TOPICS IN QUANTUM FIELD THEORY:

1. SCHWINGER'S ACTION PRINCIPLE;

2. DISPERSION RELATIONS FOR INELASTIC SCATTERING PROCESSES

by

T. W. B. Kibble

A thesis
submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
of the University of Edinburgh

October, 1958

PREFACE

I am deeply indebted to Dr. J.C. Polkinghorne for his constant guidance and encouragement, and for many stimulating discussions, as well as for suggesting the subject of the investigations presented in chapters V and VI. Certain parts of the earlier chapters, detailed in the introduction, consist of work done in collaboration with Dr. Polkinghorne.

I wish to thank Professor N. Kemmer, F.R.S., for the hospitality extended to me at the Tait Institute, and Professor Kemmer and Dr. D.L. Pursey for interesting and helpful discussions. In connection with the work of chapter IV I wish to thank Dr. R. Shaw and Dr. J.C. Taylor for a stimulating correspondence.

Finally, my thanks are due to the Department of Scientific and Industrial Research for the award of a maintenance grant.

The material here presented is asserted to be original except in so far as explicit reference is made to the work of others.

T. W. B. K.

Tait Institute of
Mathematical Physics,
University of Edinburgh

October, 1958

CONTENTS

CHAPTER I

INTRODUCTION

1. General introduction

The subject matter of this thesis falls into two distinct parts. Chapters II to IV are devoted to a discussion of Schwinger's action principle, and chapters V and VI are concerned with the proof of dispersion relations for inelastic meson-nucleon scattering.

The material of chapter II is based on some work done in collaboration with Dr. J.C. Polkinghorne, which has been published (Kibble and Polkinghorne 1957). This work was concerned with the clarification of certain points connected with the class of permissible variations in Schwinger's principle. There are, however, substantial changes in the present treatment, principally deriving from the introduction, in section II-5, of the concept of relative phases. This chapter is restricted to the case of non-relativistic quantum theory, and the discussion is extended to relativistic quantum field theory in chapter III. These chapters are devoted to a reformulation of Schwinger's action principle, and an investigation of the consequences of the new form of the action principle. Some of this material is necessarily contained in the work of Schwinger (1951, 1953a), but the treatment differs from his in several important respects. These are discussed in greater detail in section 2.

Chapter IV is devoted to a discussion of higher order spinor Lagrangians, with particular reference to the use of a two-component field satisfying a second-order equation rather than a four-component spinor satisfying a first-order equation. This procedure has been suggested by Feynman and Gell-Mann (1958) in connection with their universal Fermi interaction. The work presented in this chapter was done jointly with Dr. J.C. Polkinghorne, and has been published (Kibble and Polkinghorne 1958).

Chapters V and VI are devoted to a proof of the dispersion relations for the process in which a single meson is scattered on a nucleon into a state with several mesons. The proof follows the general lines of that by Bogolyubov, Medvedev and Polivanov (1956) for the case of elastic meson-nucleon scattering, This work has also been published (Kibble 1958).

The notation employed in the thesis is summarized in appendix A. Appendix B is devoted to a discussion of consistency conditions on the Lagrangian function.

The chapter number is omitted in references to sections or equations, except in the case of cross references between chapters.

2. Schwinger's action principle

Both in classical and quantum mechanics, Hamilton's principle of least action[1] has been widely used to provide a concise specification of a dynamical system through a single function, the Lagrangian. In quantum theory, however, a complete specification of the behaviour of the system requires knowledge not only of the equations of motion, but also of the commutation rules between the dynamical variables, which are now operators in Hilbert space. In the original formulation of quantum field theory (Heisenberg and Pauli 1929), the commutation rules had to be obtained separately, from the analogy between commutators and classical Poisson brackets (Dirac 1947). Therefore, Schwinger (1951) introduced his action principle for quantum theory, which yields both equations of motion and commutation relations. This principle differs from the classical action principle in the fact that variations are considered which do not vanish at the bounding times, and in the fact that it applies to arbitrary spacelike surfaces in place of surfaces of constant time only.

The original formulation of this principle was not entirely satisfactory. It was pointed out by Burton and Touschek (1953) that, when applied to Lagrangians linear in the derivatives, it appeared to yield commutation relations in which an undesired factor $\frac{1}{2}$ appears. Now there are quantum theories (in particular, that of the Dirac equation) which cannot easily be expressed in terms of the bilinear Lagrangians

---

1) For a discussion of the classical principle of least action see, for example, Whittaker (1937).

familiar in classical dynamics[1], whereas any theory with a bilinear Lagrangian can easily be expressed in terms of a linear one[2]. For this reason, Schwinger (1953a) presented a new and more satisfactory version of his action principle, restricted to linear Lagrangians only. He also showed (1953b) how canonical variables might be introduced for such a Lagrangian.

Nevertheless, there are some objections even to the new version of the action principle. Since the principle applies to operator variables, it is necessary, in order to obtain equations of motion, to specify the commutation properties of the variations with the dynamical variables. Schwinger considered variations which exactly commute (or in the case of Fermi variations with Fermi variables anticommute) with the dynamical variables, justifying this choice by certain heuristic arguments involving the symmetry between left and right multiplication. There are two objections to this procedure. First, it is by no means certain that the "anticommuting c-numbers", used as variations of the Fermi variables, can be consistently assumed to exist; and, second, one often wishes in practice to consider variations which are functions of the dynamical variables themselves, especially linear functions, as in the phase transformations. For this reason, Dr. Polkinghorne and the author (Kibble and Polkinghorne 1957) considered the possibility of using linear variations

---

1) This point is discussed in chapter IV.

2) This was shown to be the case in classical dynamics by Helmholtz (1886)

in the formulation of the action principle, both for classical and quantum mechanics. It was found that this could be done consistently, but that the resulting equations defining the commutation relations appeared to allow solutions more general than those corresponding to Bose and Fermi statistics. These extra solutions had already been obtained by Green (1953) from other arguments. They can only appear if the class of permissible Lagrangians is suitably restricted by conditions similar to those found by McCarthy (1955). In the course of this discussion, some light was thrown on the contradictions obtained by Burton and Touschek (1953), and a resolution of these difficulties was suggested which appears more natural than theirs. The method used to obtain a consistent theory was based on the work of Peierls (1952), and was essentially a generalization of that given by Schwinger (1953a, 1953b).

The action principle comprises two essentially distinct statements. One of these, which is closely analogous to the classical action principle, states that for a permissible variation of the dynamical variables and the space-time label, the change in the action integral must be a function of the dynamical variables at the limits of integration only. The other part is really the analogue of the classical Hamilton-Jacobi equation: it states that if the dynamical system is changed by the addition of an infinitesimal term (satisfying certain requirements) to the Lagrangian, the change in the action integral is equal to the difference of the generating functions, at the two limits of integration, of the infinitesimal canonical transformation connecting the varied and unvaried systems. The contradiction of Burton and Touschek arose from the assumption that if a canonical variation of the dynamical variables is made, the change in the action integral is again equal to the difference of the generators. This difficulty may therefore be avoided by explicitly separating the action principle into two parts, as is done here. Schwinger (1953a) evaluated the canonical variations generated by the momentum operators by assuming that the difference of their generators is given by the change in the action integral consequent upon a variation of the space-time label. This assumption is not in line with the separation of the action principle into two parts, and is in any case unnecessary. Here, these variations are evaluated by making an addition (involving the energy-momentum tensor) to the Lagrangian, and using the second part of the action principle. To do this it has been necessary to discuss pairs of systems related by an infinitesimal canonical transformation on every surface σ of a given one-parameter family of spacelike surfaces, as well as pairs so related on every spacelike surface.

The class of permissible variations used here is explicitly defined. It includes both the linear variations of Kibble and Polkinghorne (1957), and the generalized c-number variations of Schwinger (1953a) if they exist. It has seemed unsatisfactory to limit the discussion at the outset to variables which are essentially of Bose or Fermi character by requiring that the matrices A_μ which appear in the derivative term of the Lagrangian are direct sums of symmetric and antisymmetric submatrices, as was done both by Schwinger and by Kibble and Polkinghorne. In place of this assumption, the concept of "relative phase" is here introduced. The relative phase of two Bose variables would be $+1$ and of two Fermi variables -1, but it is not assumed a priori that the relative phases are all ± 1. Thus one is allowing for the possibility of commutation relations other than simple commutation or anticommutation. It should be noted that the possible generalization here considered is different in character from the generalization to Green's commutation rules, since in the latter case the variables are still either Bose or Fermi variables, although they do not have simple commutation properties. It is found, however, that requirements of consistency enable one to reject almost all the additional possibilities, and to limit the commutation rules among variables of the "same" field to the familiar Bose or Fermi rules. There is, however, a residual ambiguity in the commutation rules between kinematically independent fields. In the present formalism, the commutation rules of Green may definitely be excluded.

3. Higher order Lagrangians

It is well known that Bose fields satisfying the Klein-Gordon equation can be described either in terms of linear or bilinear Lagrangians. In the formalism presented here, one must interpret a bilinear[1] Lagrangian by associating with it a linearized Lagrangian involving auxiliary variables. This must satisfy two conditions: first, it must yield the same equations of motion on elimination of the auxiliary variables; and, second, if the auxiliary variables are eliminated from the linearized Lagrangian itself, the bilinear Lagrangian must result. This question is illustrated in section III-10 by a consideration of the familiar scalar and vector fields, with the object of providing a basis for comparison with the discussion of higher order Fermi field Lagrangians in chapter IV. It is noted, in particular, that a bilinear Lagrangian does not in general yield such complete information about the system as does its linearization.

Recently, Feynman and Gell-Mann (1958) have proposed a universal Fermi interaction which seems to arise most naturally if the Fermi fields are described in terms of two-component solutions of a second-order equation, rather than in terms of the equivalent first-order four-component Dirac equation. They introduce quantization by means of Feynman's path-integral formalism. In chapter IV, we investigate the extent to which their scheme can be reproduced in the more conventional formalism, using Schwinger's action principle to introduce quantization. This work was done in collaboration with Dr. Polkinghorne (Kibble and Polkinghorne 1958). Since the second-order equation corresponds most naturally to a

---

1) Here, "bilinear" means bilinear in the derivatives. More generally, a "second-order" Lagrangian is bilinear in ∂_μ.

bilinear rather than linear Lagrangian, the problem of linearization of bilinear spinor Lagrangians arises. It is well known that difficulties associated with an indefinite metric arise in the quantization of higher-order spinor Lagrangians[1]. These difficulties, and others associated with the reduction of the Lagrangian to two-component form, are discussed for the theory in question. It is found that third-order Lagrangians also appear, and that one has to consider improper linearizations, which are improper limits of true linearizations.

4. Dispersion relations

The purpose of the work of chapters V and VI (Kibble 1958) is to give a rigorous derivation of the dispersion relations for the process in which one incoming meson is scattered by a nucleon into a state with n outgoing mesons. Dispersion relations for this process have been derived by Polkinghorne (1956), but the derivation depends on the inversion of the order of certain integrals, which do not in general have the requisite convergence properties to justify this procedure. A similar inversion has been shown to be justifiable only in the case of forward elastic scattering (Symanzik 1957). For non-forward elastic scattering, Bogolyubov, Medvedev and Polivanov (1956) have obtained the dispersion relations

---

1) See Pais and Uhlenbeck (1950); Phillips (1955).

rigorously by a method involving analytic continuation. The relations were first derived for negative values of a parameter closely related to the square of the meson mass μ, and then certain properties of the functions involved were found which allow analytic continuation to positive values.

In section V-1 the causality condition in the form given by Bogolyubov et al. is discussed. The interaction picture is used to show that, under an adiabatic hypothesis, it follows from the more usual statements of the condition. Subsequently it will be regarded as a basic postulate. In the remaining sections of chapter V it is shown that the dispersion relations may be proved, by a method similar to that of Bogolyubov et al., on the basis of a theorem V-(6.1). An outline of the proof of this theorem is given in chapter VI. It is shown to follow from a mathematical theorem on analyticity of Lorentz-invariant functions of four-vectors. This is a generalization of a theorem stated by Bogolyubov et al. (1956, appendix) and is stated here without proof.

CHAPTER II

SCHWINGER'S ACTION PRINCIPLE IN NON-RELATIVISTIC QUANTUM THEORY

1. The dynamical variables; holonomic systems

The (non-relativistic) system will be described by a set of dynamical variables $q_a(t)$, operator functions of the time variable t. These variables will be regarded as components of a column vector $q(t)$. Without loss of generality, they may be taken to be Hermitian. The set of dynamical variables is to be chosen to include a number sufficient for a complete description of the system in the classical sense. In other words, the values of all q_a at a given time t_o are sufficient to determine their values at any other time t. In the canonical formalism, the q_a would comprise both the generalized coordinates and momenta. (In quantum theory, of course, one cannot choose all the operators $q_a(t_o)$ arbitrarily, since they must satisfy certain commutation rules.) More explicitly, the variables are to be chosen so that the equations of motion which they satisfy are explicit first-order differential equations, together with a number of equations of constraint involving no derivatives. If these equations of constraint are soluble for those variables whose derivatives do not appear, then the equations of motion are expressible in the form

$$\dot{q}_a(t) = f_a\{q(t), t\}, \tag{1.1}$$

and the system will then be said to be <u>holonomic</u>[1]. Much of the discussion will be restricted to holonomic systems.

If the f_a are not explicitly time-dependent, the system is said to be <u>closed</u>; and otherwise <u>forced</u>. It may be useful conceptually to think of the functions f_a as being commutators with a Hamiltonian function H,

$$f_a\{q(t),t\} = -i\left[q_a(t),\ H\{q(t),t\}\right], \tag{1.2}$$

and we shall in fact find at a later stage (in section 8) that they are always expressible in this form. It is not, however, necessary to assume the existence of a Hamiltonian.

The quantum analogue of the statement that the variables q_a furnish a complete description of the system (rather than of some subsystem) is the following:

> <u>Completeness assumption</u>. There exists a set of operator functions $\alpha_r\{q(t)\}$ such that the set of all α_r at any given time t constitutes a complete set of commuting observables.

The basis states may then be taken (Dirac 1947) to be simultaneous eigenstates $|\alpha' t\rangle$ of all α_r at any given time t. The development of the system with time may be completely described by means of the transformation matrix

$$\langle \alpha'' t_1 | \alpha' t_0 \rangle \tag{1.3}$$

relating the eigenstates of the α_r at t_1 to those at t_0.

---

1) It must be remarked that this does not coincide exactly with the usual definition in classical dynamics of a holonomic system, one without any non-integrable constraints. That definition cannot be applied without modification to the case of linearized equations of motion, of the form (1.1)

2. Restricted canonical transformations

Suppose that it is desired to describe the system in terms of a new set of "relabelled" variables $\bar{q}_a(t)$, given functions of $q(t)$ but not explicitly of t, and that these functions have been chosen in such a way that the commutation rules satisfied by the variables $\bar{q}$ at a fixed time t and by q at the same time t are identical. Then, for each t, there exists a unitary operator $U(t)$ such that

$$\bar{q}(t) = U(t)\, q(t)\, U^*(t). \tag{2.1}$$

Moreover, $U(t)$ must be a function of $q(t)$ but not explicitly of t. The transformation from q to $\bar{q}$ will be called a <u>restricted</u> (or <u>time-independent</u>) <u>canonical</u> <u>transformation</u>.

If the variables q and $\bar{q}$ differ infinitesimally, then U will have the form $U(t) = 1 - i\,G(t)$, where $G(t)$ is a Hermitian function $G\{q(t)\}$, and we shall in that case speak of a restricted canonical <u>variation</u>. Since the commutation relations are identical, the functions $\alpha_r\{\bar{q}(t)\}$ will constitute a complete set of commuting observables, with simultaneous eigenstates $|\bar{\alpha}'t\rangle$, say. The change in the transformation matrix (1.3) will be

$$\delta^G \langle \alpha'' t_1 | \alpha' t_0 \rangle \equiv \langle \bar{\alpha}'' t_1 | \bar{\alpha}' t_0 \rangle - \langle \alpha'' t_1 | \alpha' t_0 \rangle$$

$$= \langle \alpha'' t_1 | \, i\,G(t_1, t_0) | \alpha' t_0 \rangle, \tag{2.2}$$

where

$$G(t_1, t_0) \equiv G(t_1) - G(t_0) \tag{2.3}$$

$$= \int_{t_0}^{t_1} \frac{dG\{q(t)\}}{dt}\, dt. \tag{2.4}$$

Clearly, if $G(t_1, t_0)$ is given as a function of t_1 and t_0, then the generator $G(t)$ is uniquely determined (up to an irrelevant additive c-number) for all times t by the requirement that it is a function of $q(t)$ only.

Since the variables $\bar{q}$ are given functions of q, they satisfy equations of motion obtained from (1.1) by the appropriate substitution. It should be noted, however, that if a particular solution of the equations of motion for q is selected, not only are the equations of motion for $\bar{q}$ fixed, but also their solution is determined. This situation will be contrasted with that obtaining for general canonical transformations, in the next section.

If a Hamiltonian $H\{q\}$ exists for the system, then the equations of motion for $\bar{q}$ will be given by the Hamiltonian

$$\bar{H}\{\bar{q}(t)\} = H\{q(t)\}. \tag{2.5}$$

3. General canonical transformations

Consider now two distinct holonomic systems, respectively described by the dynamical variables $q_a(t)$ and $\bar{q}_a(t)$, which satisfy the equations of motion (1.1) and

$$\dot{\bar{q}}(t) = \bar{f}_a\{\bar{q}(t), t\}. \tag{3.1}$$

Suppose that the commutation relations satisfied by q at a fixed time t and by $\bar{q}$ at the same time t are identical. Then the variables q and $\bar{q}$ are, as before, related by a unitary transformation (2.1), but $U(t)$ is not necessarily a function of $q(t)$ only. Such transformations will be known as <u>unrestricted</u> (or <u>time-dependent</u>) <u>canonical transformations</u>. The expression "time-dependent" may be slightly misleading in that it is not true that $U(t)$ is always expressible as a function of $q(t)$ and t. For this reason, the term "unrestricted" is to be preferred. Again, if the variables q and $\bar{q}$ differ only infinitesimally, the difference between the transformation matrices will be given by (2.2) and (2.3); but now one does not have the condition that $G(t)$ is a function of $q(t)$ only, so that specification of the function $G(t_1, t_0)$ only determines the generators up to an arbitrary constant operator.

The operator $U(t)$ is of course a function of $q(t)$ and $\bar{q}(t)$ only. Since the specification of a particular solution of the equations of motion for q does not now determine a particular solution for $\bar{q}$, or in other words since $\bar{q}$ is no longer a function of q, U cannot be expressed in terms of q alone. However, we may still find a related operator which is so expressible. Using (2.1) and the equations of motion, and omitting the argument t, we have

$$\begin{aligned} f_\alpha\{q\} &= \frac{d}{dt}\left[U^* \bar{q}_\alpha U\right] \\ &= U^* \dot{\bar{q}}_\alpha U + U^* \bar{q}_\alpha \dot{U} - U^* \dot{U} U^* \bar{q}_\alpha U \\ &= U^* \bar{f}_\alpha\{\bar{q}\} U + \left[U^* \bar{q}_\alpha U,\ U^* \dot{U}\right] \\ &= \bar{f}_\alpha\{q\} + \left[q_\alpha,\ U^* \dot{U}\right]. \end{aligned} \tag{3.2}$$

This equation involves only the variables q, not $\bar{q}$, and it follows that $U^*\dot{U}$ is independent of the particular solution for $\bar{q}$. Hence $U^*\dot{U}$ is a function of q, and also explicitly of t is at least one of the systems is forced. It is uniquely determined by (3.2) up to an additive imaginary c-number[1]. This becomes particularly clear in the Hamiltonian form (1.3); for then

$$U^*\dot{U} = i\bar{H}\{q\} - iH\{q\}. \tag{3.3}$$

Similarly, $\dot{U}U^*$ is a function of $\bar{q}$ only, and indeed the same function of $\bar{q}$ as $U^*\dot{U}$ is of q.

In the infinitesimal case, the conclusion is that $\dot{G}$ is expressible as a function of q only. Thus, by (2.3),

$$G(t_1, t_0) = \int_{t_0}^{t_1} \Gamma\{q(t), t\}\, dt, \tag{3.4}$$

which differs from (2.4) in that Γ is not a total time-derivative, and in that Γ may depend explicitly on t. The ambiguity which arises in the specification of the generators by the function $G(t_1, t_0)$ may be removed by imposing a boundary condition which determines the solution for $\bar{q}$ for any given solution for q. The simplest form of boundary condition is

$$\bar{q}(t_0) = q(t_0) \tag{3.5}$$

for a fixed time t_0. This then fixes the generators by the condition

---

1) It is assumed that the variables q_α form an irreducible operator ring, so that any operator which commutes with every q_α is necessarily a c-number. Physically, this postulate is equivalent to the assumption that the system cannot be separated into two wholly non-interacting systems.

$$G(t_0) = 0, \tag{3.6}$$

whence

$$G(t_1) = G(t_1, t_0). \tag{3.7}$$

4. The action principle

The fact that, for all canonical variations, $G(t_1, t_0)$ has the form (3.4) suggests that it may be expressible as the variation, in a suitably defined manner, of a single operator of the form

$$W(t_1, t_0) = \int_{t_0}^{t_1} L\{q(t), \dot{q}(t), t\}\, dt, \tag{4.1}$$

where L is a function called the Lagrangian. $W(t_1, t_0)$ is the analogue of Hamilton's principal function in classical dynamics, but it has come to be known in quantum theory as the action[1].

In order to yield equations of motion of the required form (1.1), with or without extra equations of constraint, it will

---

1) The function usually called the action in classical dynamics is $A = 2\int T\, dt$, where T is the kinetic energy (cf. Whittaker 1937, Goldstein 1950). Some authors, however, call W the action function and A the contracted action function (cf. Joos 1951). The term "principle of stationary action" is usually restricted to the principle $\delta A = 0$, but is sometimes applied both to this and to Hamilton's principle $\delta W = 0$ (cf. Lanczos 1949). We shall use the term "action" for the function both because this is customary in quantum mechanics, and because it has the advantage of brevity.

appear that the function L must be chosen to have two terms, one bilinear in q and $\dot{q}$, and one independent of $\dot{q}$. Thus we take

$$L\{q,\dot{q},t\} = \frac{1}{4}\left[\tilde{q}(A+B)\dot{q} - \dot{\tilde{q}}(A-B)q\right] - H\{q,t\},$$

where A and B are constant c-number matrices. The function H will be called the Hamiltonian. It will be shown in a later section to have the usual properties of a Hamiltonian. The terms involving the matrix B are simply

$$\frac{1}{4}\frac{d}{dt}(\tilde{q}Bq).$$

Now the addition of a total time-derivative to L should correspond to a restricted canonical variation, so that Lagrangians differing by such terms must be regarded as describing the same system. It will be convenient therefore to take $B=0$ in the following, so that[1)]

$$L\{q,\dot{q},t\} = \frac{1}{4}\left[\tilde{q}A\dot{q} - \dot{\tilde{q}}Aq\right] - H\{q,t\}. \tag{4.2}$$

We remark, however, that an equivalent Lagrangian is

$$L\{q,\dot{q},t\} = \frac{1}{2}\tilde{q}A\dot{q} - H\{q,t\}. \tag{4.3}$$

It may readily be verified that the equations of motion and commutation relations to be found in the following sections are unchanged by taking (4.3) in place of (4.2) as the Lagrangian.

---

1) We define the Lagrangian with a factor $\frac{1}{4}$ for convenience. This is in agreement with Schwinger (1957), but represents a change from Schwinger's original notation (1953a). The argument leading to this choice of L is substantially that of Schwinger (1953a).

It is convenient for some purposes to separate off any bilinear term in H, and to write

$$H\{q,t\} = \tfrac{1}{2}\tilde{q}Mq + H_I\{q,t\}. \tag{4.4}$$

Since the generators $G(t)$ are Hermitian, the integrand of (3.4) is so. It is therefore reasonable to require that the Lagrangian function itself should be Hermitian. Then

$$\left.\begin{aligned} A^\dagger &= -A, \\ M^\dagger &= M, \\ H^*\{q,t\} &= H\{q,t\}. \end{aligned}\right\} \tag{4.5}$$

The term $-H_I$ will be known as the <u>interaction</u> Lagrangian and the other terms collectively as the <u>free</u> Lagrangian.

It will appear in the subsequent discussion that the Lagrangian function must be further restricted, to ensure consistency.

If the matrix A should be singular, then we make a real (non-singular) linear transformation of the dynamical variables in such a way that

$$q = \begin{bmatrix} q^1 \\ q^2 \end{bmatrix}, \qquad A = \begin{bmatrix} A^{11} & 0 \\ 0 & 0 \end{bmatrix}, \tag{4.6}$$

where A^{11} is non-singular. Evidently, the Lagrangian involves $\dot{q}^1$ but not $\dot{q}^2$: the variables q^1 will be termed <u>independent</u>, and the variables q^2 <u>dependent</u>.

Schwinger's action principle (1953a) comprises two essentially distinct postulates which, as remarked in chapter I, will here be stated separately, as follows.

Action principle for canonical variations. If two infinitesimally differing systems are related by a canonical variation, then the corresponding function $G(t_1, t_0)$, defined by (2.3), is equal to the difference of the action integrals (between these limits) for the two systems. Conversely, if the Lagrangian function is changed by the addition of any function of q and t which satisfies the requirements (to be specified) for a permissible term in the Lagrangian, and which does not change the equations of constraint (if any), then the new system is related to the old by a canonical variation such that the corresponding function $G(t_1, t_0)$ is equal to the change in the action integral between these limits.

Note that the condition that the equations of constraint are unchanged is necessary, if the two systems are to be related by a canonical variation. If L and $\bar{L}$ are the Lagrangian functions for the two systems, this part of the action principle requires that

$$G(t_1, t_0) = \delta W(t_1, t_0) \equiv \int_{t_0}^{t_1} \left[\bar{L}\{q, \dot{q}, t\} - L\{q, \dot{q}, t\} \right] dt. \quad (4.7)$$

If the time-interval is taken to be infinitesimal, then

$$\dot{G}(t) = \bar{L}\{q, \dot{q}, t\} - L\{q, \dot{q}, t\}. \quad (4.8)$$

It would be possible to extend the principle to cover the addition of functions of q, $\dot{q}$ and t, provided that these are chosen in such a way as not to change the matrix A (which will be found to define the commutation relations). This restriction means, however, that such additional terms can only be of the form of total time-derivatives, and can therefore only correspond to restricted canonical variations. Any non-restricted canonical variation can therefore be given by the addition to L of a function of q and t only, together with some restricted canonical variation on one of the systems.

The second form of the action principle is more closely analogous to Hamilton's classical variational principle, though its content is wider than the usual statements of that principle, since it takes into account variations which do not vanish at the bounding times. It is the following.

Action principle for permissible variations of q and t.

If an arbitrary variation of the time label

$$t \rightarrow \bar{t} = t + \Delta t \tag{4.9}$$

is made, together with an arbitrary permissible variation of the dynamical variables

$$q(t) \rightarrow \bar{q}(\bar{t}) = q(t) + \Delta q(t) \tag{4.10}$$

where $\Delta q(t)$ is a function of $q(t)$ and t only, then the resulting change in the action integral is equal to the function $G(t_1, t_0)$ corresponding to some restricted canonical variation.

It is important to remark that even if the variation (4.10) is canonical, the function $G(t_1, t_0)$ need not correspond to this variation, but only to some associated variation. What is asserted is that the change in the action integral must be a function only of the variables at t_0 and t_1. The class of permissible variations will be defined in section 6, by imposing certain restrictions on the commutation properties of the variations Δq with the dynamical variables q. The most convenient characterization of this class is in terms of the concept of "relative phase" introduced in section 5.

The symbolic statement of the action principle for permissible variations is

$$G(t_1, t_0) = \delta W(t_1, t_0) = \int_{\bar{t}_0}^{\bar{t}_1} L\{\bar{q}(\bar{t}), \dot{\bar{q}}(\bar{t}), \bar{t}\}\, d\bar{t} - \int_{t_0}^{t_1} L\{q(t), \dot{q}(t), t\}\, dt. \qquad (4.11)$$

The distinction between the two parts of the action principle is readily obvious from a comparison of (4.7) and (4.11).

5. Relative phases

In classical dynamics, the order in which the variables are written in the Lagrangian function is irrelevant; but in quantum theory this is no longer the case, and so the quantum analogue of a given classical theory is, in general, ambiguous. Moreover, it is well known that there exist quantum mechanical systems which have no classical analogue - for example, the so-called Fermi systems, which are characterized by anticommutation rather than commutation rules. It is not, therefore, surprising that a variational principle strictly analogous to the classical variational principle is not sufficient in the quantum case, and that some further assumption must be made about the commutation properties of the variations and the dynamical variables. To some extent, the choice of this assumption is a matter of taste, but the least arbitrary way of introducing the assumption seems to be through the concept of relative phase. The fundamental postulate is the following.

Relative phase assumption. To each pair of dynamical variables q_a and q_b there corresponds a relative phase α_{ab}, a c-number of modulus unity, such that

i) the expression

$$(q_a ; q_b)_- = q_a q_b - \alpha_{ab} q_b q_a \tag{5.1}$$

has a simple form as a function of the dynamical variables, as compared with any similar function with a different phase factor in place of α_{ab} (but see note below); and

ii) any operator equation may be written as a number of separate equations, each of which satisfies the condition that the relative phase of any given dynamical variable q_a with each term of the equation, defined as the product of the relative phases of q_a with each factor of the term, is the same[1].

Evidently, the symmetry of the expression (5.1) requires that

$$\alpha^*_{ab} = \alpha_{ba}. \tag{5.2}$$

The relative phase of two products of dynamical variables[1],

$$F = f q_{a_1} q_{a_2} \cdots q_{a_n} \quad \text{and} \quad G = g q_{b_1} q_{b_2} \cdots q_{b_m}, \tag{5.3}$$

where f and g are c-numbers, is defined to be

$$\alpha_{F,G} \equiv \prod_{i=1}^{n} \prod_{j=1}^{m} \alpha_{a_i b_j}. \tag{5.4}$$

1) For this purpose, a factor $\dot{q}_a$ is regarded as equivalent to q_a.

The generalized commutator and anticommutator of F and G are

$$(F;\ G)_{\mp} \equiv FG \mp \alpha_{F,G}\, GF. \tag{5.5}$$

The condition (ii) may be compared to the familiar condition that the covariant and contravariant indices of each term of a tensor equation should be the same.

Given a product of the dynamical variables, one may use the commutation rules to move a selected q_a to the extreme right hand end of the product. This will in general necessitate the addition of certain extra terms (involving the values of commutators, etc.), and the force of the assumption (i) is that, at least in general, these extra terms will have their simplest form if the reordering is done by replacing each product $q_a q_b$ in turn by $\alpha_{ab} q_b q_a$.

Note. If $q_b = q_a$, it is obvious that the simplest expression of the form (5.1) is $q_a q_a - q_a q_a = 0$. In that case, however, we still allow the possibility that α_{aa} is not unity. In fact, by (5.2), α_{aa} must be real, but it may be -1 if $q_a q_a$ is a "simple" function. A good example of variables which may be taken to have relative phases α_{aa} equal to -1 is furnished by the Pauli spin matrices[1].

It should be remarked that our relative phase assumption is not trivial. There do exist sets of operators for which two different relative phases yield a comparitively simple form (5.1), particularly those discussed in the note above. It

1) See appendix A-3.

is therefore possible to contemplate theories satisfying Hamiltonian equations of motion with a Hamiltonian, for example, linear in the Fermi variables, and to formulate the action principle in such a way as to allow for these theories[1]. However, this has several disadvantages. It does not seem possible to formulate an action principle of this kind in such a way that its results are as unambiguous as those found from the formulation presented here. Secondly, the equations of motion must in that case be obtained in Hamiltonian rather than Lagrangian form, which not only destroys part of the analogy with the classical variational principle, but also makes the extension to field theory extremely difficult. Finally, it seems to be an advantage to formulate the principle in such a way as to exclude these theories.

The relative phases discussed here are, in the first instance, to be given independently of the Lagrangian function. However, we shall find that for consistency the choice of relative phases and the choice of Lagrangian function must be closely related. It is possible to regard the Lagrangian as given, and to ask what choices of relative phases then yield a consistent theory, but it is simpler at this stage to regard the relative phases as given, and to seek consistent Lagrangians. It will be found that for many choices of relative phases there are in fact no consistent Lagrangians at all, and these choices may therefore be rejected. It must be noted that according to this treatment the equations of

---

1) I am indebted to Mr. J.L. Martin for bringing this point to my attention.

motion and commutation relations are not uniquely defined by the Lagrangian itself, but only by the Lagrangian together with the relative phases.

