t d\tau\, \dot{x}^2(\tau) < \infty \tag{VI.12}$$

and are necessarily continuous.

The identification can now be seen as follows. We can introduce projection operators $\{P_n\}$ which transform a function x into the polygonal approximations as described in Section IV.A. In view of continuity of x, $P_n \to 1$ as $n \to \infty$. Hence, if F is a completely invariant functional over the space defined by (VI.12), these approximations converge, independently of the chosen set $\{P_n\}$. We then have convergence of the Wiener integral (in the sense of Section IV.A).

We shall not consider the possibility of existence of functionals which are integrable in one sense but not in another—for example, invariant but not completely invariant.

If we make the change of variable $\dot{x} = y$, then we have $y(0)$ unrestricted, $\langle \dot{x}, \dot{x} \rangle = \int d\tau\, y^2$, and we can express the identification of the two integrals as

$$\int \mathfrak{D}_W(x)\, e^{-\langle \dot{x}, \dot{x} \rangle / 2\sigma} F\{x\} = \int \mathfrak{D}(y) e^{-\langle y, y \rangle / 2\sigma} G\{y\}, \tag{VI.13a}$$

$$G\{y\} = F\{H_W y\}, \quad (H_W y)(\tau) = \int_0^t d\tau'\, \theta(\tau - \tau') y(\tau') = \langle \theta_\tau, y \rangle, \tag{VI.13b}$$

where $\theta_\tau(\tau') = \theta(\tau - \tau')$. In Section IV we took $\sigma = \frac{1}{2}$, but now we can be more general.

Let us verify the foregoing relation for the functionals

$$F\{x\} = x(t')x(t''), \quad G\{y\} = \langle \theta_{t'}, y\rangle \langle \theta_{t''}, y\rangle.$$

For the functional G, the Hilbert space integral was evaluated in (VI.9),

$$I_{\mathcal{K}}^{\sigma=1}(G) = \langle \theta_{t'}, \theta_{t''}\rangle = \min(t', t''),$$

in agreement with Eq. (IV.14b) for $\int d\mu_W F$.

D. Some Related Integrals.

We can put the foregoing discussion into a more general context as follows. Suppose that we have a functional integral which is characterized (in some sense) by the Gaussian weight $e^{-\frac{1}{2}\langle \eta, B\eta\rangle}$, with B positive and symmetric. Then, by a linear change of variable of integration, we can transform the integral into the Hilbert space integral in canonical form:

$$\int \mathfrak{D}(\eta) e^{-\frac{1}{2}\langle \eta, B\eta\rangle} F\{\eta\} \sim \int \mathfrak{D}(\chi) e^{-\frac{1}{2}\langle \chi, \chi\rangle} F\{B^{-\frac{1}{2}}\chi\}. \qquad \text{(VI.14)}$$

This relation is only heuristic, and should be examined for each case as needed. Sometimes one can use the right hand side to define the left hand side.

Such relations arise in different ways. For example, in problems relating to noise,[25] one finds Gaussian weights which are given in terms of integrals with kernels,

$$\langle \eta, B\eta\rangle = \int d^n u\, d^n v\, \eta(u) b(u,v) \eta(v) \equiv \langle \eta, \eta\rangle_B. \qquad \text{(VI.15)}$$

Then the integral can be naturally interpreted as the integral over the Hilbert space $\mathcal{K}_B$, defined by the inner product $\langle \eta, \eta\rangle_B$. (We repeat, B is assumed positive and symmetric.)

One should not confuse an integral over $\mathcal{K}_B$ with an integral over the space $\mathcal{K}$, defined by $\langle \eta, \eta\rangle$. The two integrals differ by a factor:

$$\int_{\mathcal{K}_B} \mathfrak{D}(x) e^{-\frac{1}{2}\langle x, Bx\rangle} F(x) = \int_{\mathcal{K}} \mathfrak{D}(x) e^{-\frac{1}{2}\langle x, x\rangle} F(B^{-\frac{1}{2}}x), \qquad \text{(VI.16a)}$$

$$\int_{\mathcal{K}} \mathfrak{D}(x) e^{-\frac{1}{2}\langle x, Bx\rangle} F(x) = (\det B)^{-\frac{1}{2}} \int_{\mathcal{K}} \mathfrak{D}(x) e^{-\frac{1}{2}\langle x, x\rangle} F(B^{-\frac{1}{2}}x), \qquad \text{(VI.16b)}$$

where the left hand side of Eq. (VI.16b) has as the actual integrand $e^{-\frac{1}{2}\langle x,(B-1)x\rangle}F(x)$. The determinant is defined for operators B such that B - 1 is of the trace class [30] by the equation

$$\det B = \exp(\mathrm{Tr}\log B). \tag{VI.17}$$

Other examples of analogous functional integrals arise in quantum statistical mechanics.[31] One starts with

$$e^{-\beta V} = \exp\Big[-\beta\int d^3u\, d^3v\, \psi^*(u)\psi(u)\Phi(u-v)\psi^*(v)\psi(v)\Big],$$

where Φ is the two-body potential. It can be advantageous to transform this expression to

$$(\text{const.})\int \mathfrak{D}(\eta)e^{-\langle\eta,A\eta\rangle}\exp\Big[-\int d^3u\,\psi^*(u)\psi(u)\eta(u)\Big], \tag{VI.18a}$$

where A is an operator, in general integro-differential, satisfying

$$A\Phi = \delta. \tag{VI.18b}$$

In order to give a precise meaning to the integral (VI.18a) if A involves differentiations, one has to specify the bounday conditions. For example, for $A=-d^2/du^2$, with $\eta(0)=0$, $\eta(t)$ unrestricted, we have the Wiener integral, and the corresponding transformation was discussed in Section VI.C.

If we also impose $\eta(t)=0$, then we have a special case of the conditional integral, and in place of H_W, Eq. (VI.13b), one should use

$$(H\eta)(\tau) = \int_0^t d\tau'\Big[\theta(\tau-\tau')-\tau/t\Big]\eta(\tau'). \tag{VI.19}$$

Then HH^T has as its kernel the Green's function G for $-d^2/du^2$ and for the prescribed boundary conditions,

$$G(\tau,\tau') = \min(\tau,\tau') - \tau\tau'/t. \tag{VI.20}$$

On the other hand, for $H_WH_W^T$ we find

$$G_W(\tau,\tau') = \min(\tau,\tau'). \tag{VI.21}$$

VII. Evaluations and Transformation Laws

In this section we give formulas primarily for the Hilbert space integral. They can be adapted to the Wiener integral in accordance with Section VI.C. (See also Section IX.) The transformation laws include those for modifying an integral, and also functional Fourier transforms. Further discussion of these and related topics can be found in Reference 28.

From now on, we will usually omit the word "commutative" when referring to the Hilbert space integral of Section VI. However, in Section XII we shall acquaint ourselves with the anticommutative integral, and there the qualification will be necessary.

A. Explicit Evaluations.

We give here a list of functionals whose integrals can be given in terms of other familiar operations.

(1) Integrable cylinder functionals reduce to finite-dimensional integrals (in view of invariance, Section VI.B). In particular, for the linear exponential we find

$$\int \mathcal{D}(x) e^{-\langle x,x\rangle/2\sigma} e^{c\langle x,y\rangle} = e^{\frac{1}{2}\sigma c^2\langle y,y\rangle}. \qquad \text{(VII.1)}$$

Here c is a constant, which can be complex, and then we cannot write $cy = y' \in \mathcal{H}$.

(2) A polynomial functional is one of the form

$$P(y) = P(0) + \sum c_j y^j + \cdots + \sum c_{j_1 \ldots j_m} y^{j_1} \ldots y^{j_m}. \qquad \text{(VII.2a)}$$

If $\mathcal{H} = L^2(R^n)$, the terms of a given order k can be combined into

$$P_k\{\eta\} = \int d^n u_1 \ldots d^n u_k p_k(u_1, \ldots, u_k) \eta(u_1) \ldots \eta(u_k), \qquad \text{(VII.2b)}$$

where p_k is symmetric. The functional P_k is a sum of monomials such as shown in (VII.2a). Each monomial is a cylinder functional whose integral can be readily found.

We observe that $I_{\mathcal{H}}(P_{2\kappa+1}) = 0$. The integral $I_{\mathcal{H}}(P_{2\kappa})$ can be expressed in terms of the trace of an operator defined by the function $p_{2\kappa}$.

We may also recall that special cases of polynomial functionals were treated in Sections IV.B and VI.B.

(3) Gaussian functionals were encountered in Sections IV.B and V.C. A general integrable Gaussian functional is

$$F(x) = \exp(-\langle x, Ax\rangle/2\sigma), \qquad \text{(VII.3)}$$

where A is a symmetric trace-class operator, such that $1+A$ is strictly positive. Then

$$I_{\mathcal{H}}^{\sigma}(F) = \int \mathcal{D}(x) e^{-\langle x,(1+A)x\rangle/2\sigma} = \int \mathcal{D}(x) e^{\langle (1+A)^{\frac{1}{2}}x,(1+A)^{\frac{1}{2}}x\rangle/2\sigma}$$

$$= \left[\det(1+A)\right]^{-\frac{1}{2}}. \qquad \text{(VII.4)}$$

We obtained this result by making a linear change of variable. We shall turn to this topic presently. First, however, let us verify that the foregoing conditions on A are fulfilled for the example of Section IV.B, where

$$\langle \dot{x}, A\dot{x}\rangle = \int_0^t d\tau\, p(\tau) x^2(\tau). \qquad \text{(VII.5a)}$$

This form defines implicitly a symmetric positive operator A. We may construct A explicitly as follows. Let

$$(B\dot{x})(\tau) = p^{\frac{1}{2}}(\tau)x(\tau) = \int_0^t d\tau' K(\tau,\tau')\dot{x}(\tau'),$$

$$K(\tau,\tau') = p^{\frac{1}{2}}(\tau)\theta(\tau-\tau').$$

Then $A = B^{T}B$ (note that B is not symmetric). The kernel of B^{T} is $K^{T}(\tau,\tau') = K(\tau',\tau)$, and that of A,

$$K_A(\tau,\tau') = \int_0^t d\tau'' K^{T}(\tau,\tau'')K(\tau'',\tau') = \int_0^t d\tau'' p(\tau'')\,\theta(\tau''-\tau)\,\theta(\tau''-\tau')$$

$$= \int_{\max(\tau,\tau')}^t d\tau''\, p(\tau'').$$

We see that A is a trace-class operator:

$$\operatorname{Tr} A = \int_0^t d\tau\, K_A(\tau,\tau) = \int_0^t d\tau \int_\tau^t d\tau''\, p(\tau'') < \infty. \qquad \text{(VII.5b)}$$

B. Transformations of the Integral.

(1) Linear change of variables. Let L be a trace-class linear operator such that $(1+L)^{-1}$ exists. Then

$$I_{\mathcal{K}}^{\sigma}\Big(F[(1+L)\,\cdot]\Big) = \det|1+L|^{-1}\int \mathfrak{D}(x)e^{-\langle x,x\rangle/2\sigma}\,F(x)$$

$$\times \exp\Big\{-\Big[\langle (1+L)^{-1}x,\ (1+L)^{-1}x\rangle - \langle x,x\rangle\Big]\Big/2\sigma\Big\}. \tag{VII.6}$$

We assume that both functionals are in $\mathfrak{L}_1$. Then this equation may be confirmed in any basic system. We set $|B| = (B^*B)^{1/2}$.

(2) Translations. If $y\in\mathcal{K}$, and $F_y(x) = F(x+y)$ is in $\mathfrak{L}_1$, then translational invariance of the approximating integrals yields

$$I_{\mathcal{K}}^{\sigma}(F_y) = e^{-\langle y,y\rangle/2\sigma}\int \mathfrak{D}(x)e^{-\langle x,x\rangle/2\sigma}e^{\langle x,y\rangle/\sigma}F(x). \tag{VII.7}$$

In this case we can be quite explicit about the integrability of the second integrand. Let $F\in\mathfrak{L}_\alpha$.

First we observe that $E^{(y)}(x) = e^{\langle x,y\rangle/\sigma}$ satisfies

$$\|E^{(y)}\|_\beta < \infty \quad \text{for all} \quad \beta \geq 1. \tag{VII.8a}$$

Next, Hölder's inequality (V.11) states that

$$\|FE^{(y)}\|_\gamma \leq \|F\|_\alpha \|E^{(y)}\|_\beta \quad \text{for} \quad \alpha^{-1}+\beta^{-1} = \gamma^{-1}.$$

We can choose β^{-1} arbitrarily small but positive. It follows that

$$\|FE^{(y)}\|_\gamma < \infty \quad \text{for all} \quad \gamma, \quad 1 \leq \gamma < \alpha. \tag{VII.8b}$$

It also follows easily that $FE^{(y)}$ is γ-invariant, i.e., in $\mathfrak{L}_\gamma$.

However, in general one cannot take $\gamma = \alpha$. We give a simple counterexample for one dimension. Let

$$F(u) = e^{u^2/2\sigma}(1+u^2)^{-1}. \tag{VII.9a}$$

Then $F\in L_1$ (with reference to the Gaussian weight), but

$$\|FE^{(v)}\|_1 = \int_{-\infty}^{\infty} du(1+u^2)^{-1}e^{uv/\sigma} = \infty. \tag{VII.9b}$$

A historical remark may be of interest here. Cameron and Martin carried out extensive investigations on the Wiener integral. They showed in particular[32] that for suitable functionals one may

translate $F(x) \to F(x+y)$, provided $dy/d\tau$ is of bounded variation. (That is, provided one may write $dy/d\tau = z_+ - z_-$, with $z_\pm(\tau)$ monotonic and finite.) This was before the theory of the Hilbert space integral was developed. We now see that only a weaker condition has to be imposed,[33]

$$\int_0^t d\tau\, \dot{y}^2 < \infty.$$

In particular, $\dot{y}$ may become infinite near zero like, for example, $\tau^{-\frac{1}{4}}$, and then it is not of bounded variation.

(3) Integration by parts. The basic formula is

$$\int \mathfrak{D}(x) e^{-\langle x,x\rangle/2\sigma} (\partial/\partial x^j - x^j/\sigma) F(x) = 0, \qquad \text{(VII.10a)}$$

provided the integrands are in $\mathfrak{L}_1$. This is a consequence of integration by parts in the x^j-coordinate, and the Gaussian $\exp[-(x^j)^2/2\sigma]$ of course has to be included in the integrand. Note that $\partial_j F$ might not be bounded, and hence we cannot identify this as a Fréchet differential.

In functional notation (VII.10a) becomes, for any $h \in L_2$,

$$\int \mathfrak{D}(\eta)\, e^{-\langle \eta,\eta\rangle/2\sigma} \int d^n u\, h(u) \left[\frac{\delta}{\delta h(u)} - \frac{1}{\sigma}\eta(u)\right] F\{\eta\} = 0. \qquad \text{(VII.10b)}$$

We can adapt this for the Wiener integral as follows. Let $n=1$ and $[0,t]$ be the basic interval; let $h(\tau') = \theta(\tau - \tau')$ (with τ fixed), and replace η by $\dot{y}$. Then the inner integral becomes

$$\int_0^t d\tau'\, \theta(\tau-\tau') \left[\delta/\delta\dot{y}(\tau') - \dot{y}(\tau')/\sigma\right].$$

The second term integrates to $y(\tau)/\sigma$, and we want to convert the first to a derivative with respect to y.

The chain rule for functional derivatives takes the form

$$\frac{\delta F}{\delta \dot{y}(\tau')} = \int_0^t d\rho\, \frac{\delta F}{\delta y(\rho)} \frac{\delta y(\rho)}{\delta \dot{y}(\tau)}. \qquad \text{(VII.11)}$$

In view of the condition $x(0)=0$,

$$y(\rho) = \int_0^t d\tau''\, \theta(\rho-\tau'')\, \dot{y}(\tau''), \quad \delta y(\rho)/\delta\dot{y}(\tau') = \theta(\rho-\tau').$$

By putting all of this together we obtain

$$\int_0^t d\tau' \theta(\tau-\tau') \delta/\delta \dot{y}(\tau') = \int_0^t d\rho \Big[\int_0^t d\tau' \theta_\tau \theta_\rho \Big] \delta/\delta y(\rho),$$

and, finally,[34]

$$\int_0^t d\rho \, \min(\rho,\tau) \int d\mu_W(y) \, \delta F\{y\}/\delta y(\rho) = \sigma^{-1} \int d\mu_W(y) y(\tau) F\{y\}. \tag{VII.12}$$

We recognize the kernel $\min(\rho,\tau)$ as the Green's function associated with the Wiener integral, Eq. (VI.21).

(4) Uniform convergence. A basic result is the following. Let S be a set of real numbers of finite measure. Let F_s be functionals, $F_s \in \mathfrak{L}_\alpha$ for almost all $s \in S$. Suppose that for arbitrary basic systems $\{P_n\}$, $\{Q_m\}$ and $\epsilon > 0$ we may find N, M, and C such that

$$\| F_s(P_n \cdot) \|_\alpha < C, \tag{VII.13a}$$

$$\| F_s(P_n \cdot) - F_s(Q_m \cdot) \|_\alpha < \epsilon, \tag{VII.13b}$$

where N, M, and C are independent of s. Then:

(i) We may interchange integrations,

$$\int ds \int \mathfrak{D}(x) e^{-\langle x,x\rangle/2\sigma} F_s(x) = \int \mathfrak{D}(x) e^{-\langle x,x\rangle/2\sigma} \int ds \, F_s(x). \tag{VII.14}$$

(ii) We may differentiate a Hilbert space integral under the $\int$-sign if the resulting integrand satisfies the foregoing conditions (for some α).

An integral $I_{\mathcal{H}}(F_s)$ satisfying the foregoing conditions is appropriately called uniformly α-convergent.

C. Functional Fourier Transforms.

We shall refer in this way to functional transforms of the general type

$$F\{\eta\} \to \tilde{F}\{\chi\} = \int d\mu(\eta) \exp\Big(i \int d^n u \, \chi \eta \Big) F\{\eta\}. \tag{VII.15}$$

The differential $d\mu(\eta)$ is symbolic. It is supposed to include the basic weight, like the Gaussian, but no measure is implied here. [Perhaps a more suitable notation for such differentials would be $\mathfrak{D}\mu(\eta)$.]

A basic property of such transforms is that they allow one to replace differentiation by multiplication. If P is a polynomial functional, then one expects to conclude that under suitable conditions,

$$P\{\delta/i\delta\chi\}\,\tilde{F}\{\chi\} = \int d\mu(\eta) e^{i\langle\chi,\eta\rangle} P\{\eta\} F\{\eta\}. \tag{VII.16}$$

Let us give some formulas for the Hilbert space integral. It is convenient to replace $e^{i\langle x,y\rangle}$ by $e^{\langle x,y\rangle}$ so as to bypass the complex numbers. Then

$$F\{y\} = 1 \Rightarrow \tilde{F}(x) = e^{\frac{1}{2}\sigma\langle x,x\rangle} \tag{VII.17}$$

We can now express the integral of suitable polynomials by

$$I^{\sigma}_{\mathcal{H}}(P) = P\{\delta/\delta x\}\, e^{\frac{1}{2}\sigma\langle x,x\rangle}\Big|_{x=o}. \tag{VII.18}$$

The essential condition on P is that one can interchange differentiation and integration. Hence the theorem of Section VII.B on uniform convergence is relevant. It is easy to see that (VII.18) applies in particular to integrable cylinder polynomial functionals.

Let us return to the transform (VII.15). It can be used sometimes as follows. Suppose that we are interested in functions $f_n(u_1,\ldots,u_n)$ which are the expectation values:

$$f_n(u_1,\ldots,u_n) = \int d\mu(x) x(u_1)\ldots x(u_n). \tag{VII.19}$$

Then a generating functional F for the f_n can be constructed simply as follows,

$$F\{y\} = \int d\mu(y) e^{i\langle x,y\rangle}. \tag{VII.20}$$

These considerations apply at once to the Wiener integral. The expectations f_n are known, Eqs. (IV.14) (where $\sigma = 1$), and the generating functional can be readily found,

$$\int d\mu_W(x) e^{i\langle y,x\rangle} = \exp\Big[-\tfrac{1}{2}\int_o^t d\tau'\, d\tau''\, \min(\tau',\tau'') y(\tau') y(\tau'')\Big]. \tag{VII.21}$$

[This evaluation follows, e.g., from $I_{\mathcal{H}}(e^{i\langle\eta,\cdot\rangle}) = e^{-\frac{1}{2}\langle\eta,\eta\rangle}$.] The Fourier transform as in (VII.21) is also called the characteristic functional of $d\mu_W$.

The foregoing observations on the generating functional can also be adapted to the Hilbert space integral. Here, however, one encounters integrals which require a distribution-theoretic interpretation, e.g.,

$$\int \mathcal{D}(x) e^{-\frac{1}{2}\langle x,x\rangle} x(u)x(v) = \delta(u-v). \tag{VII.22}$$

We conclude with two remarks. First, the maps $F \to \tilde{F}$ sometimes are not very convenient, since they do not preserve the 2-norms. In general,

$$\int d\mu_W |F|^2 \neq \int d\mu_W |\tilde{F}|^2, \quad I_{\mathcal{H}}(|F|^2) \neq I_{\mathcal{H}}(|\tilde{F}|^2). \tag{VII.23}$$

We note for reference that there are related maps which do preserve the 2-norms. They are the Fourier-Wiener transform,[35]

$$F(y) \to G(y) = \int d\mu_W(x)\, F(\sqrt{2}\, x + iy), \tag{VII.24a}$$

and the Wiener transform,

$$F(y) \to F^W(y) = \int \mathcal{D}(x) e^{-\langle x,x\rangle/2\sigma} F(\sqrt{2}x + iy). \tag{VII.24b}$$

The exponential factor $\exp(-i\langle x,y\rangle/\sqrt{2}\sigma)$, or with $\langle \dot{x},\dot{y}\rangle$, appears if we make a translation, as described in Section VII.B.

As our second remark, we observe that one sometimes introduces heuristic functional Fourier transforms and δ-functions, which are supposed to satisfy the familiar relations for finite-dimensional spaces. Symbolically,

$$F(x) \longrightarrow \tilde{F}(y) \sim \int \mathcal{D}(x) e^{i\langle x,y\rangle} F(x), \tag{VII.25a}$$

$$\int \mathcal{D}(x) e^{i\langle x,y\rangle} \sim (\text{const.})\,\delta(y), \tag{VII.25b}$$

$$\int \mathcal{D}(x)\, \delta(x)\, F(x) = F(0). \tag{VII.26}$$

To the best of our knowledge, there is no rigorous basis for manipulations such as in Eqs. (VII.25a) and (VII.25b).

VIII. Functional Differential Equations

The equations in question provide an interesting field of application of functional integrals. Our present interest is in exact solutions to such equations, rather than in approximate solutions, or in manipulations such as in Sections III.C and III.D. However, the study of this subject is still in its infancy. In fact, only a few specific equations have been investigated. These studies (with very few exceptions) have not been at all systematic.

A. Simple Examples.

We start with several equations which can be discussed independently of functional integration. All except the first relate to statistical mechanics.

(1) The most elementary functional differential equation is of the form

$$\delta F\{J\}/\delta J(u) = H\{J;u\}, \tag{VIII.1}$$

where H is a given u-dependent, sufficiently smooth functional. If a twice-differentiable solution F exists, then commutativity of functional derivatives implies

$$\delta H\{J;u\}/\delta J(v) = \delta H\{J;v\}/\delta J(u). \tag{VIII.2}$$

It turns out that this condition is not only necessary but also sufficient for solubility.

To solve the equation, we start with the definition of the functional derivative and (VIII.1):

$$(d/d\epsilon)\, F\{\xi+\epsilon\eta\}\Big|_{\epsilon=0} = \int d^n v\, \eta(v)\, H\{\xi;v\}.$$

Let $\xi = sJ$, $\eta = J$. Then

$$(d/ds)\, F\{sJ\} = \int d^n v\, J(v)\, H\{sJ;v\}, \tag{VIII.3}$$

and by integration,

$$F\{J\} = F\{0\} + \int_0^1 ds \int d^n v\, J(v)\, H\{sJ;v\}. \tag{VIII.4}$$

Equation (VIII.2) is needed in order to verify that F is indeed a solution. In evaluating the functional derivative of the right hand side, we use moreover (VIII.3) with F replaced by H, and H by $\delta H/\delta sJ$, and integrate by parts.

If for suitable arguments J, H has a Volterra expansion (see (III.1)), then so does F, and we may write the solution immediately in the form

$$f_{n+1}(u_1,\ldots,u_{n+1}) = h_n(u_1,\ldots,u_n;u)\Big|_{u=u_{n+1}}. \qquad \text{(VIII.5)}$$

Symmetry of the f_{n+1} follows from (VIII.2).

(2) Our next example is due to Brittin and Chappell,[36] and refers to the vibrating string. Suppose that we have m coupled oscillators, described by the variables

$$t,\ y_\alpha,\ \dot{y}_\alpha = v_\alpha, \quad \text{where} \quad \alpha = 1,\ldots,m.$$

We can introduce a statistical distribution function f_m depending on t, y_α, v_α, and which satisfies the equation of continuity,

$$\frac{\partial f_m}{\partial t} + \sum_\alpha \left(v_\alpha \frac{\partial f_m}{\partial y_\alpha} + \dot{v}_\alpha \frac{\partial f_m}{\partial v_\alpha} \right) = 0. \qquad \text{(VIII.6)}$$

In the limit of a continuous distribution, y_α and v_α become functions y(u) and v(u) respectively, satisfying

$$\dot{y} = v, \qquad \dot{v} = c^2 \partial_u^2 y. \qquad \text{(VIII.7)}$$

Furthermore, f_m becomes a functional $f\{y,v;t\}$ which satisfies, in place of (VIII.6) and in view of (VIII.7),

$$\frac{\partial f}{\partial t} + \int_0^L du \left[v(u) \frac{\delta f}{\delta y(u)} + c^2 y''(u) \frac{\delta f}{\delta v(u)} \right] = 0. \qquad \text{(VIII.8)}$$

We now recall that for (VIII.7) the general solution is

$$y = g(u+ct) + h(u-ct),$$

where g and h are arbitrary twice-differentiable functions. One may also verify that (VIII.8) has the solution

$$f = \Omega\left\{ v(u-ct) + cy'(u-ct),\ v(u+ct) - cy'(u+ct) \right\}, \qquad \text{(VIII.9)}$$

where Ω is an arbitrary differentiable functional. It can, in fact, be shown that this is the general solution, under the additional assumption that v, y', and also f are analytic. This solution has not proven useful so far.

(3) Let us consider the generating functional $F\{\chi\}$ of Section III.D. Suppose that the potential Φ vanishes. Then Eq. (III.35)

reduces to

$$\frac{\partial F\{\chi\}}{\partial t} + \int d^6 r \sum_{\alpha=1} \frac{p^\alpha}{m} \left(\frac{\partial}{\partial q^\alpha} \frac{\delta F}{\delta \chi(r)} \right) \chi(r) = 0. \qquad \text{(VIII.10)}$$

This equation has solutions of the form

$$F\{\chi\} = \exp \int d^6 r \, f(t,r) \chi(r) \qquad \text{(VIII.11)}$$

where f is any solution of the equation of continuity,

$$\frac{\partial f}{\partial t} + \sum \frac{p^\alpha}{m} \frac{\partial f}{\partial q^\alpha} = 0. \qquad \text{(VIII.12)}$$

(Since $\Phi = 0$, the terms $\dot{p}^\alpha \partial f / \partial p^\alpha$ do not contribute.)

The reduced distribution functions implied by (VIII.11) factorize,

$$f_s\left(t; r_1, \ldots, r_s\right) = \prod_1^s f(t, r_j), \qquad \text{(VIII.13)}$$

in agreement with physical intuition. Similar factorizations have been encountered[37] in various attempts to solve equations such as (III.35).

Since Eq. (VIII.10) is linear, one can obtain other solutions as linear superpositions of the elementary solutions (VIII.11). We leave open the question, whether every solution can be obtained in this way.