6. The class of permissible variations

Let F and G be two functions of q, but not $\dot{q}$, of the form (5.3). It is useful to be able to express their commutator $(F;G)_-$, defined by (5.5), in terms of the individual commutators $(q_a;G)_-$. To this end, we define two functions, each the sum of all possible terms obtained by deleting one occurrence of a given q_a from F and replacing it by G, multiplied by a suitable phase factor. These are

$$\left.\begin{aligned} G\partial_a F &\equiv \sum_{\substack{i:\\ a_i=a}} \left\{ \prod_{j=1}^{i-1} \alpha_{G,q_{a_j}} \alpha_{a_j a} \right\} f q_{a_1}\cdots q_{a_{i-1}} G q_{a_{i+1}} \cdots q_{a_n}, \\ F\partial_a^* G &\equiv \sum_{\substack{i:\\ a_i=a}} \left\{ \prod_{j=i+1}^{n} \alpha_{a a_j} \alpha_{q_{a_j},G} \right\} f q_{a_1}\cdots q_{a_{i-1}} G q_{a_{i+1}} \cdots q_{a_n}. \end{aligned}\right\} \quad (6.1)$$

The symbols ∂_a and ∂_a^* may be regarded as modified derivation symbols acting to the right and left respectively. It is readily verified that

$$(F;G)_- = \sum_a F\partial_a^*(q_a;G)_- = \sum_a (F;q_a)_- \partial_a G,$$

or, in matrix notation[1],

$$(F;G)_- = F\partial^\dagger(q;G)_- = (F;\tilde{q})_- \partial G. \quad (6.2)$$

---

1) See appendix A-2 for a summary of the matrix notation used.

Moreover, as the notation suggests,

$$(G\,\partial_a F)^* = F^*\,\partial_a^*\,G^*. \tag{6.3}$$

Left and right derivatives of F with respect to q_a are defined by

$$\partial_a F \equiv 1\,\partial_a F \qquad \text{and} \qquad F\,\partial_a^* \equiv F\,\partial_a^*\,1. \tag{6.4}$$

They satisfy

$$F\,\partial_a^* = \alpha_{aa}\,\alpha_{q_a,F}\,\partial_a F \tag{6.5}$$

and

$$\partial_a(FG) = (\partial_a F)G + \alpha_{F,q_a}\,F(\partial_a G). \tag{6.6}$$

It is easily verified that

$$(\partial_a F)\,\partial_b^* = \partial_a(F\,\partial_b^*),$$

so that one may unambiguously define the matrix operator $\partial F \partial^\dagger$.

It is important to note that the operators $G\,\partial_a F$ and $G(\partial_a F)$ are in general different, unless G is a c-number, or generalized c-number (see below).

A set of infinitesimal Hermitian operators Δq_a is said to be a set of <u>generalized c-numbers</u> if Δq_a is independent of q and

$$q_a\,\Delta q_b = \alpha_{ab}\,\Delta q_b\,q_a \tag{6.7}$$

It is clear that if F is any product of dynamical variables the change in F consequent upon the variation $q_a \to q_a + \Delta q_a$ (with no variation of t), where the Δq_a are generalized c-numbers, is

$$\Delta F = \Delta\tilde{q}\,(\partial F) = (F\,\partial^\dagger)\,\Delta q. \tag{6.8}$$

In this relation, the brackets may be omitted or not, without affecting the value of the expressions. We do not assume, at this stage at least, that generalized c-numbers exist, and indeed it is quite likely that they do not.

We now pass to the definition of the class of permissible variations.

Permissible variations. A set of infinitesimal Hermitian operators Δq_a is said to be a set of permissible variations of q if

i) each $\Delta q_a(t)$ is a function of $q(t)$ only;

ii) the relative phase of Δq_a and q_b is α_{ab}; and

iii) the generalized commutators of Δq with q are of such a form that, for any permissible choice of the Hamiltonian, the variation ΔH may be written as the sum of two terms, one involving Δq as a factor on the left and one as a factor on the right.

Here ΔH is the variation of H consequent upon the variation $q_a \rightarrow q_a + \Delta q_a$ with no variation of t. It follows from the relative phase assumption that

$$\Delta H = \tfrac{1}{2}\left\{\Delta \tilde{q}(\partial H) + (H\partial^{\dagger})\Delta q\right\}. \qquad (6.9)$$

Indeed this relation may be taken as an alternative definition of permissible variations. The permissible Hamiltonians are those which give consistent results with the given relative phases. An explicit definition of permissible Hamiltonians will be found in section 8.

We note that generalized c-numbers (if they exist) are particular examples of permissible variations. Other examples will be found in section 10.

This section completes the statement of the action principle and the other necessary assumptions. The following sections are devoted a discussion of the conclusions which may be drawn therefrom.

7. Equations of motion

Now consider a variations (4.9) of the time label, together with a permissible variation (4.10) of q. In addition to Δq, it is useful to define

$$\delta q(t) \equiv \bar{q}(t) - q(t). \tag{7.1}$$

Note, however, that although $\Delta q(t)$ is a function of $q(t)$ and t only, this is not true of $\delta q(t)$. If f is any function of q and t, we define

$$\begin{aligned} \delta f(t) &\equiv f\{\bar{q}(t), t\} - f\{q(t), t\} \\ &= f\{q + \delta q, t\} - f\{q, t\}, \end{aligned} \tag{7.2}$$

and

$$\begin{aligned} \Delta f(t) &\equiv f\{\bar{q}(\bar{t}), \bar{t}\} - f\{q(t), t\} \\ &= f\{q + \Delta q, t + \Delta t\} - f\{q, t\}, \end{aligned} \tag{7.3}$$

whence

$$\Delta f = \delta f + \Delta t \frac{df}{dt}. \tag{7.4}$$

Moreover,

$$\frac{d}{dt}\delta f = \delta \frac{df}{dt}, \tag{7.5}$$

whence

$$\frac{d}{dt}\Delta f = \Delta \frac{df}{dt} + \frac{d\Delta t}{dt}\frac{df}{dt}. \tag{7.6}$$

Now, by (4.2), (6.9), (7.3) and (7.6),

$$\Delta L = \tfrac{1}{4}\{\Delta\tilde{q} A\dot{q} - \dot{\tilde{q}} A\Delta q\} + \tfrac{1}{4}\{\tilde{q} A \frac{d\Delta q}{dt} - \frac{d\Delta\tilde{q}}{dt} A q\}$$

$$- \tfrac{1}{4}\frac{d\Delta t}{dt}\{\tilde{q}A\dot{q} - \dot{\tilde{q}}Aq\} - \Delta H$$

$$= \tfrac{1}{2}\{\Delta\tilde{q}A\dot{q} - \dot{\tilde{q}}A\Delta q\} + \tfrac{1}{4}\frac{d}{dt}\{\tilde{q}A\Delta q - \Delta\tilde{q}Aq\}$$

$$- \tfrac{1}{4}\frac{d\Delta t}{dt}\{\tilde{q}A\dot{q} - \dot{\tilde{q}}Aq\}$$

$$- \tfrac{1}{2}\{\Delta\tilde{q}(\partial H) + (H\partial^{+})\Delta q\} - \Delta t\frac{\partial H}{\partial t}.$$

Then from (4.11)

$$\delta W(t_1,t_0) = \int_{t_0}^{t_1}\left[L\{\bar{q}(\bar{t}), \dot{\bar{q}}(\bar{t}), \bar{t}\}\frac{d\bar{t}}{dt} - L\{q(t), \dot{q}(t), t\}\right]dt$$

$$= \int_{t_0}^{t_1}\left[\Delta L + L\frac{d\Delta t}{dt}\right]dt$$

$$= \tfrac{1}{2}\int_{t_0}^{t_1}\{\Delta\tilde{q}(A\dot{q} - \partial H) - (\dot{\tilde{q}}A + H\partial^{+})\Delta q\}\,dt$$

$$+ \int_{t_0}^{t_1}\Delta t\left\{\frac{dH}{dt} - \frac{\partial H}{\partial t}\right\}dt + G(t_1) - G(t_0), \tag{7.7}$$

where

$$G(t) = \tfrac{1}{4}(\tilde{q}A\Delta q - \Delta\tilde{q}Aq) - H\Delta t. \tag{7.8}$$

Since the conditions defining the class of permissible variations are purely local in character, the variations Δq may be chosen independently at different times, as also may Δt. But the action principle requires that $\delta W(t_1; t_0)$ be of the form (2.4) appropriate to a restricted canonical variation. Hence we obtain the relations

$$\Delta\tilde{q}(A\dot{q} - \partial H) - (\dot{\tilde{q}}A + H\partial^{\dagger})\Delta q = 0 \tag{7.9}$$

for any permissible variation, and

$$\frac{dH}{dt} = \frac{\partial H}{\partial t}, \tag{7.10}$$

the usual energy-conservation equation. The equation (7.9) will be used to derive the equations of motion, using the assumptions made in sections 5 and 6.

Now let

$$f \equiv A\dot{q} - \partial H. \tag{7.11}$$

Then, by (4.5),

$$f^{\dagger} = -(\dot{\tilde{q}}A + H\partial^{\dagger}) \tag{7.12}$$

whence (7.10) becomes

$$\Delta\tilde{q}f + f^{\dagger}\Delta q = 0. \tag{7.13}$$

Now, according to the relative phase assumption, (7.11) may be written as a number of separate equations

$$f^{i} = A^{i}\dot{q} - \partial H^{i} \tag{7.14}$$

whose sum is (7.11), such that

$$\alpha_{q_c, f^i_a} = \alpha_{cb} = \alpha_{ac}\alpha_{q_c, H^i} \tag{7.15}$$

whenever A^i_{ab} and $\partial_a H^i$ are non-zero, for any q_c. However, the number of independent equations satisfied by the variables q cannot exceed the number of variables n, say. Thus the set of equations (7.14) for all different values of i must be linear combinations of n equations only. It is therefore possible to choose the variables q in such a way that there is only one f^i_a corresponding to each q_a. Then the separation of (7.11) into individual equations (7.14) is unnecessary, and the superscript i may be omitted in (7.15).

We now state a lemma which ensures the uniqueness of the equations of motion.

<u>Uniqueness lemma</u>. A set of functions f_a satisfying (7.13) and (7.15) must also satisfy

$$f^*_a = -\alpha_{q_a, f_a} f_a. \tag{7.16}$$

It is clear that if generalized c-numbers did exist, then equation (7.13) would imply

$$\Delta q_a \{ f_a + \alpha_{f_a, q_a} f^*_a \} = 0, \tag{7.17}$$

whence (7.16) follows[1]. The validity of the lemma in the case where generalized c-numbers do not exist is discussed in section 11.

Now, taking the Hermitian conjugate of (7.11) and using (6.5) and (7.16) one obtains

$$-\alpha_{q_a, f_a} f_a = \sum_b A^*_{ab} \dot{q}_b - \alpha_{aa} \alpha_{q_a, H'} \partial_a H'.$$

Thus, by (7.15),

---

1) No summation over a is required, since the generalized c-numbers Δq_a may be chosen independently for each q_a.

$$-f_a = \sum_b \alpha_{ba} A^*_{ab} \dot{q}_b - \partial_a H. \tag{7.18}$$

Comparing (7.11) and (7.18), we obtain the equations of motion

$$\sum_b \tfrac{1}{2} (A_{ab} + \alpha_{ba} A^*_{ab}) \dot{q}_b = \partial_a H, \tag{7.19}$$

from which it is clear that only that part of the matrix A which satisfies the relation

$$A^*_{ab} = \alpha_{ab} \bar{A}_{ab} = - A_{ba} \tag{7.20}$$

contributes to the equations of motion. It is thus sufficient to consider only matrices A which satisfy this condition. Then (7.19) becomes

$$A \dot{q} = \partial H. \tag{7.21}$$

We remark that the condition that the system be holonomic is that the variables q can be chosen in such a way that A is non-singular. If A is singular, then the equations of motion for the independent variables q^1, defined in (4.6), are

$$A^{11} \dot{q}^1 = \partial^1 H, \tag{7.22}$$

and the equations of constraint are the equations for the dependent variables q^2, namely

$$0 = \partial^2 H. \tag{7.23}$$

The permissible Lagrangians are clearly to be restricted by the relation (7.15), but we shall find in section 8 that they must in fact satisfy further restrictions.

We note that the matrix M may be restricted by conditions similar to (7.20). Indeed, only that part of the matrix which satisfies

$$M^*_{ab} = \alpha_{ab} M_{ab} = M_{ba} \tag{7.24}$$

contributes to $\partial_a H$, so that it is sufficient to consider matrices M satisfying this relation.

8. Bose and Fermi variables

We now consider the change in the system induced by the addition to L of the term[1)]

$$\lambda \Gamma\{q(t)\}\ \delta(t-t'),$$

where Γ is a Hermitian function satisfying the requirements for a permissible term in L, and λ is an infinitesimal real number. The equations of motion will be changed from (7.21) to

$$A\dot{q} = \partial H - \lambda\, \partial\Gamma\ \delta(t-t'). \tag{8.1}$$

Thus the equations for q^1 are

$$A^{11}\dot{q}^1 = \partial^1 H - \lambda\, \partial^1\Gamma\ \delta(t-t'), \tag{8.2}$$

and the equations of constraint are

$$0 = \partial^2 H - \lambda\, \partial^2\Gamma\, \delta(t-t'). \tag{8.3}$$

If the solutions of (7.21) and (8.1) are related by a canonical variation, then the equations of constraint must be unchanged. Thus (8.3) must be simply linear combinations of the equations

---

1) The discussion of such changes in the system is based in part on the work of Peierls (1952). It is an extension to more general cases of the treatment given by Schwinger (1953a).

(7.23). Hence Γ must be restricted by the condition

$$\partial^2 \Gamma = \mathcal{B}^{22} \partial^2 H \tag{8.4}$$

where $\mathcal{B}^{22}$ is some matrix. In particular, we may choose either to be independent of q^2 ($\mathcal{B}^{22} = 0$) or to be equal to H itself ($\mathcal{B}^{22} = 1$).

From (8.2) and (7.22) there is a discontinuity in $\delta^G q$ (where G is the generator of the canonical variation relating the solutions of these equations) at t', given by

$$A^{11} \left[\delta^G q^1 \right]_{t'-0}^{t'+0} = -\lambda\, \partial^1 \Gamma \{q(t')\}. \tag{8.5}$$

Now we take t' to lie between t_0 and t_1, so that, by the action principle (4.7),

$$G(t_1, t_0) = \delta W(t_1, t_0) = \lambda \Gamma \{q(t')\} \tag{8.6}$$

To determine the generators G we impose the boundary condition (3.5). Then $\delta^G q$ vanishes for times between t_0 and t', and, if t' is taken infinitesimally close to t_1, then[1)]

$$\begin{aligned} \delta^G q(t_1) &= i\, [q(t_1), G(t_1)] \\ &= i\, [q(t_1), \lambda \Gamma \{q(t_1)\}], \end{aligned} \tag{8.7}$$

using (3.7) and (8.6). Combining (8.5) and (8.7), we have

$$A^{11} \left[q^1(t), \Gamma\{q(t)\} \right] = i\, \partial^1 \Gamma \{q(t)\}. \tag{8.8}$$

Using the relative phase assumption again, one finds

$$\alpha_{cb} = \alpha_{ac} \tag{8.9}$$

whenever A_{ab} and $\partial_a \Gamma$ are non-zero. Since one can always

---

1) See appendix A-2, especially A-(2.4).

find some function Γ which does involve q_a, (8.9) holds whenever A_{ab} is non-zero.

It follows from (7.15) that

$$\alpha_{q_c, H} = 1 \tag{8.10}$$

for all q_c. Thus, permissible terms in H must be restricted by this condition. If A is non-singular, then this is in fact the complete specification of permissible terms. In the case where A is singular, however, there are some extra conditions, discussed in section 12, which are required for consistency of the equations of constraint. Even in that case, (8.10) is a sufficient restriction on any terms in H which do not involve the dependent variables q^2.

Since (8.9) shows that the relative phase of any variable q_c with $q_a A_{ab} q_b$ is unity (whenever A_{ab} is non-zero), the conditions (8.9) and (8.10) may be combined in the single statement

$$\alpha_{q_c, L} = 1 \tag{8.11}$$

for any variable q_c.

Equation (8.9) also shows that, if A_{ab} is non-zero, then

$$\alpha_{ab} = \alpha_{aa} = \alpha_{bb}, \tag{8.12}$$

which is necessarily real. A variable q_a will be called a <u>Bose</u> variable if $\alpha_{aa} = 1$, and a <u>Fermi</u> variable if $\alpha_{aa} = -1$. It is clear that the matrices A and M must decompose into direct sums of submatrices corresponding to the Bose and Fermi variables respectively,

$$A = A_B \oplus A_F , \qquad M = M_B \oplus M_F . \tag{8.13}$$

Using (7.20), (7.24) and (8.11), one finds that these submatrices must satisfy[1)]

$$\left.\begin{aligned} A_B^* &= A_B = -\tilde{A}_B , & M_B^* &= M_B = \tilde{M}_B , \\ -A_F^* &= A_F = \tilde{A}_F , & -M_F^* &= M_F = -\tilde{M}_F . \end{aligned}\right\} \tag{8.14}$$

We may further partition the Bose and Fermi variables into classes such that, if q_a and q_b belong to the same class, then

$$\alpha_{ac} = \alpha_{bc} \qquad \text{for all } q_c . \tag{8.15}$$

To each class there will correspond a conjugate class such that, if q_a and q_b belong to conjugate classes, then

$$\alpha_{ac}\alpha_{bc} = 1 \qquad \text{for all } q_c . \tag{8.16}$$

It follows from (8.11) that the matrices A and M have non-zero elements only between variables of conjugate classes[2)].

It is convenient to partition the vector q into subvectors q^j, each consisting of the variables belonging to a pair of conjugate classes. Then the matrices A and M are direct sums

---

1) These conditions on the matrix A were imposed as a priori assumptions by Schwinger (1953a), and by Kibble and Polkinghorne (1957).

2) This condition, together with the restrictions (8.14), specifies completely the permissible forms of the matrices A and M, except when A is singular. In that case there are additional restrictions, required by consistency of the equations of constraint. These are discussed in section 12.

of submatrices A^j and M^j acting on these subvectors, so that the equations of motion can be expressed as separate equations

$$A^j \dot{q}^j - M^j q^j = \partial^j H_I . \qquad (8.17)$$

The left side of each of these equations depends solely on the free Lagrangian, and involves the variables of one subvector only. The right side, on the other hand, depends on the interaction Lagrangian, and may involve variables of any class. The matrices A^j and M^j corresponding to any subvector have the form

$$A^j = \begin{bmatrix} 0 & A'^j \\ -A'^{j\dagger} & 0 \end{bmatrix}, \qquad M^j = \begin{bmatrix} 0 & M'^j \\ M'^{j\dagger} & 0 \end{bmatrix}, \qquad (8.18)$$

except when the two conjugate classes are identical. Note that the condition that a given class be self-conjugate is that all the relative phases of any element of the class are real.

Finally, we note that if we take $\Gamma = H$ in (8.8), then we obtain

$$\dot{q}^1 = (A^{11})^{-1} \partial^1 H = -i \left[q^1, H \right], \qquad (8.19)$$

using (7.21). These are the equations of motion in Hamiltonian form. They show that H does indeed have the correct properties of a Hamiltonian. For a holonomic system with singular A, the dependent variables q^2 are given as functions of q^1, by solving (7.23) for q^2, so that (8.19) holds for q^2 also, except when these functions are explicitly time-dependent.

9. Commutation relations

One term which is certainly a permissible addition to L is

$$\Gamma = \tfrac{1}{2}\,\tilde{q}^{1} B^{11} q^{1}, \tag{9.1}$$

provided only that the Hermitian matrix B^{11} has non-zero elements only between variables of conjugate classes, like A and M. Then, if q_a and q_b belong to conjugate classes, (8.8) yields

$$i\,[q_c, q_a q_b] = -A^{-1}_{ca} q_b - A^{-1}_{cb} \alpha_{ab} q_a . \tag{9.2}$$

Here and below, q_a, q_b and q_c are all assumed to be independent variables, components of q^{1}, and we have written A^{-1} for $(A^{11})^{-1}$, for simplicity[1]. It follows that

$$i\,[q_c, (q_a ; q_b)_-] = 0 , \tag{9.3}$$

whence $(q_a ; q_b)_-$ is a c-number. Since A^{-1}_{cb} vanishes unless $\alpha_{cb} = \alpha_{ab}$, equation (9.2) may be written as

$$i(q_c ; q_a)_- q_b - i\,q_a (q_b ; q_c)_- = -A^{-1}_{ca} q_b + A^{-1}_{bc} q_a . \tag{9.4}$$

Taking q_c to be of the same class as either q_a or q_b, one then finds that $(q_a ; q_b)_-$ must be a c-number when q_a and q_b belong to the same, as well as to conjugate, classes, and that its value is

$$(q_a ; q_b)_- = i\,A^{-1}_{ab} , \tag{9.5}$$

or, in terms of the subvectors q^{j},

---

1) No confusion can arise, since A^{-1} does not exist unless it is equal to $(A^{11})^{-1}$.

$$(q^j\,;\ \tilde{q}^j)_- = i(A^j)^{-1}. \tag{9.6}$$

There remains the question of evaluating the commutation rules between different subvectors. This cannot be done uniquely using (9.4) only. However, one may rewrite equation (8.8) in the general case as

$$(q^1;\ \tilde{q}^1)_-\,\partial^1\Gamma = i\,A^{-1}\,\partial^1\Gamma, \tag{9.7}$$

using (6.2) and the fact that the relative phases of Γ are all unity. The only general solution of (9.7) appears to be

$$(q^1\,;\ \tilde{q}^1)_- = i\,(A^{11})^{-1}, \tag{9.8}$$

which includes (9.6) as a special case. The commutation rules for the dependent variables q^2 would of course follow from (9.8) by use of the equations of constraint (7.23).

The possibility still remains that for particular choices of the relative phases, other solutions of (9.7) may exist. It was shown by Kibble and Polkinghorne (1957) that for a very restricted class of functions[1] Γ, there do exist other solutions, corresponding to the generalized commutation rules of Green (1953). The restriction there imposed on Γ is, however, impermissible here, since it disallows some terms of the form (9.1) which are definitely permissible according to our prescription.

The ambiguity in commutation rules may be definitely removed if it is assumed that generalized c-numbers do exist.

---

1) Γ was restricted to be of the form $\Gamma = \frac{1}{2}(\tilde{q}A\Delta q - \Delta\tilde{q}Aq)$, where Δq is a permissible variation of q. The allowable terms of the form (9.1) were therefore $\Gamma = \frac{1}{2}\tilde{q}(A\epsilon - \tilde{\epsilon}A)q$, where ϵ is a real matrix satisfying $A\epsilon + \tilde{\epsilon}A = 0$.

For then we may take

$$\Gamma = \tilde{q} A \xi = -\tilde{\xi} A q, \tag{9.9}$$

where the ξ_a are generalized c-numbers. Then (9.7) becomes

$$(q^1; \tilde{q}^1)_- A^{11} \xi^1 = i \xi^1, \tag{9.10}$$

whose only solution is, without ambiguity, (9.8).

Other solutions can exist however in the absence of generalized c-numbers. To illustrate this point, we consider the simplest - and most interesting - case where there is only one Bose class, and one Fermi class, each self-conjugate. The commutation rules within these classes are

$$\left.\begin{aligned} [b, \tilde{b}] &= i A_B^{-1}, \\ \{f, \tilde{f}\} &= i A_F^{-1}, \end{aligned}\right\} \tag{9.11}$$

where we have denoted the variables of the Bose class by b, and those of the Fermi class by f. For the commutation rules between b and f, equation (9.2) yields

$$\left.\begin{aligned} [b_a, f_c f_d] &= 0, \\ [f_c, b_a b_b] &= 0. \end{aligned}\right\} \tag{9.12}$$

The case where the relative phase between the Bose class and the Fermi class, α_{BF}, is $+1$ is trivial, since in that case the generalized c-numbers corresponding to the Bose variables are c-numbers in the strict sense, and therefore do exist, so that

$$[b, \tilde{f}] = 0. \tag{9.13}$$

(We note in passing that any variable, for which all the relative phases are unity, must exactly commute with any variable of another class.) However, when $\alpha_{BF} = -1$ one cannot obtain any information in addition to (9.12) from a consideration of other forms for Γ; for in that case the restriction (8.10) on a permissible term in H reduces to the requirement that H be separately even in the Bose variables, and in the Fermi variables. It is clear that (9.12) possesses other solutions besides

$$\{ b, f \} = 0, \tag{9.14}$$

which corresponds to (9.8). In particular, (9.13) is a solution, although of course this solution may be ruled out by invoking the relative phase assumption. Indeed one may use the relative phase assumption to make the choice (9.8) very plausible. For, if the commutator $(q_a ; q_b)_-$ is to be a simple function of q, it is difficult to satisfy (9.4) or more generally (9.7) with any other function than the c-number (9.8). However, the meaning of the relative phase assumption becomes somewhat obscure if the commutation rules are more complicated (in a way similar to the generalized rules of Green), so that this argument has only heuristic value.

It is satisfactory that in the "standard" case, $\alpha_{BF} = 1$, the commutation rules are uniquely defined. For other cases we shall not discuss the possible extra solutions further, but will take the commutation rules to be (9.8). We remark that the "standard" case is the one in which H is least restricted; for then the only restriction is that H be even in the Fermi variables, and this restriction is necessary in every case (since $\alpha_{H,H} = +1$), though not in other cases sufficient.

10. Associated canonical variations

It is interesting to evaluate the restricted canonical variations associated, according to the action principle, with permissible variations of q and t. For simplicity, we consider only a non-singular matrix A: if A were singular, the equations below would be restricted to the independent components q^1. The variation associated with (4.9) and (4.10) is generated by (7.8), and therefore may be separated into two parts, depending on Δq and Δt respectively, and generated by

$$\left.\begin{aligned} G_q &= \tfrac{1}{4}(\tilde{q}A\Delta q - \Delta\tilde{q}Aq), \\ G_t &= -H\Delta t. \end{aligned}\right\} \qquad (10.1)$$

The variation associated with Δt may readily be found. By (8.19), it is

$$\delta_{G_t} q = -i[q, -H\Delta t] = -\dot{q}\,\Delta t, \qquad (10.2)$$

as might be expected.

The variation associated with Δq is given by (8.8) to be

$$\begin{aligned} \delta_{G_q} q &= -i[q, G_q] = A^{-1}\partial G_q \\ &= \tfrac{1}{2}\Delta q - \tfrac{1}{4}A^{-1}(\partial\Delta\tilde{q}\,;Aq)_+ . \end{aligned} \qquad (10.3)$$

Two special cases are of interest. First, if the Δq are generalized c-numbers, then

$$\delta_{G_q} q = \tfrac{1}{2}\Delta q. \qquad (10.4)$$

The factor $\frac{1}{2}$ which appears in this relation is at the basis of the contradiction obtained by Burton and Touschek (1953), since it was pointed out by them that Schwinger's original formulation of the action principle (1951), if accepted without qualification, might lead one to suppose that $\delta_{G_q} q = \Delta q$. Their resolution of this difficulty was not, however, entirely satisfactory, since they assume that the varied dynamical variables must satisfy the same equations of motion as the unvaried ones, an assumption which it is difficult to justify. The restatement of the action principle given here and, less explicitly, in Schwinger's later formulation (1953a) circumvents the difficulty in a more self-consistent manner.

The second case of interest is that of linear variations

$$\Delta q = \epsilon q, \tag{10.5}$$

where ϵ is a real matrix which decomposes into a direct sum corresponding to the class decomposition of q. Then $\partial \Delta \tilde{q} = \tilde{\epsilon}$, whence

$$\delta_{G_q} q = \tfrac{1}{2}\left(\epsilon - A^{-1}\tilde{\epsilon} A\right) q. \tag{10.6}$$

In particular, if ϵ is chosen to satisfy

$$A\epsilon + \tilde{\epsilon} A = 0, \tag{10.7}$$

then

$$\delta_{G_q} q = \Delta q, \tag{10.8}$$

so that in this case, and in this case only, the associated canonical variation is actually identical with the original variation Δq.

We note that (10.7) is just the condition that, under a variation (10.5), the commutation relations (9.8) should be invariant; that is, the condition that (10.5) should be a canonical variation. For a general variation Δq, the corresponding condition is, by (6.2),

$$A\,\Delta q\,\partial^{\dagger} + \partial\,\Delta\tilde{q}\,A = 0. \tag{10.9}$$

One can also show that all variations of the form (10.5) are permissible[1]. To do this, we first symmetrize the Hamiltonian function H, replacing each term of degree n by $1/n!$ times the sum of all $n!$ terms obtained from the original term by permutation of the factors, each permuted term being multiplied by an appropriate phase factor. This phase factor is obtained by imposing the rule that the order of factors is to be changed from the original order to the desired permutation by replacing products $q_a q_b$ is turn by $\alpha_{ab} q_b q_a$. The additional terms introduced by this procedure are at most of degree $n-2$, since the commutators (9.8) are c-numbers. Thus each term in H may be symmetrized in turn, starting with those of highest degree. When the variation is made, we separate each term containing a Δq_a into two halves, in one of which Δq_a is to be commuted to the left, and in the other to the right, using the commutation rules (9.8). In view of (10.5), the (generalized) commutator of any Δq_a with any q_b is a c-number, so the process of commuting Δq_a to one end or the other of the term will introduce additional terms of degree $n-2$. The terms which

---

1) This argument was first given by Kibble and Polkinghorne (1957).

arise from commutation of a given Δq_a with a given q_b, and which have the other $n-2$ factors in a specified order, will be of two kinds, those in which Δq_a is moved to the left, and those in which it is moved to the right. There will be an equal number of terms of the two kinds, because of the symmetrization. Moreover, it can easily be seen that the phase factors must be such that the terms of one kind exactly cancel those of the other. Thus we can express ΔH in the form (6.9), which proves that the variation is permissible. Note that it is important for this argument that the commutators of the variations with the dynamical variables are c-numbers. One may expect that there are permissible variations which are more than linear in the dynamical variables, but they must satisfy more stringent restrictions than those on the linear variations.

11. The uniqueness lemma

We now return to the discussion of the uniqueness lemma stated in section 7. For simplicity, we again restrict the discussion to the case where A is non-singular. According to the argument of section 10, we may take the permissible variations Δq to be linear variations (10.5). Moreover, we may take ϵ_{ab} to be zero except when both q_a and q_b belong to a given class. Then (7.13) gives

$$\sum_{a,b} \epsilon_{ab} \left(q_b f_a + f_a^* q_b \right) = 0,$$

where a and b now run over the indices of variables of the given class only. Since this submatrix of ϵ is arbitrary, apart from the reality condition, it follows that

$$q_b f_a + f_a^* q_b = 0 \tag{11.1}$$

for all q_a, q_b in the given class.

It is convenient to make a real linear transformation of the variables within each class in such a way that the matrix A takes on a canonical form (Schwinger 1953b), namely the direct sum of 2×2 matrices of the forms

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \tag{11.2}$$

for the Bose and Fermi classes respectively. This can always be done for a non-singular A, since it can be shown (Schwinger 1953a) that the number of independent variables of each class is necessarily even. Then any given q_b commutes (in the generalized sense) with every variable except one, p_b say, called the conjugate of q_b. The canonical commutation rules are

$$\left.\begin{aligned} [q_b, p_b] &= i, \\ \{q_b, p_b\} &= 1, \end{aligned}\right\} \tag{11.3}$$

for the Bose and Fermi classes respectively.

Now we restrict attention to a single equation (11.1), and for simplicity drop the subscripts a and b, so that we have to consider the equation

$$qf + f^* q = 0. \tag{11.4}$$

Now we can write

$$f = \sum_i g^i \ell^i, \tag{11.5}$$

where each g^i is a function of the pair of variables q and p only, and each ℓ^i is a function of other variables only and satisfies

$$\ell^{i*} = \alpha_{q,\ell^i} \ell^i. \tag{11.6}$$

Then (11.4) becomes

$$\sum_i \left(q g^i + g^{i*} q \right) \ell^i = 0. \tag{11.7}$$

Now, in view of the independence of the variables q_a, it is reasonable to assume that the only polynomial equations satisfied by the q_a are the commutation relations (9.8) and the equations derivable therefrom. It follows that if all the ℓ^i are different functions, then

$$q g^i + g^{i*} q = 0 \tag{11.8}$$

for each i. We may now drop the superscript i, and set

$$g = k + \ell, \tag{11.9}$$

where

$$\left. \begin{aligned} k^* &= \alpha_{q,g} k, \\ \ell^* &= -\alpha_{q,g} \ell. \end{aligned} \right\} \tag{11.10}$$

Then (11.8) becomes

$$(q; k)_+ + (q; \ell)_- = 0, \tag{11.11}$$

or, using (6.2),

$$(q; k)_+ + (q; p)_- \partial_{(p)} \ell = 0. \tag{11.12}$$

In order to prove the lemma we have to show that (7.16) is satisfied; that is, that the function k must vanish. If the class is self-conjugate, this presents little difficulty; for then (11.8) must hold for p in place of q, whence

$$(p; k)_+ + (p; q)_- \partial_{(q)} \ell = 0. \qquad (11.13)$$

It is not difficult to see that for a Bose class (11.12) and (11.13) can never be satisfied simultaneously by a non-zero function k (ℓ may of course be any non-zero c-number). For a Fermi class there is in fact one non-zero solution, namely $k = pq - qp$, since this function satisfies

$$(p, k)_+ = 0 = (q, k)_+ .$$

However, one may reject this solution in most cases by invoking the condition (7.15) on the relative phases of f. This condition requires in fact that f must have the same relative phases as p ; but this function k has in fact $+1$ for each of its relative phases, so that it would be necessary to find a function h, involving variables of other classes only, whose relative phases are those of p. Clearly, this will not in general be possible, although it may be so in exceptional cases.

If, however, the class is not self-conjugate, these arguments will not suffice. It is clear that non-zero solutions of (11.12) do exist, and that one can find solutions with the correct relative phases. It is possible that one may be able to reject these extra solutions by considering more general forms for the permissible variations, for example bilinear or trilinear functions of q. However, it has not been possible to find a simple proof of this.

The lemma has therefore been proved for self-conjugate classes only. If it were not valid, the equations of motion would be arbitrary to the extent of an additive function of the form of the possible solutions for k. We remark that in the "standard" case of one Bose and one Fermi class, each self-conjugate, the lemma is certainly valid.

12. Consistency conditions for systems with singular A

We define

$$H_I' \equiv H_I + \tfrac{1}{2}\left(\tilde{q}^1 M^{12} q^2 + \tilde{q}^2 M^{21} q^1\right).$$

Then the equations of motion for q^1, (7.22), are

$$A^{11}\dot{q}^1 = M^{11} q^1 + \partial^1 H_I', \tag{12.1}$$

and the equations of constraint are

$$0 = M^{22} q^2 + \partial^2 H_I'. \tag{12.2}$$

If the system is not holonomic, then only some of the non-independent variables can be eliminated by solving (12.2). The remaining non-independent variables will appear only in (12.1), so that one is left with a number of non-differential polynomial equations (12.2) relating the independent variables only. This situation is unsatisfactory, since we have explicitly assumed in section 11 that the only polynomial equations satisfied by the independent variables are the commutation relations.

Moreover, such equations will in general be actually inconsistent with the commutation relations, and it is difficult to formulate conditions on the Lagrangian which will ensure consistency. The most straightforward way of dealing with a non-holonomic system appears to be by means of the introduction of auxiliary variables, which converts the system into a holonomic one. The solutions of the extended system corresponding to the original non-holonomic system must then be selected by means of a subsidiary condition imposed on the state-vectors. For this reason, we shall restrict our attention to the case of a holonomic system.

Actually the definition of a holonomic system requires only that (12.2) should be soluble in principle for the variables q^2. However, we shall make the further assumption that they are actually soluble explicitly. Then the matrix M^{22} must be non-singular, and H_I must be at most linear in q^2, so that (12.2) may be solved to give

$$q^2 = -(M^{22})^{-1}\, \partial^2 H_I'. \qquad (12.3)$$

Now, for consistency, it is necessary that the time-derivative of the equal-time c-number commutator (9.8) should vanish. Hence we require

$$A^{11}(\dot{q}^1 ; \tilde{q}^1)_- A^{11} + A^{11}(q^1 ; \dot{\tilde{q}}^1)_- A^{11} = 0. \qquad (12.4)$$

Using (12.1), this condition is

$$(\partial^1 H_I' ; \tilde{q}^1)_- A^{11} - A^{11}(q^1 ; H_I' \partial^{1+}) = 0, \qquad (12.5)$$

since the terms in M^{11} cancel. Now, from (6.2) and (9.8), the relation (12.5) becomes

$$(\partial^1 H_I')\,\partial^{2+}(q^2;\tilde{q}^1)_-\,A^{11} = A^{11}(q^1;\tilde{q}^2)_-\,\partial^2(H_I'\partial^{1+}). \tag{12.6}$$

But, from (12.3),

$$(q^2;\tilde{q}^1)_- = -i(M^{22})^{-1}(\partial^2 H_I'\partial^{1+})(A^{11})^{-1}, \tag{12.7}$$

using (6.2) and (9.8) again. Thus (12.6) is

$$(\partial^1 H_I')\,\partial^{2+}(M^{22})^{-1}(\partial^2 H_I'\partial^{1+}) = (\partial^1 H_I'\partial^{2+})(M^{22})^{-1}\partial^2(H_I'\partial^{1+}). \tag{12.8}$$

It can be shown that this equation is satisfied if and only if the interaction terms H_I can be symmetrized in the same way as that discussed in section 10 for systems with non-singular A (see appendix B). We shall therefore impose this condition on the permissible terms in the Hamiltonian.

13. Relative phases consistent with a class of Lagrangians

In section 5, the relative phases were assumed to be given, and the class of permissible Lagrangians has been obtained in the subsequent discussion. However it is in some ways more interesting to reverse this procedure, and to enquire which sets of relative phases are consistent with a given Lagrangian or class of Lagrangians. We shall now investigate this question.

The class of allowable Lagrangians, which is to be specified in advance, must of course be chosen in such a way that <u>some</u> set of relative phases is possible. The matrices A and M must therefore be restricted by the conditions (8.13) and (8.14). This decomposition may be taken to define the classes of Bose and Fermi variables. If we make the restriction, introduced in section 12, that the equations of constraint are to be explicitly soluble, then there will be no variables which have no non-zero elements of either A or M. If this restriction is not made, then the class (Bose or Fermi) of any variables of this kind must be specified in some other way: if they appear linearly in any term of H, then it will be possible to determine their class by the requirement that H be even in the Fermi variables. Alternatively, we may choose to specify the class of such variables _a priori_. Moreover, every term in H must be restricted to be even in the Fermi variables (whether this is regarded as defining the class of certain variables, or as a restriction on H). Thus we see that any possible choice of the class of allowable Lagrangians must be a subclass of the class restricted only by these conditions, which we shall call the "_standard_" class of Lagrangians.

It follows from the discussion of section 9 that the "standard" set of relative phases, corresponding to one Bose and one Fermi class with $\alpha_{BF} = +1$, is always a possible set for any choice of the class of Lagrangians. If the class of allowable Lagrangians is further restricted to be a proper subclass of the "standard" class, then other sets of relative phases may become possible. However, it is obvious that in no such case can the relative phases be uniquely defined by the class of allowable Lagrangians, since the "standard" set of phases in particular always remains a possible alternative.

In particular, it should be noted that the relative phases cannot in general be uniquely determined by a single Lagrangian function. It therefore appears that the only way of making the action principle a unique prescription for giving equations of motion and commutation relations from the Lagrangian function alone is to require that all Lagrangians of the standard class are permissible. Then the relative phases must be the "standard" set, and the commutation relations are uniquely determined to be (9.11) and (9.13).

In the extension to relativistic quantum field theory, discussed in chapter III, this assumption will be made at the outset, and we shall not discuss relativistic systems with sets of relative phases other than the "standard" set. There is no formal difficulty in extending the discussion to cover other cases as well, but since the difficulties which arise are exactly the same as in the non-relativistic case, it does not seem that it would be of interest to do so.

CHAPTER III

SCHWINGER'S ACTION PRINCIPLE IN RELATIVISTIC QUANTUM FIELD THEORY

1. The field variables; canonical transformations

In place of the dynamical variables $q_a(t)$, the system is now to be described by a set of field variables $\chi_a(x)$, Hermitian operator functions of a space-time point. The field equations, analogous to II-(1.1), are to be of the form

$$A_\mu \partial_\mu \chi(x) = f\{\chi(x), x\}, \tag{1.1}$$

where the A_μ are constant c-number matrices. A slightly different definition of a holonomic system is required. If σ is any given spacelike surface, then the equations of constraint on σ are those of the equations (1.1) which do not involve derivatives of χ in a direction normal to σ. A system will be said to be holonomic if the equations of constraint on any given spacelike surface can always be solved for those of the field variables whose normal derivatives do not appear in the field equations.

The analogue of the completeness assumption is the following:

Postulate of locality (and completeness). There exists a set of local[1)] operator functions $\alpha_r\{\chi(x)\}$ such that the set of all α_r for all points x on any given spacelike surface

---

1) A function $\alpha\{\chi(x)\}$ will be called local if it depends only on χ at one point x, and not on $\partial_\mu\chi$ or x (unless explicitly stated).

σ constitutes a complete set of commuting observables. The basis states are then denoted by $|\alpha'\sigma\rangle$, and exactly as in the non-relativistic case we may define the transformation matrix $\langle\alpha''\sigma_1|\alpha'\sigma_0\rangle$ connecting the eigenstates on two surfaces. In fact the postulate of locality is stronger than its non-relativistic counterpart, since it does not only assert the existence of a complete set of observables of the form of functionals of χ on σ, but of a set of the form of local functions of χ on σ. It is for this reason that the name has been changed. We note that, since there is a continuously infinite number of observables α_r, the states $|\alpha'\sigma\rangle$ cannot be normalizable. Matrix elements between these states must therefore be interpreted in the sense of distribution theory (Schwartz 1950).

If, now, the system is described in terms of a set of "relabelled" variables $\bar{\chi}_a(x)$, given local functions of $\chi(x)$ which satisfy the same commutation relations for all points on any given spacelike surface σ, then (Dirac 1947) there exist unitary operators $U(\sigma)$ such that

$$\bar{\chi}(x) = U(\sigma)\,\chi(x)\,U^*(\sigma) \quad \text{for all } x\in\sigma. \tag{1.2}$$

Clearly, $U(\sigma)$ is a functional of χ on σ only. The transformation from χ to $\bar{\chi}$ will, as before, be called a <u>restricted canonical transformation</u>. If χ and $\bar{\chi}$ differ infinitesimally, then $U(\sigma) = 1 - i\,G(\sigma)$, and the change in the transformation matrix, analogous to II-(2.2), is

$$\delta^G\langle\alpha''\sigma_1|\alpha'\sigma_0\rangle = \langle\alpha''\sigma_1|\,i\,G(\sigma_1,\sigma_0)\,|\alpha'\sigma_0\rangle, \tag{1.3}$$

where

$$G(\sigma_1, \sigma_0) \equiv G(\sigma_1) - G(\sigma_0). \qquad (1.4)$$

Here $G(\sigma)$ must be a functional of χ on σ only. Moreover, from the postulate of locality, it must be an additive functional of the form

$$G(\sigma) = \int_\sigma G(x)\, d\sigma, \qquad (1.5)$$

where $G(x)$ is a local function of χ, and of its derivatives within, but not normal to, σ[1]. The function $G(x)$ may be regarded as the normal component of a vector function $G_\mu(x)$, and, if $G(x)$ is not more than linear in $\partial_\mu \chi$, this function can be chosen in such a way that, for any unit vector n_μ, $n_\mu G_\mu(x)$ is independent of $n_\mu \partial_\mu \chi$. Then

$$G(\sigma) = \int_\sigma G_\mu(x)\, d\sigma_\mu, \qquad (1.6)$$

and

$$G(\sigma_1, \sigma_0) = \int_{\sigma_0}^{\sigma_1} \partial_\mu G_\mu(x)\, dx, \qquad (1.7)$$

which evidently closely resembles II-(2.4).

Now, as in section II-3, we consider two different systems, described by $\chi(x)$ and $\bar{\chi}(x)$ respectively. If, on every surface σ of a given continuous one-parameter family of spacelike surfaces, the commutation relations satisfied by χ and those satisfied by $\bar{\chi}$ are identical, then for each surface

1) Normal derivatives of χ cannot appear in a functional of χ on σ.

σ of this family there exists a unitary operator $U(\sigma)$ such that χ and $\bar{\chi}$ are related by (1.2). If χ and $\bar{\chi}$ differ infinitesimally, then $U(\sigma) = 1 - i\,G(\sigma)$, and the change in the transformation matrix between two surfaces of this family is again given by (1.3) and (1.4). The analogue of the statement that $\dot{G}(t)$ is a function of $q(t)$ only is evidently the statement that $G(\sigma + d\sigma) - G(\sigma)$ is a functional of χ on σ only. Invoking the postulate of locality again, we see that this expression must be an additive functional of a form similar to (1.5). Hence

$$G(\sigma_1, \sigma_0) = \int_{\sigma_0}^{\sigma_1} \Gamma(x)\, dx, \tag{1.8}$$

where $\Gamma(x)$ is a local function of χ, of its derivatives within, but not normal to, the surfaces of the given family, and possibly of x itself. This relation may be compared to equation II-(3.4). The transformation from χ to $\bar{\chi}$ will be called an <u>unrestricted canonical variation</u>.

If the two systems are related by canonical variation on every surface σ (rather than merely those of a one-parameter family), then (1.8) must still hold, but $\Gamma(x)$ must be a local function of χ and x only, not of $\partial_\mu \chi$.

2. <u>The action principle</u>

As in the non-relativistic case, the action principle expresses the function $G(\sigma_1, \sigma_0)$ as the variation, in a suitably defined manner, of the <u>action integral</u>

$$W(\sigma_1, \sigma_0) = \int_{\sigma_0}^{\sigma_1} \mathcal{L}\{\chi(x), \partial_\mu \chi(x), x\}\, dx, \tag{2.1}$$

where $\mathcal{L}$ is a Hermitian function, called the <u>Lagrangian density</u>, which has the form

$$\mathcal{L}\{\chi, \partial_\mu \chi, x\} = \tfrac{1}{4}(\tilde{\chi} A_\mu \partial_\mu \chi - \partial_\mu \tilde{\chi} A_\mu \chi) - \mathcal{H}\{\chi, x\}, \tag{2.2}$$

where the A_μ are four constant c-number matrices, and

$$\mathcal{H}\{\chi, x\} = \tfrac{1}{2}\tilde{\chi} M \chi + \mathcal{H}_I\{\chi, x\}. \tag{2.3}$$

Hermiticity requires

$$\left.\begin{aligned} A_\mu^\dagger &= -A_\mu, \\ M^\dagger &= M, \\ \mathcal{H}^*\{\chi, x\} &= \mathcal{H}\{\chi, x\}. \end{aligned}\right\} \tag{2.4}$$

The action principle (Schwinger 1953a) will again be stated in two parts, as follows.

<u>Action principle for canonical variations</u>. If two infinitesimally differing systems are related by a canonical variation on every spacelike surface of a given family, then the corresponding function $G(\sigma_1, \sigma_0)$, where σ_0 and σ_1 are surfaces of this family, is equal to the difference of the action integrals between σ_0 and σ_1 for the two systems. Conversely, if the Lagrangian density is changed by the addition of any function of χ, of its derivatives within, but not normal to, the surfaces of the given family, and of x, which satisfies the requirements (to be specified) for a permissible term in the Lagrangian,

and which does not change the equations of constraint on the surfaces of the given family, then the new system is related to the old by a canonical variation such that the corresponding function $G(\sigma_1, \sigma_0)$ is equal to the change in the action integral between these limits.

The symbolic statement of this principle is

$$G(\sigma_1, \sigma_0) = \delta W(\sigma_1, \sigma_0) \equiv \int_{\sigma_0}^{\sigma_1} \left[\bar{\mathcal{L}}\{\chi, \partial_\mu \chi, x\} - \mathcal{L}\{\chi, \partial_\mu \chi, x\} \right] dx, \tag{2.5}$$

where $\mathcal{L}$ and $\bar{\mathcal{L}}$ are the Lagrangian densities for the two systems.

<u>Action principle for permissible variations of χ and x.</u> If an arbitrary variation of the space-time label

$$x_\mu \longrightarrow \bar{x}_\mu = x_\mu + \Delta x_\mu, \tag{2.6}$$

satisfying

$$\partial_\mu \Delta x_\nu + \partial_\nu \Delta x_\mu = 0 \tag{2.7}$$

on σ_0 and σ_1, is made, together with an arbitrary permissible variation of the field variables

$$\chi(x) \rightarrow \bar{\chi}(\bar{x}) = \chi(x) + \Delta\chi(x) \tag{2.8}$$

where $\Delta\chi$ is a local function of χ and x only, then the resulting change in the action integral between σ_0 and σ_1 is equal to the function $G(\sigma_1, \sigma_0)$ corresponding to some restricted canonical variation.

The formal statement of this part of the action principle is

$$G(\sigma_1, \sigma_0) = \delta W(\sigma_1, \sigma_0) \equiv \int_{\bar{\sigma}_0}^{\bar{\sigma}_1} \mathcal{L}\{\bar{\chi}(\bar{x}), \bar{\partial}_\mu \bar{\chi}(\bar{x}), \bar{x}\}\, d\bar{x} - \int_{\sigma_0}^{\sigma_1} \mathcal{L}\{\chi(x), \partial_\mu \chi(x), x\}\, dx. \tag{2.9}$$

The relative phases may be introduced exactly as before, and the same notation will be used for commutators and derivatives with respect to field variables. In conformity with the intention expressed in section II-13, we shall assume that the relative phases are the "standard" set,

$$\alpha_{ab} = \begin{cases} -1 \text{ , if } \chi_a \text{ and } \chi_b \text{ are both Fermi variables;} \\ +1 \text{ , otherwise.} \end{cases} \qquad (2.10)$$

The class of permissible variations will be defined by exactly the same conditions as those stated in section II-6, so that in particular, for a permissible variation $\Delta\chi$ with no change of space-time label,

$$\Delta \mathcal{H} = \frac{1}{2}\{\Delta\tilde{\chi}(\partial\mathcal{H}) + (\mathcal{H}\partial^{+})\Delta\chi\}. \qquad (2.11)$$

We note that the condition (2.7), which has no analogue in the non-relativistic case, is merely the condition that the variation (2.6) should be locally a Lorentz transformation on the bounding surfaces (see appendix A-1).

3. The field equations

We now consider the variations (2.6) and (2.8). We define

$$\delta\chi(x) \equiv \bar{\chi}(x) - \chi(x), \qquad (3.1)$$

and, for any function f of χ and x,

$$\begin{aligned} \delta f(x) &\equiv f\{\bar{\chi}(x), x\} - f\{\chi(x), x\} \\ &= f\{\chi + \delta\chi, x\} - f\{\chi, x\}, \end{aligned} \qquad (3.2)$$

and

$$\Delta f(x) \equiv f\{\bar{\chi}(\bar{x}), \bar{x}\} - f\{\chi(x), x\}$$

$$= f\{\chi + \Delta\chi, x + \Delta x\} - f\{\chi, x\}. \tag{3.3}$$

Hence

$$\Delta f = \delta f + \Delta x_\mu \partial_\mu f. \tag{3.4}$$

Moreover,

$$\partial_\mu \delta f = \delta \partial_\mu f, \tag{3.5}$$

whence

$$\partial_\mu \Delta f = \Delta \partial_\mu f + (\partial_\mu \Delta x_\nu)\partial_\nu f. \tag{3.6}$$

Now, using (2.11), (3.3) and (3.6),

$$\Delta \mathcal{L} = \tfrac{1}{4}\{\Delta\tilde{\chi} A_\mu \partial_\mu \chi - \partial_\mu \tilde{\chi} A_\mu \Delta\chi\} + \tfrac{1}{4}\{\tilde{\chi} A_\mu \partial_\mu \Delta\chi - \partial_\mu \Delta\tilde{\chi} A_\mu \chi\}$$

$$- \tfrac{1}{4}(\partial_\mu \Delta x_\nu)\{\tilde{\chi} A_\mu \partial_\nu \chi - \partial_\nu \tilde{\chi} A_\mu \chi\} - \Delta\mathcal{H} \tag{3.7}$$

$$= \tfrac{1}{2}\{\Delta\tilde{\chi} A_\mu \partial_\mu \chi - \partial_\mu \tilde{\chi} A_\mu \Delta\chi\} + \tfrac{1}{4}\partial_\mu\{\tilde{\chi} A_\mu \Delta\chi - \Delta\tilde{\chi} A_\mu \chi\}$$

$$- \tfrac{1}{4}(\partial_\mu \Delta x_\nu)\{\tilde{\chi} A_\mu \partial_\nu \chi - \partial_\nu \tilde{\chi} A_\mu \chi\}$$

$$- \tfrac{1}{2}\{\Delta\tilde{\chi}(\partial\mathcal{H}) + (\mathcal{H}\partial^{+})\Delta\chi\} - \Delta x_\nu (\partial_\nu \mathcal{H})_\chi, \tag{3.8}$$

where $(\partial_\nu \mathcal{H})_\chi$ is the partial derivative of $\mathcal{H}$ with respect to x_ν, holding χ fixed[1]. This is clearly analogous to $\partial H/\partial t$ in the non-relativistic case.

Then, from (2.9),

---

1) The derivatives $\partial_\nu \mathcal{H}$ and $(\partial_\nu \mathcal{H})_\chi$ have sometimes been denoted by $d\mathcal{H}/dx_\nu$ and $\partial\mathcal{H}/\partial x_\nu$ respectively (see, for example, Schweber, Bethe and de Hoffmann 1955). This notation, however, obscures the fact that the $\partial_\nu \mathcal{H}$ are partial derivatives.

$$\delta W(\sigma_1, \sigma_0) = \int_{\sigma_0}^{\sigma_1} \left[\mathcal{L}\{\bar{\chi}(\bar{x}), \bar{\partial}_\mu \bar{\chi}(\bar{x}), \bar{x}\} \left\| \frac{\partial \bar{x}_\mu}{\partial x_\nu} \right\| - \mathcal{L}\{\chi(x), \partial_\mu \chi(x), x\} \right] dx$$

$$= \int_{\sigma_0}^{\sigma_1} \left[\Delta \mathcal{L} + \mathcal{L}\, \partial_\mu \Delta x_\mu \right] dx, \qquad (3.9)$$

whence, by (3.8),

$$\delta W(\sigma_1, \sigma_0) = \int_{\sigma_0}^{\sigma_1} \Big[\tfrac{1}{2} \Delta \tilde{\chi} (A_\mu \partial_\mu \chi - \partial \mathcal{H}) - \tfrac{1}{2} (\partial_\mu \tilde{\chi} A_\mu + \mathcal{H} \partial^+) \Delta \chi$$

$$+ \tfrac{1}{4} \partial_\mu (\tilde{\chi} A_\mu \Delta \chi - \Delta \tilde{\chi} A_\mu \chi) + (\partial_\mu \Delta x_\nu) T_{\mu\nu} - \Delta x_\nu (\partial_\nu \mathcal{H})_\chi \Big] dx,$$

where

$$T_{\mu\nu} = \mathcal{L}\, \delta_{\mu\nu} - \tfrac{1}{4} (\tilde{\chi} A_\mu \partial_\nu \chi - \partial_\nu \tilde{\chi} A_\mu \chi). \qquad (3.10)$$

Thus

$$\delta W(\sigma_1, \sigma_0) = \tfrac{1}{2} \int_{\sigma_0}^{\sigma_1} \{ \Delta \tilde{\chi} (A_\mu \partial_\mu \chi - \partial \mathcal{H}) - (\partial_\mu \tilde{\chi} A_\mu + \mathcal{H} \partial^+) \Delta \chi \} dx$$

$$- \int_{\sigma_0}^{\sigma_1} \Delta x_\nu \{ \partial_\mu T_{\mu\nu} + (\partial_\nu \mathcal{H})_\chi \} dx + G(\sigma_1) - G(\sigma_0), \qquad (3.11)$$

where

$$G(\sigma) = \int_\sigma d\sigma_\mu \{ \tfrac{1}{4} (\tilde{\chi} A_\mu \Delta \chi - \Delta \tilde{\chi} A_\mu \chi) + T_{\mu\nu} \Delta x_\nu \}. \qquad (3.12)$$

Just as in the non-relativistic case, $\delta W(\sigma_1, \sigma_0)$ must be of the form (1.7), and $\Delta \chi$ and Δx_μ may be chosen independently at different points. Hence, using an argument similar to that of section II-7, we find that the field equations, the analogues of II-(7.21), are

$$A_\mu \partial_\mu \chi = \partial \mathcal{H} , \tag{3.13}$$

and the conservation equations, the analogues of II-(7.10), are

$$\partial_\mu T_{\mu\nu} = -(\partial_\nu \mathcal{H})_\chi . \tag{3.14}$$

The argument about relative phases goes through as before, with only minor alteration. One can show that the matrices A_μ and M must decompose according to

$$A_\mu = A_{\mu B} \oplus A_{\mu F} , \qquad M = M_B \oplus M_F , \tag{3.15}$$

where

$$\left.\begin{aligned} A^*_{\mu B} &= A_{\mu B} = -\tilde{A}_{\mu B} , & M^*_B &= M_B = \tilde{M}_B , \\ -A^*_{\mu F} &= A_{\mu F} = \tilde{A}_{\mu F} , & -M^*_F &= M_F = -\tilde{M}_F . \end{aligned}\right\} \tag{3.16}$$

The requirement for a permissible term in $\mathcal{H}$ is simply evenness in the Fermi variables, except for some extra conditions associated with the consistency of the equations of constraint, which are discussed in section 4.

The separation of the equations of motion (3.13) into equations for the independent variables, and equations of constraint, is rather more complicated than in the non-relativistic case. Let σ be a spacelike surface, and let n_μ be the unit (timelike advanced) normal to σ at the point x. We set

$$A \equiv n_\mu A_\mu = \begin{bmatrix} A^{11} & 0 \\ 0 & 0 \end{bmatrix} , \qquad \chi = \begin{bmatrix} \chi^1 \\ \chi^2 \end{bmatrix} , \tag{3.17}$$

where A^{11} is non-singular. There is no loss of generality in taking the x_0-axis to be in the direction of n_μ. Then the field equations for the independent variables χ^1 are

$$A_0^{11}\partial_0\chi^1 = A_k^{11}\partial_k\chi^1 + A_k^{12}\partial_k\chi^2 + \partial^1\mathcal{H}, \tag{3.18}$$

and the equations for the dependent variables χ^2 are equations of constraint

$$0 = A_k^{21}\partial_k\chi^1 + A_k^{22}\partial_k\chi^2 + \partial^2\mathcal{H}. \tag{3.19}$$

Note that unless σ happens to be a plane surface, the independent variables at one point on σ may constitute a different set from those at another point on σ.

In order to ensure that the system is holonomic, we must impose two conditions, namely that the matrices A_k^{22} should vanish, and that the matrix M^{22} should be non-singular. For then (3.19) may be expressed in the form

$$\chi^2 = -(M^{22})^{-1}\left(A_k^{21}\partial_k\chi^1 + \partial^2\mathcal{H}_I'\right), \tag{3.20}$$

where

$$\mathcal{H}_I' \equiv \mathcal{H}_I + \tfrac{1}{2}\left(\tilde{\chi}^1 M^{12}\chi^2 + \tilde{\chi}^2 M^{21}\chi^1\right). \tag{3.21}$$

If $\mathcal{H}_I$ is not more than linear in χ^2, then (3.20) is an explicit solution for χ^2 in terms of χ^1. If, on the other hand, $\mathcal{H}_I$ is more than linear in χ^2, then (3.20) cannot in general be solved explicitly for χ^2, but it may always be solved "in principle", so that the system is certainly holonomic.

4. Commutation relations

As in section II-8, we may add to $\mathcal{L}$ the term[1)]

1) See appendix A-1 for the definition of $\delta_\mu[\sigma]$.

$$\lambda \Gamma_\mu \{ \chi(x), x \} \, \delta_\mu[\sigma], \tag{4.1}$$

where Γ_μ is a Hermitian function satisfying the requirements for a permissible term in $\mathcal{L}$, and λ is an infinitesimal real number. The new equations of motion are

$$A_\mu \partial_\mu \chi = \partial \mathcal{H} - \lambda \, \partial \Gamma_\mu \, \delta_\mu[\sigma]. \tag{4.2}$$

We must impose the condition that the equations of constraint on σ are unaltered by this addition. To do this, we assume that Γ_μ is independent of the variables χ^2 defined in (3.17). In contrast to the non-relativistic case, it is not now sufficient to assume a relation of the form II-(8.4). This is due to the presence of derivatives (within σ) in the equations of constraint. On the other hand, it is now possible to consider more general functions Γ_μ which involve both derivatives of χ within σ and the dependent variables χ^2. Such functions will be considered separately in section 6. By the same argument as in section II-8, we find that

$$n_\nu A_\nu \left[\chi(x), \int_\sigma \Gamma_\mu(x') \, d\sigma'_\mu \right] = i n_\mu \partial \Gamma_\mu(x), \text{ for } x \in \sigma, \tag{4.3}$$

whence

$$A \left[\chi(x), \Gamma(x') \right] = i \partial \Gamma(x) \, \delta^\sigma(x - x') \text{ for } x, x' \in \sigma. \tag{4.4}$$

Here n_μ is the unit normal to σ at x, and $A \equiv n_\mu A_\mu$ and $\Gamma \equiv n_\mu \Gamma_\mu$.

The commutation relations follow in exactly the same way as before, and would in the general case be subject to the same ambiguities. With the assumption (2.10), they are uniquely given to be

$$(\chi^1(x);\ \tilde{\chi}^1(x'))_- = i\,(A^{11})^{-1}\,\delta^\sigma(x-x') \qquad \text{for } x, x' \in \sigma. \tag{4.5}$$

Writing φ and ψ for the subvectors of Bose and Fermi variables respectively, the relations (4.5) are

$$\left.\begin{aligned} [\varphi^1(x), \tilde{\varphi}^1(x')] &= i(A^{11}_B)^{-1}\,\delta^\sigma(x-x'),\\ \{\psi^1(x), \tilde{\psi}^1(x')\} &= i(A^{11}_F)^{-1}\,\delta^\sigma(x-x'),\\ [\varphi^1(x), \tilde{\psi}^1(x')] &= 0, \end{aligned}\right\} \qquad \text{for } x, x' \in \sigma. \tag{4.6}$$

For a holonomic system with $\mathcal{H}_I$ not more than linear in χ^2, the commutation relations for χ^2 can be found at once from (3.20) and (4.5). Indeed, if the x_0-axis is taken to be in the direction of n_μ, then, for $x, x' \in \sigma$,

$$i\,M^{22}(\chi^2(x);\chi^1(x'))_-\,A^{11}_0 = A^{21}_k\,\partial_k\delta^\sigma(x-x') + \partial^2\mathcal{H}'_I\,\partial^{1+}\,\delta^\sigma(x-x'). \tag{4.7}$$

Using the argument of section II-12, it is then possible to show that the field equations and commutation relations are only consistent if

$$(\partial^1\mathcal{H}'_I)\,\partial^{2+}(M^{22})^{-1}(\partial^2\mathcal{H}'_I\,\partial^{1+}) = (\partial^1\mathcal{H}'_I\,\partial^{2+})\,(M^{22})^{-1}\,\partial^2(\mathcal{H}'_I\,\partial^{1+}). \tag{4.8}$$

Since this relation is identical with II-(12.8), it follows that it is satisfied if and only if $\mathcal{H}_I$ can be symmetrized[1]. We therefore assume that this can be done.

5. Behaviour under Lorentz transformation

Consider a Lorentz transformation Λ-(1.3) of the points of

1) See appendix B for a fuller discussion of this point.

space-time. It is open to us to specify in any convenient manner a corresponding relabelling of the components of χ; that is, to specify which values of $\bar{\chi}(\bar{x})$ in the new coordinate frame will be regarded as describing the same field as the $\chi(x)$. We choose this relabelling to be purely a homogeneous real linear transformation

$$\chi(x) \rightarrow \bar{\chi}(\bar{x}) = L(a_{\mu\nu})\chi(x), \tag{5.1}$$

where $L(a_{\mu\nu})$ is a real non-singular matrix depending only on the rotational part of the Lorentz transformation. This matrix will be determined by the requirements that the free Lagrangian be invariant under the simultaneous transformations A-(1.3) and (5.1), and that the matrices $L(a_{\mu\nu})$ furnish a (possibly two-valued[1]) representation of the Lorentz group. It is of course assumed that matrices satisfying these requirements exist.

One requires

$$\left.\begin{aligned} \tilde{L}(a_{\mu\nu})\, A_\mu\, L(a_{\mu\nu}) &= a_{\mu\nu} A_\nu, \\ \tilde{L}(a_{\mu\nu})\, M\, L(a_{\mu\nu}) &= M, \end{aligned}\right\} \tag{5.2}$$

and

$$L(a_{\mu\nu})\, L(a'_{\mu\nu}) = \pm L(a_{\mu\rho} a'_{\rho\nu}). \tag{5.3}$$

The reason for the choice (5.1) is merely that this is the simplest transformation under which the free Lagrangian

---

1) The representation cannot be more than two-valued, since the representative matrices are real and have determinants ± 1.

can be made to be invariant.

For an infinitesimal Lorentz transformation A-(1.5), we can write

$$L(\delta_{\mu\nu} + \alpha_{\mu\nu}) = 1 + \tfrac{1}{2}\alpha_{\mu\nu} S_{\mu\nu} \tag{5.4}$$

where the $S_{\mu\nu}$ are real matrices which may be chosen, in view of A-(1.6), to satisfy

$$S_{\mu\nu} + S_{\nu\mu} = 0. \tag{5.5}$$

The conditions (5.2) yield

$$\left.\begin{aligned} \tilde{S}_{\nu\rho} A_\mu + A_\mu S_{\nu\rho} &= \delta_{\mu\nu} A_\rho - \delta_{\mu\rho} A_\nu, \\ \tilde{S}_{\nu\rho} M + M S_{\nu\rho} &= 0, \end{aligned}\right\} \tag{5.6}$$

and (5.3) gives

$$[S_{\mu\nu}, S_{\rho\sigma}] = -\delta_{\mu\rho} S_{\nu\sigma} + \delta_{\mu\sigma} S_{\nu\rho} + \delta_{\nu\rho} S_{\mu\sigma} - \delta_{\nu\sigma} S_{\mu\rho}. \tag{5.7}$$

It can be shown (Harish-Chandra 1947) that there exists only one set of matrices $S_{\mu\nu}$ satisfying (5.5), (5.6) and (5.7).