(4) Let us take the differential equations[19),20)] for $F\{\zeta\}$ and G, as described in Section III.D. We introduce the notation

$$H_{q_1 \ldots q_m}\{\zeta\} = \delta^m H\{\zeta\} / \delta\zeta(q_1) \ldots \delta\zeta(q_m). \qquad \text{(VIII.14)}$$

For $\sigma = 0$, Eq. (III.36) can be written as

$$\frac{\partial}{\partial q^\alpha} F_q^o + \beta \int d^3 q' \, \frac{\partial \Phi(|q-q'|)}{\partial q^\alpha} \frac{\delta F_q^o\{\zeta\}}{\delta \zeta(q')} \zeta(q') = 0. \qquad \text{(VIII.15)}$$

This equation is similar to (VIII.10), and has solutions for which F_q^o has the form

$$F_q^o\{\zeta\} = (\text{const.}) \exp \int d^3 q' \, h(q,q') \zeta(q').$$

Compare now Eq. (VIII.1). The corresponding distribution functions f_s are proportional to the Boltzmann factors

$$e^{-\beta \Sigma \Phi(|q_i - q_j|)}.$$

Next, the equation for G, with $\sigma \neq 0$, is

$$\partial G_q / \partial q^\alpha + \beta \int d^3q' (\partial \Phi / \partial q^\alpha)(G_{qq'} + G_q G_{q'})[\zeta(q') + \sigma] = 0. \tag{VIII.16}$$

In this equation, we recall, the shift $\zeta = \bar{\zeta} - \sigma$ is permissible. The solutions to the zero-density equation are of the form $G^0 = \text{Log } F^0$, and for $\sigma \neq 0$ we then find

$$G\{\zeta\} = \log F^0\{a(\zeta + \sigma)\} + b. \tag{VIII.17}$$

The arbitrary constants (a, b, and those in F^0) are determined by physical considerations.

B. Applications of Functional Fourier Transforms.

Among the basic tools for the solution of ordinary and of partial differential equations are the integral transforms, especially those of Fourier and Laplace. For example, an ordinary equation with constant coefficients can be transformed in this way into an algebraic equation.

One may expect that functional Fourier transforms will become similarly important for functional differential equations. We describe several examples in the remainder of Section VIII. In two of these, a part of the problem is to find an equation which is satisfied by a given functional, as a preliminary to more general investigations.

(1) For completeness we mention here that variants of Eq. (III.26) for the time-ordered generating functional can be solved by a Fourier transformation.[38] We postpone further discussion to Sections XII.A to XII.C. There we shall also encounter another functional differential equation,[39] which however is rather trivial.

(2) Our second example is based on the approach of Hopf to hydrodynamic turbulence.[37),40] We start with a nonlinear deterministic equation,

$$\partial \eta(t, \vec{u}) / \partial t = P(\eta), \tag{VIII.18}$$

where P is a polynomial functional in η, not depending explicitly on time. We consider the space spanned by the solutions η. Suppose that we have an integral defined on this space, and that we introduce the expectations and their generating functional F as in Section VII.C:

$$F\{\chi; t\} = \int d\mu(\eta) \exp\left[i \int d^3\vec{u}\, \chi(\vec{u}) \eta(t, \vec{u}) \right]. \tag{VIII.19}$$

[We would like to have a measure which is ergodic, i.e., conserved along the motion, but for our purposes Eq. (VIII.19) is sufficient.]

Let us find the equation of motion for F. We will not concern ourselves with the detailed justification of the steps. Thus

$$\frac{\partial F}{\partial t} = \int d\mu(\eta)\left(i\int d^3\vec{u}\,\chi\,\frac{\partial \eta}{\partial t}\right)\exp\left(i\int d^3\vec{u}\,\chi\,\eta\right)$$

$$= \int d^3\vec{u}\, i\chi \int d\mu(\eta) P(\eta)\, \exp\left(i\int d^3\vec{u}\,\chi\,\eta\right),$$

and the desired equation is

$$\partial F/\partial t = \int d^3\vec{u}\, i\chi\, P(\delta/i\delta\chi)\, F\{\chi;t\}. \qquad \text{(VIII.20)}$$

For stationary turbulence, we set $\partial F/\partial t = 0$. One particular equation of interest is:

$$\int d^3k' \frac{\delta}{\delta\chi_\beta(k')} \int d^3k'' k_\beta'' \chi_\alpha^{tr}(k'+k'') \frac{\delta}{\delta\chi_\alpha(k'')} F\{\chi\} = 0, \qquad \text{(VIII.21)}$$

where k' and k" are the Fourier transform variables, $\chi = \vec{\chi}$ is a vector-valued function, and χ^{tr} its transverse part:

$$\vec{\chi}^{tr}(\vec{k}) = \vec{\chi}(\vec{k}) - \left[\vec{k}\cdot\vec{\chi}(\vec{k})/\vec{k}^2\right]\vec{k}. \qquad \text{(VIII.22)}$$

(We do not consider the familiar ambiguity which arises if χ is the gradient of a harmonic function.)

Equation (VIII.21) has solutions of the form

$$F\{\chi\} = f(Q\{\chi\}), \qquad Q\{\chi\} = \int d^3\vec{k}\,\vec{\chi}^{tr}(\vec{k})\cdot\vec{\chi}^{tr}(-\vec{k}), \qquad \text{(VIII.23)}$$

where f is any twice-differentiable function defined on $[0,\infty)$. This solution, however, does not have the singularities which should characterize turbulence.

(3) The third example is due to Donsker and Lions.[41] We follow these authors, and consider the solution $y\{q;t,v\}$ for $t>0$ to

$$\partial y/\partial t - \tfrac{1}{2}\partial^2 y/\partial v^2 = -\tfrac{1}{2}v^2 y + ivq(t)y, \qquad \text{(VIII.24a)}$$

$$y \to 0 \quad \text{as} \quad v \to \pm\infty, \qquad \text{(VIII.24b)}$$

$$y\{q;t=0,v\} = \delta(v). \qquad \text{(VIII.24c)}$$

We recognize (VIII.24a) as the diffusion equation with the harmonic oscillator potential, and with another term involving an arbitrary function q. We express y as in the Feynman-Kac formula (IV.28). The resulting Wiener integral over x has the form of a functional Fourier transform, in view of the factor $\exp(i\int d\tau xq)$.

Our plan now is to express $\delta y/\delta q$ in terms of the Wiener integral, and to transform this expression so as to obtain a functional differential equation for y. The procedure may be clarified by the following example. Consider

$$f(u) = \int_{-\infty}^{\infty} dp\, e^{-p^2} e^{ipu},$$

$$f'(u) = -\tfrac{1}{2} i \int d(e^{-p^2}) e^{ipu} = \tfrac{1}{2} i \int d(e^{ipu}) e^{-p^2} = -\tfrac{1}{2} uf(u).$$

More generally, if V(p) is a polynomial of an even degree with leading coefficient positive, and

$$g(u) = \int_{-\infty}^{\infty} dp\, e^{-V(p)} e^{ipu}, \qquad \text{(VIII.25a)}$$

then we obtain the differential equation

$$V'(i^{-1}d/du)\, g(u) = iug(u). \qquad \text{(VIII.25b)}$$

Let us return to the Wiener integral. We derive the differential equation in a heuristic fashion. We first convert the conditional integral into an unrestricted integral containing the factor

$$\delta[v-x(t)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} du\, e^{iu[v-x(t)]}, \qquad \text{(VIII.26)}$$

and obtain for the derivative $\delta y/\delta q$,

$$\delta y\{q;t,v\}/\delta q(\tau) = (2\pi)^{-1} \int_{-\infty}^{\infty} du\, e^{iuv} \int d\mu_W(x)\, ix(\tau) F\{x\},$$

$$F\{x\} = \exp\Big[-\tfrac{1}{2} \int_0^t d\rho\, x^2(\rho) + i \int_0^t d\rho\, x(\rho) q(\rho) - iux(t)\Big].$$

With the help of (VII.12), with $\sigma = 1$, $\delta y/\delta q$ can be transformed to

$$\frac{\delta y}{\delta q(\tau)} = \frac{i}{2\pi} \int_{-\infty}^{\infty} du\, e^{iuv} \int_0^t d\nu \, \min(\nu,\tau) \int d\mu_W(x) \frac{\delta F\{x\}}{\delta x(\nu)} .$$

The last derivative gives three terms, each of which may be expressed in terms of y or $\delta y/\delta q$. The following equation results:

$$\frac{\delta y\{q;t,v\}}{\delta q(\tau)} = -\Big[\int_0^t d\nu \, \min(\nu,\tau) q(\tau)\Big] y - \int_0^t d\nu \, \min(\nu,\tau) \frac{\delta y}{\delta q(\nu)}$$

$$- i\tau \, \partial y/\partial v. \qquad \text{(VIII.27)}$$

A question of interest now is the following. What further conditions must be imposed on y in order that the solution be unique, and that it be characterized by (VIII.24a)-(VIII.24c)? The cited authors give the following subsidiary conditions as a possibility:

$$\lim_{\tau \to t-} \delta y\{q;t,v\}/\delta q(\tau) = iv y, \qquad \text{(VIII.28a)}$$

and for $q = 0$,

$$\partial y/\partial t - \tfrac{1}{2} \partial^2 y/\partial v^2 = -\tfrac{1}{2} v^2 y, \qquad \text{(VIII.28b)}$$

$$y\{0;t=0,v\} = \delta(v). \qquad \text{(VIII.28c)}$$

The uniqueness can be shown in two ways: first, by employing the functional Fourier transform and, second, by utilizing a Volterra expansion, as in example (1) of Section VIII.A. The second method is in a sense the more natural one, since the problem as expressed by Eqs. (VIII.27) and (VIII.28) makes no reference at all to the Wiener integral. But the first seems easier to adapt to other situations.

For future reference we also make the following remark. Suppose that we had in (VIII.24a) in place of $-\tfrac{1}{2} v^2$ a polynomial $-V(v)$, of even degree, and V large for large $|v|$. Then Eq. (VIII.27) should be modified by replacing

$$-\delta y/\delta q(\nu) \quad \text{by} \quad -iV'\Big[i^{-1} \delta/\delta q(\nu)\Big] y. \qquad \text{(VIII.29)}$$

C. Static Model of Quantum Field Theory.

The preceding example may be summarized as follows. We started with a given differential equation; we modified it by introducing a term with an auxiliary function, and then we investigated the functional derivatives with respect to this function. This kind of

procedure was used extensively by Schwinger in quantum field theory. The auxiliary functions can then be thought of as external sources or fields.

For instance, the nucleon Green's function S, in case of interaction with pseudoscalar neutral mesons, satisfies the equation

$$\left[i\gamma^{\mu}\partial_{\mu} + m - g\gamma_5\eta(u) + ig\gamma_5\int d^4u'\,\hat{\Delta}(\eta;u,u')\,\delta/\delta\eta(u')\right] S(\eta;u,v) = \delta(u-v). \qquad \text{(VIII.30)}$$

Here η is a given external field, and $\hat{\Delta}$ is the meson Green's function.

There is also a second equation, which is derived from the equation for the meson field, and which likewise contains $\hat{\Delta}$ and S. Thus together we have a coupled system. Since S and $\hat{\Delta}$ occur multiplied together (in each equation), the system is nonlinear. The field η has no known physical interpretation, and should be set equal to zero at the end of the calculations.

We will now solve a simplified problem, which is adapted from the work of Edwards and Peierls.[42] We assume, first, that the mesons are scalar, and we neglect the nucleon recoil. Then only the time dependence is relevant for interaction, and we set $\eta(u)=\eta(u^o)$. We restrict η to be real and sufficiently smooth, e.g., continuous. Our next assumption is that

$$\hat{\Delta}(\eta;u,u') = i\delta(\vec{u}-\vec{u}')\Delta(u^o-u'^o), \qquad \text{(VIII.31)}$$

where Δ is a given function, which is sufficiently smooth.

We present two methods of solution. For the second, which has heuristic elements, we will also assume that $\Delta \in L_1$ and is real, even, and of positive type, i.e., that the Fourier transform $\tilde{\Delta}$ is nonnegative. We introduce the real Hilbert space $\mathcal{H}_{\Delta}$ with the norm

$$\langle \zeta,\zeta\rangle_{\Delta} = \int_{-\infty}^{\infty} d\omega\, \tilde{\zeta}(\omega)\tilde{\zeta}(-\omega)/\tilde{\Delta}(\omega). \qquad \text{(VIII.32)}$$

If $\tilde{\Delta}$ vanishes over an interval, then $\tilde{\zeta}(\omega)$ and $\tilde{\zeta}(-\omega)=\tilde{\zeta}^*(\omega)$ must vanish there likewise, and we set the quotient equal to zero. We will also assume, for the second method, that $\langle\eta,\eta\rangle_{\Delta}<\infty$.

The factor i in (VIII.31) eliminates oscillatory behavior in the interaction, and will enable us to use the Hilbert space integral. The net effect is analogous to what would be obtained by continuing analytically, $t\to it$, and then smearing the interaction to remove divergences.

Schwinger's equation now becomes

$$\left[i\partial_t + m - g\eta(t) - g\int_{-\infty}^{\infty} d\sigma\, \Delta(t-\sigma)\delta/\delta\eta(\sigma)\right] S(\eta;t,\tau) = \delta(t-\tau). \quad \text{(VIII.33)}$$

Ordinary and partial differential equations of this type can be simplified by familiar transformations, and this also applies to the equation at hand.

The terms mS and δ may be eliminated by introducing as a factor the free-field function

$$G^{(o)}(t-\tau) = i^{-1}\theta(t-\tau)e^{im(t-\tau)}, \quad \text{(VIII.34a)}$$

which satisfies

$$(i\partial_t + m)G^{(o)}(t-\tau) = \delta(t-\tau). \quad \text{(VIII.34b)}$$

Thus, let

$$S(\eta;t,\tau) = G^{(o)}(t-\tau)S_1(\eta;t,\tau)$$

where $S_1(\eta;\tau,\tau) = 1$ and

$$\left[i\partial_t - g\eta(t) - g\int_{-\infty}^{\infty} d\sigma\, \Delta(t-\sigma)\delta/\delta\eta(\sigma)\right] S_1 = 0.$$

Next we eliminate the term $g\eta S_1$ by setting

$$S_1 = S_2 \exp\left[-ig\int_{\tau}^{t} d\rho\, \eta(\rho)\right].$$

Then again $S_2(\eta;\tau,\tau) = 1$, and

$$\left[i\partial_t + ig^2\int_{\tau}^{t} d\sigma\, \Delta(t-\sigma) - g\int_{-\infty}^{\infty} d\sigma\, \Delta(t-\sigma)\delta/\delta\eta(\sigma)\right] S_2 = 0. \quad \text{(VIII.35a)}$$

This equation has a solution independent of η,

$$S_2 = \exp\left[-\tfrac{1}{2}g^2\int_{\tau}^{t} d\rho \int_{\tau}^{t} d\sigma\, \Delta(\rho-\sigma)\right] = S_2(t-\tau). \quad \text{(VIII.35b)}$$

We assumed here that Δ is even. Otherwise (VIII.35b) has to be trivially modified. We have nothing to say about the uniqueness of the solutions of (VIII.35a).

If we now put together the various factors, we obtain

$$S(\eta;t,\tau) = \exp\Big[-ig\int_\tau^t d\rho\,\eta(\rho)\Big]S_2(t-\tau)G^{(o)}(t-\tau). \qquad \text{(VIII.36)}$$

The previous solution[42] to this problem made extensive use of heuristic Fourier transforms, and is also of interest. There one first removes the term $g\eta S$ by introducing G_1,

$$S = G_1\exp\Big(-\tfrac{1}{2}\langle\eta,\eta\rangle_\Delta\Big). \qquad \text{(VIII.37)}$$

Next one sets

$$G_2(\zeta;t,\tau) \sim \int \mathfrak{D}(\eta)\exp\Big[i\int_{-\infty}^{\infty} d\rho\,\eta(\rho)\zeta(\rho)\Big]G_1(\eta;t,\tau).$$

Then in place of the term involving $\delta/\delta\eta$ we get another, proportional to

$$\int d\sigma\ \Delta(t-\sigma)\zeta(\sigma) \equiv \Omega(t).$$

One is led in this way to the equation

$$\Big[i\partial_t + m - g\Omega(t)\Big]G(\Omega;t,\tau) = \delta(t-\tau), \qquad \text{(VIII.38a)}$$

and G may be interpreted as the one-particle Green's function in the external field Ω. For our problem, we easily find:

$$G(\Omega;t,\tau) = \exp\Big[-ig\int_\tau^t d\sigma\,\Omega(\sigma)\Big]G^{(o)}(t-\tau). \qquad \text{(VIII.38b)}$$

One now goes back to the original function S by integrating G functionally. The result can be written as follows,

$$S(\eta;t,\tau) = \int_{\mathcal{H}_\Delta} \mathfrak{D}(\Omega)e^{-\frac{1}{2}\langle\Omega,\Omega\rangle_\Delta}G(\Omega+\eta;t,\tau). \qquad \text{(VIII.39)}$$

The functional integral can be easily evaluated, and shown to agree with (VIII.36).

IX. Measures Associated with the Integrals

We introduced both the Wiener integral and the Hilbert space integral as limits of sequences of finite-dimensional integrals. Now we are about to discuss measure-theoretic bases for these integrals.

These bases clarify some paradoxes, and are otherwise useful as well. We have already referred, for example, to the measure for the Wiener integral.

The approaches which we introduce have the virtue of simplicity, and of not requiring an extensive background. We should say that other approaches are usually adopted in the current literature.

A. The Hilbert Space Integral and Product Measures.

Let us first point out that it is impossible to construct a finite, orthogonal (or unitary) invariant measure in a Hilbert space. Consider a ball of radius $\frac{1}{2}$ around each element of an orthonormal basis. All balls should have the same measure by orthogonal invariance, and are disjoint, so the total measure is $\Sigma_1^\infty \epsilon = \infty$ if $\epsilon > 0$. (Recall that by definition a measure must be countably additive.) But $\Sigma \epsilon = \infty$ contradicts finiteness and, in particular, we cannot have $I_{\mathfrak{H}}(1) = 1$.

This non-existence of a finite and invariant measure leads to paradoxes like the following. Consider the functional

$$F(x) = e^{-\langle x,x\rangle/\rho} = \prod \exp\left[-(x^j)^2/\rho\right]. \tag{IX.1}$$

One sees that

$$I_{\mathfrak{H}}(F) = \prod_j c$$

where $c < 1$, hence this product diverges to zero. But F is smooth and positive everywhere on $\mathfrak{H}$.

One way of constructing a measure for the Hilbert space integral is the following.[26)-28)] (Of course, this measure will not be orthogonal-invariant.) We fix an orthonormal basis, and consider the space of all sequences $(x^1, x^2, \ldots)$ without the restriction $\Sigma (x^j)^2 < \infty$. This space is called a corona of $\mathfrak{H}$, and is the infinite Cartesian product of real lines,

$$\mathfrak{R} \equiv R^\infty = \mathbf{X}_{j=1}^\infty (R^1)_j. \tag{IX.2}$$

On each line $R^1 j$ we have the normalized Gaussian measure, $\mu_j(R^1 j) = 1$. Fubini's theorem gives a measure on the Cartesian product of a finite number of factors. It is significant for us that this theorem extends to an infinite number of factors, if the measure of each is unity.

We recall a few definitions. A measure space is a triple (Y, S, μ) where S is the family of measurable subsets of Y, i.e.,

$\mu(B)$ is defined for $B \in S$. The family S is a σ-ring. For us suffices the case where S is a σ-algebra, i.e., closed under complements and countable unions. We now have the following result.[43]

Theorem. If $\{(Y_j, S_j, \mu_j)\}$ is a sequence of measure spaces with $\mu_j(Y_j) = 1$, then there exists a unique measure μ on the σ-algebra $S = \mathbf{X} S_j$, having the following property: For every set of the form $A \times \mathbf{X}_{i=n+1}^{\infty} Y_i$, where A is a measurable subset of $Y_1 \times \ldots \times Y_n$,

$$\mu\left(A \times \mathbf{X}_{i=n+1}^{\infty} Y_i\right) = \left(\mu_1 \times \ldots \times \mu_n\right)(A). \qquad \text{(IX.3)}$$

Sets of this type are called cylinders, and we denote the class of all cylinders by C. Clearly, μ as given by (IX.3) is a finite, non-negative, and finitely additive function on C. For μ to be a measure, it suffices to show that μ is continuous from above at the null set ϕ. This means that if we have an arbitrary decreasing sequence $\{E^{(j)}\}$ of cylinders such that $\lim E^{(j)} = \phi$, then $\lim \mu(E^{(j)}) = 0$. An equivalent statement is, if $0 < \varepsilon \leq \mu(E^{(j)})$ for all j, then $\cap E^{(j)} \neq \phi$.

The theorem can be proved fairly easily by actually exhibiting a point $(y^1, y^2, \ldots) \in \cap E^{(j)}$, if $0 < \varepsilon < \mu(E^{(j)})$. We refer to the cited book of Halmos[43] for details.

We recall that a set A has measure zero if for every $\varepsilon > 0$, one can cover A with a denumerable family of sets whose total measure is $\leq \varepsilon$. Let us show that for the infinite product measure μ on $\mathcal{R}$, we have $\mu(\mathcal{K}) = 0$.

Given $\varepsilon > 0$, let δ be such that

$$(2\pi\sigma)^{-\frac{1}{2}} \int_{-\delta}^{\delta} du\, e^{-u^2/2\sigma} \equiv \varepsilon_0 < \min(\tfrac{1}{2}\varepsilon, \tfrac{1}{2}).$$

We construct the following cylinders,

$$C_1 = \left\{x \in \mathcal{R} : |x^1| < \delta\right\},$$

$$C_2 = \left\{x \in \mathcal{R} : |x^2|, |x^3| < \delta\right\},$$

and, in general,

$$C_n = \left\{x \in \mathcal{R} : |x^{n_1}|, |x^{n_1+1}|, \ldots, |x^{n_2}| < \delta\right\}, \qquad \text{(IX.4a)}$$

where $n_1 = \frac{1}{2}n(n-1)+1$ and $n_2 = \frac{1}{2}n(n+1)$. Since each $x \in \mathcal{K}$ has $|x^j| < \delta$ except for at most a finite number of components, $\mathcal{K} \subset \cup C_j$. But

$$\mu(C_1) = \epsilon_0, \ldots, \mu(C_n) = \epsilon_0^n, \ldots,$$

and

$$\sum \mu(C_n) = \epsilon_0 + \cdots + \epsilon_0^n + \cdots = \epsilon_0/(1 - \epsilon_0) < \epsilon. \qquad \text{(IX.4b)}$$

It follows that $\mu(\mathfrak{K}) = 0$. Note that this resolves the dilemma of the integral of $F(x) = e^{-\langle x,x\rangle/\rho}$. This functional is nonzero only if $\langle x,x\rangle < \infty$, i.e., only on a set of measure zero.

B. Measure for the Wiener Integral.

In the discussion of Section IX.A, we circumvented the possible realization of the underlying Hilbert space, as a space of functions. Moreover, we do not know if elements of the corona $\mathfrak{R}$ can be identified as functions on R^n. It is therefore instructive to see another construction of a measure, where the function-theoretic aspects of the underlying space are very much in the foreground.

The construction which we outline depends on establishing a correspondence between "quasi-intervals" of a function space and subintervals of the unit interval. Then the Lebesque measure du implies a measure for the function space. We should point out that this approach, while simple in principle, does not easily adapt to other related problems.

This construction was used by Wiener in his original work on the integral, and a detailed treatment appeared recently.[44] Here we confine ourselves to indicating the main points.

For each positive integer n, we construct the following. We first subdivide the real axis into two semi-infinite intervals and $(2n)2^n$ intervals of length 2^{-n}:

$$(-\infty, -n], (-n, -n + 2^{-n}], \ldots, (n - 2^{-n}, n], (n, \infty). \qquad \text{(IX.5)}$$

We call these intervals $A_k^{(n)}$, $1 \le k \le (2n)2^n + 2$. For the given n, we also subdivide the interval $[0,t]$ by the points $t_j^{(n)} = jt/2^n$. We now define the quasi-intervals in the space $\mathfrak{F}$ of all real-valued functions on $[0,t]$ by

$$\mathcal{J}_{jk}^{(n)} = \left\{x \in \mathfrak{F} : x(t_j^{(n)}) \in A_k^{(n)}\right\}. \qquad \text{(IX.6)}$$

We replace j and k by the single index ℓ, which ranges from one to $[(2n)2^n + 2]2^n$. For the $\mathcal{J}_\ell^{(n)}$, we have the probability function P given by integrals like (IV.6). Obviously,

$$\cup_\ell \vartheta_\ell^{(n)} = \mathfrak{F}, \quad \sum_\ell P\left(\vartheta_\ell^{(n)}\right) = 1. \tag{IX.7}$$

For the given n, we next divide the unit interval into subintervals, and establish a correspondence with the $\vartheta_\ell^{(n)}$. To $\vartheta_\ell^{(n)}$ we must associate an interval of length $P(\vartheta_\ell^{(n)})$:

$$\vartheta_\ell^{(n)} \longleftrightarrow J_\ell^{(n)} \equiv \left[u_{\ell-1}^{(n)}, u_\ell^{(n)}\right] \Rightarrow P\left(\vartheta_\ell^{(n)}\right) = u_\ell^{(n)} - u_{\ell-1}^{(n)}. \tag{IX.8}$$

The details of this correspondence are not significant, except that we must have

$$\vartheta_\ell^{(n+1)} = \cup_{\ell' \in \lambda} \vartheta_{\ell'}^{(n)} \Rightarrow J_\ell^{(n+1)} = \cup_{\ell' \in \lambda} J_{\ell'}^{(n)},$$

so that to a shrinking sequence of quasi-intervals there corresponds a shrinking sequence of intervals.

Consider now a continuous function on $[0,t]$. It defines a shrinking sequence of quasi-intervals as $n \to \infty$, hence a shrinking sequence of intervals, and hence a real number. Conversely, to every real number except for a set of measure zero, there corresponds a unique continuous function. The points to be excluded are, first, the endpoints of intervals, to which there may correspond two functions (defined by shrinking sequences of intervals from left and from right). Second, one has to exclude points for which the sequence of quasi-intervals does not define a continuous function. In fact, one can have two (completely distinct) sequences of binary rationals such that $\lim t_j = \lim \tau_k$, and a sequence of quasi-intervals such that, e.g.,

$$x(t_j) > 1, \quad x(\tau_k) < 0 \quad \text{for all } j, k.$$

The crux of the construction is to show that such points (in $[0,1]$) indeed form a set of measure zero. The precise statement that one has is the following.

Theorem. Given $\epsilon > 0$, $\exists h$ such that all functions on $[0,t]$ which satisfy

$$|x(t_2) - x(t_1)| > h|t_2 - t_1|^\mu, \tag{IX.9}$$

where the t_j are binary fractions times t, and where μ is fixed, $0 \le \mu < \frac{1}{2}$, belong to quasi-intervals whose total probability is $< \epsilon$.

Thus, the measure for the Wiener integral is concentrated on the space $\mathcal{H}^\mu$ of Hölder continuous functions. For simplicity, one often takes the space $C_{[0,t]}$ of all continuous functions on $[0,t]$

(corresponding to $\mu = 0$), and which satisfy $x(0) = 0$, as the space of integration.

We now present a sample calculation, to give an idea how the foregoing theorem can be proved. Let

$$t_1 = jt/2^n, \quad t_2 = (j+1)t/2^n, \quad \text{so} \quad t_2 - t_1 = t/2^n,$$

and consider the functions $x(\tau)$ for which

$$|x(t_2) - x(t_1)| > h(t/2^n)^{\mu}, \tag{IX.10}$$

where h also satisfies $ht^{-(\frac{1}{2}-\mu)}2^{-\frac{1}{2}} \geq 1$ (see Eq. (IX.12c) below).