Now consider the effect of Lorentz transformation on $\mathcal{H}_I$. Under the infinitesimal transformation A-(1.5) and (5.4), it transforms according to

$$\mathcal{H}_I\{\chi(x), x\} \rightarrow \mathcal{H}_I\{\bar{\chi}(\bar{x}), \bar{x}\} = \mathcal{H}_I\{(1 + \tfrac{1}{2}\alpha_{\mu\nu} S_{\mu\nu})\chi(x), \bar{x}\}. \tag{5.8}$$

The conditions for invariance under all infinitesimal Lorentz transformations may therefore be stated as

$$\left.\begin{aligned} \mathcal{H}_I\{\chi + \tfrac{1}{2}\alpha_{\mu\nu} S_{\mu\nu}\chi, x\} &= \mathcal{H}_I\{\chi, x\}, \\ \mathcal{H}_I\{\chi, x + \alpha\} &= \mathcal{H}_I\{\chi, x\}. \end{aligned}\right\} \tag{5.9}$$

If these conditions are not satisfied, it is useful to define the operator $(\partial_{\nu\rho}\mathcal{H}_I)_\chi$ by

$$\mathcal{H}_I\{\chi + \tfrac{1}{2}\alpha_{\mu\nu}S_{\mu\nu}\chi,\ x\} - \mathcal{H}_I\{\chi, x\} = \tfrac{1}{2}\alpha_{\mu\nu}(\partial_{\mu\nu}\mathcal{H}_I)_\chi. \tag{5.10}$$

Evidently, the conditions (5.9) are

$$(\partial_{\mu\nu}\mathcal{H}_I)_\chi = 0, \tag{5.11}$$

and

$$(\partial_\mu\mathcal{H}_I)_\chi = 0. \tag{5.12}$$

If both conditions are satisfied, the system will be said to be Lorentz-invariant.

6. Intrinsic and induced variations

We now return to the consideration of the general variations (2.6) and (2.8). It is convenient to separate the variation $\Delta\chi$ into two parts,

$$\Delta\chi = \Delta_0\chi + \Delta_x\chi, \tag{6.1}$$

in such a way that $\Delta_x\chi$ corresponds to the relabelling of χ components under a Lorentz transformation, discussed in section 5. This separation is a matter of convention, and may be made in any convenient manner. We choose to take

$$\Delta_x\chi = \tfrac{1}{2}(\partial_\nu\Delta x_\mu)\,S_{\mu\nu}\chi, \tag{6.2}$$

which reduces to (5.1) and (5.4) for an infinitesimal Lorentz transformation. The variation $\Delta_0\chi$ will be called the intrinsic field variation, and the variation $\Delta_x\chi$ the induced field variation.

The generator $G(\sigma)$, defined by (3.12), may correspondingly be split into two parts,

$$G(\sigma) = G_0(\sigma) + G_x(\sigma), \tag{6.3}$$

where

$$G_0(\sigma) = \frac{1}{4}\int d\sigma_\mu \left(\tilde{\chi} A_\mu \Delta_0 \chi - \Delta_0 \tilde{\chi} A_\mu \chi \right), \tag{6.4}$$

and

$$G_x(\sigma) = \int d\sigma_\mu \left[T_{\mu\nu} \Delta x_\nu + \frac{1}{2} \Sigma_{\mu\nu\rho} (\partial_\rho \Delta x_\nu) \right], \tag{6.5}$$

in which

$$\Sigma_{\mu\nu\rho} \equiv \frac{1}{4} \tilde{\chi} (A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu) \chi. \tag{6.6}$$

Note that, by (5.5),

$$\Sigma_{\mu\nu\rho} + \Sigma_{\nu\mu\rho} = 0. \tag{6.7}$$

We can now find the associated canonical variations generated by $\mathbf{G}(\sigma)$. The variation associated to the intrinsic field variation is generated by (6.4). Since the integrand of (6.4) is a function of χ only, we may use (4.3) to evaluate the variation. This gives

$$\delta_{G_0} \chi^1 \equiv -i [\chi^1, G_0] = \frac{1}{2} \Delta_0 \chi^1 - \frac{1}{4} (A^{11})^{-1} (\partial^1 \Delta \tilde{\chi}^1; A^{11} \chi^1)_+, \tag{6.8}$$

exactly as in II-(10.3). For generalized c-numbers,

$$\delta_{G_0} \chi^1 = \frac{1}{2} \Delta_0 \chi^1, \tag{6.9}$$

whereas, if

$$\Delta_0 \chi = \epsilon \chi, \tag{6.10}$$

where ϵ is a real matrix decomposing according to $\epsilon = \epsilon_B \oplus \epsilon_F$, and satisfying

$$A_\mu \epsilon + \tilde{\epsilon} A_\mu = 0, \tag{6.11}$$

then

$$\delta_{G_0} \chi^1 = \epsilon^{11} \chi^1 = \Delta_0 \chi^1. \tag{6.12}$$

One can also show, as in section II-10, that variations of the form (6.10) are always permissible.

To find the canonical variations associated with the space-time variations and induced field variations, we need to consider more general functions Γ_μ. This will be done in the next section.

7. Energy-momentum operators

We now consider the addition to $\mathcal{L}$ of a term

$$\Lambda = \lambda \Gamma_\mu \delta_\mu[\sigma], \tag{7.1}$$

where Γ_μ is a Hermitian function of χ, $\partial_\mu \chi$ and x, such that for any unit vector n_μ the function $n_\mu \Gamma_\mu$ does not involve the derivatives $n_\mu \partial_\mu \chi$. We also assume that Γ_μ is chosen in such a way that the equations of constraint on σ are unaltered. It is not easy to find a general condition on Γ_μ which will ensure that this is so, and instead we shall verify the condition for the cases considered.

We now take

$$\lambda \Gamma_\mu = T_{\mu\nu} \Delta x_\nu + \tfrac{1}{2} \Sigma_{\mu\nu\rho} \partial_\rho \Delta x_\nu. \tag{7.2}$$

To evaluate the change in the equations of motion, we have to consider the effect on Λ of a permissible variation of χ. (It is sufficient to take no space-time variation). Then, from (3.10) and (3.6),

$$\Delta T_{\mu\nu} = \tfrac{1}{2}\Delta\tilde{\chi}\left\{\delta_{\mu\nu}(A_\rho\partial_\rho\chi - \partial\mathcal{H}) - A_\mu\partial_\nu\chi\right\}$$

$$- \tfrac{1}{2}\left\{\delta_{\mu\nu}(\partial_\rho\tilde{\chi}A_\rho + \mathcal{H}\partial^+) - \partial_\nu\tilde{\chi}A_\mu\right\}\Delta\chi$$

$$+ \tfrac{1}{4}\partial_\rho\left\{\delta_{\mu\nu}(\tilde{\chi}A_\rho\Delta\chi - \Delta\tilde{\chi}A_\rho\chi) - \delta_{\nu\rho}(\tilde{\chi}A_\mu\Delta\chi - \Delta\tilde{\chi}A_\mu\chi)\right\},$$

and, from (6.6),

$$\Delta\Sigma_{\mu\nu\rho} = \tfrac{1}{4}\Delta\tilde{\chi}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho}A_\mu)\chi + \tfrac{1}{4}\tilde{\chi}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho}A_\mu)\Delta\chi.$$

Hence, using the notation A-(1.1),

$$\Delta\Lambda = \left\{\Delta T_{\mu\nu}\Delta x_\nu + \tfrac{1}{2}\Delta\Sigma_{\mu\nu\rho}\partial_\rho\Delta x_\nu\right\}\delta_\mu[\sigma]$$

$$= \tfrac{1}{2}\Delta\tilde{\chi}\Big\{\left[\delta_{\mu\nu}(A_\rho\partial_\rho\chi - \partial\mathcal{H}) - A_\mu\partial_\nu\chi\right]\Delta x_\nu\delta_\mu[\sigma]$$

$$+ \tfrac{1}{4}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho}A_\mu)\chi(\partial_\rho\Delta x_\nu)\delta_\mu[\sigma] + \delta_{\nu[\mu}A_{\rho]}\chi\,\partial_\rho(\Delta x_\nu\delta_\mu[\sigma])\Big\}$$

$$- \tfrac{1}{2}\Big\{\left[\delta_{\mu\nu}(\partial_\rho\tilde{\chi}A_\rho + \mathcal{H}\partial^+) - \partial_\nu\tilde{\chi}A_\mu\right]\Delta x_\nu\delta_\mu[\sigma]$$

$$- \tfrac{1}{4}\tilde{\chi}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho}A_\mu)(\partial_\rho\Delta x_\nu)\delta_\mu[\sigma] + \tilde{\chi}\delta_{\nu[\mu}A_{\rho]}\partial_\rho(\Delta x_\nu\delta_\mu[\sigma])\Big\}\Delta\chi$$

$$+ \tfrac{1}{2}\partial_\rho\left\{(\tilde{\chi}\delta_{\nu[\mu}A_{\rho]}\Delta\chi - \Delta\tilde{\chi}\delta_{\nu[\mu}A_{\rho]}\chi)\Delta x_\nu\delta_\mu[\sigma]\right\}.$$

Using the relation A-(1.8), the new equations of motion are, therefore,

$$A_\rho\partial_\rho\chi = \partial\mathcal{H} - \left[\delta_{\mu\nu}(A_\rho\partial_\rho\chi - \partial\mathcal{H}) - A_\mu\partial_\nu\chi\right]\Delta x_\nu\delta_\mu[\sigma]$$

$$- \left\{\tfrac{1}{4}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho}A_\mu) + \delta_{\nu[\mu}A_{\rho]}\right\}\chi(\partial_\rho\Delta x_\nu)\delta_\mu[\sigma]. \qquad (7.3)$$

Now, using (5.6),

$$\frac{1}{4}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu) + \delta_{\nu[\mu} A_{\rho]}$$

$$= \frac{1}{4}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu) + \frac{1}{2}(\delta_{\mu[\nu} A_{\rho]} + \delta_{\mu\{\nu} A_{\rho\}}) - \frac{1}{2}\delta_{\nu\rho} A_\mu$$

$$= \frac{1}{4}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu) + \frac{1}{4}(A_\mu S_{\nu\rho} + \tilde{S}_{\nu\rho} A_\mu) + \frac{1}{2}(\delta_{\mu\{\nu} A_{\rho\}} - \delta_{\nu\rho} A_\mu)$$

$$= \frac{1}{2} A_\mu S_{\nu\rho} + \frac{1}{2}(\delta_{\mu\{\nu} A_{\rho\}} - \delta_{\nu\rho} A_\mu), \tag{7.4}$$

whence (7.3) becomes

$$A_\rho \partial_\rho \chi = \partial \mathcal{H} - (A_\rho \partial_\rho \chi - \partial \mathcal{H}) \Delta x_\mu \delta_\mu[\sigma]$$
$$+ A_\mu [\Delta x_\nu \partial_\nu \chi - \frac{1}{2} S_{\nu\rho} \chi \, \partial_{[\nu} \Delta x_{\rho]}] \delta_\mu[\sigma]$$
$$- \frac{1}{2}(\delta_{\mu\{\nu} A_{\rho\}} - \delta_{\nu\rho} A_\mu) \chi (\partial_{\{\nu} \Delta x_{\rho\}}) \delta_\mu[\sigma]. \tag{7.5}$$

Thus it is clear that the equations of constraint are unchanged if and only if the condition (2.7) is satisfied on σ. When this condition is satisfied, the solutions of (3.13) and (7.5) are related by a canonical variation. Imposing a suitable boundary condition, we easily find from the action principle that the generator of this variation is $G_x(\sigma)$. We may then evaluate the discontinuity on crossing σ in the same way as before, and find the relation

$$A \delta_{G_x} \chi \equiv -i A[\chi, G_x(\sigma)] = -A\{\Delta x_\nu \partial_\nu \chi - S_{\nu\rho} \chi \, \partial_\nu \Delta x_\rho\}. \tag{7.6}$$

If the space-time transformation on σ is the infinitesimal Lorentz transformation

$$\Delta x_\nu = \alpha_\nu + \alpha_{\nu\rho} x_\rho, \tag{7.7}$$

then we can write

$$G_x(\sigma) = \alpha_\nu P_\nu(\sigma) + \frac{1}{2} \alpha_{\nu\rho} J_{\nu\rho}(\sigma), \tag{7.8}$$

where

$$\left.\begin{aligned} P_\nu(\sigma) &= \int d\sigma_\mu T_{\mu\nu}, \\ J_{\nu\rho}(\sigma) &= \int d\sigma_\mu \left(\Sigma_{\mu\nu\rho} + T_{\mu\nu}x_\rho - T_{\mu\rho}x_\nu\right). \end{aligned}\right\} \tag{7.9}$$

The operators $P_\nu(\sigma)$ and $J_{\nu\rho}(\sigma)$ are known as the linear and angular momentum operators on σ respectively. Equating coefficients in (7.6), we find

$$\left.\begin{aligned} i\left[X^1, P_\nu(\sigma)\right] &= \partial_\nu X^1, \\ i\left[X^1, J_{\nu\rho}(\sigma)\right] &= x_\rho \partial_\nu X^1 - x_\nu \partial_\rho X^1 + \sum_{j=1}^{2} S^{1j}_{\nu\rho} X^j. \end{aligned}\right\} \tag{7.10}$$

The first of these relations incorporates the equations of motion in their Hamiltonian form. Note that the correct Hamiltonian is not $\int_\sigma \mathcal{H}\, d\sigma$ as might be expected, but $-n_\mu P_\mu$; or, in other words, $\mathcal{H}$ is not a Hamiltonian density.

8. Conservation laws

From (3.14) and (7.9), one finds

$$P_\nu(\sigma_1) - P_\nu(\sigma_0) = -\int_{\sigma_0}^{\sigma_1} (\partial_\nu \mathcal{H})_X \, dx, \tag{8.1}$$

the well-known conservation law for linear momentum. The right side of (8.1) may be regarded as the momentum transferred from the source functions (the explicitly x-dependent terms in $\mathcal{L}$) to the field. Note that linear momentum is strictly conserved if and only if (5.12) is satisfied.

In order to find a similar relation for angular momentum, we consider the expression

$$\partial_\mu \Sigma_{\mu\nu\rho} + T_{\rho\nu} - T_{\nu\rho}. \tag{8.2}$$

Using the definitions (3.10) and (6.6), and the relation (5.6), (8.2) is equal to

$$\begin{aligned}
&\tfrac{1}{4}\,\partial_\mu \tilde{\chi}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu)\chi + \tfrac{1}{4}\tilde{\chi}(A_\mu S_{\nu\rho} - \tilde{S}_{\nu\rho} A_\mu)\partial_\mu \chi \\
&\quad - \tfrac{1}{4}\tilde{\chi}(\delta_{\mu\nu} A_\rho - \delta_{\mu\rho} A_\nu)\partial_\mu \chi + \tfrac{1}{4}\partial_\mu \tilde{\chi}(\delta_{\mu\nu} A_\rho - \delta_{\mu\rho} A_\nu)\chi \\
&\quad = \tfrac{1}{2}\{\partial_\mu \tilde{\chi} A_\mu S_{\nu\rho}\chi - \tilde{\chi}\tilde{S}_{\nu\rho} A_\mu \partial_\mu \chi\} \\
&\quad = -\tfrac{1}{2}\,\{(\mathcal{H}\partial^+) S_{\nu\rho}\chi + \tilde{\chi}\tilde{S}_{\nu\rho}(\partial\mathcal{H})\},
\end{aligned} \tag{8.3}$$

by the field equations (3.13). But the variation $\Delta_x \chi$, defined by (6.2), is linear, and therefore permissible. Hence, recalling the definition (5.10), one sees that

$$(\partial_{\nu\rho}\mathcal{H})_\chi = \tfrac{1}{2}\,\{\tilde{\chi}\tilde{S}_{\nu\rho}(\partial\mathcal{H}) + (\mathcal{H}\partial^+) S_{\nu\rho}\chi\}. \tag{8.4}$$

Then, from (3.14) and (7.9),

$$\begin{aligned}
J_{\nu\rho}(\sigma_1) - J_{\nu\rho}(\sigma_0) &= \int_{\sigma_0}^{\sigma_1} \left[\partial_\mu \Sigma_{\mu\nu\rho} + T_{\rho\nu} - T_{\nu\rho} + (\partial_\mu T_{\mu\nu})x_\rho - (\partial_\mu T_{\mu\rho})x_\nu\right] dx \\
&= -\int_{\sigma_0}^{\sigma_1} \left[(\partial_{\nu\rho}\mathcal{H})_\chi + x_\rho(\partial_\nu \mathcal{H})_\chi - x_\nu(\partial_\rho \mathcal{H})_\chi\right] dx.
\end{aligned} \tag{8.5}$$

This is the law of conservation of angular momentum. Clearly, if the system is Lorentz-invariant, that is if (5.11) and (5.12) are sarisfied, then angular momentum is strictly conserved.

It is useful to introduce a tensor $\Theta_{\mu\nu}$ (Belinfante 1939), in terms of which both P_ν and $J_{\nu\rho}$ can be expressed. Define

$$G_{\mu\nu\rho} = \tfrac{1}{2}\left(\Sigma_{\mu\nu\rho} + \Sigma_{\nu\rho\mu} - \Sigma_{\rho\mu\nu}\right) \tag{8.6}$$

and

$$\Theta_{\nu\rho} = T_{\nu\rho} - \partial_\mu G_{\mu\nu\rho}. \tag{8.7}$$

Then, since $G_{\mu\nu\rho}$ is antisymmetric in its first two indices,

$$\partial_\nu \Theta_{\nu\rho} = \partial_\nu T_{\nu\rho} = -(\partial_\rho \mathcal{H})_\chi\,, \tag{8.8}$$

and one may verify that

$$\Theta_{\rho\nu} - \Theta_{\nu\rho} = -(\partial_{\nu\rho}\mathcal{H})_\chi. \tag{8.9}$$

Then, using the relation A-(1.7), we find

$$\left.\begin{aligned} P_\nu(\sigma) &= \int_\sigma d\sigma_\mu \Theta_{\mu\nu}\,, \\ J_{\nu\rho}(\sigma) &= \int_\sigma d\sigma_\mu (\Theta_{\mu\nu} x_\rho - \Theta_{\mu\rho} x_\nu). \end{aligned}\right\} \tag{8.10}$$

The symmetric part of $\Theta_{\mu\nu}$ is given by

$$\begin{aligned} \Theta_{\{\mu\nu\}} &= T_{\{\mu\nu\}} - \partial_\rho \Sigma_{\{\mu\nu\}\rho} \\ &= \mathcal{L}\delta_{\mu\nu} - \tfrac{1}{4}\left(\tilde{\chi} A_{\{\mu}\partial_{\nu\}}\chi - \partial_{\{\nu}\tilde{\chi} A_{\mu\}}\chi\right) - \tfrac{1}{4}\partial_\rho\{\tilde{\chi}(A_{\{\mu}S_{\nu\}\rho} - \tilde{S}_{\{\nu\rho}A_{\mu\}})\chi\} \\ &= \tfrac{1}{4}\tilde{\chi}(A_\rho\delta_{\mu\nu} - A_{\{\mu}\delta_{\nu\}\rho} - A_{\{\mu}S_{\nu\}\rho} + \tilde{S}_{\{\nu\rho}A_{\mu\}})\partial_\rho\chi \\ &\quad - \tfrac{1}{4}\partial_\rho\tilde{\chi}(A_\rho\delta_{\mu\nu} - A_{\{\mu}\delta_{\nu\}\rho} + A_{\{\mu}S_{\nu\}\rho} - \tilde{S}_{\{\nu\rho}A_{\mu\}})\chi - \mathcal{H}\delta_{\mu\nu} \end{aligned}$$

or, using (5.6),

$$\Theta_{\{\mu\nu\}} = \tfrac{1}{2}\left(\tilde{\chi}\tilde{S}_{\{\nu\rho}A_{\mu\}}\partial_\rho\chi - \partial_\rho\tilde{\chi}A_{\{\mu}S_{\nu\}\rho}\chi\right) - \mathcal{H}\delta_{\mu\nu}. \tag{8.11}$$

In particular, for a Lorentz-invariant system, the tensor $\Theta_{\mu\nu}$ itself is given by (8.11), since by (8.9) its antisymmetric part vanishes.

9. Kinematically independent fields

We now suppose that the vector χ of field variables is partitioned into subvectors χ^j in such a way that the matrices A_μ and M decompose into direct sums of submatrices A_μ^j and M^j acting on these subvectors, and that no further partitioning of this kind is possible. Then variables belonging to different subvectors are said to be <u>kinematically independent</u>. The Lagrangian may be written in the form

$$\mathcal{L} = \sum_j \mathcal{L}_0^j + \mathcal{L}_I , \qquad (9.1)$$

where

$$\begin{aligned} \mathcal{L}_0^j &= \tfrac{1}{4}\left(\tilde{\chi}^j A_\mu^j \partial_\mu \chi^j - \partial_\mu \tilde{\chi}^j A_\mu^j \chi^j\right) - \tfrac{1}{2}\tilde{\chi}^j M^j \chi^j \\ &\simeq \tfrac{1}{2}\tilde{\chi}^j \left(A_\mu^j \partial_\mu - M^j\right)\chi^j \end{aligned} \qquad (9.2)$$

and

$$\mathcal{L}_I = -\mathcal{H}_I . \qquad (9.3)$$

Here the symbol $\simeq$ indicates equality up to an explicit divergence term.

The equations of motion (3.13) are

$$\left(A_\mu^j \partial_\mu - M^j\right)\chi^j = \partial^j \mathcal{H}_I , \qquad (9.4)$$

and the commutation relations for the independent components on a surface σ of constant x_0 are given by

$$A_0^j\left(\chi^j(\underset{\sim}{x}, x_0);\ \tilde{\chi}^j(\underset{\sim}{x}', x_0)\right)_- A_0^j = i A_0^j\, \delta(\underset{\sim}{x} - \underset{\sim}{x}') . \qquad (9.5)$$

Kinematically independent pairs of independent variables commute in the generalized sense.

The condition (5.11) for Lorentz invariance is expressible in the form of separate conditions for each χ^j, namely

$$(\tilde{\chi}^j;\ \tilde{S}^j_{\nu\rho} A^j_\mu \partial_\mu \chi^j)_+ = 0. \tag{9.6}$$

The symmetric energy-momentum tensor for a Lorentz-invariant system, given by (8.11) can be written as

$$\Theta_{\mu\nu} = \sum_j \Theta^{oj}_{\mu\nu} + \Theta^I_{\mu\nu}, \tag{9.7}$$

where

$$\Theta^{oj}_{\mu\nu} = \tfrac{1}{2}(\tilde{\chi}^j;\ \tilde{S}^j_{\{\nu\rho} A^j_{\mu\}} \partial_\rho \chi^j)_+ - \tfrac{1}{2}\tilde{\chi}^j M^j \chi^j \delta_{\mu\nu}, \tag{9.8}$$

and

$$\Theta^I_{\mu\nu} = -\mathcal{H}_I \delta_{\mu\nu}. \tag{9.9}$$

Hitherto, we have always used Hermitian field variables. It is, however, often convenient to use complex field variables, and from the present standpoint they must be interpreted by separating each variable into a Hermitian and an anti-Hermitian part. If, for example, we consider the Dirac Lagrangian

$$\mathcal{L}_o = \tfrac{1}{2}[\psi^\dagger\beta, (i\gamma_\mu\partial_\mu - m)\psi], \tag{9.10}$$

then we must write

$$\left.\begin{aligned} \psi &= \tfrac{1}{\sqrt{2}}(\psi_1 + i\psi_2), \\ \psi^\dagger &= \tfrac{1}{\sqrt{2}}(\tilde{\psi}_1 - i\tilde{\psi}_2), \end{aligned}\right\} \tag{9.11}$$

and interpret (9.10) in terms of these variables. If the Dirac matrices γ_μ are taken to be in a Majorana representation, then

$$\mathcal{L}_o = \tfrac{1}{2}\tilde{\psi}_1\beta(i\gamma_\mu\partial_\mu - m)\psi_1 + \tfrac{1}{2}\tilde{\psi}_2\beta(i\gamma_\mu\partial_\mu - m)\psi_2, \tag{9.12}$$

whereas in any other representation, there will also be terms coupling ψ_1 to ψ_2. There is no objection to writing, in place of (9.10), the Lagrangian

$$\mathcal{L}_0 = \psi^+\beta(i\gamma_\mu \partial_\mu - m)\psi. \tag{9.13}$$

The only effect of this change is that the matrices A_μ and M no longer satisfy (3.16). Since, however, only the parts of these matrices which do satisfy this relation contribute to the equations of motion or commutation relations, the change is unimportant.

10. Linear and bilinear Bose Lagrangians

It is well known that Bose fields satisfying the Klein-Gordon equation, with or without interaction, can be described either in terms of a bilinear Lagrangian which yields the Klein-Gordon equation directly, or in terms of a linear Lagrangian involving additional variables[1]. In this section we propose to examine the connection between these two Lagrangians, in order to give a basis for comparison with the discussion of higher order Fermi field Lagrangians in chapter IV.

Consider a Bose field χ of the form

$$\chi = \begin{bmatrix} \varphi \\ \pi \end{bmatrix} \tag{10.1}$$

where the subvectors φ and π are chosen in such a way that

1) See, for example, Proca (1936), Kemmer (1938).

$$A_\mu = \begin{bmatrix} 0 & -\tilde{\alpha}_\mu \\ \alpha_\mu & 0 \end{bmatrix}, \qquad M = \begin{bmatrix} \mu^2 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{10.2}$$

where the α_μ are real (non-square) matrices, and μ is a real number, called the <u>mass</u>. The free Lagrangian is

$$\mathcal{L}_0 = \tfrac{1}{2}(\tilde{\pi}\alpha_\mu \partial_\mu \varphi - \tilde{\varphi}\tilde{\alpha}_\mu \partial_\mu \pi - \mu^2 \tilde{\varphi}\varphi - \tilde{\pi}\pi), \tag{10.3}$$

so that in the absence of interaction the equations of motion are

$$-\tilde{\alpha}_\mu \partial_\mu \pi = \mu^2 \varphi, \tag{10.4}$$

$$\alpha_\mu \partial_\mu \varphi = \pi. \tag{10.5}$$

Substituting from (10.5) into (10.4) yields the equation

$$(\tilde{\alpha}_\mu \alpha_\nu \partial_\mu \partial_\nu + \mu^2)\varphi = 0 \tag{10.6}$$

for φ alone. This may be derived in the usual way from the bilinear Lagrangian

$$\mathcal{L}_0' = \tfrac{1}{2}(\partial_\mu \tilde{\varphi}\tilde{\alpha}_\mu \alpha_\nu \partial_\nu \varphi - \mu^2 \tilde{\varphi}\varphi). \tag{10.7}$$

Moreover, if (10.5) is used to eliminate π from $\mathcal{L}_0$, one obtains (10.7), apart from an irrelevant divergence term.

We shall call a linear Lagrangian $\mathcal{L}$ a <u>linearization</u> of a given bilinear Lagrangian $\mathcal{L}'$ if

i) the equations of motion given by $\mathcal{L}'$ are deducible from those given by $\mathcal{L}$, by eliminating the extra variables which appear in $\mathcal{L}$ but not in $\mathcal{L}'$; and

ii) if the extra variables are eliminated from $\mathcal{L}$ itself, then $\mathcal{L}'$ results.

The second condition is not obviously necessary. It is automatically satisfied in most cases for Bose fields. For Fermi fields, however, it will appear in the next chapter that it is by no means always satisfied. We shall discuss its necessity at that stage.

Now consider the introduction of an interaction term $\mathcal{H}_I$. We shall only consider the case where

$$-\mathcal{L}_I = \mathcal{H}_I = \tfrac{1}{2}(\tilde{\pi} f + \tilde{f}\pi) + g, \tag{10.8}$$

in which f and g may depend on φ and on any other fields, but not on π. The equations of motion become

$$\left.\begin{aligned} -\tilde{\alpha}_\mu \partial_\mu \pi &= \mu^2 \varphi + \tfrac{1}{2}\{\partial_{(\varphi)}\tilde{f}, \pi\} + \partial_{(\varphi)} g, \\ \alpha_\mu \partial_\mu \varphi &= \pi + f, \end{aligned}\right\} \tag{10.9}$$

which yield

$$(\tilde{\alpha}_\mu \alpha_\nu \partial_\mu \partial_\nu + \mu^2)\varphi = \tilde{\alpha}_\mu \partial_\mu f - \tfrac{1}{2}\{\partial_{(\varphi)}\tilde{f}, \alpha_\mu \partial_\mu \varphi\} + \partial_{(\varphi)}(\tfrac{1}{2}\tilde{f}f - g). \tag{10.10}$$

Substituting for π from (10.9) into $\mathcal{L}$ gives $\mathcal{L}' = \mathcal{L}_0' + \mathcal{L}_I'$, where $\mathcal{L}_0'$ is given by (10.7) and $\mathcal{L}_I'$ by

$$\mathcal{L}_I' = -\tfrac{1}{2}\{\tilde{f}, \alpha_\mu \partial_\mu \varphi\} + \tfrac{1}{2}\tilde{f}f - g. \tag{10.11}$$

Clearly, this Lagrangian gives the equations of motion (10.10), so that again $\mathcal{L}$ is a linearization of $\mathcal{L}'$.

The numbers of independent variables (on a surface of constant x_0, say) among φ and among π must be equal, and in fact each equal to the rank of α_0. If this is less than the number of components of φ, then the variables φ must satisfy additional restrictions besides (10.10). These extra equations will follow from (10.9), and therefore will be derivable from $\mathcal{L}$; but they will not be derivable from $\mathcal{L}'$, so that if the bilinear Lagrangian is used, one must impose such auxiliary conditions in addition to, and independently of, the equations of motion obtained from the Lagrangian. Thus it must be noted that in general $\mathcal{L}'$ does not furnish as complete information about the system as does $\mathcal{L}$.

If the mass is zero, a further complication arises in the fact that the system is non-holonomic, since the equations (10.4) cannot be solved for the dependent components of φ. To circumvent this difficulty, one has to introduce additional variables, whose presence converts the system into a holonomic one with a larger number of degrees of freedom. The quanta of the corresponding fields must then be removed by a subsidiary condition on the states of the system. Note that such a condition is different in character from the extra conditions discussed in the preceding paragraph, which are operator equations.

To illustrate these points we shall consider briefly some simple examples of well-known theories. Take first a scalar field φ. The linear Lagrangian involves an extra vector field π_μ, and is[1]

$$\left.\begin{aligned} \mathcal{L}_0 &= \tfrac{1}{2}\left(\pi_\mu \partial_\mu \varphi - \varphi \partial_\mu \pi_\mu - \mu^2 \varphi\varphi - \pi_\mu \pi_\mu\right), \\ \mathcal{L}_I &= -\tfrac{1}{2}\left\{\pi_\mu, f_\mu\right\} - g. \end{aligned}\right\} \qquad (10.12)$$

The equations of motion are

$$\left.\begin{aligned} -\partial_\mu \pi_\mu &= \mu^2 \varphi + \tfrac{1}{2}\left\{\partial_{(\varphi)} f_\mu, \pi_\mu\right\} + \partial_{(\varphi)} g, \\ \partial_\mu \varphi &= \pi_\mu + f_\mu, \end{aligned}\right\} \qquad (10.13)$$

1) See, for example, Kemmer (1938).

which yield

$$(\Box+\mu^2)\,\varphi = \partial_\mu f_\mu - \tfrac{1}{2}\{\partial_{(\varphi)} f_\mu, \partial_\mu\varphi\} + \partial_{(\varphi)}(\tfrac{1}{2} f_\mu f_\mu - g). \tag{10.14}$$

The independent components are φ and π_0. One may easily verify that, in the absence of interaction, φ satisfies the four-dimensional commutation rules

$$[\varphi(x), \varphi(x')] = i\,\Delta(x-x'), \tag{10.15}$$

where $\Delta(x)$ is the invariant singular function defined by A-(1.9). The bilinear Lagrangian corresponding to (10.12) is

$$\left.\begin{aligned} \mathcal{L}_0' &= \tfrac{1}{2}(\partial_\mu\varphi\,\partial_\mu\varphi - \mu^2\varphi\varphi), \\ \mathcal{L}_I' &= -\tfrac{1}{2}\{f_\mu, \partial_\mu\varphi\} + \tfrac{1}{2} f_\mu f_\mu - g. \end{aligned}\right\} \tag{10.16}$$

No particular difficulties arise in this case, because there are no dependent components of φ.

Now consider a vector field φ_μ. We introduce an auxiliary six-vector π^α, and define symbols $\epsilon^\alpha_{\mu\nu}$ satisfying

$$\left.\begin{aligned} \epsilon^\alpha_{\mu\nu} + \epsilon^\alpha_{\nu\mu} &= 0, \\ \epsilon^\alpha_{\mu\nu}\,\epsilon^\alpha_{\rho\sigma} &= \delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}, \\ \tfrac{1}{2}\epsilon^\alpha_{\mu\nu}\,\epsilon^\beta_{\mu\nu} &= \delta^{\alpha\beta}. \end{aligned}\right\} \tag{10.17}$$

The free Lagrangian is[1)]

$$\mathcal{L}_0 = \tfrac{1}{2}(\pi^\alpha\epsilon^\alpha_{\mu\nu}\partial_\mu\varphi_\nu - \varphi_\nu\epsilon^\alpha_{\mu\nu}\partial_\mu\pi^\alpha - \mu^2\varphi_\nu\varphi_\nu - \pi^\alpha\pi^\alpha), \tag{10.18}$$

which yields the equations of motion

$$\left.\begin{aligned} -\epsilon^\alpha_{\mu\nu}\partial_\mu\pi^\alpha &= \mu^2\varphi_\nu, \\ \epsilon^\alpha_{\mu\nu}\partial_\mu\varphi_\nu &= \pi^\alpha. \end{aligned}\right\} \tag{10.19}$$

---

1) See Proca (1936), Kemmer (1938, 1939), Duffin (1938).

Thus φ_ν satisfies

$$(\Box+\mu^2)\varphi_\nu - \partial_\nu(\partial_\mu\varphi_\mu) = 0, \tag{10.20}$$

and the corresponding bilinear Lagrangian is

$$\mathcal{L}_0' = \tfrac{1}{2}\left(\partial_\mu\varphi_\nu\partial_\mu\varphi_\nu - \partial_\mu\varphi_\nu\partial_\nu\varphi_\mu - \mu^2\varphi_\nu\varphi_\nu\right). \tag{10.21}$$

Now, if the mass μ^2 is non-zero, then from (10.19) there follows

$$\partial_\nu\varphi_\nu = 0 \tag{10.22}$$

so that φ_ν satisfies

$$(\Box+\mu^2)\varphi_\nu = 0. \tag{10.23}$$

This extra equation is derivable from $\mathcal{L}$ but not from $\mathcal{L}'$. We note that (10.23) may be derived from the bilinear Lagrangian

$$\mathcal{L}_0' = \tfrac{1}{2}\left(\partial_\mu\varphi_\nu\partial_\mu\varphi_\nu - \mu^2\varphi_\nu\varphi_\nu\right), \tag{10.24}$$

so that it appears that $\mathcal{L}'$ is ambiguous to the extent of the addition of a multiple of the left side of (10.22). The four-dimensional commutation relations obtained from $\mathcal{L}$ are

$$\left[\varphi_\mu(x),\ \varphi_\nu(x')\right] = i\left(\delta_{\mu\nu} - \frac{\partial_\mu\partial_\nu}{\mu^2}\right)\Delta(x-x'). \tag{10.25}$$

We now turn to the case of a massless vector field. If we take the Lagrangian to be (10.18) with $\mu^2 = 0$, then the system is non-holonomic, since (10.19) can not be solved for the dependent variable φ_0. Correspondingly, the equation

$$\epsilon^\alpha_{0k}\,\partial_k\pi^\alpha = 0 \tag{10.26}$$

is an extra condition on the independent components of π only. Thus one cannot in fact choose the "independent" components independently. This situation is typical of non-holonomic systems. A further difficulty is that one can no longer deduce (10.22) from (10.19), so that the equation (10.23) cannot be obtained. To remedy this situation, we introduce an auxiliary scalar field η, and take

$$\mathcal{L}_0 = \tfrac{1}{2}\big(\pi^\alpha \epsilon^\alpha_{\mu\nu} \partial_\mu \varphi_\nu - \varphi_\nu \epsilon^\alpha_{\mu\nu} \partial_\mu \pi^\alpha - \pi^\alpha \pi^\alpha$$

$$+ \eta \partial_\mu \varphi_\mu - \varphi_\mu \partial_\mu \eta - \eta\eta\big). \tag{10.27}$$

Then the equations of motion are

$$\left.\begin{aligned} -\epsilon^\alpha_{\mu\nu} \partial_\mu \pi^\alpha - \partial_\nu \eta &= 0, \\ \epsilon^\alpha_{\mu\nu} \partial_\mu \varphi_\nu &= \pi^\alpha, \\ \partial_\mu \varphi_\mu &= \eta. \end{aligned}\right\} \tag{10.28}$$

All four components of φ_μ are now independent variables, so the system is certainly holonomic. Eliminating π^α and η from (10.28) yields

$$\Box \varphi_\mu = 0. \tag{10.29}$$

The bilinear Lagrangian obtained from (10.27) by eliminating π^α and η is

$$\mathcal{L}_0' = \tfrac{1}{2} \partial_\mu \varphi_\nu \partial_\mu \varphi_\nu, \tag{10.30}$$

which obviously yields the equations of motion (10.29). The four-dimensional commutation rules are

$$[\varphi_\mu(x), \varphi_\nu(x')] = i\delta_{\mu\nu} D(x-x'), \tag{10.31}$$

where $D(x)$ is the invariant singular function $\Delta(x)$ for zero mass. The field η must now be removed by imposing a subsidiary condition

$$\eta | \rangle = 0 \tag{10.32}$$

on the states of the field. This condition is self-consistent because of the fact that η satisfies the commutation rules

$$[\eta(x), \eta(x')] = 0. \tag{10.33}$$

CHAPTER IV

HIGHER ORDER SPINOR LAGRANGIANS

1. Second order Lagrangian for four-component spinor

Feynman and Gell-Mann (1958) have recently suggested that Fermi fields should be described by means of two-component spinors satisfying second-order equations, instead of four-component spinors satisfying the Dirac equation. We now wish to consider the extent to which this can be done within the formalism of the action principle.

We first consider a four-component spinor χ, and make the reduction to two components later. In the absence of interaction, the equation to be satisfied is the Klein-Gordon equation

$$(\Box + m^2)\chi = 0, \qquad (1.1)$$

which may be derived from the second order Lagrangian

$$\mathcal{L}_0' = -\frac{1}{m}\bar{\chi}(\Box + m^2)\chi, \qquad (\bar{\chi} \equiv \chi^+\beta). \qquad (1.2)$$

In order to interpret such a Lagrangian, one must associate with it a linearized Lagrangian, exactly as in section III-10. The procedure of linearizing (1.2) may seem somewhat perverse, since the equation (1.1) was originally derived from the linear Dirac equation by eliminating two of the four components of the spinor, but it is the most systematic way of investigating its properties.

A linearized form of the equations of motion (1.1) is

$$\left.\begin{aligned} (i\gamma_\mu \partial_\mu + m)\chi &= m\psi_1, \\ (i\gamma_\mu \partial_\mu - m)\psi_1 &= 0, \end{aligned}\right\} \tag{1.3}$$

and in fact these are the equations originally used by Feynman and Gell-Mann. They may be derived from the linear Lagrangian

$$\mathcal{L}_0 = -2\bar{\chi}(i\gamma_\mu\partial_\mu + m)\chi + \bar{\chi}(i\gamma_\mu\partial_\mu + m)\psi_1 + \bar{\psi}_1(i\gamma_\mu\partial_\mu + m)\chi - m\bar{\psi}_1\psi_1. \tag{1.4}$$

This is indeed a linearization of (1.2), as may easily be verified by substituting for ψ_1.

The anticommutation relations are[1)]

$$\left.\begin{aligned} \{\chi(\underset{\sim}{x}, x_0),\ \chi^+(\underset{\sim}{x}', x_0)\} &= 0, \\ \{\psi_1(\underset{\sim}{x}, x_0),\ \psi_1^+(\underset{\sim}{x}', x_0)\} &= \delta(\underset{\sim}{x} - \underset{\sim}{x}'), \\ \{\psi_1(\underset{\sim}{x}, x_0),\ \chi^+(\underset{\sim}{x}', x_0)\} &= \tfrac{1}{2}\delta(\underset{\sim}{x} - \underset{\sim}{x}'), \\ \{\chi(\underset{\sim}{x}, x_0),\ \psi_1^+(\underset{\sim}{x}', x_0)\} &= \tfrac{1}{2}\delta(\underset{\sim}{x} - \underset{\sim}{x}'), \end{aligned}\right\} \tag{1.5}$$

all other equal-time anticommutators vanishing. Hence χ satisfies the four-dimensional anticommutation rules

$$\{\chi(x),\ \chi^+(x')\} = \tfrac{1}{2} i m \gamma_0 \Delta(x-x'). \tag{1.6}$$

To find a representation of (1.5), we write

$$\chi = \tfrac{1}{2}(\psi_1 + \psi_2). \tag{1.7}$$

The relations (1.5) are then satisfied if

---

1) Throughout this chapter we shall use the representation of the Dirac matrices in terms of the Pauli matrices which is given in appendix A-3. In particular, $\beta = \gamma_0$.

$$\{\psi_2(\underset{\sim}{x},x_0),\ \psi_1^+(\underset{\sim}{x}',x_0)\} = 0 = \{\psi_1(\underset{\sim}{x},x_0),\ \psi_2^+(\underset{\sim}{x}',x_0)\}, \tag{1.8}$$

and

$$\{\psi_2(\underset{\sim}{x},x_0),\ \psi_2^+(\underset{\sim}{x}',x_0)\} = -\delta(\underset{\sim}{x}-\underset{\sim}{x}'). \tag{1.9}$$

From (1.3), ψ_2 satisfies

$$(i\gamma_\mu\partial_\mu + m)\psi_2 = 0. \tag{1.10}$$

These results are hardly surprising, as the solutions of (1.1) are clearly connected with the Dirac equations for both positive and negative masses. If we substitute (1.7) into (1.4), we find

$$\mathcal{L}_0 = \bar{\psi}_1(i\gamma_\mu\partial_\mu - m)\psi_1 - \bar{\psi}_2(i\gamma_\mu\partial_\mu + m)\psi_2, \tag{1.11}$$

a superposition of the Dirac Lagrangians for masses $+m$ and $-m$. However, these two Lagrangians appear in (1.11) with opposite sign, and it is this fact that yields the nagative sign on the right side of (1.9), producing an indefinite metric[1] for the χ field. The corresponding negative probabilities can be removed by redefining the scalar product in Hilbert space, replacing $\langle\alpha|\beta\rangle$ by $\langle\alpha|(-1)^{n_2}|\beta\rangle$, where n_2 is the number operator for the ψ_2 field.

The Lagrangian (1.4) has the additional disadvantage that it cannot be written in terms of a purely two-component χ. If we set

$$\chi_\pm \equiv \tfrac{1}{2}(1 \pm i\gamma_5)\chi, \tag{1.12}$$

then the term $\bar{\chi}m\chi$ couples χ_+ to χ_-. If we wish to remove one of these fields, this can only be done by a subsidiary condition of the form

---

1) See Pais and Uhlenbeck (1950); Phillips (1955).

$$\left.\begin{aligned} \chi_-^{(+)} | \rangle &= 0, \\ \overline{\chi_-^{(-)}} | \rangle &= 0. \end{aligned}\right\} \tag{1.13}$$

Clearly, it would be an advantage to be able to write a Lagrangian in terms of a purely two-component χ, thus avoiding the necessity of introducing a subsidiary condition. The possibility of doing this will be considered in the next section.

2. Two-component spinor Lagrangian

Rather than impose a subsidiary condition (1.13), we might start from a two-component χ satisfying (1.1). However, it is now impossible to derive the equations from a Lagrangian of the form (1.2). In order to see what Lagrangian we may use, we consider the linearization of (1.1)

$$\left.\begin{aligned} i\sigma_\mu \partial_\mu \chi &= m\varphi, \\ i\sigma'_\mu \partial_\mu \varphi &= m\chi, \end{aligned}\right\} \tag{2.1}$$

which may be derived from the Lagrangian

$$\mathcal{L}_0 = \chi^\dagger i\sigma_\mu \partial_\mu \chi + \varphi^\dagger i\sigma'_\mu \partial_\mu \varphi - m\chi^\dagger \varphi - m\varphi^\dagger \chi. \tag{2.2}$$

This, however, is not a linearization of a second-order Lagrangian. Indeed, if we substitute for φ from (2.1) into (2.2), we obtain the third-order Lagrangian

$$\mathcal{L}'_0 = -\frac{1}{m^2}\chi^\dagger i\sigma_\mu \partial_\mu (\Box + m^2)\chi. \tag{2.3}$$

The simplest linearized Lagrangian corresponding to (2.3) is

$$\mathcal{L}_0 = \chi^+ i\sigma_\mu \partial_\mu \chi + \varphi^+ i\sigma'_\mu \partial_\mu \varphi - m\chi^+\varphi - m\varphi^+\chi$$
$$+ \lambda\left(m\xi^+\varphi + m\varphi^+\xi - \xi^+ i\sigma_\mu\partial_\mu\chi - \chi^+ i\sigma_\mu\partial_\mu\xi\right) \qquad (2.4)$$

for any non-zero value of λ. The equations of motion are then

$$\left.\begin{aligned} i\sigma_\mu\partial_\mu\chi &= m\varphi, \\ i\sigma'_\mu\partial_\mu\varphi &= m\chi + \lambda m\xi \\ i\sigma_\mu\partial_\mu\xi &= 0. \end{aligned}\right\} \qquad (2.5)$$

It is easily verified that (2.4) is a true linearization of (2.3), and that it yields commutation rules corresponding to an indefinite metric. A change from one finite value of λ to another is evidently a trivial change in the normalization of ξ. However, the Lagrangian (2.2) corresponds in some sense to taking an improper limit $\lambda \to 0$ in (2.4). The taking of this limit is of course not a strictly permissible procedure, in view of the fact that λ^{-1} appears in the commutation relations[1)]; and (2.2) will therefore be termed an <u>improper</u> linearization of (2.3). It is remarkable that the choice of the Lagrangian (2.2) avoids both the difficulties of an indefinite metric and of the extra mass-zero field ξ which occurs in (2.4) with non-zero λ.

We are now in a position to discuss the condition imposed

---

1) In this respect, it bears some resemblance to the limit $\mu \to 0$ for a vector meson field of mass μ.

in the definition of a linearization in section III-10, to the effect that a bilinear Lagrangian must be obtainable from its linearization by eliminating the extra variables. Firstly, we note that some condition in addition to requiring that the linearization give the correct equations of motion is obviously necessary; for otherwise there is no guarantee that they will yield the same commutation relations. For instance, one cannot take (2.2) to be a linearization of (1.2), since (2.2) does not give the commutation rules (1.6). Moreover, one easily verifies that the commutation rules obtained by the usual canonical method (Heisenberg and Pauli 1929, Schwinger 1951) from a bilinear Lagrangian are in fact identical with those obtained from any linearization of the Lagrangian by the use of Schwinger's principle. Thus the condition appears to be a reasonable one. Clearly, an improper linearization in the sense defined above gives only some of the solutions of the equations of motion given by the higher-order Lagrangian, and therefore could not be taken to describe the same system.

Now the Lagrangian (2.2) is merely the Dirac Lagrangian III-(9.13) written in terms of two-component spinors, and correspondingly the equations (2.1) do not give a preferred position to either χ or φ. This linearization is not at all close to the spirit of Feynman and Gell-Mann, in that they start by assigning a preferrred position to one two-component spinor. We shall therefore seek a linearized Lagrangian which gives the equations of motion in the form (1.3) but avoids the difficulties of an indefinite metric.

3. Third-order Lagrangian

A Lagrangian which satisfies the requirements stated is

$$\mathcal{L}_0 = 2\bar{\chi}(i\gamma_\mu\partial_\mu + m)\chi + \bar{\psi}_1 i\gamma_\mu\partial_\mu\psi_1 - \bar{\chi}(i\gamma_\mu\partial_\mu + m)\psi_1 - \bar{\psi}_1(i\gamma_\mu\partial_\mu + m)\chi. \quad (3.1)$$

The equations of motion are (1.3) and the commutation relations are

$$\left.\begin{aligned}
\{\chi(\underset{\sim}{x}, x_0),\ \chi^\dagger(\underset{\sim}{x}', x_0)\} &= \tfrac{1}{2}\,\delta(\underset{\sim}{x} - \underset{\sim}{x}'),\\
\{\chi(\underset{\sim}{x}, x_0),\ \psi_1^\dagger(\underset{\sim}{x}', x_0)\} &= \tfrac{1}{2}\,\delta(\underset{\sim}{x} - \underset{\sim}{x}'),\\
\{\psi_1(\underset{\sim}{x}, x_0),\ \chi^\dagger(\underset{\sim}{x}', x_0)\} &= \tfrac{1}{2}\,\delta(\underset{\sim}{x} - \underset{\sim}{x}'),\\
\{\psi_1(\underset{\sim}{x}, x_0),\ \psi_1^\dagger(\underset{\sim}{x}', x_0)\} &= \delta(\underset{\sim}{x} - \underset{\sim}{x}').
\end{aligned}\right\} \quad (3.2)$$

all other equal-time anticommutators vanishing. Equations (1.3) then yield the four-dimensional anticommutator

$$\{\chi(x),\ \chi^\dagger(x')\} = \tfrac{1}{2}\gamma_\mu\partial_\mu\gamma_0\Delta(x - x'). \quad (3.3)$$

If, as before, we write

$$\chi = \tfrac{1}{2}(\psi_1 + \psi_2),$$

then (1.8) remains unchanged, but (1.9) is replaced by

$$\{\psi_2(\underset{\sim}{x}, x_0),\ \psi_2^\dagger(\underset{\sim}{x}', x_0)\} = \delta(\underset{\sim}{x} - \underset{\sim}{x}'). \quad (3.4)$$

Again, ψ_2 satisfies the equation (1.10), and the Lagrangian (3.1) becomes

$$\mathcal{L}_0 = \bar{\psi}_1(i\gamma_\mu\partial_\mu - m)\psi_1 + \bar{\psi}_2(i\gamma_\mu\partial_\mu + m)\psi_2. \quad (3.5)$$

Since the Lagrangians for masses $+m$ and $-m$ are now combined with the same signs, there is no indefinite metric.

The Lagrangian (3.1) is an improper linearization of the four-component Lagrangian corresponding to (2.3), namely

$$\mathcal{L}_0' = -\frac{1}{m^2}\bar{\chi}\, i\gamma_\mu\partial_\mu(\Box + m^2)\chi, \tag{3.6}$$

since it may be obtained by taking the improper limit $\lambda \to 0$ in

$$\begin{aligned}\mathcal{L}_0 = 2\bar{\chi}(i\gamma_\mu\partial_\mu + m)\chi + \bar{\psi}_1 i\gamma_\mu\partial_\mu\psi_1 - \bar{\chi}(i\gamma_\mu\partial_\mu + m)\psi_1 - \bar{\psi}_1(i\gamma_\mu\partial_\mu + m)\chi \\ + \lambda\{m\bar{\psi}_1\xi + m\bar{\xi}\psi_1 - \bar{\xi}(i\gamma_\mu\partial_\mu + m)\chi - \bar{\chi}(i\gamma_\mu\partial_\mu + m)\xi\},\end{aligned} \tag{3.7}$$

which for any non-zero λ is a true linearization of (3.6) and yields an indefinite metric.

Unfortunately, although (3.6) may be written in the two-component form (2.3), its improper linearization (3.1) cannot. Thus if we wish to have a Lagrangian which gives directly the equations of motion (1.3), we must use the four-component forms (3.6) and (3.1). The two-component form (2.3) allows only the (improper) linearization (2.2), not (3.1). Therefore we shall continue to use the four-component Lagrangian, and remove the unwanted components of χ by a subsidiary condition.

The fields $\chi_\pm$ defined by (1.12) satisfy the anti-commutation relations

$$\left.\begin{aligned}\{\chi_+(x), \chi_+^\dagger(x')\} &= \tfrac{1}{4}(1 + i\gamma_5)\gamma_\mu\partial_\mu\gamma_0\Delta(x-x'), \\ \{\chi_-(x), \chi_-^\dagger(x')\} &= \tfrac{1}{4}(1 - i\gamma_5)\gamma_\mu\partial_\mu\gamma_0\Delta(x-x'), \\ \{\chi_+(x), \chi_-^\dagger(x')\} &= \{\chi_-(x), \chi_+^\dagger(x')\} = 0.\end{aligned}\right\} \tag{3.8}$$

Thus the fields $\chi_\pm$ are both kinematically and dynamically independent, and it is therefore consistent to impose the subsidiary condition (1.13) on all physical states.

We now define a field ψ by

$$m\psi = \sqrt{2}\,(i\gamma_\mu\partial_\mu + m)\chi_+, \tag{3.9}$$

The field ψ then satisfies the Dirac equation

$$(i\gamma_\mu \partial_\mu - m)\psi = 0, \tag{3.10}$$

and the anticommutation rules

$$\{\psi(x),\ \psi^\dagger(x')\} = (\gamma_\mu \partial_\mu - im)\gamma_0 \Delta(x-x'). \tag{3.11}$$

Comparing (3.8) and (3.11), we see that although χ_+ possesses a definite "handedness", ψ constructed from it does not.

One could of course define another field ψ' by

$$m\psi' = \sqrt{2}\,(i\gamma_\mu \partial_\mu + m)\chi_-, \tag{3.12}$$

but if physical states contain no quanta of the χ_- field, then they cannot contain any of the ψ' field either, and so ψ' is not of physical interest.

4. Interactions

We may consider how to form interaction terms by first considering bilinear spinor expressions of the form

$$\mathcal{H}_I = \bar{\chi} B_1 \chi + \bar{\chi} B_2 \psi_1 + \bar{\psi}_1 B_2 \chi + \bar{\psi}_1 B_3 \psi_1. \tag{4.1}$$

The operators B_j appearing in (4.1) may be functions of position, and may depend upon Bose field operators. The Lagrangian given by (3.1) and (4.1) yields the equations

$$\left.\begin{aligned} (i\gamma_\mu \partial_\mu - m)\psi_1 &= (B_1 + 2B_2)\chi + (B_2 + 2B_3)\psi_1, \\ (i\gamma_\mu \partial_\mu + m)\chi &= m\psi_1 + (B_1 + B_2)\chi + (B_2 + B_3)\psi_1. \end{aligned}\right\} \tag{4.2}$$

In order that we may obtain an equation for χ alone by elimination between the equations (4.2), we impose the condition

$$B_2 + B_3 = 0. \tag{4.3}$$

Consider first parity-conserving interactions. A necessary condition for this is that the first equation of (4.2) be an equation for ψ_1 only, that is[1])

$$B_1 + 2B_2 = 0. \tag{4.4}$$

Thus the form of interaction which we consider is simply

$$\mathcal{H}_I = 2\bar{\chi}B\chi + \bar{\psi}_1 B\psi_1 - \bar{\chi}B\psi_1 - \bar{\psi}_1 B\chi. \tag{4.5}$$

The resulting equation for χ is

$$(\Box + m^2)\chi = -(B i\gamma_\mu \partial_\mu + i\gamma_\mu \partial_\mu B)\chi + B^2\chi. \tag{4.6}$$

If this is to be consistent with the subsidiary condition removing the field χ_- from physical states, we require that B should contain a linear combination of γ_μ and $\gamma_{\mu 5}$ only. In this way, we may introduce the interaction with the electromagnetic field A_μ, and a pseudovector interaction with a pseudoscalar meson field ϕ, by writing

$$B = e\gamma_\mu A_\mu + g\gamma_{\mu 5}\partial_\mu \Phi. \tag{4.7}$$

It does not seem possible, however, to introduce a pseudoscalar coupling to ϕ and simultaneously to reduce χ to a two-component form.

---

1) It is interesting to note that in terms of the indefinite metric quantization of section 1, equation (4.4) is just the condition that there are no cross terms between the ψ_1 and ψ_2 fields, i.e. the condition that the redefinition of the scalar product removes all negative probabilities.

We now consider the possible forms for the four-Fermion interactions. It is convenient to think of these as occurring through an intermediate very heavy boson, so that the interaction may be thought of as being of the form (4.1). Now, of course, we need not necessarily impose the condition (4.4). Thus, in addition to the interaction (4.5), we may also consider a term obtained by taking $B_2 = B_3 = 0$, that is

$$\mathcal{H}_I = \bar{\chi} C \chi . \tag{4.8}$$

If we have terms of both types (4.5) and (4.8), the resulting equation for χ is

$$(\Box + m^2)\chi = -(i\gamma_\mu \partial_\mu B + B i\gamma_\mu \partial_\mu)\chi + B^2 \chi - i\gamma_\mu \partial_\mu C \chi + BC\chi . \tag{4.9}$$

Again, if this is to be consistent with the subsidiary condition removing the χ_- field, χ_+ and χ_- must be dynamically independent, whence both B and C must be linear combinations of γ_μ and $\gamma_{\mu 5}$ only. We note that when this condition is satisfied, χ is related to ψ_1 by

$$\{\chi - \tfrac{1}{2}(1 + i\gamma_5)\psi_1\}^{(+)} | \rangle = 0 . \tag{4.10}$$

The interaction term (4.9) corresponds, in the more conventional formalism, to the interaction (A - V), which was used by Feynman and Gell-Mann. This interaction corresponds to the addition of the term

$$\mathcal{H}_I = \chi_1^\dagger \sigma_\mu \chi_2 \chi_3^\dagger \sigma_\mu \chi_4 \tag{4.11}$$

to the Lagrangian (2.2). However, in terms of that Lagrangian, we are unable to suggest why φ should not appear in the weak interaction Lagrangian. One might argue that χ is to be taken as fundamental, and φ treated as "derivative", but the choice between χ and φ is arbitrary, and the parity-conserving strong interactions require the presence of both. With the Lagrangian (3.1), the position is rather better, since there is then no symmetry between χ and ψ_1, so that, regarding ψ_1 as "derivative", (4.8) is the only "non-derivative" interaction possible. Nevertheless, the strong interactions can only be written in the form (4.5), so that some explanation is required of the fact that derivatives can appear in the strong interactions but not in the weak ones.

5. Discussion

In a conventional form of theory, it is necessary that we should be able to write down a Lagrangian, and if we are to regard χ as the fundamental quantity it must be expressible in terms of χ. If the Lagrangian is linearized, it must involve some other quantity besides χ, just as the extra variables π appear in the linearized Lagrangians discussed in section III-10.

The only perfectly consistent Lagrangian which can be written down involving all interactions is that in terms of χ and φ, that is, the usual Dirac Lagrangian in terms of two-component spinors. This Lagrangian is perfectly symmetric with respect to χ and φ, so that there is no apparent reason for giving χ a preferred position. From the present point of view, therefore, a more natural choice of Lagrangian is that in terms of χ and ψ_1. However, it does not seem to be possible to express this Lagrangian in terms of a two-component χ, so that the difficulties of a subsidiary condition arise. The interaction terms which can be consistently introduced are of two types only, the parity-conserving interactions (4.5), and the parity-non-conserving ones (4.8). There is of course no reason a priori why the weak interactions should not consist of a mixture of both types, unless some reason can be found for rejecting weak derivative interactions.

CHAPTER V

DISPERSION RELATIONS FOR INELASTIC SCATTERING PROCESSES

1. Causality conditions

In this section, the interaction picture will be used to explain the postulated causality condition. The interaction operators will be defined to agree with the Heisenberg operators in the infinite future $t = +\infty$. If H is the Hamiltonian in the interaction picture, then the S-matrix may be defined, using the adiabatic hypothesis, by

$$S = S(+\infty, -\infty),$$

where

$$S(t_2, t_1) = \sum_n \frac{(-i)^n}{n!} \int_{t_1}^{t_2} dx_1 \ldots dx_n \, T\{H(x_1) \ldots H(x_n)\}.$$

The variational derivative $\delta S/\delta\varphi(x)$ with respect to the meson field function $\varphi(x)$ is defined, following Bogolyubov, Medvedev and Polivanov (1956), as the sum of all terms in which one factor $\varphi(y)$ in the expression of S as a functional of $\varphi(x)$ has been replaced by $\delta(x-y)$; in other words, it is computed as though $\varphi(x)$ were a c-number field. For fermion fields, the field variables are treated as exactly anticommuting quantities, and there is consequently a difference in sign between left and right derivatives if an even function of the fermion fields is differentiated, and a change of sign on reversing the order of double differentiation[1].

It will be assumed that S possesses variational derivatives of all orders with respect to the fields, and that their matrix elements are integrable in a generalized sense[2]. The simplest of these derivatives will now be evaluated. The variational derivative of the Hamiltonian with respect to the meson field is

$$\frac{\delta H(x)}{\delta \varphi(y)} = \delta(x-y)\, \underset{\sim}{j}(x),$$

where $\underset{\sim}{j}(x)$ is, by definition, the current in the interaction picture. Hence[3]

$$\frac{\delta S}{\delta \varphi(x)} = -i \sum_n \frac{(-i)^n}{n!} \int_{-\infty}^{+\infty} dx_1 \dots dx_n \, T\{H(x_1)\dots H(x_n)\, \underset{\sim}{j}(x)\}$$

$$= -i\, S(+\infty, x^0)\, \underset{\sim}{j}(x)\, S(x^0, -\infty),$$

and therefore

$$i \frac{\delta S}{\delta \varphi(x)} S^* = S(+\infty, x^0)\, \underset{\sim}{j}(x)\, S^*(+\infty, x^0)$$

$$= j(x), \qquad (1.1)$$

---

1) These variational derivatives are clearly related to the derivatives ∂f defined in section II-6.

2) The concept of generalized integrable functions is a refinement of that of distribution functions (Schwartz 1950). The exact definition is given by Bogolyubov and Shirkov (1955), and by Bogolyubov, Medvedev and Polivanov (1956, appendix).

3) To avoid confusion with the label distinguishing different points, the components of x are denoted in this chapter by x^μ.

the current in the Heisenberg picture. Moreover,

$$\frac{\delta S(+\infty, x^0)}{\delta \varphi(y)} = 0 \qquad \text{for} \qquad y^0 < x^0,$$

which implies that

$$\frac{\delta j(x)}{\delta \varphi(y)} = 0 \qquad \text{for} \qquad y^0 < x^0. \tag{1.2}$$

Throughout the remainder of this chapter and the next, $j_{\alpha_i}(x_i)$ will be abbreviated to j_i and $\delta/\delta\varphi_{\alpha_i}(x_i)$ to δ_i. It is convenient to define also the quantity

$$\tilde{j}_i = \tilde{j}_{\alpha_i}(x_i) = i\, S^*(\delta_i S) = S^* j_i S, \tag{1.3}$$

which satisfies the relation

$$\frac{\delta \tilde{j}(x)}{\delta \varphi(y)} = 0 \qquad \text{for} \qquad y^0 > x^0.$$

Now consider the operator

$$H^2(x_1, x_2) = (\delta_1 \delta_2 S)\, S^*.$$

It is easy to see that

$$H^2(x_1, x_2) = -i\,\delta_1 j_2 - j_2 j_1 = -i\,\delta_2 j_1 - j_1 j_2, \tag{1.4}$$

and hence that

$$\delta_1 j_2 - \delta_2 j_1 = i\,[\,j_2, j_1\,]. \tag{1.5}$$

Using (1.2), equations (1.4) and (1.5) become

$$\begin{aligned} H^2(x_1, x_2) &= -\,\theta(x_1 - x_2)\, j_1 j_2 - \theta(x_2 - x_1)\, j_2 j_1 \\ &= -\,T\{j_1 j_2\} \end{aligned} \tag{1.6}$$

and

$$i\,\delta_1 j_2 = \theta(x_1 - x_2)\,[\,j_1, j_2\,], \tag{1.7}$$

where

$$\theta(x) = \begin{cases} 1 & \text{for} \quad x^0 > 0, \\ 0 & \text{for} \quad x^0 < 0. \end{cases}$$

Since[1)]

$$[j_1, j_2] = 0 \qquad \text{for} \quad x_1 \sim x_2, \tag{1.8}$$

the causality condition is obtained from (1.7) in the form

$$\delta_1 j_2 = 0 \qquad \text{for} \quad x_1 \lesssim x_2, \tag{1.9}$$

and similarly

$$\delta_1 \tilde{j}_2 = 0 \qquad \text{for} \quad x_2 \lesssim x_1. \tag{1.10}$$

If, on the other hand, (1.9) is regarded as a basic postulate, equations (1.4) to (1.8) can obviously be derived from it. This point of view is adopted by Bogolyubov *et al.* (1956), and will be followed here. For a full discussion of a closely related condition, and the physical meaning thereof, see Bogolyubov and Shirkov (1955).

2. Radiation operators

Consider the operators defined by

$$\left.\begin{aligned} H^n(x_1 \ldots x_n) &= (\delta_1 \ldots \delta_n S)\, S^*, \\ G^n(x_1 \ldots x_{n-1} : x_n) &= -i\,(\delta_1 \ldots \delta_{n-1} j_n), \\ G^n(x_n : x_1 \ldots x_{n-1}) &= -i\, S(\delta_1 \ldots \delta_{n-1} \tilde{j}_n)\, S^*. \end{aligned}\right\} \tag{2.1}$$

---

1) We use the notation A-(1.2).

It can readily be shown, by induction starting from equations (1.6) and (1.7), that[1)]

$$H^{n+1}(x_0 \ldots x_{n+1}) = (-i)^{n+1}\, T\{j_0 \cdots j_n\}$$

$$= (-i)^{n+1} \sum \sum_{r=0}^{n} \theta(1 > \cdots > r > 0 > r+1 > \cdots > n)\, j_1 \cdots j_r j_0 j_{r+1} \cdots j_n \tag{2.2}$$

$$G^{n+1}(x_1 \ldots x_n : x_0) = -i^{n+1} \sum \theta(1 > \cdots > n > 0)\, [j_1, [j_2, \cdots [j_n, j_0] \cdots]]$$

$$= (-i)^{n+1} \sum \sum_{r=0}^{n} (-1)^{n-r}\, \theta(1 > \cdots > r > 0)\, \theta(n > \cdots > r+1 > 0)$$

$$j_1 \cdots j_r j_0 j_{r+1} \cdots j_n, \tag{2.3}$$

and

$$G^{n+1}(x_0 : x_1 \ldots x_n) = -i^{n+1} \sum \theta(0 > n > \cdots > 1)\, S[\tilde{j}_1, [\tilde{j}_2, \cdots [\tilde{j}_n, \tilde{j}_0] \cdots]]\, S^*$$

$$= -i^{n+1} \sum \theta(0 > n > \cdots > 1)\, [j_1, [j_2, \cdots [j_n, j_0] \cdots]]$$

$$= (-i)^{n+1} \sum \sum_{r=0}^{n} (-1)^{r}\, \theta(0 > r > \cdots > 1)\, \theta(0 > r+1 > \cdots > n)$$

$$j_1 \cdots j_r j_0 j_{r+1} \cdots j_n, \tag{2.4}$$

where the first summation is over all permutations of $\{1, 2, \ldots, n\}$ and

$$\theta(1 > \cdots > n) = \theta(x_1 - x_2) \ldots \theta(x_{n-1} - x_n).$$

Moreover, each of the functions G^{n+1} can easily be shown to vanish if any of the arguments is spacelike with respect to x_0[2)].