Functions satisfying (IX.10) can be classified accordingly as $x(t_1)$ or $x(t_2)$ lies in a finite or in an infinite interval $A_m^{(n)}$, for the given n (see (IX.5)). Suppose that $x(t_1)$ lies in the first finite interval from the left, and $x(t_2)$ in any finite interval. Then the probability of the quasi-intervals which include all the functions x in question, is

$$P = \left(4\pi^2 jt^2/2^{2n}\right)^{-\frac{1}{2}} \int_{-n}^{-n+2^{-n}} du_1 e^{-u_1^2 2^{n-1}/t} \int_{-n+2^{-n}k}^{n} du_2 e^{-(u_2-u_1)^2 2^{n-1}/t}, \tag{IX.11}$$

where k is the least integer such that $h(t/2^n)^{\mu} > 2^{-n}k$.

Let us obtain an estimate for the integral over u_2, which we call P_o. We set $u_2 - u_1 = u$, and observe that

$$u \geq h(t/2^n)^{\mu} - 2^{-n} \equiv u_o. \tag{IX.12a}$$

We treat the interval $[u_o, u_o + 2^{-n}]$ separately. Thus

$$P_o < \int_{u_o}^{\infty} du\, e^{-u^2 2^{n-1}/t} < f(h,n) + c_o \int_{v_o}^{\infty} dv\, e^{-v^2}, \tag{IX.12b}$$

$$f(h,n) = 2^{-n} e^{-u_o^2 2^{n-1}/t}, \quad c_o = 2^{-\frac{1}{2}(n-1)} t^{\frac{1}{2}}, \quad v_o = ht^{-(\frac{1}{2}-\mu)} 2^{-\frac{1}{2}n(\frac{1}{2}-\mu)}. \tag{IX.12c}$$

The need for $\mu < \frac{1}{2}$ is now apparent. This is needed in order to have $v_o \to \infty$ with n.

Let us find more explicitly the dependence of the last integral on h and n. In view of the condition on h made previously, $v > 1$ in the range of integration, so

$$P_o < f(h,n) + \tfrac{1}{2}\, c_o \int_{v_o^2}^{\infty} d(v^2) e^{-v^2} = f(h,n) + \tfrac{1}{2}\, c_o e^{-v_o^2}. \qquad \text{(IX.12d)}$$

We see that the second term, as well as the first, decrease very rapidly with increasing h and n. The proof of the theorem is now a matter of combinatorics and of additional calculations of this sort.

C. Consequences of the Wiener Measure.

The existence of the measure for the Wiener integral allows us to use the full apparatus of measure theory, e.g., Fubini's theorem, et cetera. We should also like to point out some more specific consequences of this measure. Various of these consequences were presupposed in the previous sections.

(1) One can define a measure for the conditional Wiener integral by a similar construction, or in terms of the Radon-Nikodým derivative $d\mu_W / d[x(t)]$. Then the Wiener measure is the product measure, cf. Eq. (IV.24).

(2) It follows directly from the foregoing construction that a Wiener cylinder functional is Wiener measurable. [We refer here to functionals $f\{x(t_1), \ldots, x(t_m)\}$ such that f is measurable in R^m.] Moreover, one can show easily that any continuous functional $g\{x\}$ on $C_{[0,t]}$, i.e., depending continuously on continuous functions $x(\tau)$, is Wiener measurable.[45] The second assertion, or both, applies in particular to the following functionals which we encountered previously:

$$x(t_1) \ldots x(t_k), \qquad e^{-\int d\tau V\{x\}}, \qquad e^{i\int d\tau xy}. \qquad \text{(IX.13)}$$

(3) By way of recapitulation, we observe that we have available three ways for defining the Wiener integral: (i) directly as a limit, as in Section IV.A; (ii) as a Hilbert space integral over the space $\mathfrak{K}_W$, and (iii) as a measure-theoretic integral. Still another way, (iv) in terms of semigroups, will be indicated in Section XI.B.

For some functionals, these definitions are not equivalent. However, in various important cases, at least the first three are. Consider, e.g., methods (i) and (ii). For a given functional F, they define approximating functionals F^s. Suppose that all such F^s are bounded by a functional which is integrable in the measure-theoretic sense. Then the dominated convergence theorem implies that

$$\int d\mu_W F^{sm} \longrightarrow \int d\mu_W F$$

if $F^{sm}\{x\} \to F\{x\}$ for all $x \in C_{[0,t]}$, and the definition (iii) agrees with (i) and (ii). This applies obviously to the last two functionals in (IX.13), for which $|F|, |F^s| \leq 1$ (if $V \geq 0$), and for which a variety of other approximating sequences could also be employed. See, e.g., Section IV.B. For the functional $x(t_1)\ldots x(t_k)$, we may argue similarly.

However, one can give an example[45] of a functional for which (i) and (iii) give unambiguous but different results. Therefore, the term Wiener integral is reserved by convention for the measure-theoretic integral.

(4) A general theorem[46] implies that the characteristic functional $\int d\mu_W(x) e^{i\langle x,y\rangle}$ depends continuously on y, where y may range over the dual to $C_{[0,t]}$. (This dual includes all measures over $[0,t]$.) This fact is confirmed by the evaluation (VII.21). Another result is the following.

If F is bounded and continuous on $C_{[0,t]}$ and

$$\int d\mu_W(x) e^{i\langle x,y\rangle} F\{x\} = 0 \qquad \text{(IX.14)}$$

for every $y \in C_{[0,t]}$, then[41] $F = 0$ (on $C_{[0,t]}$). We note that if we were to replace in (IX.14) the Wiener integral by the Hilbert space integral, then the functional $e^{-\langle x,x\rangle/\rho}$ would give a counterexample.

X. Some Related Measures

A. Examples Based on the Smoluchowski Equation.

In the remainder of these notes we shall modify the two functional integrals in various ways. The first possibility that we discuss is to generalize the construction of the Wiener integral. In particular, we proceed as in Section IV.A, but do not insist on the connection with the diffusion equation.

Thus, for the probability that $a_i \leq x(\tau_i) \leq b_i$, we generalize (IV.6) to

$$\int_{a_1}^{b_1} du_1 \cdots \int_{a_m}^{b_m} du_m P\big(0|u_1;\tau_1\big) P\big(u_1|u_2;\tau_2-\tau_1\big) \cdots$$

$$P\big(u_{m-1}|u_m;\tau_m-\tau_{m-1}\big). \qquad \text{(X.1)}$$

The function P defines the conditional probabilities, and it will be assumed positive and continuous for $\tau_{j+1} - \tau_j > 0$. It must, moreover, satisfy the compatibility condition for all τ, $0 < \tau < t$,

$$P(v|u;t) = \int_{-\infty}^{\infty} d\xi\, P(v|\xi;\tau)\, P(\xi|u;t-\tau). \tag{X.2}$$

This equation is known as the Smoluchowski equation.[22]

If we have such a function P, we can compute expectations of integrable cylinder functionals, and can also try to define a functional integral. See the remarks at the end of this section (X.A), and also Section X.B.

We now give two examples of functions P. First, let

$$P_\alpha(v|u;t) = (2\pi)^{-1} \int_{-\infty}^{\infty} d\xi\, \exp\left[i\xi(u-v) - t|\xi|^\alpha \right]. \tag{X.3}$$

One can show[47] that P_α is positive for $0 < \alpha \leq 2$, and (X.2) follows if we recognize the Fourier integral for the δ-function. For $\alpha = 2$, we get the solution to the diffusion equation and also the Wiener integral.

For $\alpha = 1$, the integral (X.3) is elementary,

$$P_1(v|u;t) = t/\pi\left[t^2 + (u-v)^2 \right]. \tag{X.4}$$

The cylinder functional $x(t_1)x(t_2)$ is not integrable now. For

$$f\{x\} = \left| e^{i\beta_1 x(t_1)} - e^{-i\beta_2 x(t_2)} \right|^2, \tag{X.5a}$$

we easily find the expectation E. If $t_1 < t_2$,

$$E(f) = 2 - 2\exp\left[-|\beta_2|(t_2 - t_1) - |\beta_2 - \beta_1| t_1 \right]. \tag{X.5b}$$

The second example that we mention corresponds to the Brownian motion of the one-dimensional harmonic oscillator.[48] In the presence of an external force, the diffusion process is described by the Fokker-Planck equation, here

$$\partial\psi/\partial t = D\partial^2\psi/\partial u^2 - f^{-1}\partial(F\psi)/\partial u, \tag{X.6}$$

where F is the external force and $-fV$ is the force of friction, V being the velocity. For $F = -\omega^2 mu$, we may write

$$\partial\psi/\partial t = D\partial^2\psi/\partial u^2 + k\partial(u\psi)/\partial u, \qquad (X.7)$$

where $k=\omega^2 m/f$, $k>0$. For the fundamental solution P^k satisfying

$$P^k(v|u;t=0) = \delta(u-v),$$

we find

$$P^k(v|u;t) = k^{\frac{1}{2}}\left[2\pi D(1-e^{-2kt})\right]^{-\frac{1}{2}}$$

$$\times \exp\left[-(k/2D)(u-ve^{-kt})^2(1-e^{-2kt})^{-1}\right]. \qquad (X.8)$$

We see a Gaussian distribution. The mean is also the expectation value of $x(t_1)$,

$$f\{x\} = x(t_1) \Rightarrow E(f) = ve^{-kt}, \qquad (X.9)$$

provided we redefine the probabilities [(X.1), et cetera], so that the paths satisfy $x(0) = v$.

In order to verify (X.2), we first set $\rho = \xi e^{-k\tau}$. Then we observe that the integrand can be put into the same form as for the diffusion equation, with the time t replaced by e^{-2kt}.

Now let us consider the functional

$$f\{x\} = \left[x(t_1) - ve^{-kt_1}\right]\left[x(t_2) - ve^{-kt_2}\right], \qquad (X.10a)$$

i.e., we subtracted the mean values. To find the expectation of f we make some obvious changes of variables, and obtain

$$E(f) = (D/k)\left(e^{-k|t_2-t_1|} - e^{-k(t_2+t_1)}\right). \qquad (X.10b)$$

Note that as $k\to 0$ this reduces to $2D\min(t_1,t_2)$, in agreement with (IV.14b).

This example relates to the solution of a model for a one-dimensional many-body system. There the basic potential is $-e^{-\gamma|u_1-u_2|}$ but modified by a hard core.[49]

One would still like to know if the P_α and the P^k in fact define measure-theoretic integrals. The answer is yes. For P_α, $\alpha<2$, one integrates over the space of functions with jump discontinuities.[22] Note, however, that the expectation (X.5b) does not reflect

any discontinuity. On the other hand, for the P^k one can easily relate the integral to the Wiener integral, and it suffices to take the Wiener space[50] $C_{[0,t]}$.

B. Product Measures. A Theorem of Kolmogorov.

The theorem of Section IX.A on infinite product measures allows us to construct, quite generally,

$$d\mu(x) = \prod dx^j \rho_j(x^j), \tag{X.11}$$

where $\rho \geq 0$ and $\int dx^j \rho_j = 1$. We give two examples.

First, let

$$\rho_j^\sigma(x^j) = (2\pi\sigma)^{-\frac{1}{2}} \exp\left[-(x^j - a^j)^2/2\sigma\right]. \tag{X.12}$$

If each $a^j = 0$, then we obtain the measure of Section IX.A. This example relates to the question of unitary equivalence of the two sets of canonical systems of operators:[28),33)]

$$p^1, q^1; p^2, q^2; \ldots \quad \text{and} \quad p^1, q'^1; \ldots \quad \text{where} \quad q'^j = q^j + a^j \tag{X.13a}$$

and where

$$\left[p^j, q^k\right] = \left[p^j, q'^k\right] = i^{-1}\delta^{jk}. \tag{X.13b}$$

For the second example, we take

$$\rho_j(x^j) = \delta(x^j - a^j), \tag{X,14a}$$

which in fact corresponds to the limit $\sigma \to 0$ of ρ_j^σ. Then

$$\int d\mu(x) F(x) = F(a), \tag{X.14b}$$

where $a = (a^1, a^2, \ldots)$, and a is not restricted $\mathcal{K}$.

The following theorem of Kolmogorov can be considered a generalization of the theorem on product measures, and can be proved in a similar way:[51)]

Theorem. Suppose that we have an index set S (which need not be denumerable), and for any finite set of indices $i_1, \ldots, i_m \in S$ there exists a real-valued function $F_{i_1 \ldots i_m}(u^1, \ldots, u^m)$ on R^m. These functions must be non-negative, non-decreasing, continuous from the right, and must satisfy:

(i) $F_{i_k}(u) \to 1$ as $u \to \infty$ and $F_{i_1 \dots i_{k+1}}\left(u^1, \dots, u^k, u^{k+1}\right) \to 0$

as $u^{k+1} \to -\infty$,

(ii) $F_{i_1 \dots i_{k+1}}\left(u^1, \dots, u^k, u^{k+1}\right) \to F_{i_1 \dots i_k}\left(u^1, \dots, u^k\right)$

as $u^{k+1} \to \infty$,

(iii) $F_{i_1 \dots i_m}\left(u^1, \dots, u^m\right) = F_{i_{\sigma(1)} \dots i_{\sigma(m)}}\left(u^{\sigma(1)}, \dots, u^{\sigma(m)}\right)$,

where σ is any permutation on $\{1, \dots, m\}$. Then there exists a measure $d\mu(x)$ on the (perhaps non-denumerably infinite) product space

$$\mathfrak{R} = \mathbf{X}_{i \in S} (R^1)_i \tag{X.15}$$

such that

$$F_{i_1 \dots i_m}\left(u^1, \dots, u^m\right) = \int_{\mathfrak{P}} d\mu(y), \tag{X.16a}$$

where

$$\mathfrak{P} = \left\{ y \in \mathfrak{R} : y^{i_j} \le u^j \quad \text{for} \quad 1 \le j \le m \right\}. \tag{X.16b}$$

The product measure corresponds to the factorizations

$$F_{i_1 \dots i_m}\left(u^1, \dots, u^m\right) = F_{i_1}\left(u^1\right) \dots F_{i_m}\left(u^m\right). \tag{X.17}$$

This theorem also applies to the integrals of $P(v|u;\tau)$ as in (X.1),

$$F_{\tau_1 \dots \tau_m}\left(b_1, \dots, b_m\right) = \int_{-\infty}^{b_{i_1}} dv_1 \dots \int_{-\infty}^{b_{i_m}} dv_m \, P\left(0|v_1;\tau_{i_1}\right) \dots$$

$$P\left(v_{m-1}|v_m;\tau_{i_m} - \tau_{i_{m-1}}\right), \tag{X.18}$$

where the $\tau_{i_1}, \dots, \tau_{i_m}$ are now in chronological order, $\tau_{i_j} \le \tau_{i_{j+1}}$. In this case, (ii) is equivalent to the Smoluchowski equation. However, the measure implied in the present theorem is then over the space of all functions, without regard to continuity [but satisfying $x(0) = 0$]. Often, and for us, this is not the measure of interest.

We make a final observation. Let us take the integrals of $P(v|u;\tau)$ for the case of the Wiener integral, and restrict τ to binary rationals, multiplied by t. Kolmogorov's theorem then allows us to bypass the correspondence with subintervals of the unit interval, in the construction of the Wiener measure.

XI. Representation of Operators of Evolution

The interrelation between the diffusion equation and the Wiener integral has been prominent in the previous sections. We also saw in Section X.A that a functional integral may be associated with the Fokker-Planck equation for the case of a spring force.

In this section we will discuss the possibility of relating certain other partial differential equations to functional integrals. Our special interest is in the Schrödinger equation.

A. Parabolic Equations of Higher Order.[52]

The complications which occur with the Schrödinger equation can be seen, perhaps more easily, if one examines

$$\partial\psi/\partial t = (-1)^{r+1}\partial^{2r}\psi/\partial u^{2r}, \tag{XI.1}$$

where $r \geq 2$. The fundamental solution p, defined by the initial condition

$$p(t=0, u) = \delta(u),$$

can be expressed as follows,

$$p(t,u) = (2\pi)^{-1}\int_{-\infty}^{\infty} d\xi \exp\left(i\xi u - t\xi^{2r}\right), \tag{XI.2}$$

and satisfies, in view of the Fourier integral for the δ-function,

$$\int_{-\infty}^{\infty} du\, p(t,u) = 1, \quad \text{all } t \geq 0. \tag{XI.3}$$

We could try to use p to define measures of quasi-intervals, as in Section IX.B:

$$\mu(\mathcal{I}) = \int_{a_1}^{b_1} du_1 \cdots \int_{a_m}^{b_m} du_m \prod_{k=1}^{m} p_k, \tag{XI.4a}$$

$$p_k = p\left(t_k - t_{k-1},\, u_k - u_{k-1}\right), \tag{XI.4b}$$

and we set $t_0 = 0$, $u_0 = 0$. We could then hope to express solutions to

equations like

$$\partial\psi/\partial t = (-1)^{r+1}\partial^{2r}\psi/\partial u^{2r} + V\psi \tag{XI.5}$$

by means of functional integrals constructed in this way.

However, the difficulty is that p can take on negative values. In fact, the integral (XI.2), with ξ^{2r} replaced by $|\xi|^{\alpha}$, is positive for $0<\alpha\leq 2$, but not for $\alpha>2$ (cf. Section X.A). We do not know of any elementary proof of this fact, but one can easily see that for sufficiently large α or r, the integral will become negative. Indeed, if t is fixed and $r\to\infty$, then $\exp(-t\xi^{2r})$ becomes the step-function $\theta(1-\xi^2)$, whose Fourier transform is

$$\int_{-1}^{1} d\xi\, e^{i\xi u} = 2(\sin u)/u. \tag{XI.6}$$

This evaluation will be changed by arbitrarily little if r is sufficiently large.

One might try to remedy this situation by writing (see (XI.4a))

$$\mu(\mathcal{J}) = \mu_+(\mathcal{J}) - \mu_-(\mathcal{J}), \tag{XI.7}$$

where the two contributions correspond respectively to the two regions of R^m where Πp_k is positive and negative. Then, if μ_+ and μ_- should define independent functional integrals, the procedure could be essentially unchanged.

We could use such a decomposition if each of the two measures $\mu_\pm$ defined a finite measure on the space of all functions. (Then we would say that μ is of bounded variation, in analogy with the finite-dimensional theory.) But this is not the case, as the following argument shows.

Consider $\mu_+(\mathcal{J}) + \mu_-(\mathcal{J})$. This quantity results if we replace Πp_k by $\Pi|p_k|$ in (XI.4a). Furthermore, let $\mathcal{J}$ now be the space $\mathfrak{F}$ of all functions, i.e., let $(a_j, b_j)\to(-\infty,\infty)$ for all j. Then the integral over each u_j yields the factor

$$\int_{-\infty}^{\infty} du_j\,|p(t,u_j)| = (2\pi)^{-1}\int_{-\infty}^{\infty} dv\left|\int_{-\infty}^{\infty} d\eta\, e^{-\eta^2+i\eta v}\right| \equiv c>1. \tag{XI.8}$$

We changed variables, $v=ut^{-1/2r}$ and $\eta=\xi t^{1/2r}$ and have thus shown that c does not depend on t. Moreover, $c>1$, since $\int d\mu_j p = 1$, and p becomes negative. Thus

$$\mu_+(\mathfrak{F}) + \mu_-(\mathfrak{F}) = c^m, \tag{XI.9}$$

and $c^m\to\infty$ as $m\to\infty$. Consequently, our previous constructions of

functional integrals cannot be easily adapted to the present circumstance.

A similar argument applies to the modified Schrödinger equation,[53]

$$\partial\psi/\partial t = (i+\epsilon)\partial^2\psi/\partial u^2. \tag{XI.10}$$

We included the real and positive increment ϵ. At one time it was hoped that the presence of ϵ would enable us to use an integral like the Wiener integral, and we could let $\epsilon \to 0$ at the end of the manipulations.

Let us return to Eq. (XI.5) with the potential V. It is possible to construct the analogue to the Feynman-Kac formula (IV.28), and we turn to this problem presently. However, one can also study the "generalized measure" defined by (XI.1) and show, e.g., that it is concentrated on a space of Hölder continuous functions.

B. Semigroups of Operators.

First, we give an example from ordinary differential equations. Suppose that we have a system of equations

$$dy/dt = A(t)y, \tag{XI.11}$$

where A(t) is a matrix depending smoothly on time t. If

$$[A(\tau_1), A(\tau_2)] = 0 \quad \text{for} \quad 0 \le \tau_1, \tau_2 \le t, \tag{XI.12a}$$

then we have the solution

$$y(t) = \exp\Big[\int_0^t d\tau\, A(\tau)\Big]y(0). \tag{XI.12b}$$

In the general case, we can use a familiar procedure. We break the interval $[0,t]$ into m equal parts, set $\Delta t = t/m$, and the solution is given by

$$y(t) = \Big\{\lim_{m\to\infty}\Big[1 + \Delta t\, A(t)\Big]\Big[1 + \Delta t\, A((m-1)\Delta t)\Big]\cdots\Big[1 + \Delta t A(\Delta t)\Big]\Big\}\, y(0)$$

$$= \Big(\exp\int_0^t d\tau\, A(\tau)\Big)_+ y(0), \tag{XI.13}$$

(cf. (III.16) and (III.17)). This time-ordered exponential is also called a multiplicative integral.[54] It has various useful properties analogous to those of, say, the Riemann integral.

The remainder of Section XI.B is based largely on an article by Nelson.[55]

In the case of the Schrödinger equation,

$$\partial_t \psi = i\left[(2\mu)^{-1}\Delta_u - V\right]\psi, \quad \psi(t=0, \vec{u}) = f(\vec{u}), \tag{XI.14}$$

the situation is somewhat different from (XI.11). For the two simpler equations,

$$\partial_t x = i(2\mu)^{-1}\Delta_u x, \quad \partial_t y = -iV(\vec{u})y, \tag{XI.15a,b}$$

we can write the solution immediately,

$$\tilde{x}(t,\vec{k}) = e^{-it\vec{k}^2/2\mu}\tilde{x}(0,\vec{k}), \quad y(t,\vec{u}) = e^{-itV}y(0,\vec{u}), \tag{XI.15c,d}$$

where $\tilde{x}$ is the Fourier transform of x.

Since Δ and V do not commute, we again have to use a limiting relation. We have the following. If

$$P^t = \exp(it\Delta/2\mu), \quad Q^t = \exp(-itV), \tag{XI.16a,b}$$

then, under suitable restrictions on V,

$$\psi_t = \lim_{m\to\infty}\left(P^{t/m}Q^{t/m}\right)^m \psi_{t=0}. \tag{XI.16c}$$

We have here an example of Trotter's formula. Its applicability to the Schrödinger equation will be discussed in Section XI.C.

Let us now give a more precise and also more general formulation of such relations. A contraction semigroup on a Banach space $\mathfrak{B}$ is a semigroup $\{K^t\}$, $t\geq 0$, of bounded operators mapping $\mathfrak{B}\to\mathfrak{B}$ and satisfying

$$K^0 = 1, \quad K^tK^s = K^{t+s}, \quad \|K^t\| \leq 1, \quad \lim_{t\to 0} K^t\psi = \psi, \tag{XI.17}$$

for all $\psi\in\mathfrak{B}$. The infinitesimal generator A of $\{K^t\}$ is defined by

$$A\psi = \lim_{t\to 0} t^{-1}(K^t\psi - \psi), \tag{XI.18}$$

and its domain $D(A)$ consists of all $\psi\in\mathfrak{B}$ for which this limit exists.

In particular, if A is a self-adjoint operator on a Hilbert space, then the $U^t = e^{itA}$ form a unitary group, and a fortiori a contraction semigroup. However, semigroups rather than groups

ordinarily apply to the diffusion equations, where, e.g., the initial condition $\psi_{t=0}(u) = \delta(u)$ does not define a solution for $t < 0$.

We can now state a weak form of Trotter's theorem.

Theorem. Let A, B, and A+B be the infinitesimal generators of the contraction semigroups $\{X^t\}$, $\{Y^t\}$, and $\{Z^t\}$, respectively, acting on the Banach space $\mathfrak{B}$. Then for all $\psi \in \mathfrak{B}$,

$$Z^t\psi = \lim_{m\to\infty}\left(X^{t/m}Y^{t/m}\right)^m\psi. \tag{XI.19}$$

This theorem can be proved in an elementary fashion. The key step is to use the principle of uniform boundedness in order to show that

$$\lim_{h\to 0} h^{-1}\left\|\left(X^hY^h - Z^h\right)Z^\tau\psi\right\| = 0,$$

the limit being uniform in τ for $0 \le \tau \le t$.

This theorem applies in particular to the diffusion equation with a potential $V(\vec{u}) \ge 0$. If we take $L_2(R^n)$ as the basic space, then clearly $V \ge 0$, $H_o \equiv -\Delta \ge 0$, and $V + H_o \ge 0$. Recall that

$$A \ge 0 \iff \langle\psi, A\psi\rangle \ge 0 \quad \text{for all } \psi \in L_2. \tag{XI.20}$$

Hence

$$\left\|e^{-tV}\right\|, \quad \left\|e^{-tH_o}\right\|, \quad \left\|e^{-t(H_o+V)}\right\| \le 1. \tag{XI.21a-c}$$

However, we can also employ the space $L_1(R^n)$. Then (XI.21a) remains obvious, while (XI.21b-c), in fact $\left\|e^{-tH_o}\right\| = 1$, are implied in Section IV.

For the diffusion equation with potential, the successive approximations in (XI.19) correspond just to the approximations used in the construction of the Wiener integral in Section IV.A (except for integration over the final endpoint). Now, as we implied in Section IX.C, Trotter's formula might serve to give still another definition of the Wiener integral. Such a definition would be much more limited in scope than the previous ones. Moreover, it might possibly lead to different results, e.g., for some functionals which depend on $\dot{x}$ as well as on x.

Finally, we note that the foregoing theorem also applies to the higher order parabolic equations like (XI.5).

C. The Path Integral.[25),55)]

Let us now apply Trotter's formula to the Schrödinger equation. The solution (XI.15c) in coordinate space, in n dimensions, takes

the form

$$x(t,\vec{u}) = (2\pi it/\mu)^{-\frac{1}{2}n} \int d^n\vec{v} \exp\left(\tfrac{1}{2} i\mu |\vec{u}-\vec{v}|^2/t\right) x(0,\vec{v}) ,$$

where $i^{\pm\frac{1}{2}} = e^{\pm\frac{1}{4}i\pi}$. This equation is valid for $x \in L_1 \cap L_2$. However, the operator here represented is unitary, and so can be extended to all of L_2. The m-th approximation in (XI.16c) becomes

$$\psi^{(m)}(t,\vec{u}) = \left(2\pi i\, \Delta t/\mu\right)^{-\frac{1}{2}nm} \int d^n\vec{v}_1 \ldots d^n\vec{v}_m e^{iS_m} \psi(0,\vec{v}_1),$$

(XI.22a)

$$S_m = \sum_{j=1}^{m} \left[\tfrac{1}{2}\mu \frac{|\vec{v}_{j+1}-\vec{v}_j|^2}{(\Delta t)^2} - V(v_j)\right] \Delta t, \qquad \text{(XI.22b)}$$

with $\vec{v}_{m+1} = \vec{u}$.

In the limit $m \to \infty$, S_m tends to the classical action $S\{x\}$, Eq. (I.5b) (we suppress the t-dependence). We now denote the limit of the approximation (XI.22a) as follows:

$$\psi(t,\vec{u}) = \int d^n\vec{v} \int_{C(\vec{v},\vec{u})} \mathfrak{D}_F(\vec{x}) e^{iS\{\vec{x}\}} \psi(0,\vec{v}). \qquad \text{(XI.23)}$$

The integral here introduced will be called the Feynman path integral (or simply, the path integral). We continue to use the symbol $C(\vec{v},\vec{u})$, but no relation to continuous functions is implied here.