---

1) The multiple-commutator expression for G^{n+1} is due to Screaton (1957).

2) In the published version of this work (Kibble 1958) it was incorrectly stated that these functions vanish if any two of the arguments are relatively spacelike.

The functions H^n are the radiation operators, in terms of which the S-matrix elements can be expressed; and the functions G^n will be used in the construction of "causal" comparison functions.

The following relations for G^3 will be required. They can easily be derived from the definition (2.1) by using (1.3).

$$G^3(x_1 x_2 : x_3) - G^3(x_1 : x_2 x_3) = -[j_2, \delta_1 j_3],$$

and, consequently,

$$\begin{aligned} G^3(x_1 x_2 : x_3) - G^3(x_3 : x_1 x_2) &= -[j_1, \delta_3 j_2] - [j_2, \delta_1 j_3] + [j_3, \delta_1 j_2] \\ &= -[j_1, \delta_2 j_3] - [j_2, \delta_3 j_1] + [j_3, \delta_2 j_1]. \qquad (2.5) \end{aligned}$$

Note that the permutation of subscripts is not cyclic.

The relation for Hermitian conjugation, which follows from the Hermiticity of j_i, is

$$G^{n+1\,*}(x_1 \ldots x_n : x_0) = G^{n+1}(x_1 \ldots x_n : x_0). \qquad (2.6)$$

3. Kinematics

In the process to be considered, the incoming particles are a nucleon of four-momentum p and spin and isobaric spin variables β and a meson of four-momentum $-k_0$ and isobaric spin α_0. The outgoing particles are a nucleon (p', β') and n mesons (k_j, α_j). The kinematical description employed here is essentially a special case of that given by Polkinghorne (1956).

It is convenient to define the four-vectors

$$Q = \tfrac{1}{2}(p + p') \qquad \text{and} \qquad q = \tfrac{1}{2}(p - p'). \tag{3.1}$$

In view of the mass relations

$$p^2 = p'^2 = M^2,$$

these vectors must satisfy

$$Q^2 + q^2 = M^2, \qquad Q \cdot q = 0. \tag{3.2}$$

The condition of conservation of momentum is then

$$2q = \sum_{j=0}^{n} k_j \tag{3.3}$$

Define the quantities E, ω and ν_j by

$$\left.\begin{aligned} Q^2 &= E^2, \\ Q \cdot k_j &= E\omega\nu_j, \end{aligned}\right\} \tag{3.4}$$

with the condition $\nu_0 = -1$. Note that (3.3) implies

$$\sum_{j=0}^{n} \nu_j = 0. \tag{3.5}$$

The mass relations for the mesons are

$$k_j^2 = \mu^2. \tag{3.6}$$

These quantities have a particularly simple meaning in the Breit frame for the nucleons, in which Q has zero spatial components and q, being orthogonal to Q, has zero time component. In this frame

$$Q = (E, \underset{\sim}{0}), \qquad q = (0, \underset{\sim}{q}),$$

$$p = (E, \underset{\sim}{q}), \qquad p' = (E, -\underset{\sim}{q}),$$

and

$$k_j = (\omega \nu_j,\ \underset{\sim}{k}_j).$$

The kinematical description may now be completed by specifying the magnitudes of $2n-2$ of the δ-vectors introduced by Polkinghorne, and defined as linear combinations of the meson momenta with vanishing time-components in the Breit frame.

For the discussion of the next section, it is convenient to define the vectors

$$\left.\begin{aligned} K_{j+} &\equiv (\omega_{j+},\ \underset{\sim}{K}_{j+}) \equiv \sum_{i=1}^{j} k_i\,, \\ K_{j-} &\equiv (\omega_{j-},\ \underset{\sim}{K}_{j-}) \equiv \sum_{i=j}^{n} k_i\,. \end{aligned}\right\} \tag{3.7}$$

Now, restricting attention to the case $n=2$, the most convenient description is obtained by fixing the values of the scalar products $k_0 \cdot k_1$ and $k_0 \cdot k_2$. Then, by squaring equation (3.3), it can be seen that $k_1 \cdot k_2$ is also fixed, independent of ω. Hence by (3.6) the squares and scalar products of all linear combinations of meson momenta are fixed: in particular, those of the vectors q and

$$\delta \equiv (0,\ \underset{\sim}{\delta}) \equiv \nu_2 k_1 - \nu_1 k_2. \tag{3.8}$$

It is easy to verify that all δ-vectors lie in the plane of q and δ, which will be called the δ-plane. Evidently

$$\underset{\sim}{k}_j = \underset{\sim}{\ell}_j + \nu_j \lambda \underset{\sim}{e} \tag{3.9}$$

where $\underset{\sim}{\ell}_j$ is a fixed component in the δ-plane, and $\underset{\sim}{e}$ is a unit vector perpendicular to it. The relations (3.6) yield

$$\omega^2 - \lambda^2 = \frac{\mu^2 + \underset{\sim}{\ell}_j^2}{\nu_j^2} \qquad (j = 0, 1, 2). \tag{3.10}$$

The four linearly independent vectors may be taken to be Q, q, δ and

$$K = k_1 + k_2 = (\omega,\ \mathbf{L} + \lambda \mathbf{e}). \tag{3.11}$$

Then

$$\left.\begin{aligned} k_0 &= 2q - K, \\ k_1 &= \nu_1 K + \delta, \\ k_2 &= \nu_2 K - \delta. \end{aligned}\right\} \tag{3.12}$$

Introduce a parameter τ_0 defined by

$$K^2 = \omega^2 - \mathbf{L}^2 - \lambda^2 = \tau_0. \tag{3.13}$$

Then the mass relations (3.6) yield

$$\tau_0 = \frac{\mu^2 + \boldsymbol{\delta}^2}{\nu_1 \nu_2}, \tag{3.14}$$

and

$$\left.\begin{aligned} 4\,\mathbf{q}\cdot\mathbf{L} &= \mu^2 + 4\mathbf{q}^2 - \tau_0, \\ 2\,\boldsymbol{\delta}\cdot\mathbf{L} &= (\nu_1 - \nu_2)\tau_0. \end{aligned}\right\} \tag{3.15}$$

Thus we can write

$$\mathbf{L} = \mathbf{a} + \mathbf{b}\tau_0, \tag{3.16}$$

where

$$\left.\begin{aligned} \mathbf{q}\cdot\mathbf{a} &= \tfrac{1}{4}\mu^2 + \mathbf{q}^2, & \mathbf{q}\cdot\mathbf{b} &= -\tfrac{1}{4}, \\ \boldsymbol{\delta}\cdot\mathbf{a} &= 0, & \boldsymbol{\delta}\cdot\mathbf{b} &= \tfrac{1}{2}(\nu_1 - \nu_2). \end{aligned}\right\} \tag{3.17}$$

Finally, one may verify that

$$(k_j - q)^2 = \nu_j\left(\tfrac{1}{2} - \nu_j\right)\tau_0 + \gamma_j, \tag{3.18}$$

where the γ_j are constants expressible in terms of q^2, $q \cdot \delta$, δ^2, μ^2 and the ν_j only.

4. The comparison functions

In this section, the S-matrix element

$$\langle p'\beta' ; k_1 \ldots k_n | S | -k_0 ; p\beta \rangle$$

for the process discussed in the last section will be related to certain "causal" amplitudes. We use the well-known relations[1)]

$$[S, a^*_\alpha(k)] = \frac{1}{(2\pi)^{3/2}} \int dx \frac{e^{-ik.x}}{(2k_0)^{1/2}} \frac{\delta S}{\delta \varphi_\alpha(x)}$$

$$[a_\alpha(k), S] = \frac{1}{(2\pi)^{3/2}} \int dx \frac{e^{ik.x}}{(2k_0)^{1/2}} \frac{\delta S}{\delta \varphi_\alpha(x)}$$

where $a^*_\alpha(k)$ and $a_\alpha(k)$ are creation and annihilation operators for a meson of four-momentum k and isobaric spin α. Then, assuming translational invariance, we find

$$\langle p'\beta' ; k_1 \ldots k_n | S | -k_0 ; p\beta \rangle = \langle p'\beta' | a_{\alpha_1}(k_1) \ldots a_{\alpha_n}(k_n) S \, a^*_{\alpha_0}(-k_0) | p\beta \rangle$$

$$= \frac{\int dx_0 \ldots dx_n \, e^{i \Sigma k_j \cdot x_j} \langle p'\beta' | \delta_0 \ldots \delta_n S | p\beta \rangle}{\{ -(2\pi)^{3(n+1)} 2^{n+1} k_0^0 \ldots k_n^0 \}^{\frac{1}{2}}}$$

$$= \frac{(2\pi)^4 \delta(p' + \Sigma k_j - p)}{\{ -(2\pi)^{3(n+1)} 2^{n+1} k_0^0 \ldots k_n^0 \}^{1/2}} F_{\beta'\beta}(k_0 \ldots k_n), \qquad (4.1)$$

1) See Bogolyubov <u>et</u> <u>al</u>. (1956) for a discussion of the validity of these relations.

where

$$F_{\beta'\beta}(k_0 \ldots k_n) = \int dx_1 \ldots dx_n \, e^{i\sum_1^n k_j \cdot x_j} \langle p'\beta' | H^{n+1}(x_1 \ldots x_n 0) | p\beta \rangle \quad (4.2)$$

The comparison functions M will be defined by

$$M_{\beta'\beta}(k_1 \ldots k_n : k_0) = \int dx_1 \ldots dx_n \, e^{i\sum_1^n k_j \cdot x_j} \langle p'\beta' | G^{n+1}(x_1 \ldots x_n : 0) | p\beta \rangle \quad (4.3)$$

and

$$M_{\beta'\beta}(k_0 : k_1 \ldots k_n) = \int dx_1 \ldots dx_n \, e^{i\sum_1^n k_j \cdot x_j} \langle p'\beta' | G^{n+1}(0 : x_1 \ldots x_n) | p\beta \rangle \quad (4.4)$$

Now introduce a complete set of eigenstates $|i\rangle$ of four-momentum, normalized according to

$$\langle i' | i \rangle = \delta(\underset{\sim}{P}_{i'} - \underset{\sim}{P}_i)\, \delta_{s_i' s_i}$$

where $P_i = (E_i, \underset{\sim}{P}_i)$ is the momentum of the state $|i\rangle$ and s_i represents all other variables necessary to characterize it completely. Substituting (2.2) and (2.3) in (4.2) and (4.3), and using the relations

$$\int dx^0 \, e^{i\omega x^0} \theta(\pm x^0) = \frac{i}{\pm\omega + i\epsilon},$$

to perform the space-time integrations, one obtains

$$\left\{ \begin{matrix} F(k_1 \ldots k_n k_0) \\ M(k_1 \ldots k_n : k_0) \end{matrix} \right\} = -i \sum_{s_i} \sum_i \sum_{r=0}^{n} \langle p'\beta' | j_1(0) | 1^- \rangle \ldots$$

$$\ldots \langle r-1^- | j_r(0) | r^- \rangle \langle r^- | j_0(0) | r+1^+ \rangle \langle r+1^+ | j_{r+1}(0) | r+2^+ \rangle \ldots$$

$$\ldots \langle n^+ | j_n(0) | p\beta \rangle$$

$$\left\{ (\eta_{1-} + i\epsilon) \ldots (\eta_{r-} + i\epsilon)(\eta_{r+1,+} \pm i\epsilon) \ldots (\eta_{n+} \pm i\epsilon) \right\}^{-1} \quad (4.5)$$

on the assumption that all the momenta are real, so that momentum δ-functions are meaningful. The momenta $\underset{\sim}{P}_{i\pm}$ of the intermediate states are given by

$$\left.\begin{aligned} \underset{\sim}{P}_{i-} &= -\underset{\sim}{q} + \underset{\sim}{K}_{i-}, \\ \underset{\sim}{P}_{i+} &= \underset{\sim}{q} - \underset{\sim}{K}_{i+}, \end{aligned}\right\} \tag{4.6}$$

and the quantities $\eta_{i\pm}$ appearing in the denominators of (4.5) are defined by

$$\left.\begin{aligned} \eta_{i-} &= E_{i-} - E - \omega_{i-}, \\ \eta_{i+} &= E_{i+} - E + \omega_{i+}. \end{aligned}\right\} \tag{4.7}$$

Evidently, an equation similar to (4.5) holds for $M(k_0 : k_1 \dots k_n)$ also, the sign of $i\epsilon$ being opposite in each factor to that in $M(k_1 \dots k_n : k_0)$.

The <u>comparison</u> <u>theorem</u>, which states that <u>for a real physical process</u>

$$F(k_1 \dots k_n k_0) = M(k_1 \dots k_n : k_0) \tag{4.8}$$

will now be proved. For such processes, all the vectors K_{i+} are timelike advanced, being sums of real-particle momentum vectors. The intermediate state momenta P_{i+} are also timelike advanced. Now F and M differ only in the signs of the infinitesimal imaginary part $i\epsilon$ in the later factors of (4.5). Such a difference is only significant when the corresponding η_{i+} vanishes, that is, when

$$E = E_{i+} + \omega_{i+}. \tag{4.9}$$

But

$$E^2 = M^2 + \underset{\sim}{q}^2 = M^2 + (\underset{\sim}{P}_{i+} + \underset{\sim}{K}_{i+})^2,$$

and hence

$$(E_{i+} + \omega_{i+})^2 - E^2 = P_{i+}^2 + 2P_{i+}\cdot K_{i+} + K_{i+}^2 - M^2$$

$$> P_{i+}^2 - M^2$$

$$= \kappa_i^2 - M^2,$$

where κ_i is the mass of the intermediate state $|i+\rangle$. Assuming that $\kappa_i \geqslant M$ for a real intermediate state (by conservation of nucleon number), (4.9) can never be satisfied, and so the theorem is proved.

It is interesting to note that a similar theorem holds for the process in which there are n incoming mesons of momenta $-k_j$ and one outgoing meson of momentum k_0. In that case,

$$F(k_0 k_1 \ldots k_n) = M(k_0 : k_1 \ldots k_n).$$

5. <u>Derivation of dispersion relations for imaginary mass</u>

In this section, the method of Bogolyubov <u>et</u> <u>al</u>. (1956) will be extended to derive dispersion relations for the process described in section 3, restricting attention to the case $n = 2$. These relations will, in the first instance, be derived for negative values of a parameter τ closely related to the square of the meson mass. They will then be extended analytically to real positive values of τ in section 6. The advantage of this procedure is that the required analyticity properties of M are much simpler when there is no real value of ω for which the momenta become complex, because complex momenta give rise to real exponential factors in the integrands. In addition, momentum δ-functions occur which are not well-defined for such complex momenta.

With $n=2$, the function (4.3) is

$$M(k_1 k_2 : k_0) = \int dx_1\, dx_2\, e^{ik_1.x_1 + ik_2.x_2} \langle p'\beta' | G^3(x_1 x_2 : 0) | p\beta \rangle$$

$$= \int dx\, dy\, e^{i\delta.x + iK.y} \langle p'\beta' | G^3(y+\nu_2 x, y-\nu_1 x : 0 | p\beta \rangle \quad (5.1)$$

Note that by the causality condition the integrand vanishes unless $y + \nu_2 x > 0$ and $y - \nu_1 x > 0$. By convexity of the light-cone, it therefore vanishes unless $y > 0$. The exponent in (5.1) is

$$-i\underset{\sim}{\delta}.\underset{\sim}{x} + i\omega y^0 - i\underset{\sim}{L}.\underset{\sim}{y} - i\lambda\, \underset{\sim}{e}.\underset{\sim}{y}$$

and hence (5.1) is an analytic function of ω, $\underset{\sim}{L}$ and λ, regarded as independent variables, in the region

$$\mathcal{I}\omega > | \mathcal{I}\sqrt{\underset{\sim}{L}^2 + \lambda^2} |$$

provided that Q, q and δ are fixed real vectors. Now in fact $\underset{\sim}{L}$ and λ are completely fixed by the value of ω. It is however useful to regard them as functions of two independent variables, ω, and τ, replacing the τ_0 of section 3. τ_0 is, of course, fixed by (3.14), but τ will not be required to satisfy this relation. $\underset{\sim}{L}$ and λ will be given as functions of ω and τ by the analogues of (3.13) and (3.16), namely

$$\left.\begin{aligned} \underset{\sim}{L} &= \underset{\sim}{a} + \underset{\sim}{b}\tau, \\ \lambda &= \sqrt{\omega^2 - \tau - \underset{\sim}{L}^2}. \end{aligned}\right\} \tag{5.2}$$

The vectors $\underset{\sim}{a}$ and $\underset{\sim}{b}$ are to be treated as constants given by (3.17).

Now, writing (5.1) as a function, $M^r(\omega, \tau)$ say, of ω and τ, we see that

$$M(k_1 k_2 : k_0) = M^r(\omega, \tau_0),$$

and that $M^r(\omega, \tau)$ would be an analytic function of ω and τ in the region

$$\mathfrak{D}_+ : \quad \mathcal{I}\omega > |\mathcal{I}\sqrt{\omega^2 - \tau}|$$

but for the fact that branch-points may occur due to the ambiguity of sign of the square root in (5.2). To eliminate these, we define, for any function $f(\lambda)$,

$$\left.\begin{aligned} \mathfrak{S} f(\lambda) &= \{f(\lambda) + f(-\lambda)\}/2, \\ \mathfrak{A} f(\lambda) &= \{f(\lambda) - f(-\lambda)\}/2\lambda. \end{aligned}\right\} \tag{5.3}$$

Then it is clear that $\mathfrak{S} M^r$ and $\mathfrak{A} M^r$ are analytic in $\mathfrak{D}_+$. In the discussion below, either $\mathfrak{S}$ or $\mathfrak{A}$ will be tacitly assumed to act on all functions of λ, but they will not be explicitly indicated.

By a similar argument,

$$M(k_0 : k_1 k_2) = M^a(\omega, \tau_0)$$

and $M^a(\omega, \tau)$ is an analytic function in

$$\mathfrak{D}_- : \quad \mathcal{I}\omega < -|\mathcal{I}\sqrt{\omega^2 - \tau}|.$$

Now consider a fixed real negative value of τ. The regions $\mathcal{D}_{\pm}$ become the half-planes

$$\mathcal{R}_{\pm}: \quad \pm \mathcal{I}\omega > 0.$$

Define, for real ω,

$$\left.\begin{aligned} \mathfrak{M}(\omega,\tau) &= M^{r}(\omega,\tau) - M^{a}(\omega,\tau), \\ \overline{\mathfrak{M}}(\omega,\tau) &= M^{r}(\omega,\tau) + M^{a}(\omega,\tau). \end{aligned}\right\} \qquad (5.4)$$

Then from equation (4.5)

$$\mathfrak{M}(\omega,\tau) = -i(2\pi)^4 \sum_{s,s'} \Big[\langle p'\beta' | j_1(0) | 1^- \rangle \langle 1^- | j_2(0) | 2^- \rangle$$

$$\langle 2^- | j_0(0) | p\beta \rangle \left\{ \frac{\delta(\eta_{1-})}{\eta_{2-}} + \frac{\delta(\eta_{2-})}{\eta_{1-}} \right\} + \text{five similar terms} \Big].$$

Noting the relation (Gross 1951, 1956)

$$\delta(f(\omega)) = \sum_j \frac{\delta(\omega - \omega_j)}{|f'(\omega_j)|},$$

where ω_j are the (simple) zeros of $f(\omega)$, one finds

$$\delta(\eta_{1-}) = \frac{E + \omega_{\kappa 1} \nu_1}{E \nu_1} \, \delta(\omega - \omega_{\kappa 1}),$$

$$\delta(\eta'_{2-}) = \frac{E + \omega_{\kappa 2} \nu_2}{E \nu_2} \, \delta(\omega - \omega_{\kappa 2}),$$

$$\delta(\eta_{2-}) = \delta(\eta'_{1-}) = \frac{E - \omega_{\kappa 3}}{E} \, \delta(\omega + \omega_{\kappa 3}),$$

$$\delta(\eta_{1+}) = \delta(\eta'_{2+}) = \frac{E + \omega_{\kappa 3}}{E} \, \delta(\omega - \omega_{\kappa 3}),$$

$$\delta(\eta'_{1+}) = \frac{E - \omega_{\kappa 1}\nu_1}{E\nu_1}\,\delta(\omega + \omega_{\kappa 1}),$$

$$\delta(\eta_{2+}) = \frac{E - \omega_{\kappa 2}\nu_2}{E\nu_2}\,\delta(\omega + \omega_{\kappa 2}),$$

in which $\eta'_{j\pm}$ are the analogues of $\eta_{j\pm}$ for the terms in which the indices 1 and 2 are interchanged, and

$$\omega_{\kappa j} = \frac{\kappa^2 - E^2 - (k_j - q)^2}{2E|\nu_j|} . \tag{5.5}$$

Note that, in virtue of (3.18), the $\omega_{\kappa j}$ are independent of ω, and depend linearly on τ.

It will be assumed that the state of lowest mass which contributes to the sum over intermediate states has $\kappa = M$, and that no state has a mass κ between M and $M+\mu$. Then

$$\mathfrak{M}(\omega,\tau) = 2\pi i \sum_{j=1}^{3} \left\{ \mathfrak{G}_j(\tau)\,\delta(\omega - \omega_{Mj}) + \mathfrak{H}_j(\tau)\,\delta(\omega + \omega_{Mj}) \right\} + \mathfrak{F}(\omega,\tau), \tag{5.6}$$

where

$$\mathfrak{F}(\omega,\tau) = 0 \qquad \text{for} \qquad -\omega_A < \omega < \omega_A, \tag{5.7}$$

in which

$$\omega_A = \min(\omega_{M+\mu,j}).$$

As in the standard derivations for elastic processes[1], it is necessary to assume that the fixed momenta have sufficiently small values for ω_A to be positive. Then, since M^r is analytic in R_+ and M^a in R_-, and their difference $\mathfrak{M}$ vanishes in the interval (5.7), except at the six points $\omega = \pm\omega_{Mj}$, they form together a single analytic function

---

1) See Bogolyubov et al. (1956); Capps and Takeda (1956).

$M(\omega,\tau)$ in the entire ω-plane with cuts from $-\infty$ to $-\omega_A$ and from ω_A to $+\infty$, and with poles of the first order at the points $\pm\omega_{Mj}$. It will be assumed that M tends to infinity not faster than some polynomial of fixed degree ν in the region $|\mathcal{I}\omega| \geqslant \delta > 0$. Hence Cauchy's theorem can be applied to

$$\frac{M(\omega,\tau)}{(\omega-\omega_0)^{\nu+1}}$$

with vanishing integral round an infinite semi-circle. Using the symbolic identity

$$\lim_{\epsilon\to 0+} \frac{1}{(a+i\epsilon)^{\nu+1}} = P\,\frac{1}{a^{\nu+1}} - \frac{i\pi(-1)^{\nu}}{\nu!}\,\delta^{(\nu)}(a),$$

it is easy to derive, for real ω_0, the relations

$$\frac{(\omega-\omega_0)^{\nu+1}}{2\pi i}\,P\int\frac{M^r(\omega',\tau)\,d\omega'}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)} + \tfrac{1}{2}M^r(\omega_0,\tau)+\cdots$$

$$\cdots + \frac{(\omega-\omega_0)^{\nu}}{2\,\nu!}\,M^{r(\nu)}(\omega_0,\tau) = \begin{cases} M^r(\omega,\tau) & \text{for } \mathcal{I}\omega>0, \\ 0 & \text{for } \mathcal{I}\omega<0, \end{cases}$$

and

$$-\frac{(\omega-\omega_0)^{\nu+1}}{2\pi i}\,P\int\frac{M^a(\omega',\tau)\,d\omega'}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)} + \tfrac{1}{2}M^a(\omega_0,\tau)+\cdots$$

$$\cdots + \frac{(\omega-\omega_0)^{\nu}}{2\,\nu!}\,M^{a(\nu)}(\omega_0,\tau) = \begin{cases} 0 & \text{for } \mathcal{I}\omega>0, \\ M^a(\omega,\tau) & \text{for } \mathcal{I}\omega<0. \end{cases}$$

Hence, adding and using the definitions (5.4), one gets, for all $\mathcal{I}\omega \neq 0$,

$$M(\omega,\tau) = \frac{(\omega-\omega_0)^{\nu+1}}{2\pi i} P\int \frac{\mathfrak{M}(\omega',\tau)\,d\omega'}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)} + \tfrac{1}{2}\mathcal{P}_\nu(\omega,\tau) \qquad (5.8)$$

where $\mathcal{P}_\nu(\omega,\ \tau)$ is the polynomial in ω,

$$\mathcal{P}_\nu(\omega,\tau) = \overline{\mathfrak{M}}(\omega_0,\tau) + \cdots + \frac{(\omega-\omega_0)^\nu}{\nu!}\overline{\mathfrak{M}}^{(\nu)}(\omega_0,\tau). \qquad (5.9)$$

Using equation (5.6),

$$M(\omega,\tau) = N(\omega,\tau) + \frac{P(\omega,\tau)}{\prod_j(\omega_{Mj}^2-\omega^2)}, \qquad (5.10)$$

where

$$N(\omega,\tau) = \frac{(\omega-\omega_0)^{\nu+1}}{2\pi i} P\int \frac{\mathfrak{F}(\omega',\tau)\,d\omega'}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)} \qquad (5.11)$$

and

$$\frac{P(\omega,\tau)}{\prod_j(\omega_{Mj}^2-\omega^2)} = \sum_{j=1}^{3}\left[\left\{\frac{\omega-\omega_0}{\omega_{Mj}-\omega_0}\right\}^{\nu+1}\frac{\mathcal{G}_j(\tau)}{\omega_{Mj}-\omega} - \left\{\frac{\omega_0-\omega}{\omega_0+\omega_{Mj}}\right\}^{\nu+1}\frac{\mathfrak{H}_j(\tau)}{\omega_{Mj}+\omega}\right] + \tfrac{1}{2}\mathcal{P}_\nu(\omega,\tau), \qquad (5.12)$$

so that $P(\omega,\tau)$ is a polynomial of degree $\nu+6$ in ω.

6. Analytic continuation of the dispersion relations

The following theorem will be assumed. An outline of the proof will be given in chapter VI, on the basis of a mathematical theorem analogous to those of Bogolyubov et al. (1956, appendix).

Theorem. For real values of ω, τ and $\sqrt{\omega^2-\tau}$, the function $\mathfrak{F}(\omega,\tau)$ can be written in the form

$$\mathfrak{F}(\omega,\tau) = \sum_{j=1}^{3} \left\{ \mathfrak{F}_{j+}(\xi_{j+},\tau) + \mathfrak{F}_{j-}(\xi_{j-},\tau) \right\} \tag{6.1}$$

where[1)]

$$\xi_{j\pm} = \left(\tfrac{1}{2} - |\nu_j|\right)\tau \pm 2E\omega$$

provided only that $\tau < \tau_0 + \rho\mu^2$, where ρ is a positive constant. The functions $\mathfrak{F}_{j\pm}(\xi,\tau)$ are analytic functions of the complex variable τ for real values of ξ, in the region

$$\xi: \begin{cases} \mathcal{R}\tau < \tau_0 + \rho\mu^2, \\ |\mathcal{I}\tau| < \rho\mu^2, \end{cases}$$

and are generalized integrable with respect to ξ, the vectors p, p', δ, $\underset{\sim}{a}$ and $\underset{\sim}{b}$ being held real and fixed. Moreover,

$$\mathfrak{F}_{j\pm}(\xi,\tau) = 0 \quad \text{for} \quad \nu_j\xi < (M+\mu)^2 - E^2 - \gamma_j, \tag{6.2}$$

where the γ_j are the constants defined in (3.18). The quantity τ_0 is the value of τ for real processes, namely (3.14).

---

1) In the published version of this work (Kibble 1958), it was incorrectly stated that $\xi_{j\pm} = \beta_j\tau \pm 2E\omega$ where the β_j are real and positive. Actually, $\beta_j = \frac{1}{2} - |\nu_j|$, which are evidently not all positive.

To use this theorem, we choose a fixed real value of ω_0 in the region $|\omega_0| < \omega_A$, not coincident with any of the poles of M. Then the first denominator in (5.11) never vanishes (except when the numerator is zero), so that, substituting (6.1) in (5.11), we obtain

$$N(\omega,\tau) = \frac{(\omega-\omega_0)^{\nu+1}}{2\pi i} \sum_{j=1}^{3} \left\{ \int \frac{\mathcal{F}_{j+}(\xi,\tau)\,(2E)^{\nu+1}\,d\xi}{[\xi-(\frac{1}{2}-|\nu_j|)\tau-2E\omega_0]^{\nu+1}\,[\xi-(\frac{1}{2}-|\nu_j|)\tau-2E\omega]} \right.$$

$$\left. + \int \frac{\mathcal{F}_{j-}(\xi,\tau)\,(2E)^{\nu+1}\,d\xi}{[-\xi+(\frac{1}{2}-|\nu_j|)\tau-2E\omega_0]^{\nu+1}\,[-\xi+(\frac{1}{2}-|\nu_j|)\tau-2E\omega]} \right\}, \qquad (6.3)$$

which defines an analytic function of ω and τ in the region $\mathcal{E} \cap \mathcal{J}$, where $\mathcal{J}$ is defined as

$$\mathcal{J}: \quad \beta\,|\mathcal{I}\tau| < 2E\,|\mathcal{I}\omega|,$$

with[1])

$$\beta = |\tfrac{1}{2} - \nu_1| = |\tfrac{1}{2} - \nu_2|$$

except when $\nu_1 = \nu_2 = \frac{1}{2}$. In the latter case,

$$\beta = \tfrac{1}{2}.$$

But the function $M(\omega,\tau)$ is analytic in the region

$$\mathcal{D}: \quad |\mathcal{I}\omega| > |\mathcal{I}\sqrt{\omega^2-\tau}|.$$

Hence the function

$$\{M(\omega,\tau) - N(\omega,\tau)\} \prod_j (\omega_{M_j}^2 - \omega^2) \qquad (6.4)$$

---

1) The definition of β given previously (Kibble 1958) was incorrect. See footnote on page 120. It is assumed here that ν_1 and ν_2 are both positive.

is analytic in the intersection $\mathcal{E} \cap \mathcal{J} \cap \mathcal{D}$. But by (5.10) it is equal to the polynomial in ω,

$$P(\omega, \tau) = \sum_r c_r(\tau)\, \omega^r,$$

say, for real negative values of τ; and therefore it must be equal to a polynomial P throughout the region of analyticity. In other words, the coefficients $c_r(\tau)$ can be extended by analytic continuation throughout the region of the τ-plane spanned by $(\omega, \tau) \in \mathcal{E} \cap \mathcal{J} \cap \mathcal{D}$.