If we recall the definition of the propagation kernels,

$$\psi(t,\vec{u}) = \int d^n\vec{v}\, K(\vec{v}|\vec{u};t)\, \psi(0,v), \qquad \text{(XI.24a)}$$

we see that

$$K(\vec{v}|\vec{u};t) = \int_{C(\vec{v},\vec{u})} \mathfrak{D}_F(\vec{x}) e^{iS\{\vec{x}\}} . \qquad \text{(XI.24b)}$$

A basic question now is, for what potentials are these formulas valid. In view of the fact that Eqs. (XI.15a-d) already define unitary groups, and in particular contraction semigroups, this will be the case if $(2\mu)^{-1}\Delta - V$ is self-adjoint. We state one set of criteria due to Kato. Let V be a real multiplicative operator, with $V(\vec{u})$ measurable. Then V will be self-adjoint on a domain D(V).

Theorem. If $D(V) \supseteq D(\Delta)$ and there exist constants $a < 1$ and b such that

$$\|V\psi\| \leq a(2\mu)^{-1}\|\Delta\psi\| + b\|\psi\|, \tag{XI.25}$$

for all $\psi \in D(\Delta)$, then $(2\mu)^{-1}\Delta - V$ is self-adjoint.

Note that if V is bounded, then we can take $b = \max|V|$, and the theorem will apply.

However, this theorem does not apply to various functionals of interest. For example, if $V = -k/r^2$, with k sufficiently large, then $(2\mu)^{-1}\Delta - V$ is not unambiguously defined. In particular, it might not be a self-adjoint operator. In such a case, Z^t in (XI.19) is not unitary, and so cannot be approximated (in the sense of strong convergence) by unitary operators. Moreover, some functionals, like $F\{x\} = x(t_0)$, are not associated with contraction semigroups in a natural way.

We therefore mention two other possibilities for defining the path integral. First, we may start with the measure-theoretic integral for an imaginary mass, $i/\mu > 0$, and continue analytically in μ. Second, we may define the path integral directly as a limit,[53] as in Section IV.A. We do not consider the questions, which functionals would then be integrable and when do the various definitions agree.

For evaluating the path integral, the following formula may be useful:

$$\int \mathcal{D}_F(x) e^{i\frac{1}{2}\mu\langle\dot{x},\dot{x}\rangle} F\{x\} = \int \mathcal{D}_W(x) e^{-\frac{1}{2}\mu\langle\dot{x},\dot{x}\rangle} F\{i^{\frac{1}{2}}x\}. \tag{XI.26}$$

This formula is valid if $x(0) = 0$, $x(t)$ is unrestricted, or if $x(0) = x(t) = 0$. Other conditional integrals may be brought to this latter form by a change of variables. (Suitable restrictions on F are understood.)

In the remainder of Section XI we describe some applications of the path integral.

Suppose first that ΔV is a perturbing potential, and let $f = \exp(-i\int d\tau\, \Delta V)$. Then the probability of transition from ψ at time $t = 0$ to χ at time t will be the following matrix element,

$$\langle f\rangle_{\psi\chi} \equiv \langle\psi|f|\chi\rangle$$

$$= \int d^n\vec{u}\,\chi(\vec{u}) \int d^n\vec{v}\,\psi(\vec{v}) \int_{C(\vec{v},\vec{u})} \mathcal{D}_F(x) e^{iS\{x\}} f\{x\}. \tag{XI.27}$$

If we expand f in powers of ΔV, we can easily obtain the familiar perturbation series. Equation (XI.27) also can give a convenient representation for other matrix elements, e.g., of x, p, et cetera.

We now indicate some heuristic manipulations. For suitable functionals F we should have, as in the examples of Section VII.B,

$$\int_{C(\vec{v},\vec{u})} \mathcal{D}_F(x) \frac{\delta}{\delta x^\alpha(\tau)} \left[e^{iS\{x\}} F\{x\} \right] = 0, \tag{XI.28a}$$

and

$$\langle \delta F/\delta x^\alpha(\tau) \rangle_{\psi\chi} = -i \langle F \delta S/\delta x^\alpha(\tau) \rangle_{\psi\chi}. \tag{XI.28b}$$

One can interpret $\delta S/\delta x^\alpha(\tau)$ with the help of Green's functions, as in (VII.12), or by using the rules of Section I.B.

Equation (XI.28b) is a special case of Schwinger's action principle.[56] This principle asserts that, quite generally,

$$\langle \delta F \rangle_{\psi\chi} = -i \langle F \delta S \rangle_{\psi\chi}. \tag{XI.29}$$

For example, let F and S depend on a suitably-behaving function of position η. Then also $\eta(\tau) = \eta(\vec{x}(\tau))$, and the chain rule

$$\frac{\delta}{\delta \eta(\rho)} = \sum_\alpha \int d\tau \frac{\delta x^\alpha(\tau)}{\delta \eta(\rho)} \frac{\delta}{\delta x^\alpha(\tau)} \tag{XI.30}$$

implies (XI.29) for variations with respect to η.

We can also give an example from quantum field theory. If we have an electromagnetic field which depends (at least partially) on a given external current J_μ, and

$$S = S^{(o)} + \int d^4v' A^\mu(v') J_\mu(v'), \tag{XI.31}$$

where $S^{(o)}$ does not depend on J_μ, then

$$\left\langle \frac{\delta A^\nu(u)}{\delta J_\mu(v)} \right\rangle = -i \left\langle A^\nu(u) \frac{\delta S}{\delta J_\mu(v)} \right\rangle = -i \langle A^\nu(u) A^\mu(v) \rangle. \tag{XI.32}$$

The expectations refer to arbitrarily chosen states.

D. Classical Limit of Quantum Mechanics.[57]

The path integral elucidates the nature of the stationary phase (or WKB or JWKB) approximation in quantum mechanics. We start with Eq. (XI.24b) for the propagation kernel, but we do not set $\hbar = 1$:

$$K(\vec{v}|\vec{u};t) = \int_{C(\vec{v},\vec{u})} \mathcal{D}_F(x) e^{iS\{x\}/\hbar} . \qquad \text{(XI.33)}$$

We are interested in deriving an approximation for K, which is valid if $\hbar$ is small compared with other parameters of the system. Now, for small $\hbar$, the exponential will oscillate rapidly, and the principal contribution will correspond to minimal oscillations. These are the paths near the classical path $\vec{X}(\tau)$, which of course satisfies

$$\delta S\{\vec{x}\}/\delta x^{\alpha}(\tau)\Big|_{\vec{x}=\vec{X}} = 0. \qquad \text{(XI.34)}$$

We therefore proceed as follows. Let

$$\vec{x} = \vec{X} + \hbar^{\frac{1}{2}} \vec{\chi} , \qquad \text{(XI.35a)}$$

where $\vec{x}(0) = \vec{X}(0) = \vec{v}$ and $\vec{x}(t) = \vec{X}(t) = \vec{u}$. Thus

$$\vec{\chi}(0) = \vec{\chi}(t) = 0. \qquad \text{(XI.35b)}$$

Let us assume a Volterra expansion for S and take the first two non-vanishing terms:

$$S\{\vec{x}\} \approx S\{\vec{X}\} + \tfrac{1}{2}\hbar S_2\{\chi\}, \qquad \text{(XI.36a)}$$

$$S_2\{\vec{\chi}\} = \sum_{\alpha,\beta} \int_0^t d\tau \left(\frac{\partial^2 T}{\partial \dot{x}^{\alpha} \partial \dot{x}^{\beta}}\Big|_{x=X(\tau)} \dot{\chi}^{\alpha}\dot{\chi}^{\beta} - \frac{\partial^2 V}{\partial x^{\alpha}\partial x^{\beta}}\Big|_{x=X(\tau)} \chi^{\alpha}\chi^{\beta} \right)$$

$$= \int_0^t d\tau \left[\mu \dot{\vec{\chi}}^2(\tau) - q_{\alpha\beta}(\tau)\chi^{\alpha}(\tau)\chi^{\beta}(\tau) \right] . \qquad \text{(XI.36b)}$$

Here $q_{\alpha\beta}$ is a known function insofar as we can solve the classical problem, i.e., find $X(\tau)$.

We now insert (XI.36a) into (XI.33) and define the approximate kernel,

$$K_2 = e^{iS\{\vec{X}\}/\hbar} \int_{C(\vec{0},\vec{0})} \mathcal{D}_F(\vec{\chi}) e^{i\frac{1}{2}S_2\{\vec{\chi}\}} \equiv e^{iS\{\vec{X}\}/\hbar} J. \qquad \text{(XI.37)}$$

Note that the two integrals, in (XI.33) and (XI.37), refer to different variances, $\hbar\mu^{-1}$ and μ^{-1} respectively (aside from factors i). Still, the formal manipulations lead to a correct expression.

The integral J is manifestly independent of $\hbar$. There appears to be a correction to K_2 of order $\hbar^{1/2}$, but this is an integral of an

odd functional, hence zero. Consequently, $K-K_2$ is of order $\hbar$.

Let us look at the integral J more closely. Equation (XI.26) implies

$$J = \int_{C(\vec{0},\vec{0})} \mathfrak{D}_W(\vec{\chi})\, e^{-\frac{1}{2}\mu\langle\vec{\chi},\vec{\chi}\rangle} \exp\Big(\tfrac{1}{2}\int_0^t d\tau\, q_{\alpha\beta}\chi^\alpha\chi^\beta\Big). \tag{XI.38a}$$

For simplicity, let us now consider one-dimensional motion, and let

$$J^{(1)} = \int_{C(0,0)} \mathfrak{D}_W(\chi) e^{-\frac{1}{2}\mu\langle\dot{\chi},\dot{\chi}\rangle} \exp\Big(\tfrac{1}{2}\int_0^t d\tau\, q\chi^2\Big). \tag{XI.38b}$$

As a typical example, we may have

$$V = \lambda u^4, \qquad q(\tau) = 12\lambda X^2(\tau) \geq 0. \tag{XI.39}$$

This is in contrast to (IV.15), where we specified $q=-p\leq 0$. However, $S\{x\}$ has a minimum[3] at $x=X$, so S_2 is positive definite. With reference to (VII.3)-(VII.4), this means that $1+A>0$, as required there. The fact that perhaps $A<0$ is of no consequence.

In Section IV.B we described the evaluation of an integral like $J^{(1)}$ in (XI.38b), but unrestricted. The method used there can also be adapted to the conditional integral, and the answer is as follows.[58] If

$$f''(\tau)+q(\tau)f(\tau) = 0, \qquad f(t) = 0, \qquad f'(t) = -1, \tag{XI.40a}$$

then

$$J^{(1)} = [\pi f(0)]^{-\frac{1}{2}}. \tag{XI.40b}$$

The expression (XI.37) for K_2 is in agreement with the results of the stationary phase approximation. There one writes, with an error of order $\hbar$,

$$\psi_2 = e^{iS\{\vec{X}\}/\hbar} e^{S_1}. \tag{XI.41a}$$

In the cited reference,[57] indirect arguments were given that $J=e^{S_1}$. It does not seem easy to compare the evaluation (XI.40b) with the standard time-independent solution,

$$(e^{S_1})(u) = \text{const.}\,|E-V(u)|^{-\frac{1}{4}}. \tag{XI.41b}$$

XII. Two Special Functional Integrals

In this section we describe the Feynman history integral and an anticommutative integral. Both of these have been introduced primarily to deal with situations which so far are peculiar to quantum field theory. However, it is not at all unusual to introduce an integral which may be of a rather limited applicability. We may quote, in this connection, the title of a recent review article:[59] "Integrals devised for special purposes."

A. A Simple Model.

We are interested in constructing a functional integral representation for time-ordered vacuum expectation values. Here we present a trivial model, and the treatment that follows can easily be made rigorous.

We consider a system with a denumerable number of excitations, e.g., an oscillator. We employ the standard annihilation and creation operators,

$$a^- = \begin{bmatrix} 0 & 1 & & 0 \\ & 0 & \sqrt{2} & \\ & & 0 & \sqrt{3} \\ 0 & & \ddots & \ddots \end{bmatrix}, \qquad a^+ = \begin{bmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \sqrt{2} & 0 & \\ 0 & & \ddots & \ddots \end{bmatrix}, \tag{XII.1a}$$

$$[a^-, a^+] = 1, \tag{XII.1b}$$

which are supposed to relate to the excitations in the usual way. We now define

$$\varphi(\tau) = a^- - \tau a^+. \tag{XII.2}$$

Here τ has the interpretation of imaginary time, $i^{-1}t$, and various standard field-theoretic relations should be modified accordingly. As a result, we will be able to use the Wiener integral rather than the Feynman path integral, and for the former a more extensive theory is available.

We see that $\dot{\varphi} = -a^+$, and that the equation of motion and the "equal time" commutation relations are as follows,

$$\ddot{\varphi} = 0, \quad [\dot{\varphi}, \varphi] = 1. \tag{XII.3a,b}$$

The τ-dependence, or evolution, of φ can be described in terms of a "Euclidean Hamiltonian,"

$$H = \tfrac{1}{2}\dot{\varphi}^2 = \tfrac{1}{2}(a^+)^2, \tag{XII.4a}$$

$$\varphi(\tau) = e^{\tau H}\varphi(0)e^{-\tau H}. \qquad \text{(XII.4b)}$$

Of course, we are dealing here with unbounded operators, and questions of domain should be carefully considered.

Let us determine some vacuum expectation values:

$$\langle\varphi(\tau_1)\varphi(\tau_2)\rangle_0 = \langle a^-\tau_2 a^+\rangle_0 = \tau_2, \qquad \text{(XII.5a)}$$

and with reference to "Euclidean time ordering,"

$$\langle\big(\varphi(\tau_1)\varphi(\tau_2)\big)_+\rangle = \min(\tau_1,\tau_2)$$

$$= \int d\mu_W(\eta)\eta(\tau_1)\eta(\tau_2), \qquad \text{(XII.5b)}$$

in view of (IV.14b). (There and here we take $\sigma = 1$.) One may verify the corresponding relation for higher order vacuum functions. We conclude that for functionals restricted to a suitable class,

$$\langle(F\{\varphi\})_+\rangle_0 = \int d\mu_W(\eta)F\{\eta\}. \qquad \text{(XII.6)}$$

We note that we may take $[0,\infty)$ as the basic interval for the Wiener integral.

As in Section III.C, we can construct the generating functional,

$$\Omega\{J\} \equiv \langle\big(e^{i\int d\tau\varphi J}\big)_+\rangle_0 = \exp\Big[-\tfrac{1}{2}\int_0^\infty\int_0^\infty d\tau' d\tau''\min(\tau',\tau'')J(\tau')J(\tau'')\Big] \qquad \text{(XII.7a)}$$

(see (VII.21)). This functional satisfies an equation analogous to (III.26),

$$\big(d^2/d\tau^2\big)\big[\delta\Omega\{J\}/\delta J(\tau)\big] = J(\tau)\Omega\{J\}. \qquad \text{(XII.7b)}$$

B. The History Integral.

We will now adapt the foregoing considerations to the free scalar field. We will be led in this way to an integral, whose mathematical structure is as yet imperfectly understood.

Here we present a heuristic discussion, which is based on the approach of Bogolubov and Shirkov.[39] Our short summary should be regarded as a supplement to the clear and concise exposition in the cited text.

First we evaluate the vacuum expectation value which is the free field generating functional, cf. Section III:

$$T^{(0)}\{J\} = \langle \left(\exp i \int d^4 p\, \varphi^{(0)}(p) J(-p)\right)_+ \rangle_0 . \qquad \text{(XII.8)}$$

(We suppress the tildes which would indicate Fourier transforms.) In general, a time-ordered product can be reduced to the normal form by replacing, in momentum space,

$$\varphi^{(0)}(k) \to \varphi^{(0)}(k) - i\Delta^c(k)\,\delta/\delta\varphi^{(0)}(-k), \qquad \text{(XII.9a)}$$

$$\Delta^c(k) = -\left(k^2 - \mu^2 + i\varepsilon\right)^{-1} = \Delta^c(-k). \qquad \text{(XII.9b)}$$

More precisely, if $(F_1\{\varphi^{(o)}\})_+ = :F_2\{\varphi^{(o)}\}:$, then, in an obvious notation,

$$F_2\{\zeta_k\} = F_1\left\{\zeta_k - i\Delta_k^c \delta/\delta\zeta_{-k}\right\}\cdot 1. \qquad \text{(XII.9c)}$$

This rule may be verified directly for polynomial functionals.

Such rules express the combinatorial features of the ordering prescriptions. We can give another example. Suppose that we have a free field whose annihilation and creation parts are so adjusted that they satisfy

$$\left[\chi^{(-)}(\vec{u}),\, \chi^{(+)}(\vec{v})\right] = \delta(\vec{u} - \vec{v}). \qquad \text{(XII.10a)}$$

Let $F\{\chi^{(-)}, \chi^{(+)}\}$ be an arbitrary functional. If we want to reduce F to the normal form, $F = :H:$, then we set

$$H\left\{\xi^-, \xi^+\right\} = F\left\{\xi^- + \delta/\delta\xi^+,\, \xi^+\right\}\cdot 1. \qquad \text{(XII.10b)}$$

Let us return to $T^{(o)}$. We introduce

$$\Phi_J\{\zeta;\lambda\} = \exp\left\{i\lambda \int d^4p\left[\zeta(p) - i\Delta^c(p)\,\delta/\delta\zeta(-p)\right]J(-p)\right\}\cdot 1, \qquad \text{(XII.11a)}$$

and $T^{(o)}\{J\} = \Phi_J\{0;1\}$ (since, in general, $\langle :F\{\varphi^{(o)}\}:\rangle_0 = F\{0\}$). We differentiate with respect to λ:

$$\frac{\partial}{\partial\lambda}\Phi_J\{\zeta;\lambda\} = i\int d^4p\left[\zeta(p) - i\Delta^c(p)\frac{\delta}{\delta\zeta(-p)}\right]J(-p)\,\Phi_J\{\zeta;\lambda\}. \qquad \text{(XII.11b)}$$

We see a differential equation, which resembles those that we solved in Section VIII. We find Φ_J, by trial and error, or otherwise, and obtain

$$T^{(o)}\{J\} = \exp\left[\tfrac{1}{2} i \int d^4p\, J(p)\, \Delta^c(p)\, J(-p)\right]. \qquad \text{(XII.12)}$$

Thus the vacuum expectation value of a linear exponential gives a Gaussian. This is in analogy with the formulas for the two basic integrals. We therefore write $T^{(o)}\{J\}$ (heuristically) as a functional integral, with a suitable Gaussian weight. Explicitly,

$$\langle \left(e^{i\int d^4p \varphi^{(o)}(p) J(-p)} \right)_+ \rangle_o \sim \int \mathfrak{D}(\eta) e^{iA^{(o)}\{\eta\}} e^{i\int d^4p\, \eta(p) J(-p)}, \qquad \text{(XII.13)}$$

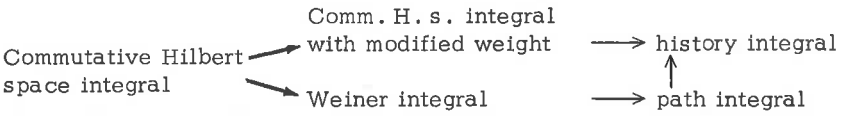
where $A^{(o)}$ is the free-field action,

$$A^{(o)}\{\eta\} = \int d^4p\, \eta(p)(p^2 - \mu^2)\, \eta(-p). \qquad \text{(XII.14)}$$

As in (XII.9b), the prescription $\mu^2 \to \mu^2 - i\varepsilon$ near $p^2 = \mu^2$ is implied. Moreover, the obvious normalization of the functional integral is to be understood, i.e., to unity for $J = 0$. We call the new functional integral the Feynman history integral. More generally,

$$\langle \left(F\{\varphi^{(o)}\} \right)_+ \rangle_o \sim \int \mathfrak{D}(\eta) e^{iA^{(o)}\{\eta\}} F\{\eta\}. \qquad \text{(XII.15)}$$

The history integral has the basic form of the Hilbert space integral with a modified weight, as in the examples of Section VI.D. To see more clearly the interrelation of the various integrals, the following diagram may be helpful:



The two arrows on the left indicate specialization, and we refer to Section VI. The horizontal arrows indicate analytic continuation. See Sections XI.C and XII.C, where we mention continuation in μ and in t, respectively. Finally, one can try to interrelate the two Feynman integrals as follows.[60] We choose an orthonormal basis for the one-particle sector of $\varphi^{(o)}$, and in this way separate the degrees of freedom. For each degree of freedom, we describe the time

evolution by means of the path integral. When the path integrals are combined, one obtains a history integral. But this approach remains to be worked out more fully.

C. Time-Ordered Functions and Functionals.[39),61)] We can now derive easily (but of course heuristically) some interesting formulas for interacting fields. We consider the Lagrangian

$$L(v) = L^{(o)}(v) + L_1(v), \tag{XII.16a}$$

where, e.g., $L_1(v) = \lambda\varphi^4(v)$. All questions of renormalization will be ignored here. The total action is

$$A = A^{(o)} + A_1 = \int d^4v\, L^{(o)} + \int d^4v\, L_1. \tag{XII.16b}$$

We also write $L(\varphi)$, $A\{\varphi\}$, et cetera. The S-matrix corresponding to L_1 is given by

$$S = \Big(\exp i \int d^4v\, L_1(v)\Big)_+. \tag{XII.17}$$

For the time-ordered function $\tau_2 = \langle(\varphi\,\varphi)_+\rangle_o$ we have the expression

$$\tau_2(u-v) = \langle\Big(\varphi^{(o)}(u)\varphi^{(o)}(v)S\Big)_+\rangle_o / \langle S\rangle_o \tag{XII.18a}$$

$$= N^{-1}\int \mathcal{D}(\eta) e^{iA\{\eta\}}\eta(u)\eta(v), \tag{XII.18b}$$

in view of (XII.15)-(XII.17). Here $N = \langle S\rangle_o$. More generally, we see that we can construct the generating functional T, Eq. (III.16), as follows,

$$T\{J\} = N^{-1}\int \mathcal{D}(\eta) e^{iA\{\eta\}} \exp\Big[i\int d^4v\, \eta(v)J(v)\Big]. \tag{XII.19}$$

If we combine the two exponents, we can think of the sum as determined by a new Lagrangian,

$$L_J(v) = L^{(o)}(v) + L_1(v) + \varphi(v)J(v). \tag{XII.20a}$$

It follows that $T\{J\}$ is the vacuum-to-vacuum transition function for the field which satisfies the equation

$$(\Box - \mu^2)\varphi(v) = (L_1'(\varphi))(v) + J(v). \tag{XII.20b}$$

We thus see that the function J, which was introduced merely as the argument for a functional, in fact has the natural interpretation as an external source.

It is remarkable that the strictly heuristic expression (XII.19) can be given a rigorous meaning (for, e.g., $L_1 = \lambda\varphi^4$) if three modifications are made:[38]

(i) A transition is made to the Euclidean region, $t \to it$, so that the integral becomes a commutative Hilbert space integral, with a modified weight.

(ii) The Klein-Gordon operator is replaced by a product of such operators so that the Green's function

$$G(u-v) = \frac{1}{(2\pi)^4} \int d^4k \, \frac{e^{ik(u-v)}}{\Pi(k^2+\mu_i^2)} \qquad \text{(XII.21)}$$

is everywhere finite. The Lagrangian and the action must be modified accordingly.

(iii) The system is restricted to a finite volume of space-time.

Under these modifications, the integral analogous to (XII.19) converges to a generating functional for Euclidean time-ordered functions. This functional satisfies the analogue to Eq. (III.26). Hopefully, one could continue analytically the Euclidean functions and functionals to the Minkowski space-time. But at present this is of secondary interest, since the solution to the Euclidean problem is still so incomplete. [In particular, one would like to be able to dispense with modifications (ii) and (iii).]

We make another comment about the integral (XII.19). This integral is the Fourier transform of an exponentiated polynomial, if L_1 is a polynomial in φ. It may be worth noting that we encountered several times before the Fourier transforms of such functions or functionals and the corresponding differential equations. Compare with (VIII.25), (VIII.29), and (XI.2).

Let us now consider, again in (XII.19), the two terms in $A\{\eta\}$. By interchanging operations, we can obtain

$$T\{J\} = N^{-1} \exp A_1\{\delta/i\delta J\}\, T^{(0)}\{J\}. \qquad \text{(XII.22)}$$

See (XII.12) for $T^{(0)}$. We have here one form of solution to the field theory defined by L_1. Unfortunately, this solution still contains all the standard divergences, and it is not easily manageable otherwise (except by an expansion in series).

For the case of the trilinear boson-fermion couplings, it is not difficult to derive the following expression for the two-point fermion time-ordered (i.e., Green's) function:

$$S(u-v) = N_1^{-1} \int \mathcal{D}(\Omega) e^{iA^{(o)}\{\Omega\}} G(\Omega;u,v) \langle S \rangle_\Omega , \qquad \text{(XII.23a)}$$

$$N_1 = \int \mathcal{D}(\Omega) e^{iA^{(o)}\{\Omega\}} \langle S \rangle_\Omega . \qquad \text{(XII.23b)}$$

Here $G(\Omega;u,v)$ is the one-particle Green's function, and $\langle S \rangle_\Omega$ is the vacuum expectation value of the S-matrix, both in the presence of the external field Ω.

If one assumes a given external field η, then he should replace Ω by $\Omega+\eta$ in G and in $\langle S \rangle_\Omega$. If, moreover, there are no antifermions in the theory, then $\langle S \rangle_{\Omega+\eta} = 1$, and we obtain an expression analogous to (VIII.39).

D. Exterior Algebras.[62]

It is natural to try to extend the functional approach of the preceding section to fermi fields. We can only make a very modest beginning in this direction. To start with, we will develop some techniques for handling antisymmetric tensors.

We recall that the exterior algebra (or Grassmann algebra) over R^n, which we denote by $E_r(R^n)$, is the algebra of skew-symmetric tensors. If $(e_1,\ldots,e_n)$ is an orthonormal basis for R^n, and $i_1,\ldots,i_k$ are any integers with $1 \le i_j \le n$, we define

$$e_{i_1} \wedge \ldots \wedge e_{i_k} = \sum (k!)^{-1} \operatorname{sgn} \sigma \; e_{\sigma(i_1)} \otimes \ldots \otimes e_{\sigma(i_k)} , \qquad \text{(XII.24a)}$$

where we sum over all permutations. We also adjoint the unity to the algebra. Then such elements for

$$1 \le i_1 < \ldots < i_k \le n; \quad k = 0,1,\ldots,n, \qquad \text{(XII.24b)}$$

where the unity corresponds to $k=0$, form a basis for the exterior algebra. It is obvious that

$$e_1 \wedge e_2 = -e_2 \wedge e_1, \quad e_1 \wedge e_1 = 0, \text{ etc.}$$

The normalization of the vectors (XII.24a) is as follows,

$$\langle e_{i_1} \wedge \ldots \wedge e_{i_k}, e_{i_1} \wedge \ldots \wedge e_{i_k} \rangle = (k!)^{-1} . \qquad \text{(XII.24c)}$$

The total dimension of the algebra is

$$1 + n + \tfrac{1}{2} n(n-1) + \cdots + n + 1 = 2^n . \qquad \text{(XII.25)}$$

The algebra $E_r(R^n)$ has a natural complex extension, which we denote by $E(R^n)$.