Now choose any $\tau_{\pm} = \tau_r \pm i\eta$ in the regions $\mathcal{E}_{\pm}$ defined by

$$\mathcal{E}_{\pm} : \begin{cases} \mathcal{R}\tau < \tau_0 + \rho\mu^2, \\ 0 < \pm \mathcal{J}\tau < \rho\mu^2. \end{cases}$$

For any real ω_r satisfying $\omega_r < M$ and $\omega_r^2 - \tau_r - (\eta^2/4\omega_r^2) > 0$ one has $(\omega_{\pm}, \tau_{\pm}) \in \mathcal{E} \cap \mathcal{J} \cap \mathcal{D}$, where $\omega_{\pm} = \omega_r \pm i\eta/2\omega_r$. Hence, $c_r(\tau)$ can be continued analytically throughout $\mathcal{E}_{\pm}$, that is, throughout $\mathcal{E}$ with a cut along the real τ-axis. This cut can be shown to be unnecessary however, since, by (6.3),

$$\lim_{\eta \to 0+} N(\omega_{\pm}, \tau_{\pm}) = \lim_{\epsilon \to 0+} N(\omega_r \pm i\epsilon, \tau_r),$$

and so

$$\lim_{\eta \to 0+} \left\{ N(\omega_+, \tau_+) - N(\omega_-, \tau_-) \right\}$$

$$= \frac{(\omega_r - \omega_0)^{\nu+1}}{2\pi i} \sum_{j=1}^{3} \left\{ \int \frac{\mathfrak{F}_{j+}(\xi, \tau_r)\, (2E)^{\nu+1}\, d\xi}{[\xi - (\tfrac{1}{2} - |\nu_j|)\tau_r - 2E\omega_0]^{\nu+1}}\, 2\pi i\, \delta[\xi - (\tfrac{1}{2} - |\nu_j|)\tau_r - 2E\omega_r] \right.$$

$$\left. + \int \frac{\mathfrak{F}_{j-}(\xi, \tau_r)\, (2E)^{\nu+1}\, d\xi}{[-\xi + (\tfrac{1}{2} - |\nu_j|)\tau_r - 2E\omega_0]^{\nu+1}}\, 2\pi i\, \delta[-\xi + (\tfrac{1}{2} - |\nu_j|)\tau_r - 2E\omega_r] \right\}$$

$$= \sum_{j=1}^{3} \left\{ \mathfrak{F}_{j+}\left[(\tfrac{1}{2}-|\nu_j|)\tau_r + 2E\omega_r, \tau_r\right] + \mathfrak{F}_{j-}\left[(\tfrac{1}{2}-|\nu_j|)\tau_r - 2E\omega_r, \tau_r\right] \right\}$$

$$= \mathfrak{F}(\omega_r, \tau_r)$$

$$= \mathfrak{M}(\omega_r, \tau_r) \quad \text{for } |\omega_r| \geq \omega_A .$$

But also, since $(\omega_\pm, \tau_\pm) \in \mathfrak{D}$,

$$\lim_{\eta \to 0+} M(\omega_+, \tau_+) = M^r(\omega_r, \tau_r),$$

$$\lim_{\eta \to 0+} M(\omega_-, \tau_-) = M^a(\omega_r, \tau_r),$$

and therefore

$$\lim_{\eta \to 0+} \{ M(\omega_+, \tau_+) - M(\omega_-, \tau_-) \} = \mathfrak{M}(\omega_r, \tau_r).$$

Using the fact that $P(\omega, \tau)$ is equal to (6.4), these relations together imply

$$\lim_{\eta \to 0+} \{ P(\omega_+, \tau_+) - P(\omega_-, \tau_-) \} = 0,$$

or, since P is a polynomial,

$$\lim_{\eta \to 0+} c_\nu(\tau_+) = \lim_{\eta \to 0+} c_\nu(\tau_-),$$

which shows that $c_\nu(\tau)$ are actually analytic throughout $\mathcal{E}$.

Now the equation (5.10) holds throughout $\mathcal{E} \cap \mathcal{J} \cap \mathfrak{D}$. But the right-hand side has been shown to be analytic in the wider region $\mathcal{E} \cap \mathcal{J}$, so the left-hand side can be extended analytically throughout this region, and this extension will equal the right-hand side. In particular, since $(\omega, \tau_0) \in \mathcal{E} \cap \mathcal{J}$ for any ω with $\mathcal{J}\omega \neq 0$, equation (5.10) holds for $\tau = \tau_0$ and $\mathcal{J}\omega \neq 0$. To show that one may take the limit $\mathcal{J}\omega \to 0$, the relations

$$\left.\begin{aligned} \lim_{\epsilon\to 0+} M(\omega_r + i\epsilon, \tau_r) &= M^r(\omega_r, \tau_r), \\ \lim_{\epsilon\to 0+} M(\omega_r - i\epsilon, \tau_r) &= M^a(\omega_r, \tau_r), \end{aligned}\right\} \qquad (6.5)$$

are needed. First, we note that from the form of (5.10) one may deduce

$$\lim_{\epsilon\to 0+} M(\omega_r \pm i\epsilon, \tau_r) = \lim_{\epsilon\to 0+} M(\omega_r \pm i\epsilon, \tau_r \pm i\tau_i), \qquad (6.6)$$

provided only that $0 < \tau_i < \rho\mu^2$ and $\beta\tau_i < 2E\epsilon$. If $\beta\tau_i$ is chosen as $2\omega_r\epsilon$ if $|\omega_r| \leq M$ and as $2M\epsilon$ if $|\omega_r| > M$, then for sufficiently small positive ϵ both conditions are satisfied (because $M < E$). Also $(\omega_r \pm i\epsilon, \tau_r \pm i\tau_i) \in \mathcal{D}$, as can easily be verified. Hence the right-hand side of (6.6) is equal to the right-hand side of the appropriate equation of (6.5). This proves the relations (6.5).

Thus one may let $\mathfrak{I}\omega \to 0$ in (5.10). Omitting the argument τ which is now equal to τ_0, one gets for real ω, using (5.11) and (5.12),

$$\tfrac{1}{2}\,\mathfrak{M}(\omega) = \frac{(\omega-\omega_0)^{\nu+1}}{2\pi i}\, P\int \frac{\mathfrak{F}(\omega')\,d\omega'}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)}$$

$$+\sum_{j=1}^{3}\left[\left\{\frac{\omega-\omega_0}{\omega_{Mj}-\omega_0}\right\}^{\nu+1}\frac{\mathfrak{G}_j}{\omega_{Mj}-\omega_0} - \left\{\frac{\omega_0-\omega}{\omega_0+\omega_{Mj}}\right\}^{\nu+1}\frac{\mathfrak{H}_j}{\omega_{Mj}+\omega_0}\right] + \tfrac{1}{2}\,\mathfrak{P}_\nu(\omega), \qquad (6.7)$$

where the principal value sign now applies to the second denominator only.

It is important to notice that there is a region of integration over which $\mathfrak{F}(\omega)$ cannot be interpreted as $\mathfrak{M}(\omega)$, but must be considered as defined by the analytic continuation (6.1). In this region $\mathfrak{M}(\omega)$ has no direct meaning.

To obtain a more useful form of (6.7), introduce a set of constants c_k and ω_k satisfying

$$\sum_k c_k \omega_k^r = 0 \qquad (r = 0, 1, \ldots, \nu),$$

so that for any polynomial of degree ν, $P_\nu(\omega)$, one has

$$\sum_k c_k P_\nu(\omega_k) = 0.$$

Since the expression

$$\frac{(\omega-\omega_0)^{\nu+1}}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)} - \frac{1}{\omega'-\omega} = \frac{(\omega-\omega_0)^{\nu+1} - (\omega'-\omega_0)^{\nu+1}}{(\omega'-\omega_0)^{\nu+1}(\omega'-\omega)}$$

is a polynomial of degree ν in ω, (6.7) yields

$$\sum_k \tfrac{1}{2} c_k \mathfrak{M}(\omega_k) = \frac{1}{2\pi i} P\int d\omega' \, \mathfrak{F}(\omega') \sum_k \frac{c_k}{\omega'-\omega_k}$$

$$+ \sum_{j=1}^{3} \left\{ \mathfrak{G}_j \sum_k \frac{c_k}{\omega_{Mj}-\omega_k} - \mathfrak{H}_j \sum_k \frac{c_k}{\omega_{Mj}+\omega_k} \right\}. \qquad (6.8)$$

If in particular the ω_k are chosen as the set $\pm\omega_i$, with c_i satisfying

$$\sum_i c_i \omega_i^r = 0 \qquad \left(r = 0, 2, 4, \ldots \left\{ \begin{matrix} \nu \\ \nu-1 \end{matrix} \right\}\right),$$

then (6.8) can be rewritten as

$$\sum_i \tfrac{1}{2} c_i \left\{ \overline{\mathfrak{M}}(\omega_i) + \overline{\mathfrak{M}}(-\omega_i) \right\}$$

$$= \frac{1}{\pi i} P \int_0^\infty d\omega' \left\{ \mathfrak{F}(\omega') - \mathfrak{F}(-\omega') \right\} \sum_i \frac{c_i \, \omega'}{\omega'^2 - \omega_i^2}$$

$$+ \sum_{j=1}^{3} 2 \left\{ (\mathfrak{G}_j - \mathfrak{H}_j) \sum_i \frac{c_i \, \omega_{Mj}}{\omega_{Mj}^2 - \omega_i^2} \right\} \tag{6.9}$$

and

$$\sum_i \tfrac{1}{2} c_i \left\{ \overline{\mathfrak{M}}(\omega_i) - \overline{\mathfrak{M}}(-\omega_i) \right\}$$

$$= \frac{1}{\pi i} P \int_0^\infty d\omega' \left\{ \mathfrak{F}(\omega') + \mathfrak{F}(-\omega') \right\} \sum_i \frac{c_i \, \omega_i}{\omega'^2 - \omega_i^2}$$

$$+ \sum_{j=1}^{3} 2 \left\{ (\mathfrak{G}_j + \mathfrak{H}_j) \sum_i \frac{c_i \, \omega_i}{\omega_{Mj}^2 - \omega_i^2} \right\}. \tag{6.10}$$

These are the dispersion relations. They have now been proved, provided that the theorem stated in the following chapter is valid. They are not, however, expressed in terms of physically meaningful quantities. They can be put in such a form by making use of the symmetry properties possessed by the comparison functions. The problem is rather more complicated than in the case of elastic scattering, because of the restricted symmetry of the event studied, but it can be solved by using the well-known properties of the operators under both space and time inversion, as well as charge-symmetry. These relations have been obtained independently by Logunov and Tavkhelidze (1958a, 1958b), who also discuss in detail the extent of the unphysical region.

It seems probable that a proof similar to that given here could be found for processes involving a larger number of outgoing mesons, but processes with more than one incoming, and more than one outgoing, meson cannot be treated in the same way, as no simple comparison function has been found. This point is discussed by Polkinghorne (1956) and Screaton (1957).

CHAPTER VI

OUTLINE OF PROOF FOR THEOREM OF SECTION V-6

1. Variational derivatives for Fermi fields

The variational derivatives with respect to the Fermi field functions are now required. Because of the anti-commutativity of such functions, left and right derivatives differ in their properties. Defining

$$\vec{\Delta}_i = \frac{\vec{\delta}}{\delta \bar{\psi}_{\beta_i}(x_i)}, \qquad \overleftarrow{\Delta}_i = \frac{\overleftarrow{\delta}}{\delta \bar{\psi}_{\beta_i}(x_i)},$$

one easily finds that[1)]

$$\vec{\Delta}_i \vec{\Delta}_j f = - \vec{\Delta}_j \vec{\Delta}_i f, \qquad (i \neq j),$$

$$\vec{\Delta}_i f = (-1)^{n-1} f \overleftarrow{\Delta}_i$$

$$\vec{\Delta}_i (fg) = (\vec{\Delta}_i f) g + (-1)^n f (\vec{\Delta}_i g),$$

where n is the number of Fermi field functions entering multiplicatively into f. Hence also

$$\vec{\Delta}_i (f \overleftarrow{\Delta}_j) = (\vec{\Delta}_i f) \overleftarrow{\Delta}_j \qquad (i \neq j),$$

so that one may write unambiguously $\vec{\Delta}_i f \overleftarrow{\Delta}_j$. For definiteness, only the left derivative

$$\Delta_i = \vec{\Delta}_i$$

---

1) These properties are entirely analogous to those of the derivatives $\partial_r f$ and $f \partial_r^*$ defined in section II-6.

and its Dirac conjugate

$$\bar{\Delta}_i = \frac{\overleftarrow{\delta}}{\delta \psi_{\beta_i}(x_i)}$$

will be considered. We define the Fermi currents, analogous to V-(1.1) by[1)]

$$\left.\begin{aligned} J_i &= i\,(\Delta_i S)S^* = -i\,S\,(\Delta_i S^*), \\ \bar{J}_i &= i\,(S\bar{\Delta}_i)\,S^* = -i\,S\,(S^*\bar{\Delta}_i). \end{aligned}\right\} \tag{1.1}$$

The analogues of V-(1.5) are

$$\left.\begin{aligned} i(\delta_1 J_2 - \Delta_2 j_1) &= [j_1, J_2], \\ i(\delta_1 \bar{J}_2 - j_1 \bar{\Delta}_2) &= [j_1, \bar{J}_2], \\ i(\Delta_1 \bar{J}_2 - J_1 \bar{\Delta}_2) &= \{J_1, \bar{J}_2\}, \end{aligned}\right\} \tag{1.2}$$

and, corresponding to the relation

$$\delta_1 \tilde{j}_2 = \delta_1\,(S^* j_2 S) = S^*(\delta_2 j_1)\,S \tag{1.3}$$

one has

$$\left.\begin{aligned} \Delta_1 \tilde{j}_2 &= S^*(\delta_2 J_1)\,S, \\ \tilde{j}_2\,\bar{\Delta}_1 &= S^*(\delta_2 \bar{J}_1)\,S. \end{aligned}\right\} \tag{1.4}$$

Moreover, the causality condition V-(1.2) is valid for the derivative of any current with respect to any field function.

---

1) These definitions differ from those given by Bogolyubov <u>et al</u>. (1956) by a change of sign.

2. Lemmas on functions with causal transforms

Consider now a function $f(x_1 \dots x_n)$ of n four-vector variables. Let

$$F(k_1 \dots k_n) = \int dx_1 \dots dx_n \, e^{i\Sigma k_j \cdot x_j} f(x_1 \dots x_n). \tag{2.1}$$

The following lemmas will be required.

Lemma 1. If $f = \langle 0 | j_1(x_1) \Omega(x_2 \dots x_n) | p \rangle$, where Ω is independent of x_1, then F vanishes when $k_1^2 < (3\mu)^2$.

To prove this result, we introduce a complete set of eigenstates of momentum. Then

$$F = \int d\underset{\sim}{P} \sum_s \int dx_1 \dots dx_n \, e^{i\Sigma k_j \cdot x_j - iP\cdot x_1} \langle 0 | j_1(0) | Ps \rangle \langle Ps | \Omega | p \rangle.$$

But $\langle 0 | j_1(0) | Ps \rangle$ vanishes if the state $|Ps\rangle$ is the vacuum or a one-meson state, by stability of these states, and if it is a two-meson state, by pseudoscalarity. Hence, assuming that no bound states of the nucleon-meson system exist with a mass less than 3μ, the integrand vanishes for $P^2 < (3\mu)^2$. The result follows, in virtue of the fact that F depends on k_1 through the factor $\delta(k_1 - P)$.

Lemma 2. If $f = \langle 0 | J_1(x_1) \Omega(x_2 \dots x_n) | p \rangle$, then F vanishes when $k_1^2 < (M+\mu)^2$.

The proof is similar to that of lemma 1, with the assumption that no bound states of mass less than $M+\mu$ exist.

Lemma 3. If $f = \langle 0 | \{ \delta_1(x_1) J_2(x_2) \} \Omega(x_3 \dots x_n) | p \rangle$, then F vanishes when $(k_1+k_2)^2 < (M+\mu)^2$, except possibly for the value $(k_1+k_2)^2 = M^2$. The same result holds if $\delta_1 J_2$

is replaced by $\delta_1 \bar{J}_2$, $\Delta_2 j_1$ or $j_1 \bar{\Delta}_2$.

Noting the fact that

$$\int dx_2\, e^{ik_2 \cdot x_2} \langle 0 | \delta_1(x_1)\, J_2(x_2) | Ps \rangle$$

$$= \int d\xi\, e^{ik_2\cdot(x_1+\xi) - iP\cdot x_1} \langle 0 | \delta_1(0)\, J_2(\xi) | Ps \rangle,$$

the result may be proved in the same way as lemma 1.

Lemma 4. If f vanishes for $x_i \lesssim 0$ $(i = 1, \ldots, r)$, where $r \leqslant n$, then F is the boundary value for real values of the arguments of a function $F(p_1 \ldots p_n)$ of r complex and $n-r$ real four-vectors, analytic in its complex arguments in the region

$$\mathscr{I} p_i > 0 \quad (i = 1, \ldots, r),$$

$$\mathscr{I} p_i = 0 \quad (i = r+1, \ldots, n),$$

for all finite values of $\mathscr{R}\, p_i$. Moreover, if $F(k_1 \ldots k_n)$ vanishes whenever $(k_{r+1}, \ldots, k_n)$ belongs to a given region, then so does $F(p_1 \ldots p_n)$.

The first part of the lemma follows immediately from the definition (2.1)[1]. The second part is a consequence of the well-known properties of analytic functions[2].

From this point onwards, only functions of five four-vector variables will be considered.

---

1) See also Bogolyubov and Parasyuk (1956).

2) See, for example, Titchmarsh (1939).

Lemma 5. If $f(x_1 \ldots x_5)$ vanishes whenever $x_1 \lesssim x_2$ or $x_3 \lesssim x_5$ or $x_4 \lesssim x_5$, then F is the boundary value for real values of the arguments of a function $F(p_1 \ldots p_5)$ analytic in p_1, p_2, p_4 for fixed real values of $(p_1 + p_2)$ and $(p_3 + p_4 + p_5)$ in the region

$$\mathcal{I} p_1 > 0, \quad \mathcal{I} p_3 > 0, \quad \mathcal{I} p_4 > 0.$$

By (2.1), one has

$$F(k_1 \ldots k_5) = \int dy_1 \ldots dy_5 \, e^{ik_1 \cdot y_1 + i(k_1+k_2)\cdot y_2 + ik_3 \cdot y_3 + ik_4 \cdot y_4 + i(k_3+k_4+k_5)\cdot y_5}$$

$$f(y_1 + y_2,\ y_2,\ y_5 + y_3, y_5 + y_4,\ y_5)$$

$$= F'(k_1, k_1 + k_2,\ k_3,\ k_4,\ k_3 + k_4 + k_5), \text{ say},$$

where $f'(y_1, \ldots y_5) = 0$ for $y_1 \lesssim 0$, $y_3 \lesssim 0$ or $y_4 \lesssim 0$. The lemma follows, by lemma 4.

If f is a translational invariant, then F contains the factor $\delta(\Sigma k_j)$, and it is convenient to write

$$F(k_1 \ldots k_n) = (2\pi)^4 \, \delta(\Sigma k_j) \, \mathfrak{F}(k_1 \ldots k_5). \tag{2.2}$$

Lemma 6. If f is a translational-invariant function satisfying the conditions of lemma 5, then $\mathfrak{F}(k_1 \ldots k_5)$ is the boundary value for real values of the arguments of a function $\mathfrak{F}(p_1 \ldots p_5)$ analytic in p_1, p_3, p_4 for fixed $(p_1 + p_2)$ in the region

$$\left.\begin{aligned} &\mathcal{I} p_1 > 0, \quad \mathcal{I} p_3 > 0, \quad \mathcal{I} p_4 > 0, \\ &\mathcal{I}(p_1 + p_2) = 0, \\ &\Sigma p_j = 0. \end{aligned}\right\} \tag{2.3}$$

This follows at once from (2.2) and lemma 5.

3. Analytic properties of the function $\mathfrak{M}$

Now consider the function $\mathfrak{M}(\omega, \tau)$ defined by V-(5.4) and V-(4.3,4). It is convenient to write this explicitly as a function of four-vectors, $\mathfrak{M}(k_1, k_2, k_0, p', -p)$.

Then, in the notation (2.2), and using V-(2.5),

$$M(k_1 \ldots k_5) = i\int dx_1 dx_2 dx_3\, e^{i\sum_1^3 k_j \cdot x_j} \langle k_4 | \, [j_1, \delta_2 j_3]$$

$$+ [j_2, \delta_3 j_1] - [j_3, \delta_2 j_1] \, | -k_5 \rangle$$

$$= M^+_{123} + M^+_{231} - M^+_{321} - M^-_{123} - M^-_{231} + M^-_{321}, \tag{3.1}$$

where

$$\left.\begin{aligned} M^+_{123} &= i\int dx_1 dx_2 dx_3\, e^{i\sum_1^3 k_j \cdot x_j} \langle k_4 | j_1 \delta_2 j_3 | -k_5 \rangle, \\ M^-_{123} &= i\int dx_1 dx_2 dx_3\, e^{i\sum_1^3 k_j \cdot x_j} \langle k_4 | (\delta_2 j_3) j_1 | -k_5 \rangle. \end{aligned}\right\} \tag{3.2}$$

Also, by the second of equations V-(2.5), equation (3.1) could be written with the second and third suffices reversed in every term.

Consider first the term M^+_{123}. The matrix element in the integrand is

$$\langle k_4 | j_1 \delta_2 j_3 | -k_5 \rangle = \frac{1}{(2\pi)^{3/2}} \int dx_4\, e^{ik_4 \cdot x_4}\, \bar{u}(k_4) \langle 0 | \Delta_4 (j_1 \delta_2 j_3) | -k_5 \rangle$$

$$= \frac{1}{(2\pi)^{3/2}} \int dx_4\, e^{ik_4 \cdot x_4}\, \bar{u}(k_4) \langle 0 | (\Delta_4 j_1)(\delta_2 j_3) + j_1 (\Delta_4 \delta_2 j_3) | -k_5 \rangle.$$

Using lemma 1, one sees that the contribution of the second term to M^+_{123} vanishes if $k_1^2 < (3\mu)^2$. The region over which the k_j are allowed to vary will now be restricted by the conditions

$$\left.\begin{aligned} k_j^2 &< (3\mu)^2 \quad (j=1,2,3), \\ k_j^2 &< (M+\mu)^2 \quad (j=4,5). \end{aligned}\right\} \tag{3.3}$$

Then only the first term contributes. Moreover, by lemma 3, $M^+_{123} = 0$ if

$$\left.\begin{aligned} (k_1+k_4)^2 &< (M+\mu)^2, \\ (k_1+k_4)^2 &\neq M^2. \end{aligned}\right\} \tag{3.4}$$

The dependence on k_5 may be expressed similarly, and therefore, writing

$$\mathfrak{M}^+_{123} = \frac{1}{(2\pi)^3}\, \bar{u}(k_4)\, \mathcal{B}^+_{123}\, u(-k_5), \tag{3.5}$$

one obtains, in the notation (2.1) and (2.2),

$$b^+_{123} = \langle 0 | \{(\Delta_4 j_1)(\delta_2 j_3)\} \bar{\Delta}_5 | 0 \rangle$$
$$= c^+_{123} + d^+_{123},$$

where

$$c^+_{123} = \langle 0 | (\Delta_4 j_1)(\delta_2 j_3 \bar{\Delta}_5) | 0 \rangle,$$
$$d^+_{123} = \langle 0 | (\Delta_4 j_1 \bar{\Delta}_5)(\delta_2 j_3) | 0 \rangle.$$

Evidently, both $\mathcal{B}^+_{123}\, u(-k_5)$ and C^+_{123} vanish if (3.4) is satisfied. Hence, so does their difference $D^+_{123}\, u(-k_5)$. Using the causality condition, $c^+_{123} = 0$ for $x_4 \lesssim x_1$, $x_2 \lesssim x_3$, or $x_5 \gtrsim x_3$. Thus the conditions of lemma 6 are satisfied, and hence $\mathfrak{C}^+_{123}$ is the boundary value of a function

$\mathfrak{C}^+_{123}(p_1 \dots p_5)$ analytic in p_2, p_4, p_5, for fixed (p_1+p_4) in the region

$$\left.\begin{aligned} \mathcal{I}p_2 > 0, \quad \mathcal{I}p_4 > 0, \quad \mathcal{I}p_5 > 0, \\ \mathcal{I}(p_1+p_4) = 0, \\ \Sigma p_j = 0, \end{aligned}\right\} \tag{3.6}$$

and vanishing if (3.4) is satisfied. Similarly, $\mathfrak{D}^+_{123}$ is the boundary value of a function analytic in p_2, p_4, p_5, for fixed (p_2+p_3), in the analogous region. But now $\mathfrak{D}^+_{123}\, u(-k_5)$ is analytic in (p_1+p_4) and vanishes for real values of (p_1+p_4) satisfying (3.4). Hence it vanishes everywhere, and the contribution of $\mathfrak{D}^+_{123}$ to $\mathfrak{M}^+_{123}$ is zero.

Now, to avoid the possibility that $(k_1+k_4)^2 = M^2$, we consider, in place of $\mathfrak{M}^+_{123}$, the function

$$\mathfrak{N}^+_{123} = \left\{(k_1+k_4)^2 - M^2\right\} \mathfrak{M}^+_{123}. \tag{3.7}$$

Let

$$\mathfrak{N}^+_{123} = \frac{1}{(2\pi)^3}\, \bar{u}(k_4)\, \mathfrak{E}^+_{123}\, u(-k_5), \tag{3.8}$$

where clearly

$$e^+_{123} = -\left[\left\{\frac{\partial}{\partial x_4} + \frac{\partial}{\partial x_1}\right\}^2 + M^2\right] c^+_{123}. \tag{3.9}$$

The statements about analyticity of $\mathfrak{C}^+_{123}$ evidently apply to $\mathfrak{E}^+_{123}$ also, Moreover, $\mathfrak{E}^+_{123}$ vanishes if

$$(k_1+k_4)^2 < (M+\mu)^2.$$

In view of the stability of the one-nucleon states, one may write

$$\langle k_4 | j_1 \delta_2 j_3 | -k_5 \rangle = \langle k_4 | S^* j_1 S S^* (\delta_2 j_3) S | -k_5 \rangle$$

$$= \langle k_4 | \tilde{j}_1 \delta_3 \tilde{j}_2 | -k_5 \rangle,$$

using (1.3) and V-(1.3). Hence

$$c^+_{123} = \langle 0 | (\Delta_4 \tilde{j}_1)(\delta_3 \tilde{j}_2 \bar{\Delta}_5) | 0 \rangle,$$

so that, using the causality condition in the form V-(1.10), one sees that $\mathfrak{C}^+_{123}$ is also the boundary value of a function analytic in p_3, p_4, p_5, for fixed $p_1 + p_4$, in the region

$$\left.\begin{array}{c} \mathcal{I} p_3 \leq 0, \quad \mathcal{I} p_4 < 0, \quad \mathcal{I} p_5 < 0, \\ \mathcal{I}(p_1 + p_4) = 0, \\ \Sigma p_j = 0. \end{array}\right\} \qquad (3.10)$$

Also, c^+_{123} could be replaced by

$$c'^+_{123} = \langle 0 | (\delta_1 J_4)(\delta_2 j_3 \bar{\Delta}_5) | 0 \rangle,$$

since, by (1.2),

$$c'^+_{123} - c^+_{123} = \langle 0 | i [J_4, j_1] (\delta_2 j_3 \bar{\Delta}_5) | 0 \rangle,$$

whose contribution to $\mathfrak{C}^+_{123}$ vanishes under the restriction (3.3), by lemmas 1 and 2. Using this fact, the condition $\mathcal{I} p_4 > 0$ of (3.6) could be replaced by $\mathcal{I} p_4 < 0$. Similarly, one may replace $\mathcal{I} p_4 < 0$ in (3.10) by $\mathcal{I} p_4 > 0$. The results may be summarized as follows.

The function $\mathfrak{C}^+_{123}$, related to $\mathfrak{M}^+_{123}$ by (3.7) and (3.8), is the boundary value for real values of the arguments satisfying (3.3) of each of four functions analytic with respect to p_4, p_5 and either p_2 or p_3, for fixed values of

$p_1 + p_4$, in the respective regions $\mathcal{A}_{ij}$ $(i,j = r, a)$, defined by the following conditions:

$$\mathcal{A}_{ij}: \left\{ \begin{array}{c} \mathcal{G} p_2 > 0, \qquad \mathcal{G} p_3 < 0, \\ \mathcal{G}(p_1 + p_4) = 0, \\ \sum p_j = 0; \end{array} \right\} \tag{3.11}$$

$$\left. \begin{array}{ll} \mathcal{A}_{aj}: \ \mathcal{G} p_4 > 0; & \mathcal{A}_{rj}: \ \mathcal{G} p_4 < 0; \\ \mathcal{A}_{ia}: \ \mathcal{G} p_5 > 0; & \mathcal{A}_{ir}: \ \mathcal{G} p_5 < 0. \end{array} \right\} \tag{3.12}$$

Each of these functions vanishes for $(p_1 + p_4)^2 < (M+\mu)^2$. For $\mathfrak{E}^-_{123}$, the same results hold with the condition $\mathcal{G}(p_1 + p_4) = 0$ of (3.11) replaced by $\mathcal{G}(p_1 + p_5) = 0$. Similar results also hold for the other permutations of the suffixes $(1\,2\,3)$.

4. <u>Expression of $\mathfrak{M}$ in terms of scalar products</u>

It follows from the Lorentz invariance of the S-matrix that $\mathfrak{E}^+_{123}$ may be expressed in the form

$$\mathfrak{E}^+_{123} = \sum_{r,s=0}^{1} (\gamma \cdot k_4)^r \gamma_5 (\gamma \cdot k_5)^s \, \mathfrak{E}^{+rs}_{123}, \tag{4.1}$$

where each $\mathfrak{E}^{+rs}_{123}$ is a Lorentz-invariant scalar function.

Any Lorentz-invariant function can be expressed as a function of an independent set of scalar products, and it has been shown (Hall and Wightman 1957) that if the invariant function of the vectors is analytic in a certain region, the function of scalar products is analytic in the region of scalar-product space spanned by vectors in the given region. The form in which this theorem will be used (stated below) differs from that given by Hall and Wightman in two respects. First, some of the scalar products will be held real, and analyticity required in the remainder; and secondly, as in the theorems of Bogolyubov et al. (1956, appendix), the existence of four functions analytic in certain regions with a common limit for real values must be used to infer the existence of a single analytic function in an extended domain[1].

To obtain a convenient independent set of scalar products, it is useful to reintroduce the vectors defined in section V-3, and given by V-(3.1), V-(3.9) and V-(3.11). It must be emphasized that the ν_j are no longer to be defined by

$$Q \cdot k_j = E\omega\nu_j ,$$

which would imply that they were complex, but rather as arbitrary real parameters satisfying

$$\nu_1 + \nu_2 = 1 = -\nu_3 .$$

The connexion with the earlier definition will be made at a later stage by requiring the vanishing of certain scalar products.

The following theorem will be assumed without proof. It is closely analogous to the theorems of Bogolyubov et al.

---

1) This fact is closely related to the theory of analytic completion (Bochner and Martin 1948). The theorem stated here may perhaps be proved by using the "edge of the wedge" theorem given by Bremermann, Oehme and Taylor (1958).

(1956, appendix).

<u>Theorem</u>. If the Lorentz-invariant function $\mathfrak{F}(k_1 \dots k_5)$, generalized integrable in the sense of Bogolyubov and Shirkov (1955)[1], is the boundary value for real values of the arguments satisfying (3.3) of each of four functions $\mathfrak{F}^{ij}(p_1 \dots p_5)$, where $i, j = r, a$, analytic with respect to three of the four-vectors p_j for fixed real values of $p_1 + p_4$ in the respective regions $\mathcal{A}_{ij}$ (in virtue of the vanishing of f^{ij} in regions of x-space); then there exists a function Θ of the independent set of scalar products $\{Q^2, q^2, \delta^2, Q\cdot q, Q\cdot\delta, q\cdot\delta, p_1^2, p_2^2, p_3^2, (p_1+p_4)^2\}$; equal to $\mathfrak{F}^{ij}$ whenever the latter is defined; analytic with respect to the three variables p_j^2 in the region

$$\left.\begin{aligned} \mathcal{R}\, p_j^2 &< \mu^2 + \rho_1\mu^2, \\ |\mathcal{I} p_j^2| &< \rho_1 \mu^2, \end{aligned}\right\} \tag{4.2}$$

where ρ_1 is a positive constant; and generalized integrable with respect to its last (real) argument for fixed real values of the other arguments belonging to a certain region, which includes, in particular, all values satisfying the conditions

$$\left.\begin{aligned} Q^2 > M^2, &\qquad Q^2 + q^2 = M^2, \\ Q\cdot\delta = 0, &\qquad Q\cdot q = 0, \\ (q\cdot\delta)^2 &\leq q^2\,\delta^2, \end{aligned}\right\} \tag{4.3}$$

---

1) See also Bogolyubov and Parasyuk (1956), and Bogolyubov <u>et al</u>. (1956, appendix).

and possible certain restrictions on the magnitudes of q and δ. Moreover, if $\mathfrak{F}$ vanishes when

$$(p_1+p_4)^2 < (M+\mu)^2,$$

then so does Θ.

Each of the functions $\mathfrak{F}^{+vs}_{123}$ satisfies the conditions of the theorem, and one may therefore infer the existence of functions Θ^{+vs}_{123} with the stated properties. The first six arguments are to be regarded as fixed parameters, chosen to satisfy (4.3). This guarantees that all scalar products*of the form $Q.(\nu_i p_j - \nu_j p_i)$ vanish, so that the components of p_j in the direction of Q are proportional to ν_j, and hence one may again define E and ω by V-(3.4).

It is now convenient to make a further transformation of variables, replacing the three complex arguments by the equivalent set K^2, $K.q$, $K.\delta$, the conditions (4.3) being replaced by suitable conditions on these variables. It is unnecessary to state these conditions in full, since for the present purposes the three independent variables may be restricted to one by certain conditions. As in section V-5, K^2 will be denoted by τ, and $K.q$ and $K.\delta$ will be expressed as linear functions of τ, of the form V-(3.15). Then the functions Θ^{+vs}_{123} are analytic in τ in the region

$$\left.\begin{aligned} \mathcal{R}\tau &< \tau_0 + \rho\mu^2, \\ |\mathcal{I}\tau| &< \rho\mu^2, \end{aligned}\right\} \qquad (4.4)$$

where $\rho = \rho_1/\nu_1\nu_2$, and vanish for $(p_1+p_4)^2 < (M+\mu)^2$.

By V-(3.1),

$$(p_1+p_4)^2 = (p_1-q)^2 + Q^2 + 2Q.p_1$$

$$= (p_1-q)^2 + E^2 + 2\nu_1 E\omega.$$

But, using (3.18),

$$(p_1+p_4)^2 = \nu_1(\tfrac{1}{2}-|\nu_1|)\tau + \gamma_1 + E^2 + 2\nu_1 E\omega$$

$$= \nu_1 \xi_{1+} + \gamma_1 + E^2.$$

The vectors p_4 and p_5 may now be regarded as fixed, so that by (4.1) these conclusions apply to the function $\mathfrak{S}^{+}_{123}$, as well as to each $\mathfrak{S}^{+\nu s}_{123}$.