Suppose that we have an antisymmetric function,

$$F(s,t) = y_1(s)y_2(t) - y_1(t)y_2(s). \tag{XII.26}$$

A natural way to look at F is as a functional, which is also a skew-symmetric tensor in the space of functions y_1. Let us restrict ourselves for the time being to an n-dimensional space of functions, $n<\infty$. Then we may identify F as an element of the algebra $E(R^n)$. The most general such functional depending on $x = \Sigma x^i e_i$ is

$$A(x) = A(0) + \sum A_i x^i e_i + \sum_{i<j} A_{ij} x^i x^j e_i \wedge e_j + \cdots$$

$$+ A_{1\ldots n} x^1 \ldots x^n e_1 \wedge \ldots \wedge e_n, \tag{XII.27a}$$

or, if $x_i = x^i e_i$ (do not sum), then

$$A(x) = A(0) + \sum A_i x_i + \sum_{i<j} A_{ij} x_i \wedge x_j + \ldots + A_{1\ldots n} x_1 \wedge \ldots \wedge x_n. \tag{XII.27b}$$

We have here a very close analogy with the polynomial functionals (VII.2a). In fact, the latter may be identified with the symmetric tensors over $\mathfrak{K}$, or over its complex extension. One has the natural correspondence (do not sum):

$$y^j y^k \longleftrightarrow \frac{1}{2}\left(y^j e_j \otimes y^k e_k + y^k e_k \otimes y^i e_j\right). \tag{XII.28}$$

We now introduce three new operations. First, we define a linear map on $E(R^n)$ by its action on the elements (XII.24) of the basis,

$$\left(e_{i_1} \wedge e_{i_2} \wedge \ldots \wedge e_{i_k}\right)' = e_{i_k} \wedge \ldots \wedge e_{i_2} \wedge e_{i_1} = (-1)^{k+1} e_{i_1} \wedge \ldots \wedge e_{i_k}, \tag{XII.29}$$

for $k>0$, and $(1)' = 1$ for $k=0$.

Second, we define left and right differentiation:

$$\partial_p x_q = x_q \overleftarrow{\partial}_p = \delta_{pq}, \quad \partial_p 1 = 1 \overleftarrow{\partial}_p = 0, \tag{XII.30a}$$

$$\partial_p \left[x_i \wedge A(x) \right] = \delta_{pi} A(x) - x_i \wedge \partial_p A(x), \quad \left(A \wedge x_j \right) \overleftarrow{\partial}_p = \delta_{pj} A - \left(A \overleftarrow{\partial}_p \right) x_j . \tag{XII.30b}$$

It follows that

$$\partial_i \partial_j A = -\partial_j \partial_i A, \quad A \overleftarrow{\partial}_j = \left(\partial_j A' \right)'. \tag{XII.31}$$

Third, let $x, y \in R^n$, let $(x,y) = (x^1, y^1, x^2, \ldots, y^n)$. We define Gaussian functionals as elements of $E(R^{2n})$ by their expansion:

$$\exp(x,y) = \exp\left(\sum x_j \wedge y_j \right) = \exp(-y,x)$$

$$= 1 + x_1 \wedge y_1 + \cdots + x_1 \wedge y_1 \wedge x_2 \wedge y_2 + \ldots + \bigwedge_{j=1}^{n} (x_j \wedge y_j) \tag{XII.32a}$$

$$= \bigwedge_j \exp(x_j \wedge y_j). \tag{XII.32b}$$

We may observe that the functional $\exp(\frac{1}{2} \Sigma_{i,j} c_{ij} u_i \wedge u_j)$ can always be brought to the form (XII.32) by a suitable choice of basis.

E. The Anticommutative Integral.[62]

We outline here the construction of the anticommutative analogue to the commutative Hilbert space integral. Various extensions, like to the history integral, must await future developments.

Our discussion is based on exterior algebras, and follows the presentation of Berezin. Another possibility is to use Clifford algebras[63] (see below). There has also been an attempt to construct a functional integral for spinor fields in terms of the more familiar commuting functions.[64]

We now define the formal rules of integration for antisymmetric functionals as in (XII.27):

$$\int x_j dx_j = \int dx_j x_j = 1, \tag{XII.33a}$$

$$\int A(x)dx_k = 0 \quad \text{if } A \text{ is independent of } x_k. \tag{XII.33b}$$

We emphasize that these rules are formal, and no summations or limits are implied here. The first of these means, in particular,

$$\int A_{ij}x_i \wedge x_j dx_j = A_{ij}x_i = -A_{ij}\int dx_j x_i \wedge x_j, \text{ etc.}$$

Note that it is consistent to set, for $i \neq j$,

$$dx_i dx_j = -dx_j dx_i, \qquad dx_i x_j = -x_j dx_i. \tag{XII.34}$$

(One could also write $dx_i \wedge x_j$, et cetera, if he desired.) With reference to the expansions (XII.27), we have

$$\int A d^n\vec{x} \equiv \int A dx_n \ldots dx_2 dx_1 = A_{1\ldots n}. \tag{XII.35}$$

We do not investigate here the detailed properties of the integral, e.g., the transformation implied by a change of basis. However, we give a rule for integration by parts to indicate the analogy with the commutative integral:

$$\int A(x)\left[\partial_p B(x)\right] d^n\vec{x} = \int \left[A(x)\overleftarrow{\partial}_p\right] B(x) d^n\vec{x}. \tag{XII.36}$$

Let us now consider Gaussian integrals,

$$\int \exp(x,y)\, C(x,y) dy_n dx_n \ldots dy_1 dx_1. \tag{XII.37a}$$

It is easy to see that those terms of C will contribute which contain, for each j, either (i) both x_j and y_j, or (ii) neither x_j nor y_j. A consequence of this is the formula

$$\int \exp(x,y)\left[A^*(x)\right]' B(y) \prod dy_j dx_j = \sum A^*_{i_1\ldots i_k} B_{i_1\ldots i_k}. \tag{XII.37b}$$

The final sum is the scalar product of the elements in $E(R^n)$ which we may associate with the functionals A and B respectively, e.g.,

$$A(x) \longrightarrow a = A(0) + \sum A_i e_i + \sum_{i<j} A_{ij}(2!)^{\frac{1}{2}} e_1 \wedge e_j + \cdots$$

$$+ (n!)^{\frac{1}{2}} A_{1\ldots n} e_1 \wedge \cdots \wedge e_n. \qquad \text{(XII.38)}$$

(Compare the normalization (XII.24c).)

The exterior algebra, as well as the integral, can be extended to Hilbert spaces. We will not discuss such an extension in detail, except to point out the following. For the product of the Gaussian factor and a cylinder functional, the integral will become finite-dimensional, as in the case of the commutative integral. The Gaussian factor alone will integrate to unity.

We will now make some remarks about the Clifford algebra $C(R^{2\nu})$. For the moment we assume even dimensionality 2ν and $\nu < \infty$ for the basic space. The algebra is to be over the complex numbers.

We recall that this algebra is generated by elements $f_1, \ldots, f_{2\nu}$, corresponding to an orthonormal basis for $R^{2\nu}$. The products of these elements obey the rule

$$\left[f_i, f_j\right]_+ = 2\delta_{ij}. \qquad \text{(XII.39)}$$

The elements of $C(R^{2\nu})$ can be represented irreducibly by $2^\nu \times 2^\nu$ matrices.

In terms of such a representation, we can define on $C(R^{2\nu})$ the following linear form: if $g \in C(R^{2\nu})$,

$$T(g) = 2^{-\nu} \operatorname{tr} g. \qquad \text{(XII.40a)}$$

This functional satisfies, independently of the specific representation (if irreducible):

$$T(1) = 1, \quad T(f_{i_1} \ldots f_{i_k}) = 0 \qquad \text{(XII.40b)}$$

if the i_j are all different. Thus, T picks out from g the coefficient of the unit element.

Suppose that we now have an antisymmetric functional A(x) as in (XII.27). There A was a map into $E(R^n)$. We may also introduce a map into $C(R^n)$:

$$A_c(x) = A(0) + \sum A_i x^i f_i + \sum_{i<j} A_{ij} x^i x^j f_i f_j + \cdots$$

$$+ A_{1\ldots n} x^1 \ldots x^n f_1 \ldots f_n \qquad \text{(XII.41)}$$

(n can be assumed even, but this is not essential here).

We shall not attempt to construct rules for integrating $A_c(x)$; however, we note the following. If, in analogy to $a \in E(R^n)$, (XII.38), we define

$$a_c = A(0) + \sum A_i f_i + \ldots + A_{1\ldots n} f_1 \ldots f_n, \qquad \text{(XII.42a)}$$

and b_c refers similarly to $B(x)$, then we can recover the evaluation (XII.37b) by utilizing the trace form:

$$T(a_c^* b_c) = T(b_c a_c^*) = \sum A^*_{i_1\ldots i_k} B_{i_1\ldots i_k}. \qquad \text{(XII.42b)}$$

As in the case of the exterior algebra and the anticommutative integral, the foregoing construction extends to infinite-dimensional Hilbert spaces. In particular, the trace form T satisfying (XII.40b) can be defined.

F. Free-Field Vacuum Expectation Values.

We first consider the free boson field $\varphi^{(o)}$. For definiteness of functional form, let $\varphi^{(o)}$ depend on three-vectors $\vec{p}$, so that

$$\varphi^{(o)} = \varphi^{(o,+)} + \varphi^{(o,-)}, \quad \left[\varphi^{(o,-)}(\vec{p}), \varphi^{(o,+)}(\vec{q})\right] = \delta(\vec{p} - \vec{q}). \qquad \text{(XII.43)}$$

An arbitrary vector in the Fock space of $\varphi^{(o)}$ can now be written in two ways,

$$F\left\{\varphi^{(o,+)}\right\} |0\rangle \qquad \text{and} \qquad f\left\{\varphi^{(o)}\right\} |0\rangle. \qquad \text{(XII.44)}$$

The correspondence of the two functional forms, for boson and for fermion fields, is analyzed in detail in two theorems of Segal.[63),65)]

For the boson fields, the scalar products corresponding to the two forms in (XII.44) may be conveniently represented in terms of the commutative Hilbert space integral.[28),62)] For the first form,

$$\langle F_1\left\{\varphi^{(o,-)}\right\} F_2\left\{\varphi^{(o,+)}\right\}\rangle_o = \int \mathfrak{D}(\eta^-)\mathfrak{D}(\eta^+) e^{-(\eta^-,\eta^+)} F_1\{\eta^-\} F_2\{\eta^+\}. \qquad \text{(XII.45)}$$

Here we integrate over a complex Hilbert space, i.e., separately over the real and over the imaginary components. Furthermore, $\eta^- = (\eta^+)^*$, so that

$$(\eta^-, \eta^+) = \int d^3\vec{p}\, |\eta^\pm|^2 > 0, \qquad \text{for } \eta^\pm \neq 0.$$

Second, we have

$$\langle f\{\varphi^{(o)}\}\rangle_o = \int \mathfrak{D}(\eta) e^{-\frac{1}{2}\langle \eta,\eta\rangle} f\{\eta\}. \qquad \text{(XII.46a)}$$

The quantity $\langle f\rangle_o$ can also be expressed as follows,

$$\langle f\{\varphi^{(o)}\}\rangle_o = f\{\eta + \delta/\delta\eta\}\Big|_{\eta=0}, \qquad \text{(XII.46b)}$$

in view of (XII.10b). It is a remarkable fact that in (XII.44) the function f can be obtained from F by a Weiner transform, Eq. (VII.24b).

For the free spinor field, the situation is analogous. We consider a neutral field,

$$\psi^{(o)}(\vec{p}) = \psi^{(o,+)}(\vec{p}) + \psi^{(o,-)}(\vec{p}) = \left[\psi^{(o)}(\vec{p})\right]^*.$$

The discussion of Sections XII.D and XII.E does not include the functional formalism, and restricts us to using coordinates. Therefore, we assume an orthonormal basis for $L_2(R^3)$, and then

$$\psi_j^{(o)} = \psi_j^{(o,+)} + \psi_j^{(o,-)}, \qquad \left[\psi_j^{(o)}, \psi_k^{(o)}\right]_+ = 2\delta_{jk}, \qquad \text{(XII.47a,b)}$$

$$\left[\psi_j^{(o,-)}, \psi_k^{(o,+)}\right]_+ = \delta_{jk}, \quad \left[\psi_j^{(o,\pm)}, \psi_k^{(o,\pm)}\right]_+ = 0. \qquad \text{(XII.47c,d)}$$

Again, a vector in the Fock space of $\psi^{(o)}$ may be written in two ways,

$$H\{\psi^{(o,+)}\}|0\rangle \quad \text{and} \quad h\{\psi^{(o)}\}|0\rangle. \qquad \text{(XII.48a)}$$

The corresponding scalar products may be represented as follows,

$$\langle H_1\{\psi^{(o,-)}\} H_2\{\psi^{(o,+)}\}\rangle_o = \int \exp(\zeta^-,\zeta^+) H_1\{\zeta^-\} H_2\{\zeta^+\} \prod d\zeta_j^+ d\zeta_j^-, \qquad \text{(XII.48b)}$$

$$\langle h\{\psi^{(o)}\}\rangle_o = T(h). \qquad \text{(XII.48c)}$$

Note that, in view of (XII.47b), $h\{\psi^{(o)}\}$ can be taken as an element of a Clifford algebra over $L_2(R^3)$. In (XII.48b), it is natural to identify $\zeta^- = (\zeta^+)^*$, as in (XII.45).

The two functionals H and h can be related simply as A and A_c respectively in Eqs. (XII.27a) and (XII.42a). There is also an analogue to (XII.46b),

$$\langle h\{\psi_1^{(o)}, \psi_2^{(o)}, \ldots\}\rangle_o = h\{\zeta_1 + \vec{\partial}_1, \zeta_2 + \vec{\partial}_2, \ldots\}\cdot 1\Big|_{\text{all } \zeta_j = 0}. \tag{XII.49}$$

The last equation is an example of the realization of a Clifford algebra, say $C(R^{2\nu})$, as operators on functionals in $E(R^\nu)$. In particular, we take the (left or right) multiplicative and differentiation operators, and observe that

$$\left[\vec{x}_j, \vec{\partial}_k\right]_+ = \delta_{jk} = \left[\overleftarrow{x}_j, \overleftarrow{\partial}_k\right]_+. \tag{XII.50a}$$

Thus we may realize (XII.39) by setting, for $1 \le j \le \nu$,

$$f_j = \vec{x}_j + \vec{\partial}_j, \qquad f_{j+\nu} = i^{-1}(\vec{x}_j - \vec{\partial}_j). \tag{XII.50b}$$

References

1. V. Volterra, Theory of functionals and of integral and integro-differential equations (Dover Publications, Inc., New York, 1959).
2. K. O. Friedrichs, Mathematical aspects of the quantum theory of fields (Interscience Publishers, Inc., New York, 1953), Section II.
3. H. Goldstein, Classical mechanics (Addison-Wesley Publishing Co., Inc., Reading, Mass., 1950), Chapter 2.
4. R. Courant and D. Hilbert, Methods of mathematical physics, Vol. I (Interscience Publishers, Inc., New York, 1953), Chapter IV.
5. F. Riesz and B. Sz.-Nagy, Functional analysis (F. Ungar Publishing Co., New York, 1955).
6. N. I. Muskhelishvili, Singular integral equations (P. Noordhoff N. V., Groningen, Holland, 1953).
7. L. Schwartz, Théorie des distributions (Hermann and Cie., Paris, 1957-59, Vol. I-II).
8. E. Hille and R. S. Phillips, Functional analysis and semigroups (American Math. Society, Providence, R.I., Second Edition, 1957).
9. C. Lovelace, Commun. Math. Phys. 4, 261 (1967).

10. L. V. Kantorovich and G. P. Akilov, Functional analysis in normed spaces (Pergamon Press, New York, 1964), Chapters XVII and XVIII.
11. E. D. Rainville, Special functions (Macmillan Co., New York, 1960), Section 72.
12. S. S. Schweber, An introduction to relativistic quantum field theory (Harper & Row, New York, 1964), especially Chapter 18.
13. H. Araki, Ann. Phys. (New York) 11, 260 (1960).
14. K. Symanzik, in Lectures on high energy physics (at Hercegnovi, 1961), Vol. II, ed. by B. Jaksić (Federal Nuclear Energy Commission of Yugoslovia, 1962).
15. W. Zimmermann, Commun. math. Phys. 6, 161 (1967).
16. B. Zumino, "Quantum theory of fields" (lecture notes, New York University, 1960-61).
17. N. N. Bogolubov, The dynamical theory in statistical physics (Hindustan Publishing Corp., Delhi, 1965).
18. D. Ruelle, Ann. Phys. (New York) 25, 109 (1963).
19. I. P. Bazarov, Zh. Eksp. Teor. Fiz. U.S.S.R. 32, 1065 (1957) [transl.: Soviet Phys.-JETP 5, 872].
20. B. Zumino, Phys. Fluids 2, 20 (1959).
21. I. M. Gelfand and A. M. Yaglom, J. Math. Phys. 1, 48 (1960).
22. M. Kac, Probability and related topics in physical sciences (Interscience Publishers, Inc., New York, 1959), Chapter IV.
23. R. Courant, Differential and integral calculus, Vol. II (Blackie and Son, Ltd., London, 1936), p. 326.
24. K. Symanzik, J. Math. Phys. 6, 1155 (1965).
25. R. P. Feynman and A. R. Hibbs, Quantum mechanics and path integrals (McGraw-Hill Book Co., New York, 1965).
26. K. O. Friedrichs and H. N. Shapiro, Proc. Natl. Acad. Sci. (U.S.) 43, 336 (1957).
27. K. O. Friedrichs, H. N. Shapiro, et al., Integration of functionals (Courant Institute, New York University, seminar notes, 1957).
28. J. Tarski, lectures in Winter School in Karpacz, Poland, 1967 (to be published).
29. I. Segal. Bull. Amer. Math. Soc. 71, 419 (1965).
30. D. Shale, Trans. Amer. Math. Soc. 103, 149 (1962).
31. S. F. Edwards, in Analysis in function space, ed. by W. T. Martin and I. Segal (M.I.T. Press, Cambridge, Mass., 1964), p. 167 ff.
32. R. H. Cameron and W. T. Martin, Ann. of Math. 45, 386 (1944).
33. I. E. Segal, Trans. Amer. Math. Soc. 88, 12 (1958).
34. R. H. Cameron, Proc. Amer. Math. Soc. 2, 914 (1951).
35. I. M. Kovalchik, Russian Math. Surveys 18, No. 1, p. 97 ff (1963).

36. W. E. Brittin and W. R. Chappell in Lectures in Theoretical Physics, Vol. VIII-A, ed. by W. E. Brittin (University of Colorado Press, 1966), p. 101 ff.
37. T. Nakayama and J. M. Dawson, J. Math. Phys. 8, 553 (1967).
38. K. Symanzik, in Analysis in function space (see Reference 31), p. 197 ff; New York University report IMM-NYU 327 (1964).
39. N. N. Bogolubov and D. V. Shirkov, Introduction to the theory of quantized fields (Interscience Publishers, Inc., New York, 1959), Chapter VII.
40. E. Hopf, J. Ratl. Mech. Analysis 1, 87 (1952); E. Hopf and E. W. Tiff, ibid. 2, 587 (1953).
41. M. D. Donsker and J. L. Lions, Acta. Math. 108, 147 (1962).
42. S. F. Edwards and R. E. Peierls, Proc. Roy. Soc. (London) 224A, 24 (1954).
43. P. R. Halmos, Measure theory (D. van Nostrand Co., Inc., New York, 1950).
44. N. Wiener, A. Siegel, B. Rankin, and W. T. Martin, Differential space, quantum systems, and prediction (M.I.T. Press, Cambridge, Mass., 1966), Chapter 2.
45. G. E. Shilov, Russian Math. Surveys 18, No. 2, p. 99 ff (1963).
46. L. Gross, Mem. Amer. Math. Soc. No. 46 (1963).
47. W. Feller, An introduction to probability theory and its applications (John Wiley & Sons, Inc., New York, 1966), Vol. II, p. 540 ff.
48. G. E. Uhlenbeck and L. S. Ornstein, Phys. Rev. 36, 823 (1930).
49. M. Kac, G. E. Uhlenbeck, and P. C. Hemmer, J. Math. Phys. 4, 216, 229 (1963).
50. J. L. Doob, Ann. of Math. 43, 351 (1942).
51. A. N. Kolmogorov, Foundations of the theory of probability (Chelsea Publishing Co., New York, 1956).
52. V. Yu. Krylov, Doklady Akad. Nauk (U.S.S.R.) 132, 1254 (1960) [transl.: Soviet Math. 1, 760].
53. R. H. Cameron, J. Math. and Phys. 39, 126 (1960).
54. F. R. Gantmacher, Applications of the theory of matrices (Interscience Publishers, Inc., New York, 1959), Chapter IV.
55. E. Nelson, J. Math. Phys. 5, 332 (1964).
56. J. Schwinger, Phys. Rev. 82, 914 (1951); ibid. 91, 713 (1953).
57. C. Morette, Phys. Rev. 81, 848 (1951).
58. E. W. Montroll, Comm. Pure Appl. Math. 5, 415 (1952).
59. E. J. McShane, Bull. Amer. Math. Soc. 69, 597 (1963).
60. S. S. Schweber, J. Math. Phys. 3, 831 (1962).
61. Yu. V. Novozhilov and A. V. Tulub, The method of functionals in the quantum theory of fields (Gordon and Breach, Inc., New York, 1961).

62. F. A. Berezin, The method of second quantization (Academic Press, New York, 1966), Chapter I.
63. I. E. Segal, Ann. of Math. 63, 160 (1956).
64. J. R. Klauder, Ann. Phys. (New York) 11, 123 (1960).
65. I. E. Segal, Trans. Amer. Math. Soc. 81, 106 (1956).

LECTURES ON NONLINEAR RATE EQUATIONS, ESPECIALLY THOSE WITH QUADRATIC NONLINEARITIES[†‡]

Elliott W. Montroll
Department of Physics
University of Rochester
Rochester, New York

I. Introduction

This is a discussion of a number of rate equations which appear in a variety of disciplines, chemistry, physics, population dynamics, fluid dynamics, et cetera, and which have been selected because of their nonlinear character. In trying to develop some experience in the relation of nonlinear equations to qualitative process such as growth to saturation, competition, diffusion, and stability, it was decided to emphasize quadratic nonlinearities with the hope that, after enough examples are investigated, some features of a general theory will become evident.

While our main interest will be in partial differential equations and integral equations governing rate processes, the presentation starts with some examples which are ordinary differential equations. Some interesting ideas already appear in these cases, and in some solvable partial differential equations the solution finally evolves from solving ordinary differential equations which come from Fourier transforms or after other schemes are used.

A prototype of a nonlinear rate process of the class which will interest us is that which describes a chain of chemical reactions

$$A_i + A_j \rightarrow A_\ell \qquad i = 1,2,3 \ldots$$

where i, j, and ℓ run over an appropriate set of integers which are used to identify the various species. The appropriate rate equations might be written

---

[†]Presented at the THEORETICAL PHYSICS INSTITUTE, University of Colorado, Summer 1967.

[‡]This work was partially supported by the Air Force Office of Scientific Research Grant No. AF OSR 1314-67.

$$\frac{dc_j}{dt} = -k^{(j)}c_j + \sum_{\ell,m}{}' k^{(j)}_{\ell,m} c_\ell c_m - c_j \sum_{\ell} k^{(j)}_{\ell} c_\ell \qquad (1)$$

where c_j is the concentration of the j^{th} species. The first term on the right hand side represents the rate at which species A_j disappears without participating in the chain of reactions. The second term represents the growth of species j through the interaction of all species ℓ and m which might lead to the formation of j, and the third term is the loss of species j through the combination of species ℓ.

Once one has the confidence that equations such as (1) are appropriate to describe some process of interest, there are two problems. One is to find the appropriate rate constants and the other is to solve the nonlinear equations. The two are sometimes intimately connected because experimental data might be given in terms of some function of the solution of the equations so that without the solution the rate constants cannot be deduced from the data.

In a few cases the rate constants can be determined from first principles. For example, the Navier-Stokes equation for the flow of a viscous fluid is

$$u_t + (u\cdot\nabla)u = -\rho^{-1}\nabla p + \nu\nabla^2 u \qquad (2a)$$

where $u(r,t)$ is the velocity vector at point r at time t, p the pressure at space-time point, ρ the density and ν the viscosity. This equation is to be combined with the continuity equation which, in the case of an incompressible fluid, has the form

$$\nabla\cdot u = 0. \qquad (2b)$$

In the case of an incompressible fluid, the pressure term can be removed by taking the curl of (2a). The resulting equation has a quadratic nonlinearity and the viscosity is the only quantity which must be determined from an auxiliary experiment or calculated by statistical mechanics. In the final rate equation the appropriate "rate constants" are completely determined.

The lectures will be concerned with both the construction of the nonlinear equation to fit various situations, as well as with the solution of some of them. It should be emphasized at the beginning that it is seldom that the equations are derived from first principles. They are usually produced out of thin air to be used as a model of some phenomenon.

II. Some Ordinary Differential Equations of Population Growth and Demise

We trust that the physical scientist will allow us to start our mathematical discussion with ideas concerning population growth. Actually many of the nonlinear equations of physics describe the variation of occupation numbers in various energy levels associated with some system and the chemist is often interested in the population of various atomic or molecular species in the course of some reaction.

The pessimist, Malthus, in 1798 was concerned with the exponential population growth (and linear growth of resources) and, therefore, with the rate equation which describes such a growth

$$dn/dt = kn.$$

The more optimistic Belgian, Verhulst, introduced a "coefficient of retardation" which would limit the population growth to a saturation level N so that the growth rate of the population would be given by

$$\frac{dn}{dt} = kn(N - n) \tag{3}$$

which has the solution

$$n(t)/n(0) = N\left\{n(0) + [N-n(0)]\, e^{-kt}\right\}^{-1}. \tag{4}$$

In words, Eq. (3) states that the growth rate is proportional to the population and to the resources that remain for the future population.

The population growth of the United States is fitted by a logistic curve in Figure 1. Now Eq. (3) can be considered as an equation for the first moment in some random process which would be derived from the as yet unknown dynamics of human behavior. In experiments on the growth of bacteria, fruit flies, rats, et cetera, on limited media a good agreement with (4) is also obtained. However, there are occasional experiments in which the population starts to oscillate in time as saturation is reached. One can argue that the members of the population are not identical. As saturation is reached, it is harder for the individual to find sustenance. Aggressive members of the society are worried and stock pile resources; more timid members give up the fight. The population then declines and oscillation might start.

The more detailed equations of population growth should involve the competitive process between individuals or groups. A number of models of competition have been investigated by Volterra[1] and Lotka.[2] Volterra's work was motivated by observations made by D'Ancona in his statistical study of the Adriatic fisheries over the

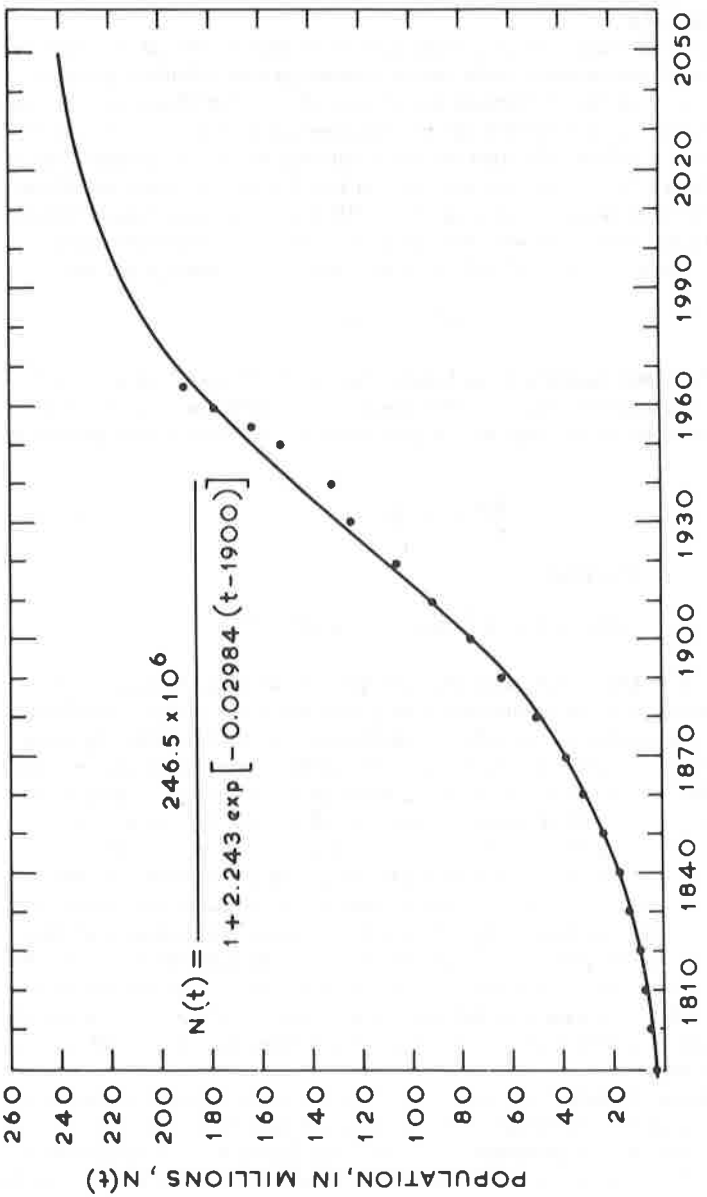


Figure 1. Population of U.S. logistic curve fitted so that observed points at 1840, 1900 and 1960 are exact. Points represent census data.

period 1905-1923. It seemed that members of a voracious species preyed on smaller fish to the extent that the survival of the larger depended on a plentiful supply of the smaller. As this supply diminished, the population of the larger species would diminish so that more of the smaller prey could grow and reproduce. Volterra characterized this process in the following manner.

If $N_1(t)$ is the population of the prey and $N_2(t)$ that of the predator, then this process can be described by the rate equations

$$dN_1/dt = k_1 N_1 - \lambda_1 N_2 N_1 \tag{5a}$$

$$dN_2/dt = \lambda_2 N_1 N_2 - k_2 N_2. \tag{5b}$$

In the absence of predator ($N_2 = 0$), the prey population would grow exponentially while, in the absence of prey ($N_1 = 0$), the predator population would decrease in an exponential manner with rate constant k_2. Since the eating process is the result of a binary encounter, the rate of loss of N_1 is proportional to $N_1 N_2$ as is the rate of increase of the predator.

For convenience we let

$$f_1(t) = \lambda_2 N_1(t)/k_2 \; ; \; f_2(t) = \lambda_1 N_2(t)/k_1. \tag{6}$$

Then

$$df_1/dt = k_1 f_1 (1-f_2) \tag{7a}$$

$$df_2/dt = k_2 f_2 (1-f_1) \tag{7b}$$

so that

$$dt = \frac{df_1}{k_1 f_1 (1-f_2)} = \frac{-df_2}{k_2 f_2 (1-f_1)} \tag{8}$$

from which we see that

$$\left(f_1 e^{-f_1}\right)^{1/k_1} \left(f_2 e^{-f_2}\right)^{1/k_2} = \text{constant}, \tag{9}$$

the value of the constant depending on the initial conditions.

The periodic character of the solution of Volterra's equations can easily be seen in terms of the relation between f_1 and f_2 according to (9) through the aid of the four diagrams in Figure 2.

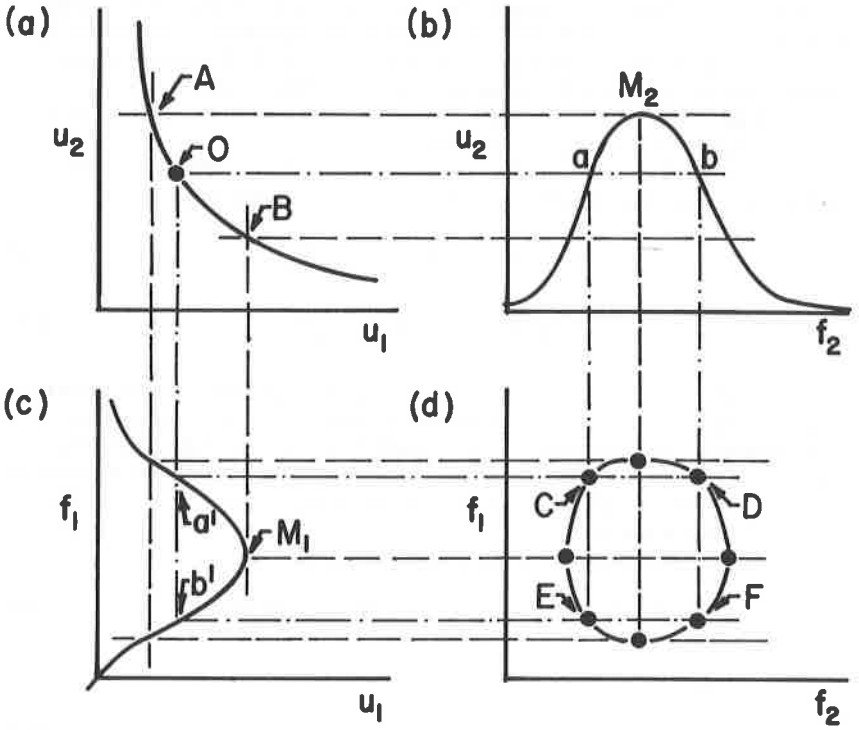


Figure 2. Geometric scheme used to prove the periodicity of the solution (10) of Volterra's equations (5). The discussion of the scheme is given in the text.

Let

$$u_j = \left(f_j \exp -f_j\right)^{1/k_j} \quad \text{with} \quad j = 1,2. \tag{10}$$

Then (9) is equivalent to the hyperbola

$$u_1 u_2 = \text{constant} \tag{11}$$

which is plotted in Figure 2a. Figures 2b and 2c show the behavior of u_j plotted as a function of f_j. An important feature of Eq. (10) and these two figures is that u_1 and u_2 attain maximum values which are identified by M_1 and M_2 in the figures. Hence the relevant region of the hyperbola (11) in Figure 2a is bounded by points A and B in the

figure. Note that a typical point 0 between A and B corresponds to two values of f_2 (a and b), and to two values of f_1, a' and b'. Hence, on the graph (2d) which relates f_1 and f_2, it is clear that the point 0 corresponds to the four points C, D, E, F. As one goes from A to B in (2a), one traces out the closed curve in (2d). The end points A and B correspond respectively to extrema in f_2 and f_1, respectively, on (2d).

A family of closed curves exists, each member of which is related to a possible set of initial values $f_1(0)$ and $f_2(0)$. The points on the closed curve can be identified with the time by integrating (8):

$$t = \int_{f_1(0)}^{f_1(t)} df_1/k_1 f_1(1-f_2). \tag{12}$$

Since f_2 is given as a function of f_1 on the curve (3d), one could start at the initial values $(f_1(0), f_2(0))$, integrate numerically along the curve a short distance to a pre-chosen point $f_1(t)$ and, from the value of t determined by (12), one would identify the time appropriate for that value $f_1(t)$. The period would be determined as the time required to return to the starting point $(f_1(0), f_2(0))$.

Lanchester's[3] equations of combat between two opposing forces

$$dN_1/dt = -\alpha_1 N_2 - \lambda_1 N_1 N_2 \tag{13a}$$

$$dN_2/dt = -\alpha_2 N_1 - \lambda_2 N_1 N_2 \tag{13b}$$

are similar to the Volterra equations except that the strength of both forces [$N_i(t)$ being the strength of the i^{th} force] diminishes with time. Lanchester emphasized two cases. When each side is visible to the other and every man on each side is able to fire on any opposing individual, the loss rate is proportional to the number of individuals firing. Hence, if reinforcements arrive at rates $P_j(t)$, these equations reduce to

$$dN_1/dt = -\alpha_1 N_2 + P_1(t); \quad dN_2/dt = -\alpha_2 N_1 + P_2(t). \tag{14}$$

Some examples of this case are aerial dogfights, naval battles and open land battles of the classical type. When each side is invisible to the other and each fires into the area occupied by the other, the loss rate on one side is proportional to the number of men on the other and to the number of men in the area under fire.

$$dN_1/dt = -\lambda_1 N_1 N_2 \quad \text{and} \quad dN_2/dt = -\lambda_2 N_1 N_2. \tag{15}$$

Artillery attacks are examples of this type of combat.

Mixed cases can also be described in which part of the force is hidden and the remainder is in the open. The ambush situation in which one side is hidden and administers an attack from ambush to the other who is in the open leads to the equations

$$dN_1/dt = -\lambda_1 N_1 N_2 \quad \text{and} \quad dN_2/dt = -\alpha_2 N_1. \tag{16}$$

The relationship between N_1 and N_2 for these various combat models can be obtained by equating various expressions for dt. For example, in the first case in the absence of reinforcements

$$dN_1/\alpha_1 N_2 = dN_2/\alpha_2 N_1 \tag{17}$$

so that

$$\alpha_2 N_1^2 - \alpha_1 N_2^2 = \alpha_2 N_{10}^2 - \alpha_1 N_{20}^2 = \text{constant}, \tag{18}$$

N_{jo} being the initial strength of the j^{th} combat force. This equation is represented in Figure 3a. It is to be noted that force 2 is annihilated when $\alpha_2 N_{10}^2 > \alpha_1 N_{20}^2$.

In Lanchester's second case characterized by (15), the population of the two forces are linearly related:

$$\lambda_1 N_2 - \lambda_2 N_1 = \lambda_1 N_{20} - \lambda_2 N_{10} = \Delta = \text{constant} \tag{19}$$

which is plotted in Figure 3b.

The mixed case (16) has been discussed by S. Y. Deitchman.[4] A detailed analysis of the Battle of Iwo Jima has been made by J. Engel.[5]

Autocatalytic chemical reactions[20),21)] are also characterized by equations such as (5). Let an infinite reservoir of material S_o exist and suppose that it decomposes slowly into S_1 with S_1 being a catalyst necessary for its own formation. Let us also suppose that S_1 is connected to S_2 with S_2 also being the catalyst required for its own formation. Finally, suppose that S_2 is withdrawn from the system at a rate proportional to its concentration. Then variation in the concentration of S_1 and S_2 is governed by (5).

A good review of the topological theory (i.e., stability, periodicity, et cetera) of equations in two variables

$$\dot{x} = P(x,y) \quad , \quad \dot{y} = Q(x,y) \, ,$$

of which the above are examples, is given in Reference 6. L. Markus[7] has developed an algebraic theory for the case of $P(x,y)$ and $Q(x,y)$ being quadratic forms and has sketched extensions to more than two variables.

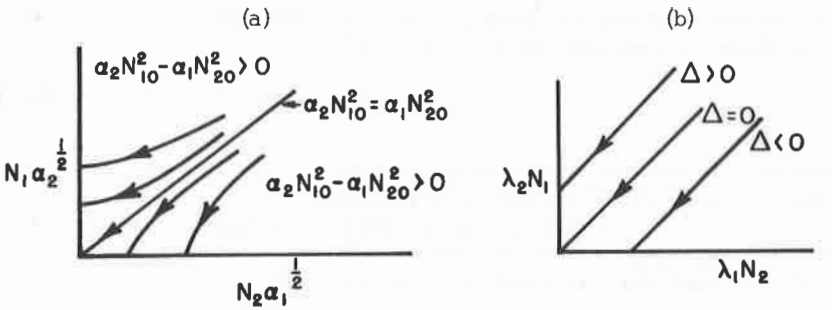


Figure 3. Schematic solutions of Lanchester's equation of combat. The numbers N_1 and N_2 represent the strength of forces on each side; the α's and λ's are parameters which depend on firepower, accuracy of firing, et cetera. Case (a) corresponds to the situation in which each side is invisible to the other and the firing is random into a region. Case (b) corresponds to the situation in which each side is visible to the other and each man on one side can fire on any opposing individual.

III. Some Examples of Partial Differential Equations with Quadratic Nonlinearities

1. Population Growth and Diffusion.

In this section we list a number of typical nonlinear partial differential equations which correspond to a variety of physical situations.

We start with a generalization of the growth equation (3). One way a growing population can fight the saturation which is inherent in (3) is to diffuse into neighboring unoccupied space. The equation which describes the combination of population growth and diffusion was first proposed by R. A. Fisher[8] (see, also, Kolmogorov, Petrovsky, and Piscounoff[9]).

$$\frac{\partial n}{\partial t} = D\frac{\partial^2 n}{\partial x^2} + kn(\theta - n)/\theta, \tag{20}$$

θ representing the saturation population per unit length. This equation was introduced in the context of a virile mutant which appears in a population and finally saturates it. Of course, the two-dimensional and three-dimensional generalization of (20) are

$$\frac{\partial n}{\partial t} = D\nabla^2 n + kn(\theta - n)/\theta \tag{20a}$$

where θ is now the saturation population per unit area (or volume in the 3D case) and D the diffusion constant.

If the two competing species whose population variation develops according to Volterra's equation (5) are allowed to roam, a diffusion term should be added to (5a) and (5b)

$$\frac{\partial N_1}{\partial t} = D_1\nabla^2 N_1 + k_1 N_1 - \lambda_1 N_1 N_2 \tag{21a}$$

$$\frac{\partial N_2}{\partial t} = D_2\nabla^2 N_2 + \lambda_2 N_1 N_2 - k_2 N_2. \tag{21b}$$

2. Clannish Random Walkers.

An equation similar to (20) can be derived for the motion of two interacting populations which tend to be clannish and wish to live near those of their own kind. We discuss only the one-dimensional case here.

Consider two species A and B of random walkers whose walks are limited to a line. Let ρ be a uniform stationary distribution of walkers (of either species) on the line. The line might be divided into a set of equal length segments, each of length a. The number of walkers per segment is then $N = \rho a$. If τ is the average time required per step of length a, we define $N(ja; \nu\tau)$ as the number of walkers of species A on the j^{th} segment at a time $\nu\tau$ after the walk begins and $f(ja; \nu\tau)$ as the fraction of walkers on the j^{th} segment which are A's.

We postulate the "home" of the A's to be segments of large positive j and that of the B's to correspond to large negative j's. The clannishness is characterized by a parameter ψ, which on our segmented line we write as

$$\psi = \alpha a/\tau. \tag{22}$$

We choose α so that the probability of an A walker at segment k stepping to segment $k \pm 1$ is

$$\frac{1}{2}\left\{1 \pm \alpha\left[1 - N^{-1}N(ka;\nu\tau)\right]\right\}.$$

A walker who finds himself completely surrounded by members of his own group [such that $N(ka;\nu\tau)=N$] steps with equal probability in either direction. One in the midst of strangers steps toward his home region with a probability $\frac{1}{2}(1+\alpha)$ and in the opposite direction with probability $\frac{1}{2}(1-\alpha)$. The number of A's on the k^{th} segment after $(\nu+1)$ steps is

$$N\left(ka;[\nu+1]\tau\right) = \tfrac{1}{2}N\left([k-1]a;\nu\tau\right)\left\{1+\alpha\left[1-N^{-1}N([k-1]a;\nu\tau)\right]\right\} + \tfrac{1}{2}N\left([k+1]a;\nu\tau\right)\left\{1-\alpha\left[1-N^{-1}N([k+1]a;\nu\tau)\right]\right\}, \tag{23}$$

$$\left\{N\left(ka;[\nu+1]\tau\right) - N(ka;\nu\tau)\right\}\Big/\tau = \left(a^2/2\tau\right)\left\{N\left([k-1]a;\nu\tau\right) - 2N(ka;\nu\tau) + N\left([k+1]a;\nu\tau\right)\right\}\Big/a^2 - (\alpha a/\tau)\left\{N\left([k+1]a;\nu\tau\right)\left(1 - N^{-1}N\left([k+1]a;\nu\tau\right)\right) - N\left([k-1]a;\nu\tau\right)\left(1 - N^{-1}N([k-1]a;\nu\tau)\right)\right\}\Big/2aN.$$

This equation becomes a differential equation in the limit $\tau = dt \to 0$, $a = dx \to 0$, $\alpha \to 0$ if the limit is taken so that D and ψ defined by

$$D = a^2/2\tau \qquad \text{and} \qquad \psi = \alpha a/\tau \tag{24}$$

are finite. It is convenient to introduce the density of A's at x and T

$$f(x,T) = N^{-1}N(x,T) \tag{25}$$

to obtain the nonlinear partial differential equation

$$\frac{\partial f}{\partial T} = \frac{\partial}{\partial x}\left\{D\,\frac{\partial f}{\partial x} + \psi f(f-1)\right\}. \tag{26}$$

A symmetry exists between the walk of species A and that of B. The density of B's, 1-f, can easily be shown to satisfy exactly the same equation as f provided that x is replaced by -x.

This equation can be reduced to one without parameters by letting

$$t = T\psi^2/D, \quad y = x\psi/D, \quad \text{and} \quad g = f - \tfrac{1}{2}. \tag{27}$$

Then

$$\partial g/\partial t = \partial/\partial y\left[\partial g/\partial y + g^2\right]. \tag{28}$$

3. Separation Cascades.

A physical application of (26) is in the theory of separation cascades.[10)-12)] Such a cascade is sketched schematically in Figure 4. A mixture of two materials A and B is separated by an

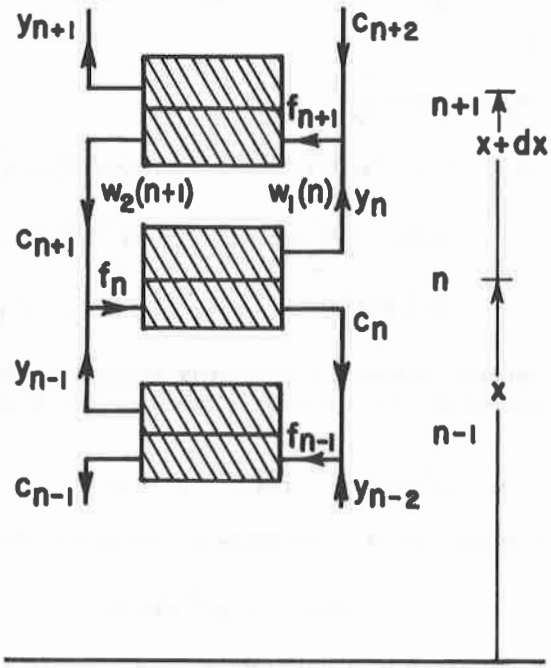


Figure 4. Schematic flow chart for cascade separation process. The input mixture of concentration f_n of component A flows into the n-th black box and is separated into two streams; the enriched one of concentration y_n of A flows upward in the cascade and the depleted one of concentration c_n in A flows downward. The w's represent flow rates in the directions indicated.

operation performed in the black boxes of the figure. If, for example, the mixture is fed into the n^{th} black box through the feed inlet f_n, two streams of material are emitted from the black box, one somewhat enriched in A in concentration in y_n, the other somewhat depleted in A in concentration c_n. When components A and B are very similar, as is the case in isotopic mixtures, one stage of separation does not do very much so that a cascade of separators must be employed. In the cascade of Figure 4, the depleted material from the (n+1)-st separator is combined with the enriched material from the (n-1)-st to feed into the n^{th} one. I believe the first cascade separation of gases was done by Lord Rayleigh employing a triangular array rather than a linear set of separators to separate argon from nitrogen in his investigation of the gases of the atmosphere. His elementary separation process was the passage of the gas through dried clay or rubber membranes, his clay membranes being flower pots and clay tobacco pipes and the lighter gas component passing more easily through the membrane. The membranes used in isotope separation processes at Oak Ridge and elsewhere are undoubtedly somewhat more sophisticated.

We will use Rayleigh's separation formula to relate c_n to y_n:

$$\left[y_n/(1-y_n)\right]\Big/\left[c_n/(1-c_n)\right] = \exp \Psi. \tag{29a}$$

When the separation factor Ψ is very small, we find

$$y_n = c_n + \Psi c_n(1-c_n) + 0(\Psi^2). \tag{29b}$$

We also let $w_1(n)$ represent the flow rate upward from the n^{th} stage to the (n+1)-st, and, $w_2(n)$ that of the depleted stream from the n^{th} stage downward to the (n-1)-st. If material is withdrawn at a rate P from the top of the cascade, then the net flow upward past each stage is also P so that

$$P = w_1(n) - w_2(n+1). \tag{30}$$

The rate of accumulation of component A at the n^{th} stage is then

$$H_n \partial c_n/\partial T = y_{n-1}w_1(n-1) + c_{n+1}w_2(n+1) - c_n w_2(n) - y_n w_1(n). \tag{31}$$

We limit ourselves to the special case in which all flow rates w_1 and w_2 and all stage holdups H_n are independent of n and in which the separation factor Ψ is very small so that many stages are required for the separation process. The length of a stage becomes short compared to the total length of the cascade. If the n^{th} stage is at point

x, the (n+1)'st is at x+a and we let

$$(c_n - c_{n-1})/a \to \partial c(x,T)/\partial x \qquad \text{as} \quad a \to 0.$$

Then, after substituting (29) and (30) into (31) with $w_1(n) \equiv w_1$ for all n, we find

$$\partial c/\partial T = D(\partial/\partial x)\left\{\partial c/\partial x - \psi' c(1-c) - pc\right\} \tag{32a}$$

where

$$D^{-1} = H/w_1 a^2, \quad \psi = \psi'/a \quad \text{and} \quad p = P/aw_1. \tag{32b}$$

This can be transformed into (26) by making the substitutions

$$c = f(p+\psi')/\psi' \quad \text{and} \quad \psi = D(\psi' + p). \tag{33}$$

Equation (26) is also valid for continuous separation processes such as thermal diffusion and distillation processes. Similar equations exist for centrifugal separation and chromatographic analysis.

4. Equation of Burger's Model of Turbulence and Those of Molecular Recombination.

The main equation of Burger's model of turbulence[14] also has the form (28) He considers the one-dimensional version of the Navier-Stokes equation (2a) after omitting the pressure term and the continuity equation (2b):

$$u_t + uu_x = \nu u_{xx} \tag{34}$$

which can be transformed into (28) by the substitutions

$$x = y\nu^{\frac{1}{2}} \quad \text{and} \quad u = -2g\nu^{\frac{1}{2}}. \tag{35}$$

The growth equation (20) can be turned into an equation for spreading of particles by diffusion and loss by recombination. Let $\alpha = k/\theta$ and take the limits $k \to 0$ and $\theta \to 0$ in such a manner that the ratio remains α. Then (20a) becomes[15]

$$\frac{\partial n}{\partial t} = D\nabla^2 n - \alpha n^2. \tag{36}$$

An example of the recombination processes is the recombination of free radicals. If an electric arc is passed through a gas such as O_2

or N_2, the free radicals 0 or N are formed which diffuse away from the arc and which also recombine by collision. Equation (36) gives a description of the process if $n(r,t)$ represents the free radical density at point r at time t.

5. Introduction of Mobility into Lanchester's Equations of Combat.

Mobility can also be introduced into Lanchester's equations (13a) and (13b) by adding a diffusion term. The existence of specified points that are to be attacked and defended can be treated through attractions and repulsions from fixed points or regions. A form of generalization of (13) is then

$$\partial N_1/\partial t = \text{div } D_1\left[\text{grad } N_1 + \beta_1 N_1 \text{ grad } U_1\right] - \alpha_1 N_2 - \lambda_1 N_1 N_2 + P_1(t,r) \tag{37a}$$

$$\partial N_2/\partial t = \text{div } D_2\left[\text{grad } N_2 + \beta_2 N_2 \text{ grad } U_2\right] - \alpha_2 N_1 - \lambda_2 N_1 N_2 + P_2(t,r). \tag{37b}$$

The diffusion "constants" may vary with position of population in a more complicated model. For ground or naval warfare, these equations would have two space variables. Military strategy is the proper choice of the quantities $\beta_1 \text{grad } U_1$ and $\beta_2 \text{grad } U_2$ and choice of the reinforcement functions $P_j(t,r)$.

6. Two-Dimensional Navier-Stokes Equation.

The last example of a nonlinear partial differential equation which we shall exhibit in this section is the two-dimensional version of the Navier-Stokes equation (2). The velocity has two components

$$u = (u_1, u_2) \tag{38a}$$

which satisfy three equations derivable from Eq. (2):

$$\partial u_1/\partial t + u_1 \partial u_1/\partial x_1 + u_2 \partial u_1/\partial x_2 = -\rho^{-1} \partial p/\partial x_1 + \nu \nabla^2 u_1 + \mathcal{F}_1(r,t) \tag{38b}$$

$$\partial u_2/\partial t + u_1 \partial u_2/\partial x_1 + u_2 \partial u_2/\partial x_2 = -\rho^{-1} \partial p/\partial x_2 + \nu \nabla^2 u_2 + \mathcal{F}_2(r,t) \tag{38c}$$

$$\partial u_1/\partial x_1 + \partial u_2/\partial x_2 = 0 \tag{38d}$$

where $\mathcal{F}_i$ is a driving force acting in the i^{th} direction.

The pressure term can be eliminated by taking the derivative of the first equation with respect to x_2 and the second with respect to x_1 and subtracting the second resulting equation from the first. These equations can be simplified through the introduction of the scalar potential ϕ which is related to the flow velocities by

$$u_1 = -\partial\phi/\partial x_2 \quad \text{and} \quad u_2 = \partial\phi/\partial x_1 . \tag{39}$$

Then

$$-\frac{\partial}{\partial t}\nabla^2\phi + \nu\nabla^2\left(\nabla^2\phi\right) = \begin{vmatrix} \frac{\partial\nabla^2\phi}{\partial x_1} & \frac{\partial\nabla^2\phi}{\partial x_2} \\ \frac{\partial\phi}{\partial x_1} & \frac{\partial\phi}{\partial x_2} \end{vmatrix} + \frac{\partial \not{g}_1}{\partial x_2} - \frac{\partial \not{g}_2}{\partial x_1} . \tag{40}$$

This equation can be put into a form which looks more like equations with which one has had experience by defining a new function f such that

$$\nabla^2\phi = f \tag{41}$$

which has the appearance of the Poisson equation of electromagnetic theory while (40) has the form

$$\frac{\partial f}{\partial t} = \nu\nabla^2 f + \frac{\partial\phi}{\partial x_2}\frac{\partial f}{\partial x_1} - \frac{\partial\phi}{\partial x_1}\frac{\partial f}{\partial x_2} + \frac{\partial \not{g}_1}{\partial x_2} - \frac{\partial \not{g}_2}{\partial x_1} . \tag{42}$$

When the viscosity ν is very large, one would expect the first term on the right hand side of (42) to dominate over the second so that in the absence of a driving force, the resulting equation would be the diffusion equation. Its solution would be substituted into (41) to find ϕ as the solution of "Poisson's equation."

IV. A Cheap Trick for Manufacturing Nonlinear Differential Equations with Built-In Solutions

1. Transformation of Dependent Variables.

One way of solving certain nonlinear equations is to construct them from linear equations. For example, let us generate several classes of nonlinear equations from the diffusion equation

$$f_t = Df_{xx} . \tag{43}$$

Let g(u) be an arbitrary function of a new dependent variable u whose

appropriate derivatives exist. Then let

$$f = g(u) \tag{44}$$

so that

$$f_t = g'(u)u_t, \quad f_x = g'(u)u_x \tag{45}$$

$$f_{xx} = g'(u)u_{xx} + g''(u)u_x^2 \tag{46}$$

where

$$g'(u) \equiv dg/du\,. \tag{47}$$

Then the diffusion equation becomes

$$u_t = Du_{xx} + Du_x^2 g''(u)/g'(u). \tag{48}$$

If we wished to solve the differential equation

$$u_t = Du_{xx} + u_x^2 F(u), \tag{49}$$

we would set

$$Dg''(u) = g'(u)F(u) \tag{50}$$

and solve for $g(u)$. For example, if $F(u) = \lambda$, then

$$f = g(u) = (D/\lambda)\exp(\lambda u/D) \tag{51}$$

so that

$$u = (D/\lambda)\log(\lambda f/D). \tag{52}$$

Generally,

$$f = g(u) = \int^u \exp\left\{\int^{u_1} D^{-1}F(u_2)du_2\right\} du_1\,. \tag{53}$$

We see then that the solution of the nonlinear equation with quadratic nonlinearity u_x^2:

$$u_t = Du_{xx} + \lambda u_x^2 \tag{54}$$

is given by (52) where $f(x,t)$ is the solution of the diffusion equation.