Thus, it has been shown that the function $\mathfrak{N}^{+}_{123}$, given by (3.8), may be written as a function of ξ_{1+} and τ, analytic in τ in the region (4.4), and vanishing when

$$\nu_1 \xi_{1+} < (M+\mu)^2 - E^2 - \gamma_1 .$$

It is clear that $\mathfrak{N}^{+}_{132}$ behaves in exactly the same way, and that similar statements apply to the other $\mathfrak{N}$-functions.

It remains to show that $\mathfrak{F}(\omega,\tau)$ which is equal to $\mathfrak{M}(\omega,\tau)$ except at the singular points $\pm\omega_{Mj}$ is expressible in terms of the $\mathfrak{N}$-functions. Now from (3.1)

$$\mathfrak{M} = \sum_{j=1}^{3} (\mathfrak{M}_{j+} + \mathfrak{M}_{j-}),$$

where

$$\mathfrak{M}_{1\pm} = \tfrac{1}{2}(\mathfrak{M}^{\pm}_{123} + \mathfrak{M}^{\pm}_{132}),$$

etc. A simple calculation shows that $\mathfrak{M}_{1+}$ has a singularity only at $+\omega_{M1}$, so that, by (3.7),

$$\tfrac{1}{2}(\mathcal{R}^+_{123} + \mathcal{R}^+_{132}) = \{(k_1+k_4)^2 - M^2\}\,\mathcal{M}_{1+},$$
$$= \{(k_1+k_4)^2 - M^2\}\,\mathfrak{F}_{1+},$$

since the "bound-state" term does not contribute, owing to the presence of the other factor. Thus,

$$\mathfrak{F}_{1+} = \frac{\tfrac{1}{2}(\mathcal{R}^+_{123} + \mathcal{R}^+_{132})}{(k_1+k_4)^2 - M^2},$$

which is analytic in τ for real ξ_{1+}, since the denominator only vanishes when the numerator is continuously zero.

A similar argument can be applied to the other $\mathfrak{F}$-functions. This completes the proof of the theorem of section V-6, provided that the theorem stated in this section is valid.

APPENDIX A

NOTATION

1. Relativistic notation

The points of space-time are labelled by $x = (x_0, \underset{\sim}{x})$, where $\underset{\sim}{x} = (x_1, x_2, x_3)$. Roman subscripts $k, \ell, \ldots$ run from 1 to 3, and Greek subscripts $\mu, \nu, \ldots$ from 0 to 3. Summation is implied over repeated suffices, in the sense that

$$a \cdot b = a_\mu b_\mu = a_0 b_0 - \underset{\sim}{a} \cdot \underset{\sim}{b},$$

$$\underset{\sim}{a} \cdot \underset{\sim}{b} = a_k b_k = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

The metric tensor is

$$\delta_{\mu\nu} = \begin{cases} 1, & \mu = \nu = 0, \\ -1, & \mu = \nu \neq 0, \\ 0, & \mu \neq \nu, \end{cases}$$

and the totally antisymmetric tensor is $\epsilon_{\mu\nu\rho\sigma}$, with $\epsilon_{0123} = +1$. The determinant of a tensor $a_{\kappa\lambda}$ is

$$\| a_{\kappa\lambda} \| = -\epsilon_{\mu\nu\rho\sigma} a_{\mu 0} a_{\nu 1} a_{\rho 2} a_{\sigma 3},$$

so that, in particular, $\| \delta_{\kappa\lambda} \| = +1$. For any tensor $a_{\mu\nu}$, we define

$$\left. \begin{aligned} a_{\{\mu\nu\}} &= \tfrac{1}{2}(a_{\mu\nu} + a_{\nu\mu}), \\ a_{[\mu\nu]} &= \tfrac{1}{2}(a_{\mu\nu} - a_{\nu\mu}). \end{aligned} \right\} \tag{1.1}$$

We write

$$\left.\begin{array}{lll} x > y & \text{for} & (x-y)^2 \geqslant 0, \quad x_0 > y_0; \\ x \sim y & \text{for} & (x-y)^2 < 0; \\ x < y & \text{for} & (x-y)^2 \geqslant 0, \quad x_0 < y_0. \end{array}\right\} \qquad (1.2)$$

A surface σ is <u>spacelike</u> if $x \sim y$ for every pair of distinct points $x, y \in \sigma$.

A Lorentz transformation is a relabelling of the points of space-time, under which

$$x_\mu \rightarrow \bar{x}_\mu = a_{\mu\nu} x_\nu + a_\mu, \qquad (1.3)$$

where the $a_{\mu\nu}$ and a_μ are real constants, and

$$a_{\mu\rho} a_{\nu\rho} = \delta_{\mu\nu} = a_{\rho\mu} a_{\rho\nu}. \qquad (1.4)$$

This relation implies that $\| a_{\mu\nu} \| = \pm 1$, and the transformation is called <u>proper</u> if $\| a_{\mu\nu} \| = +1$. It is <u>orthochronous</u> if $a_{00} > 0$, and <u>homogeneous</u> if $a_\mu = 0$. Infinitesimal Lorentz transformations are given by

$$\left.\begin{array}{l} a_{\mu\nu} = \delta_{\mu\nu} + \alpha_{\mu\nu}, \\ a_\mu = \alpha_\mu, \end{array}\right\} \qquad (1.5)$$

where the $\alpha_{\mu\nu}$ and α_μ are infinitesimal real constants, and, by (1.4),

$$\alpha_{\mu\nu} + \alpha_{\nu\mu} = 0. \qquad (1.6)$$

Space-time derivatives are denoted by $\partial_\mu = \partial/\partial x_\mu$, and the d'Alembertian operator $\partial_\mu \partial_\mu$ by $\Box$. The four-dimensional volume element is $dx = dx_0 \, d\underset{\sim}{x}$, with $d\underset{\sim}{x} = dx_1 \, dx_2 \, dx_3$. The three-dimensional surface element on a spacelike surface σ is $d\sigma$, and the directed surface element is $d\sigma_\mu = n_\mu \, d\sigma$, where n_μ is the unit timelike advanced normal to σ, satisfying $n^2 = 1$ and $n > 0$.

The four-dimensional δ-function is defined by

$$\int f(x') \, \delta(x - x') \, dx' = f(x),$$

and the three-dimensional δ-function on the surface σ by

$$\int_\sigma f(x') \, \delta^\sigma(x - x') \, d\sigma' = f(x), \qquad x \in \sigma.$$

On a surface σ of constant x_0, $\delta^\sigma(x)$ is denoted by $\delta(\underset{\sim}{x})$. A one-dimensional δ-function selecting the particular surface σ is defined by

$$\int f(x) \, \delta_\mu[\sigma] \, dx = \int_\sigma f(x) \, d\sigma_\mu .$$

In view of the fact that

$$\int_\sigma \{ n_\mu \partial_\nu f(x) - n_\nu \partial_\mu f(x) \} \, d\sigma = 0, \tag{1.7}$$

the function $\delta_\mu[\sigma]$ satisfies

$$\partial_\nu \delta_\mu[\sigma] = \partial_\mu \delta_\nu[\sigma]. \tag{1.8}$$

The invariant singular function $\Delta(x)$ for mass μ is defined by

$$\left.\begin{aligned} (\Box + \mu^2)\, \Delta(x) &= 0, \\ \Delta(0, \underset{\sim}{x}) &= 0, \\ \partial_0 \Delta(0, \underset{\sim}{x}) &= -\delta(\underset{\sim}{x}). \end{aligned}\right\} \tag{1.9}$$

It is given by

$$\Delta(x) = \frac{-i}{(2\pi)^3} \int dk \, e^{-ik.x} \, \delta(k^2 - \mu^2) \, \varepsilon(k), \tag{1.10}$$

where

$$\varepsilon(k) = \begin{cases} 1, & k_0 > 0, \\ -1, & k_0 < 0. \end{cases}$$

We also define

$$\theta(k) = \begin{cases} 1, & k_0 > 0, \\ 0, & k_0 < 0. \end{cases}$$

2. Operator and matrix notation

If A is any matrix, A^* denotes the complex conjugate matrix, or, if A is also an operator, the operator Hermitian conjugate. The transposed matrix[1)] is denoted by $\tilde{A}$, and the Hermitian conjugate matrix by $A^\dagger = \tilde{A}^*$. If A and B are two matrices, or two operators, then

$$\{A, B\} = AB + BA,$$

$$[A, B] = AB - BA.$$

The Heisenberg picture, in which the states of the system are time-independent, will be used throughout, except when the contrary is explicitly stated. Let the operators α_r form a complete set of commuting observables. Then the states of the system may be expressed in terms of the simultaneous

1) In chapters V and VI, the symbol $\sim$ is used with a different meaning. No confusion can arise, however, since transposition of matrices is not discussed in these chapters.

eigenstates $|\alpha'\rangle$ of all α_r as basis states. Here α' denotes the set of all eigenvalues α_r'. An operator U is unitary if

$$U^* U = 1 = U U^*.$$

Any unitary operator U induces a transformation of states,

$$|\gamma\rangle \rightarrow |\gamma^U\rangle = U|\gamma\rangle,$$

under which the basis states transform into eigenstates $|\alpha'^U\rangle$ of the operators

$$\alpha_r^U = U\alpha_r U^*,$$

with the eigenvalues α_r'. If we regard the transformation as a transformation of the operators, it is natural to define the transform A^U of any operator A by

$$\langle\alpha''^U|A^U|\alpha'^U\rangle = \langle\alpha''|A|\alpha'\rangle.$$

Then

$$A^U = UAU^*. \tag{2.1}$$

If, on the other hand, we regard the transformation as a relabelling transformation of the basis states, then it is natural to define the transform A_U of A by

$$\langle\alpha''|A_U|\alpha'\rangle = \langle\alpha''^U|A|\alpha'^U\rangle,$$

whence

$$A_U = U^*AU. \tag{2.2}$$

If G is any infinitesimal Hermitian operator, then an infinitesimal unitary transformation is induced by the operator $U = 1 - iG$. The operator G is called the generator of this transformation. It is convenient to define the increment of a state $|\xi\rangle$ by

$$\delta^G|\xi\rangle = |\xi^U\rangle - |\xi\rangle = -iG|\xi\rangle, \tag{2.3}$$

and that of an operator A by

$$\left.\begin{aligned} \delta^G A &= A^U - A = i[A, G], \\ \delta_G A &= A_U - A = -i[A, G]. \end{aligned}\right\} \tag{2.4}$$

Finally, if c is any c-number, the real and imaginary parts of c are denoted by $\mathcal{R}c$ and $\mathcal{I}c$ respectively.

3. The Dirac matrices

A set of Dirac matrices is a set of four 4×4 matrices γ_μ satisfying

$$\{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu}. \tag{3.1}$$

We define also

$$\left.\begin{aligned} \gamma_5 &= \gamma_0\gamma_1\gamma_2\gamma_3, \\ \gamma_{\alpha\beta} &= \tfrac{1}{2}i[\gamma_\alpha, \gamma_\beta], \end{aligned}\right\} \tag{3.2}$$

where α, β run over $0, 1, 2, 3, 5$. The sixteen linearly independent 4×4 matrices γ_A are $1, \gamma_\alpha, \gamma_{\alpha\beta}$. There exists a matrix β satisfying

$$\beta^\dagger = \beta, \qquad \beta\gamma_A = \gamma_A^\dagger\beta, \tag{3.3}$$

for each of the sixteen γ_A.

A Majorana representation of the Dirac matrices is defined by the extra conditions

$$\left.\begin{aligned} \gamma_\mu^* &= -\gamma_\mu, \\ \beta^* &= -\beta. \end{aligned}\right\} \tag{3.4}$$

In a Majorana representation,

$$\left.\begin{aligned} (\beta\gamma_A)^* = (\beta\gamma_A)^\sim &= -\beta\gamma_A \qquad \text{for} \quad \gamma_A = 1, \gamma_5, \gamma_{\mu 5}; \\ (\beta\gamma_A)^* = (\beta\gamma_A)^\sim &= \beta\gamma_A \qquad \text{for} \quad \gamma_A = \gamma_\mu, \gamma_{\mu\nu}. \end{aligned}\right\} \tag{3.5}$$

The Pauli spin matrices σ_k are three Hermitian 2×2 matrices satisfying

$$\left.\begin{aligned} \sigma_k\sigma_k &= 1 \qquad \text{(no sum)}, \\ -\sigma_\ell\sigma_k = \sigma_k\sigma_\ell &= i\sigma_m \qquad (k, \ell, m \text{ cyclic}). \end{aligned}\right\} \tag{3.6}$$

Then a (non-Majorana) representation of the Dirac matrices is given by

$$\gamma_\mu = \begin{bmatrix} 0 & \sigma'_\mu \\ \sigma_\mu & 0 \end{bmatrix} \tag{3.7}$$

where the matrices σ_μ and σ'_μ are defined by

$$\sigma'_0 = \sigma_0 = 1, \qquad \sigma'_k = -\sigma_k. \tag{3.8}$$

This representation is employed in chapter IV. In this representation,

$$\gamma_5 = -i\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \beta = \gamma_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \beta\gamma_\mu = \begin{bmatrix} \sigma_\mu & 0 \\ 0 & \sigma'_\mu \end{bmatrix}. \tag{3.9}$$

APPENDIX B

CONSISTENCY CONDITIONS ON THE LAGRANGIAN FUNCTION

1. Consistency conditions for a general holonomic system

We now consider a holonomic relativistic system, but we no longer impose the extra condition that $\mathcal{H}_I$ be linear in χ^2. The equation III-(3.20) is therefore not an explicit equation for χ^2, but it may always be solved "in principle", for example by iteration (if the relevant series converges).

Now we can evaluate the commutator of χ^2 with χ^1 from III-(3.19) and III-(4.5). Indeed, using II-(6.2),

$$i A_k^{21}(A_0^{11})^{-1}\partial_k\delta(\underset{\sim}{x}-\underset{\sim}{x}') + M^{22}(\chi^2(x);\tilde{\chi}^1(x'))_-$$
$$+(\partial^2\mathcal{H}_I'(x)\partial^{1+})\,i(A_0^{11})^{-1}\delta(\underset{\sim}{x}-\underset{\sim}{x}') + (\partial^2\mathcal{H}_I'(x))\partial^{2+}(\chi^2(x);\tilde{\chi}^1(x'))_- = 0. \qquad (1.1)$$

Let

$$(\chi^2(x);\tilde{\chi}^1(x'))_- = i\,C^{21}(x)\,\delta(x-x') + i\,C_k^{21}(x)\,\partial_k\delta(\underset{\sim}{x}-\underset{\sim}{x}'). \qquad (1.2)$$

Then (1.1) yields

$$\left.\begin{aligned} M^{22}C^{21} + (\partial^2\mathcal{H}_I'\partial^{1+})(A_0^{11})^{-1} + (\partial^2\mathcal{H}_I')\partial^{2+}C^{21} &= 0,\\ A_k^{21}(A_0^{11})^{-1} + M^{22}C_k^{21} + (\partial^2\mathcal{H}_I')\partial^{2+}C_k^{21} &= 0. \end{aligned}\right\} \qquad (1.3)$$

If we set

$$\left.\begin{aligned} M^{22}C^{21}A_0^{11} &= K^{21},\\ M^{22}C_k^{21}A_0^{11} &= -A_k^{21} - K^{22}(M^{22})^{-1}A_k^{21}, \end{aligned}\right\} \qquad (1.4)$$

then it may easily be verified that (1.3) is satisfied if the matrix operator K satisfies the implicit equation

$$K^{ij} + \partial^i \mathcal{H}_I' \partial^{j\dagger} + (\partial^i \mathcal{H}_I')\partial^{2\dagger}(M^{22})^{-1} K^{2j} = 0, \quad (i,j = 1,2). \tag{1.5}$$

Corresponding to the iterative solution for χ^2, there is an iterative solution for K, namely

$$\begin{aligned} K^{ij} = &- \partial^i \mathcal{H}_I' \partial^{j\dagger} + (\partial^i \mathcal{H}_I')\partial^{2\dagger}(M^{22})^{-1}(\partial^2 \mathcal{H}_I' \partial^{j\dagger}) \\ &- (\partial^i \mathcal{H}_I')\partial^{2\dagger}(M^{22})^{-1}\{(\partial^2 \mathcal{H}_I')\partial^{2\dagger}(M^{22})^{-1}(\partial^2 \mathcal{H}_I' \partial^{j\dagger})\} + \cdots \end{aligned} \tag{1.6}$$

Note that from (1.2) one may deduce

$$\begin{aligned} (\chi^1(x);\, \tilde{\chi}^2(x'))_- &= -i\, C^{21\dagger}(x')\, \delta(\underset{\sim}{x} - \underset{\sim}{x}') + i\, C_k^{21\dagger}(x')\partial_k \delta(\underset{\sim}{x} - \underset{\sim}{x}') \\ &= i\{-C^{21\dagger}(x) + \partial_k C_k^{21\dagger}(x)\}\delta(\underset{\sim}{x} - \underset{\sim}{x}') + i\, C_k^{21\dagger}(x)\,\partial_k \delta(\underset{\sim}{x} - \underset{\sim}{x}'). \end{aligned} \tag{1.7}$$

Now, for consistency, one must require, as in II-(12.4) that

$$A_0^{11}(\partial_0 \chi^1(x);\, \tilde{\chi}^1(x'))_-\, A_0^{11} + A_0^{11}(\chi^1(x);\, \partial_0 \tilde{\chi}^1(x'))_-\, A_0^{11} = 0. \tag{1.8}$$

Using III-(3.18) to express $A_0^{11}\partial_0 \chi^1$ in terms of χ, and evaluating the commutators by means of (1.2), (1.4), (1.7), II-(6.2) and III-(4.5), one can show by a straighforward but somewhat lengthy calculation that (1.8) is satisfied if and only if

$$\left.\begin{aligned} &A_k^{12}(M^{22})^{-1}\partial_k(K^{21} - K^{12\dagger}) - (K^{11} - K^{11\dagger}) = 0, \\ &A_k^{12}(M^{22})^{-1}(K^{21} - K^{12\dagger}) + (K^{12} - K^{21\dagger})(M^{22})^{-1}A_k^{21} \\ &\qquad - A_\ell^{12}(M^{22})^{-1}\partial_\ell(K^{22} - K^{22\dagger})(M^{22})^{-1}A_k^{21} = 0, \\ &A_k^{12}(M^{22})^{-1}(K^{22} - K^{22\dagger})(M^{22})^{-1}A_\ell^{21} \\ &\qquad + A_\ell^{12}(M^{22})^{-1}(K^{22} - K^{22\dagger})(M^{22})^{-1}A_k^{21} = 0. \end{aligned}\right\} \tag{1.9}$$

These consistency conditions are satisfied if

$$K^{\dagger} = K, \tag{1.10}$$

and in general this is also a necessary condition for consistency.

It may be noted that if $\mathcal{H}_I$ is linear in χ^1 the series (1.6) for K terminates at the second term, and the consistency condition (1.10) reduces to III-(4.8). This is a fairly simple condition. It will be discussed further in section 3. We shall not however discuss further the consistency conditions for a general system, which are often very complicated.

2. A theorem on symmetrized functions

We shall now state and prove an interesting theorem on symmetrized functions, which is used in the discussion in the following section.

Let $a_1, \ldots, a_n$ be n given operators. Their symmetrized product is defined to be

$$S[a_1 \ldots a_n] = \frac{1}{n!} \sum_P \alpha(P)\, a_{P1} \ldots a_{Pn},$$

where the sum is over all permutations P of $(1 \ldots n)$, and the phase factor $\alpha(P)$ is given in terms of the relative phases of the a_j by the prescription described in section II-10.

Theorem. If $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$ are operators with well-defined relative phases such that

$$(a_i ; b_j)_- = c_{ij}, \tag{2.1}$$

where each c_{ij} is a c-number which vanishes unless a_i and b_j are variables of conjugate classes, then

$$S[a_1 \dots a_n S[b_1 \dots b_m]] = S[S[a_1 \dots a_n] b_1 \dots b_m]. \tag{2.2}$$

Proof. Let P and Q denote respectively permutations of $(1 \dots n)$ and $(1 \dots m)$. Let the relative phase of a_{Pi} and b_{Qj} be $\alpha_{Pi,Qj}$, etc. Then

$$f \equiv S[a_1 \dots a_n S[b_1 \dots b_m]]$$

$$= \frac{1}{(n+1)!} \frac{1}{m!} \sum_P \sum_Q \alpha(P) \alpha(Q) \{ a_{P1} \dots a_{Pn} b_{Q1} \dots b_{Qm}$$

$$+ (\alpha_{Pn,Q1} \dots \alpha_{Pn,Qm}) a_{P1} \dots a_{P(n-1)} b_{Q1} \dots b_{Qm} a_{Pn}$$

$$+ (\alpha_{P(n-1),Q1} \dots \alpha_{P(n-1),Qm})(\alpha_{Pn,Q1} \dots \alpha_{Pn,Qm})$$
$$a_{P1} \dots a_{P(n-2)} b_{Q1} \dots b_{Qm} a_{P(n-1)} a_{Pn} + \dots$$

$$+ (\alpha_{P1,Q1} \dots \alpha_{P1,Qm}) \dots (\alpha_{Pn,Q1} \dots \alpha_{Pn,Qm}) b_{Q1} \dots b_{Qm} a_{P1} \dots a_{Pn}. \tag{2.3}$$

Now, in virtue of the fact that c_{ij} vanishes unless a_i and b_j belong to conjugate classes,

$$\alpha_{ik} c_{ij} = \alpha_{kj} c_{ij}. \tag{2.4}$$

Thus

$$\sum_Q \alpha(Q) (\alpha_{i,Q1} \dots \alpha_{i,Qm}) b_{Q1} \dots b_{Qm} a_i$$

$$= \sum_Q \alpha(Q) \{ a_i b_{Q1} \dots b_{Qm} - c_{i,Q1} b_{Q2} \dots b_{Qm}$$
$$- (\alpha_{i,Q1}) c_{i,Q2} b_{Q1} b_{Q3} \dots b_{Qm} - \dots - (\alpha_{i,Q1} \dots \alpha_{i,Q(m-1)}) c_{i,Qm} b_{Q1} \dots b_{Q(m-1)} \}$$

$$= \sum_Q \alpha(Q) \{ a_i b_{Q1} \dots b_{Qm} - c_{i,Q1} b_{Q2} \dots b_{Qm}$$
$$- (\alpha_{Q1,Q2}) c_{i,Q2} b_{Q1} b_{Q3} \dots b_{Qm} - \dots - (\alpha_{Q1,Qm} \dots \alpha_{Q(m-1),Qm}) c_{i,Qm} b_{Q1} \dots b_{Q(m-1)} \}$$

$$= \sum_Q \alpha(Q) \{ a_i b_{Q1} \dots b_{Qm} - m\, c_{i,Q1} b_{Q2} \dots b_{Qm} \},$$

by making a permutation of the indices (Qj) in each term, and absorbing the appropriate phase factor. Hence, each term in (2.3) may be expressed in terms of the preceding term, giving

$$\begin{aligned} f = \frac{1}{(n+1)!} \frac{1}{m!} \sum_P \sum_Q \alpha(P)\alpha(Q) \{ &(n+1) a_{P1} \dots a_{Pn} b_{Q1} \dots b_{Qm} \\ &- m [n\, a_{P1} \dots a_{P(n-1)} b_{Q2} \dots b_{Qm} c_{Pn,Q1} \\ &\quad + (n-1)(\alpha_{Pn,Q1} \dots \alpha_{Pn,Qm}) a_{P1} \dots a_{P(n-2)} b_{Q2} \dots b_{Qm} a_{Pn} c_{P(n-1),Q1} \\ &\quad + \dots + (\alpha_{P2,Q1} \dots \alpha_{P2,Qm}) \dots (\alpha_{Pn,Q1} \dots \alpha_{Pn,Qm}) \\ &\qquad b_{Q2} \dots b_{Qm} a_{P2} \dots a_{Pn} c_{P1,Q1}] \} \end{aligned}$$

$$\begin{aligned} = S[a_1 \dots a_n]\, S[b_1 \dots b_m] - \frac{1}{(n+1)!} \frac{1}{m!} \sum_P \sum_Q \alpha(P)\alpha(Q)\, m\, c_{Pn,Q1} \\ \{ n\, a_{P1} \dots a_{P(n-1)} b_{Q2} \dots b_{Qm} \\ + (n-1)(\alpha_{P(n-1),Q2} \dots \alpha_{P(n-1),Qm}) a_{P1} \dots a_{P(n-2)} b_{Q2} \dots b_{Qm} a_{P(n-1)} \\ + \dots + (\alpha_{P1,Q2} \dots \alpha_{P1,Qm}) \dots (\alpha_{P(n-1),Q2} \dots \alpha_{P(n-1),Qm}) \\ b_{Q2} \dots b_{Qm} a_{P1} \dots a_{P(n-1)} \}, \end{aligned}$$

using (2.4) again and making suitable permutations of the indices (Pi). The process may be repeated, and each term again expressed in terms of the preceding one. We get

$$\begin{aligned} f = S[a_1 \dots a_n]\, S[b_1 \dots b_m] - \frac{1}{(n+1)!} \frac{1}{m!} \sum_P \sum_Q \alpha(P)\alpha(Q) \\ \{ m \binom{n+1}{2} a_{P1} \dots a_{P(n-1)} b_{Q2} \dots b_{Qm} c_{Pn,Q1} \\ - m(m-1) \binom{n+1}{3} a_{P1} \dots a_{P(n-2)} b_{Q3} \dots b_{Qm} c_{P(n-1),Q2} c_{Pn,Q1} + \dots \} \end{aligned}$$

$$= S[a_1 \ldots a_n]\, S[b_1 \ldots b_m] + \sum_P \sum_Q \alpha(P)\alpha(Q)$$

$$\sum_{i=1}^{\min(m,n)} (-1)^i \; \frac{a_{P1} \ldots a_{P(n-i)}}{(n-i)!} \; \frac{b_{Q(i+1)} \ldots b_{Qm}}{(m-i)!} \; \frac{c_{P(n-i+1),Qi} \ldots c_{Pn,Q1}}{(i+1)!} . \tag{2.5}$$

In view of the obvious symmetry of this expression between the operators $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$, it is clear that the right side of (2.2) is also equal to (2.5). This completes the proof.

3. Symmetrization of $\mathcal{H}$

We shall now show that, if $\mathcal{H}_I$ is linear in χ^2, the consistency condition II-(12.8) or III-(4.8) is equivalent to the requirement that $\mathcal{H}_I$ be symmetrized in the sense of section 2.

If $\mathcal{H}_I$ is symmetrized, then each element of the matrix $\partial^2 \mathcal{H}_I' \partial^{1\dagger}$ will be a symmetrized function of χ^1 only. Consider the term

$$(\partial_r^1 \mathcal{H}_I')\, \partial_s^{2\dagger} (M^{22})^{-1}{}_{st} (\partial_t^2 \mathcal{H}_I' \partial_u^{1\dagger}) \tag{3.1}$$

from the left side of III-(4.8). It is clear that, apart from the c-number factor $(M^{22})^{-1}{}_{st}$, this is equal to the left side of (2.2) with $a_1 \ldots a_n$ replaced by $\partial_r^1 \mathcal{H}_I' \partial_s^{2\dagger}$ and $b_1 \ldots b_n$ by $\partial_t^2 \mathcal{H}_I' \partial_u^{1\dagger}$. The corresponding term

$$(\partial_r^1 \mathcal{H}_I' \partial_s^{2\dagger})(M^{22})^{-1}{}_{st}\, \partial_t^2 (\mathcal{H}_I' \partial_u^{1\dagger}) \tag{3.2}$$

from the right side of III-(4.8) is, apart from the same c-number factor, equal to the right side of (2.2). But, since these expressions are functions of χ^1 only, the relations (2.1) are satisfied. Hence (3.1) and (3.2) are equal, by the theorem of section 2.

The "symmetrized" form of any theory is obtained by replacing each term of $\mathcal{H}$ by the corresponding symmetrized term. Evidently, a theory which is equivalent to its symmetrized form is consistent, and it may be conjectured that the converse is also true, at least for relativistic systems; for any non-symmetrized terms in $\mathcal{H}_I$ will contribute to the equations of motion terms involving generalized commutators as factors, and these terms may be replaced by expressions which either vanish or contain the infinite factor $\delta(\underset{\sim}{0})$. Such infinite terms must generally give rise to inconsistencies, and it may readily be verified that in such cases the consistency condition is usually violated.

It is not clear whether a similar criterion of symmetrization can be applied for theories in which $\mathcal{H}_I$ is more than linear in χ^2, but it seems plausible that it may do so. It is remarkable that the condition of symmetrization here imposed is precisely the condition of "ordering of operators" imposed by Pauli (1955) in his proof of invariance under SR. It may therefore be conjectured that the consistency condition is equivalent to the condition for SR-invariance.

REFERENCES

Belinfante, F.J. 1939 Physica 6, 887.

Bochner, S. and Martin, W.T. 1948 Several complex variables, chap. 4. Princeton University Press.

Bogolyubov, N.N., Medvedev, B.V. and Polivanov, M.K. 1956 Lecture notes issued in translation by the Institute for Advanced Study, Princeton; now published as chap. 9 of Introduction to the theory of quantized fields by N.N. Bogolyubov and D.V. Shirkov (Moscow: Gostekhizdat, 1958). The appendix has not been published, and I am indebted to Dr. G. Källén for letting me see a copy.

Bogolyubov, N.N. and Parasyuk, O.S. 1956 Dokl. Akad. Nauk SSSR 109, 717.

Bogolyubov, N.N. and Shirkov, D.V. 1955 Uspekhi fiz. Nauk 55, 149.

Bremermann, H.J., Oehme, R. and Taylor, J.G. 1958 "A proof of dispersion relations in quantized field theories" (preprint).

Capps, R.H. and Takeda, G. 1956 Phys. Rev. 103, 1877.

Dirac, P.A.M. 1947 The principles of quantum mechanics, 3rd ed. Oxford: Clarendon Press.

Duffin, R.J. 1938 Phys. Rev. 54, 1114 (letter).

Feynman, R.P. and Gell-Mann, M. 1958 Phys. Rev. 109, 193.

Goldstein, H. 1950 Classical mechanics. Cambridge, Mass.: Addison-Wesley Press.

Green, H.S. 1953 Phys. Rev. 90, 270.

Gross, B. 1951 Z. Naturf. 6a, 676.

Gross, B. 1956 Nuovo Cim. 3 Suppl., 235.

Hall, D. and Wightman, A.S. 1957 Mat.-fys. Medd. 31, no. 5.

Harish-Chandra 1947 Phys. Rev. 71, 793.

Heisenberg, W. and Pauli, W. 1929 Z. Phys. 56, 1; 59, 168.

von Helmholtz, H. 1886 J. fur Math. 100, 137.

Joos, G. 1951 *Theoretical physics*, 2nd ed. London: Blackie and Son.

Kemmer, N. 1938 Proc. Roy. Soc. A 166, 127.

Kemmer, N. 1939 Proc. Roy. Soc. A 173, 91.

Kibble, T.W.B. 1958 Proc. Roy. Soc. A 244, 355.

Kibble, T.W.B. and Polkinghorne, J.C. 1957 Proc. Roy. Soc. A 243, 252.

Kibble, T.W.B. and Polkinghorne, J.C. 1958 Nuovo Cim. 8, 74.

Lanczos, C. 1949 *The variational principles of mechanics*. University of Toronto Press.

Logunov, A.A. and Tavkhelidze, A.N. 1958a "On the theory of dispersion relations" (preprint).

Logunov, A.A. and Tavkhelidze, A.N. 1958b "General dispersion relations" (preprint).

McCarthy, I.E. 1955 Proc. Camb. Phil. Soc. 51, 131.

Pais, A. and Uhlenbeck, G.E. 1950 Phys. Rev. 79, 145.

Pauli, W. 1955 *Niels Bohr and the development of physics*, pp. 30-51, esp. p. 34. London: Pergamon Press.

Peierls, R.E. 1952 Proc. Rpy. Soc. A 214, 143.

Phillips, R.J.N. 1955 Nuovo Cim. 1, 822.

Polkinghorne, J.C. 1956 Nuovo Cim. 4, 216.

Proca, A. 1936 J. Phys. Radium 7, 347.

Schwartz, L. 1950 *Théorie des distributions* (Actualités scientifiques et industriels, no. 1091). Paris: Hermann et Cie.

Schweber, S.S., Bethe, H.A. and de Hoffmann, F. 1955 Mesons and fields, vol. 1. Evanston, Illinois: Row, Peterson and Co.

Schwinger, J. 1951 Phys. Rev. 82, 914.

Schwinger, J. 1953a Phys. Rev. 91, 713.

Schwinger, J. 1953b Phil. Mag. 44, 1171.

Schwinger, J. 1957 Annals of Physics 2, 407.

Screaton, G.R. 1957 Nuovo Cim. 5, 1398.

Symanzik, K. 1957 Phys. Rev. 105, 743.

Titchmarsh, E.C. 1937 The theory of functions, 2nd ed., p. 157. London; Oxford University Press.

Whittaker, Sir Edmund 1937 Analytical dynamics, 4th ed. Cambridge University Press.