If the initial conditions of $u(x,t)$ are known in an unbounded space, then

$$f(x,0) = (D/\lambda)\exp\left\{\lambda u(x,0)/D\right\} \tag{55}$$

and

$$f(x,t) = \frac{1}{2(\pi tD)^{\frac{1}{2}}} \int_{-\infty}^{\infty} (D/\lambda)\exp\left[\lambda u(x',0)/D - (x-x')^2/4Dt\right]dx' \tag{56}$$

so that

$$u(x,t) = (D/\lambda)\log\left\{\frac{1}{2(\pi tD)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left[\lambda u(x',0)/D - (x-x')^2/4Dt\right] dx'\right\}. \tag{57}$$

2. Exponential Integral Transformation of Dependent Variable.

Another class of nonlinear equations can be generated from the diffusion equation by the transformation

$$f = \exp \int^x g(u)dx \tag{58}$$

where $g(u)$ is again an arbitrary function. Here

$$f_t = f\int^x u_t g'(u)dx \quad \text{and} \quad f_x = fg(u) \tag{59}$$

$$f_{xx} = fg'(u)u_x + fg^2(u). \tag{60}$$

Hence

$$\int^x u_t g'(u)dx = Dg'(u)u_x + g^2(u). \tag{61}$$

If this is differentiated with respect to x, one finds

$$u_t = D\left[u_{xx} + g''(u)u_x^2/g'(u)\right] + 2g(u)u_x. \tag{62}$$

As a special case let

$$g(u) = \tfrac{1}{2}u. \tag{63a}$$

Then we have the equation with quadratic nonlinearity

$$u_t = Du_{xx} + uu_x \tag{63b}$$

which is equivalent to the equation which described our friendly random walkers as well as the Burgers equation. The linearization of Burgers equation was first accomplished by E. Hopf[16] and J. Cole.[17] That of the thermal diffusion equation by Majumdar[18] and of separation cascades with production by Montroll and Newell.[12]

The solution of this equation can be expressed in terms of f, the solution of the diffusion equation by observing that (58) implies

$$u = 2(\log f)_x. \tag{64}$$

Consider first the case of an infinite range of x and an initial relation between u and x:

$$u(x,0) = u_o(x). \tag{64a}$$

Then the initial dependence of f on x is

$$f(x,0) = \exp \int_{x_o}^{x} u_o(x')dx'. \tag{64b}$$

Since f(x,t) is a solution of the diffusion equation, on our unbounded line

$$f(x,t) = \frac{1}{2(\pi tD)^{\frac{1}{2}}} \int_{-\infty}^{\infty} f(x'',0)\left[\exp - (x-x'')^2/4Dt\right]dx'' \tag{65a}$$

$$= \frac{1}{2(\pi tD)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left\{-(x-x'')^2/4Dt + \int_{x_o}^{x''} u_o(x')dx'\right\}dx'' \tag{65b}$$

and

$$u(x,t) = 2\frac{\partial}{\partial x}\left[\log\left(\frac{1}{2(\pi tD)^{\frac{1}{2}}} \int_{-\infty}^{\infty} \exp\left\{-(x-x'')^2/4Dt + \right.\right.\right.$$

$$\left.\left.\left. + \int_{x_o}^{x''} u_o(x')dx'\right\} dx''\right)\right]. \tag{66}$$

Notice that x_o can have any convenient value since, if it is changed from x_o to x_o', the change in the value of the integral is a quantity which is independent of x, say $I(x_o', x_o)$. It yields a factor

$\exp\{I(x_o', x_o)\}$ in the integrand of the x integral. After one takes the logarithm and differentiates with respect to x, one sees that $u(x,t)$ is independent of $I(x_o', x_o)$ and, therefore, of x.

3. Population Growth and Diffusion.

Still another class of equations follows from transforming the diffusion equation with[19]

$$f = e^{-kt} g(u) \tag{67}$$

so that

$$f_t = -ke^{-kt} g(u) + g'(u) e^{-kt} u_t \tag{67a}$$

$$f_{xx} = e^{-kt}\left\{g'(u)u_{xx} + g''(u)u_x^2\right\}. \tag{67b}$$

Then

$$u_t = D\left[u_{xx} + g''(u)u_x^2/g'(u)\right] + kg(u)/g'(u). \tag{67c}$$

If we wish the term proportional to k to have the form $F(u)$, then

$$u_t = D\left\{u_{xx} + \left[\frac{1-F'(u)}{F(u)}\right]u_x^2\right\} + kF(u) \tag{68a}$$

with

$$g(u) = \int^u du/F(u). \tag{68b}$$

By choosing[19]

$$F(u) = u(\theta-u)/\theta, \tag{69}$$

the right hand term becomes the same as Fisher's, but the full equation has the form

$$u_t = D\left\{u_{xx} + 2u_x^2/(\theta-u)\right\} + ku(\theta-u)/\theta \tag{70}$$

where $g(u)$ was chosen to be

$$g(u) = u/(\theta-u) \tag{71}$$

and u is related to the function f which satisfies the diffusion equation through

$$u(x,t) = \theta f(x,t)/\left[e^{-kt} + f(x,t)\right]. \tag{72}$$

It is interesting to note that while Eq. (70) differs from Fisher's equation by the term involving u_x^2, one might expect initial distributions of a certain type to develop in essentially the same way according to (70) as they would have according to the Fisher equation.

Suppose that initially $u(x,0) \ll \theta$ for all x. Then as $\theta \to \infty$ both the Fisher equation and ours have the form $u_t = Du_{xx} + ku$, which corresponds to unlimited growth. On the other hand, we know that as $t \to \infty$ the solution $u(x,t)$ (72) of our equation approaches saturation at all points, as would be the case with (20). Hence, at both early and late times (70) and (20) are equivalent under the initial condition given above. We will show at the end of this section that for certain initial conditions, the solution of (70) is practically the same as Fisher's equation.[19]

We now solve Eq. (70) subject to several initial conditions. If $u(x,0)$ is known, one can employ (72) to find $f(x,0)$ (the function $f(x,t)$ is, of course, the solution of the diffusion equation (43)).

The first initial distribution which we discuss is one in which saturation exists on the negative half line to within a distance of about $1/\beta$ from the origin and no mutant population exists at points beyond a distance of about $1/\beta$ to the right of the origin. A representation of such an initial distribution is

$$u(x,0) = \theta/(1+e^{\beta x}). \tag{73}$$

Then

$$f(x,0) = e^{-\beta x}. \tag{74}$$

If this is substituted in (65a) one finds that

$$f(x,t) = e^{-\beta(x-\beta Dt)} \tag{75}$$

and

$$u(x,t) = \theta/\left\{1 + \exp\beta[x - (\beta D + k/\beta)t]\right\}. \tag{76}$$

The rate at which the mutant front propagates is

$$dx/dt = (\beta D + k/\beta). \tag{77}$$

In this case in which saturation exists behind the mutant front, the shape of the diffusion front remains invariant as the wave propagates. It is interesting to note that the propagation rate is determined by the

initial slope at $x=0$. If the initial front is sharp (β large), the propagation is dominated by diffusion while, if it is broad (small β), it is dominated by saturation development. When both processes make the same contribution, the propagation velocity has a minimum value which is

$$(dx/dt)_{min} = (2kD)^{\frac{1}{2}} \tag{78}$$

and which corresponds to $\beta = (k/D)^{\frac{1}{2}}$. A similar result was noted by Fisher[8] in connection with Eq. (20). He found that no solution of the form $p=f(x+vt)$ could be found for (20) for which $|v| < (2kD)^{\frac{1}{2}}$.

As a second example we discuss an initial distribution of mutant which is peaked at the origin at a fraction η of the saturation level θ, and which extends over a length a. Then

$$u(x,0) = \eta\theta/\left[\eta + (1-\eta)\exp(x^2/2a^2)\right], \tag{79}$$

so that

$$f(x,0) = \left[\eta/(1-\eta)\right]\exp(-x^2/2a^2), \tag{80a}$$

and from (65a)

$$f(x,t) = \frac{\eta a \exp[-x^2/2(a^2+2Dt)]}{(1-\eta)(a^2+2Dt)^{\frac{1}{2}}}. \tag{80b}$$

Therefore, from (72)

$$u(x,t) = \theta\eta a\left[\eta a + (1-\eta)(a^2+2Dt)^{\frac{1}{2}}\exp\left\{-kt + x^2/2(a^2+2Dt)\right\}\right]^{-1}. \tag{81}$$

When $|x| \ll \{2kt(a^2+2Dt)\}^{\frac{1}{2}}$ the exponential term in the denominator can be neglected and one has saturation. The propagating population front at any time can be identified with the value of x which makes the argument of the exponential vanish, i.e.,

$$x = \pm\left\{2kt(a^2+2Dt)\right\}^{\frac{1}{2}} \tag{82a}$$

so that with increasing t the velocity of propagation approaches

$$dx/dt = \pm 2(kD)^{\frac{1}{2}} \tag{82b}$$

independently of a; i.e., of the dispersion of initial distribution.

This case differs from (77) in which a saturation region exists initially, for in that case the propagation velocity depends on the

initial shape of the wave while in the everywhere unsaturated case it becomes independent of the initial distribution. It is interesting to note that (82b) is just the minimum propagation velocity which can be achieved by an initial distribution which has a saturated region (see Eq. (78)).

The manner in which the initial distribution (79) propagates seems to be typical of that of any initial distribution of $h(x,0)$ whose Fourier transform exists. For further discussion of this, see Reference 19.

4. Using the Wrong Equation to Find the Right Solution.

Even when a transformation cannot be found to linearize a nonlinear equation, one can sometimes find another nonlinear equation which has an extra term which does not appear in the equation of interest, but which can be transformed to a linear one. This new equation might have the feature that, for certain initial conditions, the extra term is always very small so that the solution of the new equation under those initial conditions is in all space time a good approximation to the original one.

While no one has succeeded in finding a scheme for linearizing Fisher's equation, the solution of our more complicated Eq. (70) mimics that of Fisher's equation when the initial condition is such that no region is near saturation. We demonstrate this possibility through a detailed examination of the solution (81) which corresponds to the almost Gaussian initial condition (79).

Equation (81) will be substituted into the right hand side of (70) and the contribution of each term to the rate u_t of increase in population will be calculated. In all space-time regimes, our extra term $2Du_x^2/(\theta-u)$ will be small compared to the sum of the other two contributions to u_t.

First, at early times in the growth process before saturation develops, the linearized versions of both our equation and Fisher's are valid. However, the linearized versions are identical so, in the early stages of the population buildup, the solutions of both equations are identical.

Now let us consider the later stages of the process when $2Dt \gg a^2$. We examine the contribution of each of the three terms on the right hand side of (70) as determined from (81) to $\partial u/\partial t$, in the three regimes

$$x^2 - 4Dkt^2 \gg 4Dt; \quad |x^2 - 4Dkt^2| \ll 4Dt; \quad x^2 \ll 4Dt^2.$$

In all three regimes at all times

$$\text{(a)}\quad u_{xx} = -F\left\{\eta a\left[1+\frac{x^2}{a^2+2Dt}\right] + (1-\eta)(a^2+2Dt)^{\frac{1}{2}} \times \left[1-\frac{x^2}{a^2+2Dt}\right]E\right\}, \tag{83a}$$

$$\text{(b)}\quad 2u_x^2/(\theta-u) = F\left\{2a\eta x^2/(a^2+2Dt)\right\}, \tag{83b}$$

$$\text{(c)}\quad ku(\theta-u)/\theta = F\left\{(a^2+2Dt)\left[\eta a + (1-\eta)(a^2+2Dt)^{\frac{1}{2}}E\right]k/D\right\}, \tag{83c}$$

where E and the common factor F are defined by

$$E \equiv \exp\left\{-kt + x^2/2(a^2+2Dt)\right\} \tag{83d}$$

and

$$F \equiv \frac{a\theta\eta(1-\eta)E}{(a^2+2Dt)^{\frac{1}{2}}}\left\{\eta a + (1-\eta)(a^2+2Dt)^{\frac{1}{2}}E\right\}^{-3}. \tag{83e}$$

Since we postulate $2Dt \gg a^2$, the a^2 can be neglected compared with $2Dt$ in all the formulae above. In the first regime with $x^2 - 4Dkt^2 \gg 4Dt$

$$E \sim \exp\left\{(x^2-4Dkt^2)/4Dt\right\} \tag{84}$$

is very large so the non-exponential terms in (a), (b) and (c) can all be neglected. Since our extra diffusion term (b) does not contain E, it is negligible compared with the regular diffusion term (a). The exponential term in (a) has a factor $x^2/4Dt^2$ to be compared with $(4Dt)^{3/2}k/D$ in (c). The definition of our regime then implies that the regular diffusion term is the most important contribution to $\partial u/\partial t$. Actually, since F is proportional to $E^{-2}/(2Dt)^2$, $\partial u/\partial t$ is very small in this regime.

As we move toward the center of the population wave with $|x^2-4Dt^2| \ll 4Dt$, $E \simeq 1$ so that

$$\text{(a)}\qquad u_{xx} \simeq F(1-\eta)(2t)^{3/2}kD^{1/2}, \tag{85a}$$

$$\text{(b)}\qquad 2u_x^2/(\theta-u) \simeq 4\eta aktF, \tag{85b}$$

$$\text{(c)}\qquad ku(\theta-u)/\theta \simeq (1-\eta)Fk(2t)^{3/2}D^{1/2}. \tag{85c}$$

Hence, in the regime where the action is, (a) and (c) contribute the same amount to $\partial u/\partial t$. Since both exceed our extra term by a factor $t^{\frac{1}{2}}$, our extra diffusion term becomes less and less important with increasing time.

Finally, let us examine the influence of our extra term in our saturation region behind the population wave. In this regime with $x^2 \ll 4Dkt^2$ the exponential term E is negligible. The main contribution comes from the growth term (c) which is of order $2F\eta aDt$ compared with $F\eta a$, the main term in (a), and $Fa\eta x^2/Dt$ in our extra diffusion term. However, all terms are small in the regime.

In conclusion, we see that in each of the regimes considered the extra diffusion term contributes very little to u_t compared with the effect of either normal diffusion or population generation. Hence, it seems that as long as a region is not initially saturated, the generation and diffusion of an initial distribution as described by our rate equation (70) should be essentially the same as that which would be derived by solving Fisher's Equation (20).

V. Integral Equations with Quadratic Nonlinearities

We now show how certain of the partial differential equations discussed in the last section can be converted to integral equations. Those differential equations which are solvable then correspond to integral equations which are also solvable.

Let us consider the two-dimensional analogue of (54)

$$u_t = D\nabla^2 u + \left(u_x^2 + u_y^2\right) \tag{86a}$$

whose solution

$$u(r,t) = \log f(r,t) \tag{86b}$$

is related through f to the solution of the 2D diffusion equation

$$f_t = D\nabla^2 f. \tag{86c}$$

If we define

$$u(r,t) = \int_{-\infty}^{\infty}\int U(k,t)e^{-ik\cdot r} d^2k, \tag{87}$$

then

$$u_x^2(r,t) = -\int_{-\infty}^{\infty}\int k_x' k_x'' U(k',t)U(k'',t)e^{-i(k''+k')\cdot r}\, d^2k''\, d^2k'$$

$$= -\int\int k_x'(k_x - k_x')U(k',t)U(k-k',t)e^{-ik\cdot r}\, d^2k'\, d^2k \qquad (88)$$

while

$$u_{xx}(r,t) = -\int\int k_x^2 U(k,t)e^{-ik\cdot r}\, d^2k. \qquad (89)$$

If we substitute these expressions and the corresponding y derivatives into (86a), we find

$$\int_{-\infty}^{\infty}\int e^{-ik\cdot r}\, d^2k\left\{U_t + Dk^2U + \int_{-\infty}^{\infty}\int k'\cdot(k-k')U(k,t)U(k-k',t)d^2k'\right\} = 0. \qquad (90)$$

Hence, our original differential equation (86a) is equivalent to the integral equation

$$\frac{\partial}{\partial t}U(k,t) = -Dk^2U(k,t) - \int_{-\infty}^{\infty}\int k'\cdot(k-k')U(k,t)U(k-k',t)d^2k'. \qquad (91)$$

The solution of the integral equation is then

$$U(k,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int\left[\log f(r,t)\right]\exp ik\cdot r\, d^2r \qquad (93)$$

where $f(r,t)$ is the solution of the diffusion equation (86c). The 1D and 3D generalizations of this equation are clear.

An equation similar to (91), but with a simpler kernel is that which occurs in Smoluchowski's theory of coagulation of colloids and Schumann's theory of cloud formation. In these theories it is assumed that two particles, one of mass k and the other of mass k', collide and form a new one of mass $(k+k')$. If $n(k)dk$ is the number of particles of mass between k and $k+dk$ at time t, then the rate equation, which describes the growth of particles, is

$$\frac{\partial n(k,t)}{\partial t} = -2\alpha n(k,t)N(t) + \alpha\int_0^k n(k-k',t)n(k',t)dk'. \qquad (94)$$

The first term on the right of (94) represents the rate at which particles of mass k are lost through collision and combination of other particles. The collision rate is proportional to the total number of particles

$$N(t) = \int_0^\infty n(k,t)dk \tag{95}$$

with which any one of mass k might collide. The factor 2 is included because either of the two particles which collide might be of mass k. The integral term on the right of (94) represents the gain in the number of particles of mass k due to collisions of those of mass k' and k-k'. The constant α is a rate constant for the process.

Instead of solving (94), we solve a more complicated equation from whose solution that of (94) can be deduced. Let k now be a vector, say a 2D one, and let $\alpha(t)$ be an arbitrary function of t. Then consider (with λ constant)

$$\frac{\partial n(k,t)}{\partial t} = -\alpha(t)n(k,t) + \lambda \int_{-\infty}^{\infty}\int n(k',t)n(k-k',t)d^2k'. \tag{96}$$

If initially $n(k,0)=0$ unless $k_x>0$ and $k_y>0$, then, through collisions, the resulting k's can only increase so that at all times $n(k,t)=0$ if $k_x>0$ and $k_y>0$. Under these circumstances, the integration limits in (96) could be taken over the range $(0,0;k_x,k_y)$. In view of the Faltung form of the integrand of (96), we multiply both sides of (96) by

$$\exp ik\cdot\theta \equiv \exp[i(k-k')\cdot\theta]\exp(ik'\cdot\theta).$$

Then

$$\frac{\partial N(\theta,t)}{\partial t} = -\alpha(t)N(\theta,t) + \lambda N^2(\theta,t) \tag{97}$$

where

$$N(\theta,t) = \int_{-\infty}^{\infty}\int n(k,t)e^{ik\cdot\theta}\,d^2k. \tag{98}$$

If we let

$$g(\theta,t) = 1/N(\theta,t), \tag{99a}$$

then

$$-dg/dt + \alpha(t)g = \lambda. \tag{99b}$$

Since

$$-\frac{d}{dt}\left\{g\exp\left[-\int_0^t\alpha(t')dt'\right]\right\}=\left\{-\frac{dg}{dt}+\alpha(t)g\right\}\exp\left[-\int_0^t\alpha(t')dt'\right]$$

$$g(\theta,t)=g(\theta,0)\exp\int_0^t\alpha(\tau)d\tau-\lambda\int_0^t\left[\exp\int_\tau^t\alpha(t')dt'\right]d\tau$$

so that finally

$$N(\theta,t)=\frac{N(\theta,0)}{\exp\int_0^t\alpha(\tau)d\tau-\lambda N(\theta,0)\int_0^t[\exp\int_\tau^t\alpha(t')dt']d\tau} \tag{100a}$$

and

$$n(k,t)=\frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int\frac{N(\theta,0)e^{-i\theta\cdot k}d^2k}{[\exp\int_0^t\alpha(\tau)d\tau]-\lambda N(\theta,0)\int_0^t[\exp\int_\tau^t\alpha(t')dt']d\tau}. \tag{100b}$$

The solution of the Smoluchowski equation is obtained by first integrating (94) with respect to k so that

$$\frac{dN(t)}{dt}=-2\alpha N^2(t)+\alpha\int_0^\infty dk\int_0^k n(k-k',t)n(k',t)dk'$$

$$=-2\alpha N^2(t)+\alpha\int_0^\infty n(k',t)dk'\int_{k'}^\infty n(k-k',t)dk=-\alpha N^2(t)$$

and

$$N(t)=N(0)/[1+\alpha tN(0)]. \tag{101}$$

Then (94) is the 1D equivalent of (96) if one sets

$$\alpha(t)=2\alpha N(0)/[1+\alpha tN(0)]\quad\text{and}\quad\lambda=\alpha. \tag{102}$$

From the 1D analogue of (100) we find the solution of Smoluchowski's equation to be

$$n(k,t)=\frac{1}{(2\pi)}\int_{-\infty}^{\infty}\frac{N(\theta,0)\exp(-i\theta k)dk}{[1+t\alpha N(0)]^2-\alpha tN(\theta,0)[1+t\alpha N(0)]}. \tag{103}$$

Several special initial size distributions are of interest. Schumann used the initial Poisson distribution in his theory of cloud formation[24),25)]

$$n(k,0) = \begin{cases} N(0)e^{-\beta k} & \text{if } k>0 \\ 0 & \text{if } k<0 \end{cases}. \tag{104a}$$

Then

$$N(\theta,0) = \beta N(0)/(\beta - i\theta) \tag{104b}$$

so that

$$n(k,t) = \frac{\beta N(0)}{2\pi[1+\alpha t N(0)]^2}\int_{-\infty}^{\infty}\frac{e^{i\theta k}d\theta}{\beta[1+\alpha t N(0)]^{-1}-i\theta} \tag{105a}$$

$$= \begin{cases} \dfrac{\beta N(0)\exp[-\beta k(1+\alpha t N(0)^{-1}]}{2\pi[1+\alpha t N(0)]^2} & \text{if } k>0 \\ 0 & \text{if } k<0 \end{cases}. \tag{105b}$$

Hence an initial Poisson distribution retains its Poisson character at all times. Similarly, one can show that the initial exponential distribution

$$n(k,0) = \tfrac{1}{2}\beta N(0)e^{-\beta|k|} \tag{106a}$$

also retains its character. One finds[25),26)]

$$n(k,t) = \frac{\beta N(0)}{2[1+\alpha t N(0)]^{3/2}}\exp[-\beta|k|/(1+\alpha t N(0))^{\frac{1}{2}}]. \tag{106b}$$

Of course, since particle sizes cannot be negative (106a) is not a possible initial distribution for the Smoluchowski problem. On the other hand, one can define a stochastic process which is characterized by the integral equation (94) such that variable k ranges from $-\infty$ to ∞.

It can also be shown that any initial distribution such that if $n(k,0)=n(-k,0)$, with the initial second moment[26)]

$$\mu_2 = [N(0)]^{-1}\int_{-\infty}^{\infty}k^2 n(k,0)dk \tag{107a}$$

then, as $t\to\infty$

$$n(k,t) \sim (\alpha t)^{-3/2}\left[2N(0)\mu_2\right]^{-\frac{1}{2}}\exp\left\{-|k|\left[\tfrac{1}{2}\alpha t N(0)\mu_2\right]^{-\frac{1}{2}}\right\}. \tag{107b}$$

An initially unsymmetrical distribution with first moment

$$\mu_1 = [N(0)]^{-1}\int_{-\infty}^{\infty} kn(k,0)dk \neq 0 \tag{108a}$$

approaches the Poisson distribution as $t\to\infty$:

$$n(k,t) \sim \begin{cases} [\alpha t(1+\alpha t N(0))]^{-1}\exp[-k/\alpha t\mu_1 N(0)] & \text{if } k>0 \\ 0 & \text{if } k<0. \end{cases} \tag{108b}$$

The momentum representation of Burger's equation has two equivalent forms

$$\partial U(k,t)/\partial t = -\nu k^2 U(k,t) + i\int_{-\infty}^{\infty}(k-k')U(k',t)U(k-k',t)dk' \tag{109a}$$

or

$$\partial U(k,t)/\partial t = -\nu k^2 U(k,t) + i\int_{-\infty}^{\infty} k'U(k't)U(k-k',t)dk'. \tag{109b}$$

The solution of these equations is known because it is essentially the Fourier transform of (66).

The momentum representation of the 2D Fisher equation is

$$\frac{\partial}{\partial t}U(k,t) = -(Dk^2-\lambda)U(k,t) - (\lambda/\theta)\int_{-\infty}^{\infty}\int U(k',t)U(k-k',t)d^2k'. \tag{110}$$

VI. 2D Integral Equation for Incompressible Flow

The Navier-Stokes equations for a 2D incompressible fluid were written in Eqs. (41) and (42) in the form

$$\frac{\partial f}{\partial t} = \nu\nabla^2 f + \frac{\partial\phi}{\partial x_1}\frac{\partial f}{\partial x_1} - \frac{\partial\phi}{\partial x_1}\frac{\partial f}{\partial x_2} + \frac{\partial f_1}{\partial x_2} - \frac{\partial f_2}{\partial x_1}. \tag{111a}$$

$$\nabla^2\phi = f. \tag{111b}$$

In the case of an unbounded fluid, these equations can be converted to an integral equation through the introduction of the Fourier transforms

$$F(k,t) = \int f(r,t)e^{ik\cdot r}\,d^2r \tag{112a}$$

$$\Phi(k,t) = \int \phi(r,t)e^{ik\cdot r}\,d^2r \tag{112b}$$

$$\mathfrak{F}(k,t) = \int \mathcal{f}(r,t)e^{ik\cdot r}\,d^2r \tag{112c}$$

so that

$$f(r,t) = \frac{1}{(2\pi)^2}\int F(k,t)e^{-ik\cdot r}\,d^2k \tag{113a}$$

$$\phi(r,t) = \frac{1}{(2\pi)^2}\int \phi(k,t)e^{-ik\cdot r}\,d^2k \tag{113b}$$

$$\mathcal{f}(r,t) = \frac{1}{(2\pi)^2}\int \mathfrak{F}(k,t)e^{-ik\cdot r}\,d^2k. \tag{113c}$$

Then (111b) is equivalent to

$$-k^2\Phi(k,t) = F(k,t), \tag{114a}$$

and, if we substitute (112) into (111a), we find

$$\int_{-\infty}^{\infty}\int e^{-ik\cdot r}\,d^2k\left\{\frac{\partial F(k)}{\partial t} + \nu k^2F(k) + \right.$$

$$\left. + \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int (k_2k_1' - k_1k_2')F(k')\,\Phi(k-k')d^2k' + i(\mathfrak{F}\times k)_z\right\} = 0$$

so that

$$\frac{\partial F(k)}{\partial t} + \nu k^2F(k) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int [k\times k']_z F(k')\Phi(k-k')d^2k' + i(k\times \mathfrak{F})_z. \tag{114b}$$

If we employ (114a) the resulting equation for the potential is

$$\frac{\partial}{\partial t}\Phi(k,t) = -\nu k^2\Phi(k,t) +$$

$$+ \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int (k'^2/k^2)[k\times k']_z\Phi(k',t)\,\Phi(k-k',t)d^2k' - i(k\times \mathfrak{F})_z/k^2 \tag{115}$$

which, in the absence of the driving force $\mathfrak{F}$, is a continuum example of the chemical rate equation discussed in the introduction. The "rate constants" are $(k'^2/k^2)[k\times k']_z/2\pi^2$.

In the case of a driven disturbance, it is useful to extend the definition of the various relevant Fourier transforms so that

$$F(k,\omega) = \int f(r,t)e^{i(k\cdot r-\omega t)}\,d^2r\,dt \tag{116}$$

$$\Phi(k,\omega) = \int \phi(r,t)e^{i(k\cdot r-\omega t)}\,d^2r\,dt \tag{117}$$

$$\mathfrak{F}(k,\omega) = \int \mathcal{f}(r,t)e^{i(k\cdot r-\omega t)}\,d^2r\,dt \tag{118}$$

with the appropriate inverses so that

$$f(r,t) = \frac{1}{(2\pi)^3}\int F(k,\omega)e^{-i(k\cdot r-\omega t)}\,d^2k\,d\omega, \text{ etc.} \tag{119}$$

Then the analogue of (115) is

$$\Phi(k,\omega) = -\frac{i(k\times\mathfrak{F})_z}{k^2(i\omega+\nu k^2)} + \frac{1}{(2\pi)^3}\int_{-\infty}^{\infty}\int\int \frac{[k\times k']_z k'^2}{k^2(i\omega+\nu k^2)}$$

$$\times\,\Phi(k',\omega')\,\Phi(k-k',\omega-\omega')d^2k'\,d\omega'. \tag{120}$$

VII. 2D Incompressible Fluid Driven by Two Fourier Exponentials

Let us now see if we can find some driving force for which we can solve this integral equation.

The simplest driven disturbance which we might consider is

$$\mathcal{f}(r,t) = A\exp i(t\omega_o - r\cdot q). \tag{121}$$

However, as we see from (41) and (42) the response of the scalar potential to such a disturbance is

$$\phi = iA\cdot(q_y - q_x)k^{-2}(i\omega_o + \nu q^2)^{-1} \exp i(t\omega_o - r\cdot q)$$

with

$$f = -\phi k^2.$$

Since this driven disturbance does not create a shear flow, the nonlinear term in (42) vanishes identically under its influence.

The simplest driving force which induces a shear flow is the sum of two Fourier exponentials

$$\mathcal{F}(r,t) = A \sum_{j=1}^{2} \alpha_j \exp i(t\omega_o - r\cdot q_j) \qquad (122a)$$

which, in (ω,k) space, yields the term

$$(k\times\mathcal{F})_z = (k\times A)_z\, \delta(\omega-\omega_o)\left[\alpha_1\delta(k-q_1) + \alpha_2\delta(k-q_2)\right]. \qquad (122b)$$

Through the nonlinearity in (116) we expect to be able to generate vectors of the form $\ell_1 q_1 + \ell_2 q_2$ by successive combinations of q_1 and q_2. The first combination of q_1 and q_2 is (q_1+q_2) which corresponds to the lattice point $(\ell_1,\ell_2)=(1,1)$ in Figure 5. Then, (q_1+q_2) can combine with q_1 to form $(2,1)$ and with q_2 to form $(1,2)$. By continuing this process the whole lattice of wave vectors (ℓ_1,ℓ_2) with $\ell_1 = 1, 2, \ldots$ and $\ell_2 = 1, 2, \ldots$ can be generated. The lattice

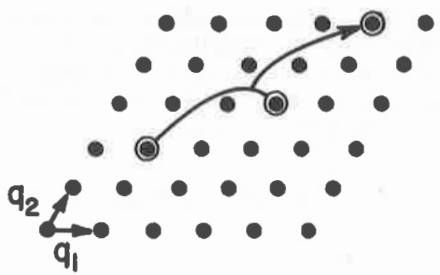


Figure 5. Lattice of wave numbers $(\ell_1 q_1, \ell_2 q_2)$ for excitation of a 2D fluid driven by two Fourier exponentials. Through the nonlinearity in the Navier-Stokes equation, any two wave vectors combine and form a new one which is the sum of the two.

points $(\ell,0)$ and $(0,\ell)$ with $\ell > 1$ do not correspond to possible wave vectors because of the term $[k \times k']_z$ in the integrand of (120); two vectors in the same direction cannot combine since

$$[\ell q_1 \times \ell' q_1] = 0.$$

The frequency of the wave vector $(\ell_1 q_1 + \ell_2 q_2)$ is

$$\omega = (\ell_1 + \ell_2)\omega_o. \tag{123}$$

Notice that a given wave vector $\ell_1 q_1 + \ell_2 q_2$ can only be obtained by combining pairs of wave vectors closer to the origin since no negative wave vectors are included in our driving force. Hence we can expect to be able to find the amplitude associated with a given wave vector in terms of those closer to the origin through a recursion formula. This will be done in detail in the next section.

Before proceeding with that analysis it should be pointed out that the driving force (122) does not lead to a real physical flow since contrary to the technique used in the investigation of linear equations we cannot associate the real part of the response potential of the fluid with the real part of the driving force. To obtain a physical flow we would have to use a driving force

$$(k \times \mathfrak{F})_z = (k \times A)_z \left\{ \delta(\omega - \omega_o)\left[\alpha_1 \delta(k - q_1) + \alpha_2 \delta(k - q_2)\right] \right.$$
$$\left. + \delta(\omega + \omega_o)\left[\alpha_1 \delta(k + q_1) + \alpha_2 \delta(k + q_2)\right]\right\}. \tag{124}$$

Such a driving force would yield a feedback mechanism such that one would no longer have a simple recurrence formula for the amplitude of the various wave numbers. We will make a few remarks about this more complicated case after our investigation of the response to (124).

On the basis of our observation of the formation of large wave vectors from smaller ones we seek a solution of (116) (with (122) inserted into the driving term) of the form

$$\Phi(k,\omega) = \sum_{\ell_1=0}^{\infty} \sum_{\ell_2=0}^{\infty} a(\ell_1,\ell_2)\, \delta(k - \ell_1 q_1 - \ell_2 q_2)\, \delta\left(\omega - [\ell_1 + \ell_2]\omega_o\right). \tag{125}$$

Then our problem will be to determine the amplitudes $a(\ell_1,\ell_2)$. Let us define

$$\beta_j = \alpha_j (q_j \times A)_z / q_j^2 \tag{126}$$

and

$$\lambda = \left(q_1 \times q_2\right)_z / (2\pi)^3 . \tag{127}$$

In order to reduce later complications we make the following abbreviations for a number of combinations of $\ell_1, \ell_2, q_1, q_2, \nu$ and ω_o will appear at every stage of our calculations.

$$f_{\ell_1 \ell_2} \equiv \left(\ell_1 q_1 + \ell_2 q_2\right)^2 \tag{128a}$$

$$g_{\ell_1, \ell_2} \equiv i\left(\ell_1 + \ell_2\right)\omega_o + \nu\left(\ell_1 q_1 + \ell_2 q_2\right)^2 \tag{128b}$$

$$h_{\ell_1, \ell_2} \equiv q_2^2 - \left[\ell_1 q_1 + (\ell_2 - 1) q_2\right]^2 = q_2^2 - f_{\ell_1, \ell_2 - 1} \tag{128c}$$

$$k_{\ell_1, \ell_2} \equiv q_1^2 - \left[(\ell_1 - 1) q_1 + \ell_2 q_2\right]^2 = q_1^2 - f_{\ell_1 - 1, \ell_2} . \tag{128d}$$

Then if we substitute (125) into our fundamental integral equation (120) with the driving term (122), we find

$$\begin{aligned}
&\sum_{\ell} a\left(\ell_1, \ell_2\right) \delta\left(k - \ell_1 q_1 - \ell_2 q_2\right) \delta\left(\omega - [\ell_1 + \ell_2]\omega_o\right) \\
&\quad = -i(k \times A)_z \left[\alpha_1 \delta(k - q_1) + \alpha_2 \delta(k - q_2)\right] / k^2 (i\omega + \nu k^2) \\
&\quad + \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} d^2 k' \, d\omega' \sum_{\ell_1', \ell_2'}' a\left(\ell_1', \ell_2'\right) \delta\left(k' - \ell_1' q_1 - \ell_2' q_2\right) \\
&\quad \times \delta\left(\omega' - [\ell_1' + \ell_2']\omega_o\right) a\left(\ell_1'', \ell_2''\right) \delta\left(k - k' - \ell_1'' q_1 - \ell_2'' q_2\right) \\
&\quad \times \delta\left(\omega - \omega' - [\ell_1'' + \ell_2'']\omega_o\right) (k \times k')_z k'^2 / k^2 (i\omega + \nu k^2) .
\end{aligned} \tag{129}$$

If we set $k = \ell_1 q_1 + \ell_2 q_2$ and $\omega = (\ell_1 + \ell_2)$ and equate the coefficients of $\delta(k - \ell_1 q_1 - \ell_2 q_2)$ on both sides of (129), we find, remembering the definitions of the β's and λ (Eqs. (126) and (127),

$$a(\ell_1,\ell_2) = -i\beta_1\delta_{\ell_1,1}\delta_{\ell_2,0}/g_{1,0} - i\beta_2\delta_{\ell_1,0}\delta_{\ell_2,0}/g_{0,1}$$

$$+\lambda \sum_{n_1=0}^{\ell_1}\sum_{n_2=0}^{\ell_2} a(n_1,n_2)a(\ell_1-n_1,\ell_2-n_2)$$

$$\times(\ell_1 n_2 - \ell_2 n_1)(n_1 q_1 + n_2 q_2)^2 / f_{\ell_1,\ell_1} g_{\ell_1,\ell_2}. \qquad (130)$$

We have used the fact that negative values of neither ℓ_1 nor ℓ_2 can be achieved from our initial conditions.

Certain scaling factors are immediately apparent from the quadratic form of (130). Suppose that (ℓ_1,ℓ_2) is neither $(1,0)$ nor $(0,1)$. Then it is clear that we can express $a(\ell_1,\ell_2)$ in the form

$$a(\ell_1,\ell_2) = [a(1,0)]^{\ell_1}[a(0,1)]^{\ell_2} / b(\ell_1,\ell_2) \qquad (131)$$

where

$$b(1,0) = b(0,1) = 1. \qquad (132)$$

For (ℓ_1,ℓ_2) vectors other than $(0,1)$ and $(1,0)$, $b(\ell_1,\ell_2)$ satisfies

$$b(\ell_1,\ell_2) = \lambda \sum_{n_1=0}^{\ell_1}\sum_{n_2=0}^{\ell_2} b(n_1,n_2)\, b(\ell_1-n_1,\ell_2-n_2)$$

$$\times(\ell_1 n_2 - \ell_2 n_1) f_{n_1 n_2} / f_{\ell_1,\ell_2} g_{\ell_1,\ell_2}. \qquad (133)$$

An alternative form for this equation can be obtained by making the change in variable $n_i' = \ell_i - n_i$ in the summation, then removing the prime and averaging the resulting expression with (133). Then

$$b(\ell_1,\ell_2) = \frac{1}{4} \sum_{n_1=0}^{\ell_1}\sum_{n_2=0}^{\ell_2} b(n_1,n_2) b(\ell_1-n_1,\ell_2-n_2)$$

$$\times[\ell_1(2n_2-\ell_2) - \ell_2(2n_1-\ell_1)][(2n_1-\ell_1)q_1 + (2n_2-\ell_2)q_2]$$

$$\times[(\ell_1 q_1 + \ell_2 q_2)] / g_{\ell_1,\ell_2} \qquad (134)$$

which has the vector form

$$b(\ell) = \frac{1}{4}\sum_{0}^{\ell} b(n)b(\ell-n)\left[\ell \times (2n-\ell)\right]_z \hat{\ell}\cdot(2\hat{n}-\hat{\ell})/\hat{\ell}\cdot\hat{\ell}\; g(\ell) \tag{135}$$

where

$$\hat{\ell} \equiv \ell_1 q_1 + \ell_2 q_2 \quad \text{and} \quad \ell \equiv (\ell_1, \ell_2) \quad \text{so that} \quad [\ell \times m]_z = \ell_1 m_2 - \ell_2 m_1 . \tag{136}$$

For later reference we note that

$$a(1,0) = -i\beta_1/g_{1,0} = -i\beta_1/(i\omega_o + \nu q_1^2) = -(\beta_1/\omega_o)/(1 - iR_1^{-1}) \tag{137a}$$

$$a(0,1) = -i\beta_1/g_{0,1} = -i\beta_2/(i\omega_o + \nu q_2^2) = -(\beta_2/\omega_o)/(1 - iR_2^{-1}) \tag{137b}$$

where R_j is the Reynolds number defined by

$$R_j = (\omega_o/|q_j|)(1/|q_j|)/\nu = \omega_o/\nu q_j^2 \tag{138}$$

when $\ell_1 = \ell_2 = 1$, (133) and (132) yield

$$b(1,1) = \left\{\frac{q_2^2 - q_1^2}{2i\omega_o + \nu(q_1+q_2)^2}\right\}(q_1 \times q_2)_z/(2\pi)^3 . \tag{139}$$

When $\ell_1 > 1$ and $\ell_2 > 1$, Eq. (133) can be written in a manner in which its being a recursion formula is clear. We must remember that $b(\ell,0)$ and $b(0,\ell)$ both vanish when $\ell > 1$. Then

$$b(\ell_1,\ell_2) =$$

$$= \lambda\left\{\ell_1 b(\ell_1,\ell_2-1)h_{\ell_1,\ell_2} - \ell_2 b(\ell_1-1,\ell_2)k_{\ell_1,\ell_2}\right\}\Big/ f_{\ell_1,\ell_2} g_{\ell_1,\ell_2}$$

$$+ \lambda \sum_{n_1=1}^{\ell_1-1} \sum_{n_2=1}^{\ell_2-1}$$

$$\times b(n_1,n_2)\, b(\ell_1-b_1,\ell_2-n_2)(n_1q_1+n_2q_2)^2(\ell_1n_2-\ell_2n_1)\Big/ f_{\ell_1,\ell_2} g_{\ell_1,\ell_2}. \tag{140}$$

This is a recurrence formula from which we can systematically find each $a(\ell_1,\ell_2)$ from those which are closer to the origin of our lattice of Figure 5.

A systematic scheme for the analytical solution of the recurrence formula is discussed by the lecturer in Reference 27. The numerical solution of the formula has been programmed by B. Parekh for an IBM 360-50 computer. Some qualitative idea of the results can be obtained from Figures 6a and 6b. The choices of parameters in these calculations are, respectively,

$$[a(1,0)]^4 = [a(0,1)]^4 = k = \begin{cases} 0.4 \times 10^{-1} & \text{case a} \\ 0.3 \times 10^{-1} & \text{case b,} \end{cases}$$

$$q_1^2 = 10, \quad q_2^2 = 20, \quad R_1 = \omega_o/\nu q_1^2 = 10, \quad \nu = 5\times10^{-5}\,\text{meter}^2/\text{sec},$$

$$\omega = R\nu q_1^2 = 5 \times 10^{-3}/\text{sec}$$

and the angle chosen between q_1 and q_2 is 90^o. The absolute values of the amplitudes $a(\ell_1,\ell_2)$ are exhibited in the figures through the lines of constant $|a(\ell_1,\ell_2)|$.

Certain topological features of these lines of constant $|a(\ell_1,\ell_2)|$ are immediately apparent and can be understood in terms of the manner in which various wave vectors combine to form new wave vectors. First we know that as $|\ell_1+\ell_2| \to \infty$ for any finite Reynolds number $|a(\ell_1,\ell_2)| \to 0$ since viscosity and dissipation into heat takes over at the very small eddy sizes, i.e., large wave numbers. Hence, if a minimum exists near the origin and $|a(\ell_1,\ell_2)|$, gets small for large $|a(\ell_1,\ell_2)|$, there must be at least one maximum

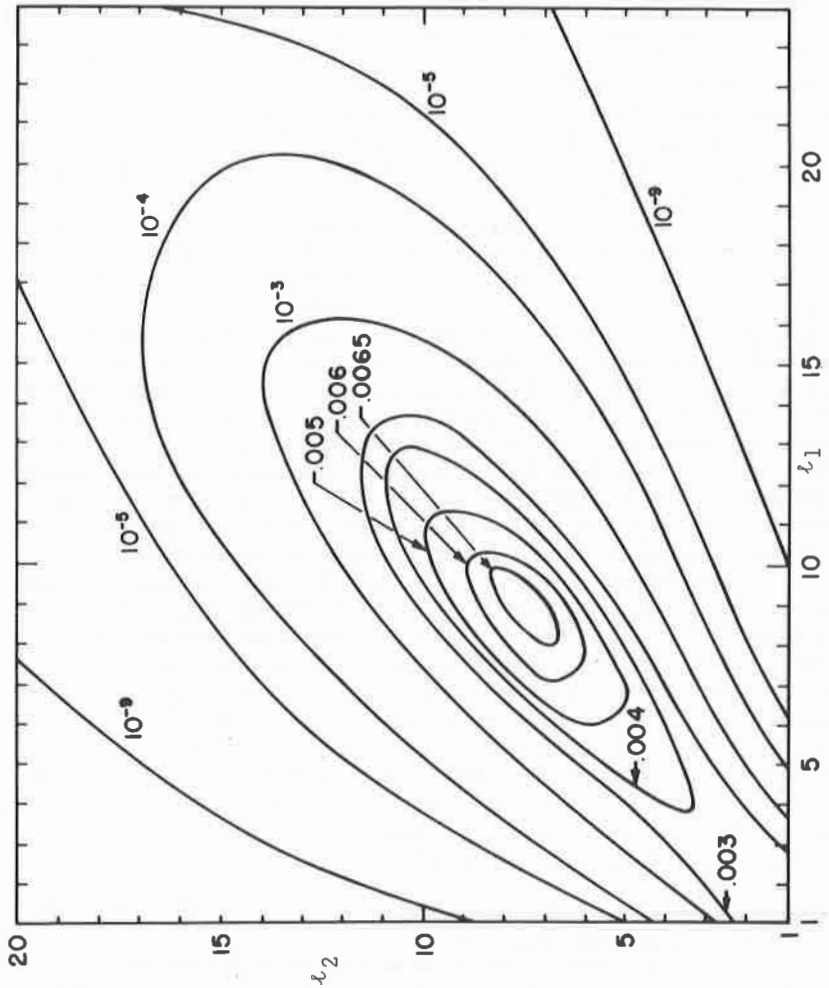


Figure 6a. An example of curves of constant amplitude associated with various wave numbers, choice of parameters given in text.

of $|a(\ell_1,\ell_2)|$ in some intermediate range. Such a maximum is clear in Figures 6a and 6b.

Now, in order to put a semblance of rigor into further remarks, we should consider $|a(q_1\ell_1+q_2\ell_2)|$ to be a continuous function of continuous variables $x_1=\ell_1|q_1|$ and $x_2=\ell_2|q_2|$. This would correspond to a physical situation in which the driving amplitudes α_1 both become very large while the lengths of the driving vectors $|q_1|$ and $|q_2|\to 0$ so that the lattice points become dense or, alternatively, the viscosity ν becomes very small (large Reynolds numbers) so that peaks occur at such great distances from the origin compared with $(|q_1+q_2|^2)^{\frac{1}{2}}$ that for all practical purposes our lattice can be considered to be a continuum.

As is evident from Eq. (140), the lattice point (ℓ_1,ℓ_2) can be occupied through the combination of any other pair of lattice points (n_1,n_2) and (ℓ_1-n_1,ℓ_2-n_2). Indeed, the weight $b(\ell_1,\ell_2)$ is a sum over all "marriages" of pairs of wave numbers which add up to (ℓ_1,ℓ_2). We can ask which "marriages" are most fruitful in their contribution to $b(\ell_1,\ell_2)$: Are they those involving almost like couples $(n_1,n_2)\simeq(\frac{1}{2}\ell_1,\frac{1}{2}\ell_2)$ or are they those of unlike couples such as $(2,\ell_2-2)$ and $(\ell_1-2,2)$? An analytical discussion of this question is given in Reference 27 where it is shown that when ℓ_1 and ℓ_2 are large, the greatest contribution comes from almost alike pairs (identical pairs do not contribute since, in that case, the cross-section factor $(\ell_1n_2-\ell_2n_1)$ in (140) would vanish if $n_2=\ell_2/2$ and $n_1=\ell_1/2$), with very dissimilar pairs contributing practically nothing.

This optimum pairing rule implies that if a first peak appears at (ℓ_1,ℓ_2), another subsidiary peak should appear at $(2\ell_1,2\ell_2)$, another at $(4\ell_1,4\ell_2)$, et cetera. The peaks more distant from the origin than the first may not be noticeable if the viscosity and dissipation cause such a rapid drop in the amplitude that, while the second peak exists, it is so small that it cannot be noticed. This is the case in the choice of the parameters used to calculate the curves in Figure 6b. The first peak in Figure 6b is near the point (11,11) and the second near (21,21) as expected. In Figure 6a, the peak at (8,7) is very flat and one does not yet seem to be in the "large" (ℓ_1,ℓ_2) regime in which our argument is valid.

VIII. Some Remarks on the Influence of Two Periodic Driving Forces on a 2D Incompressible Fluid

While we have not yet made any detailed calculations or analysis of the response to the physically interesting driving force (124), we can make a few conjectures about the nature of the lines of constant energy in our wave number space (k-space). In this case we must consider the amplitude $a(\ell_1,\ell_2)$ at all lattice points on the infinite lattice. Here

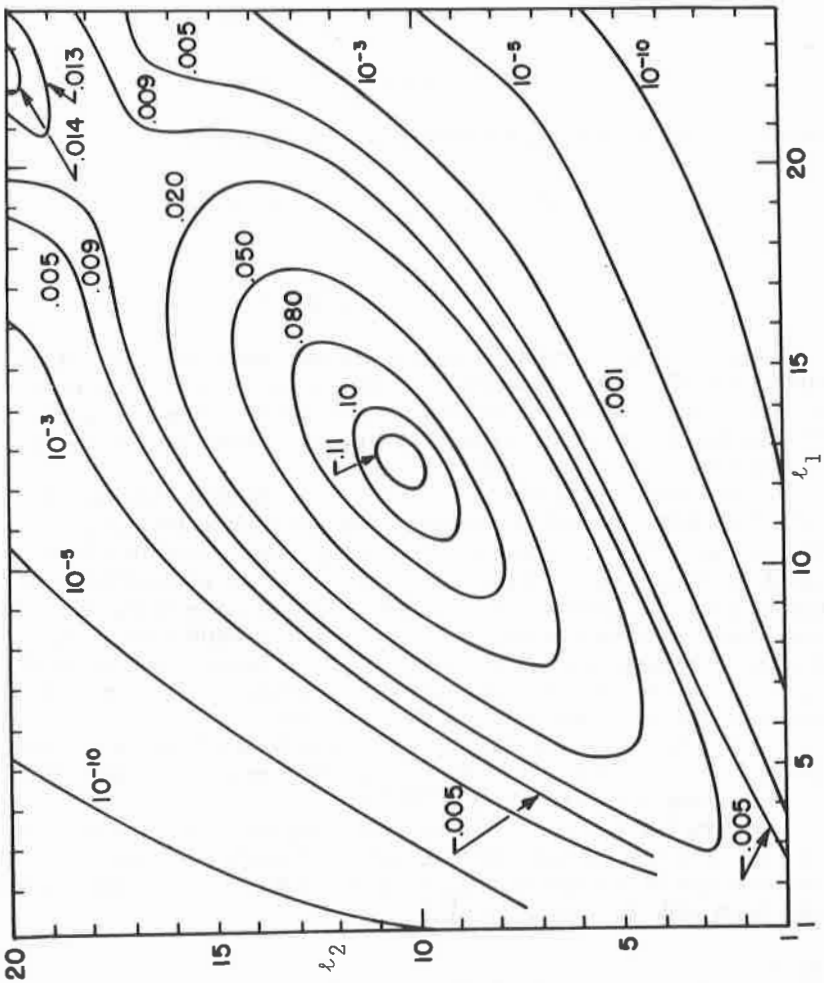


Figure 6b. A second example of curves of constant amplitude associated with various wave numbers, choice of parameters, given in text.

$$a(\ell_1,\ell_2) = a^*(-\ell_1,-\ell_2)$$

the (ℓ_1,ℓ_2) component of the velocity is related to $a(\ell_1,\ell_2)$ by

$$U_1(\ell,\omega) = -i\ell_2 q_2 \Phi(\ell,\omega)$$

$$U_2(\ell,\omega) = i\ell_1 q_1 \Phi(\ell,\omega).$$

Hence the energy associated with that wave number is

$$(\ell,\omega) = \frac{1}{2}\left(\ell_1^2 q_1^2 + \ell_2^2 q_2^2\right)|\Phi(\ell,\omega)|^2$$

$$= \frac{1}{2}\left(\ell_1^2 q_1^2 + \ell_2^2 q_2^2\right)|a(\ell_1,\ell_2)|^2.$$

The curves of constant energy are the objects we now examine rather than $a(\ell_1,\ell_2)$. However, the qualitative remarks about both are essentially the same. We can imagine two possible situations developing, one in which some isolated peaks develop in the manner in which they did in Figure 6b.

If a peak exists at (ℓ_1,ℓ_2) then, if our optimum pairing rule is still valid here, several higher wave number peaks at $(2\ell_1,2\ell_2)$, $(4\ell_1,4\ell_2)$, et cetera, again would develop. One would also have saddle points as are exhibited in Figure 6b. A second possibility is that of a ring of points where one finds a closed curve of maxima. In this case a sequence of further rings of maxima would appear, as in the case discussed in the last section. These higher wave numbers and frequency maxima might not show themselves when dissipation into heat developed just after the first maxima.

The discussion of the topology of the lines of constant energy given above is somewhat similar to that which appears in the theory of lattice vibrations of crystals.[28),30)]

Detailed analyses of these questions, as well as a search for a Kolmogorov type regime, will be discussed elsewhere. An investigation of various aspects of the 2D Navier-Stokes equation has also been made by Kraichnan.[31)]

References

1. V. Volterra, Léçon sur la théorie mathematiques de la lutte pour la vie, Paris 1931.
2. A. J. Lotka, Elements of Mathematical Biology 1924 (Dover Reprint, 1956).

3. F. W. Lanchester, Aircraft in Warfare; The Dawn of the Fourth Arm, Constable and Company, Ltd. (London 1916).
4. S. J. Deitchman, J. Oper. Res., 10, 818 (1962).
5. J. Engel, J. Oper. Res., 2, 163 (1954).
6. H. T. Davis, Introduction to Nonlinear Differential and Integral Equations (Dover 1962).
7. L. Markus, Contributions to Theory of Nonlinear Oscillations, V 185 (Princeton 1960).
8. R. A. Fisher, Ann. Eugen., 7, 355 (1937).
9. A. Kolmogorov I. Petrovsky and N. Piscounov, Bull. de l'Univ. d'Etat a Moscou (Ser. Internat.), A I, 1 (1937).
10. M. Benedict, Encyclopedia of Chemical Technology (Interscience Press, New York, 1950), Vol. V, p. 76.
11. K. Cohen, Science and Engineering of Nuclear Power (Addison-Wesley Press, Inc., Cambridge, Mass., 1949), Vol. 2, Chapter 2, p. 19.
12. E. Montroll and G. Newell, J. Appl. Phys., 23, 184 (1952).
13. Lord Rayleigh, Collected works IV, 261 (Cambridge, 1903).
14. J. M. Burgers, Proc. Acad. Sci., Amsterdam, 43, 2 (1940); 53, 247 (1950).
15. Cf. E. P. Gray and D. E. Kerr, Ann. of Phys., 17, 276 (1959).
16. E. Hopf, Comm. Pure Ap. Math., 3, 201 (1950).
17. J. Cole, Quar. Appl. Math., 9, 225 (1951).
18. C. D. Majumdar, Phys. Rev., 81, 844 (1951); 72, 390 (1947).
19. E. Montroll, J. Appl. Prob., 4, 281 (1967).
20. A. J. Lotka, J. Am. Chem. Soc., 42, 1595 (1920).
21. M. J. Moore, Trans. Far. Soc., 45, 1098 (1949).
22. M. Smoluchowski, Z. Physik Chem., 92, 129 (1918).
23. S. Chandrasekhar, Rev. Mod. Phys. 15, 1 (1943).
24. W. O. Schumann, Quart. J. Royal Meteor. Soc. 66, 195 (1940).
25. M. Tsuji, Memoir of the Faculty of Sci. Kyusu Univ. Ser. B. 1, 74 (1953).
26. E. Montroll, Energetics in Metalurg. Phen. 3, 123 (Gordon and Breach, 1967).
27. E. Montroll, J. Math. Phys. (In press.)
28. L. van Hove, Phys. Rev. 89, 1189 (1953).
29. E. Montroll, J. Chem. Phys. 15, 575 (1947); Am. Math. Monthly 67 (Supp. 7), 46 (1954).
30. J. C. Phillips, Phys. Rev. 104, 1263 (1956).
31. R. Kraichnan, Phys. of Fluids 10, 1417 (1967).
32. G. H. Weiss and D. A. Yphantis, J. Chem. Phys. 42, 2117 (1965).