

Quantum Theory of Fields and Elementary Particles

W. HEISENBERG

Max Planck Institut für Physik, Göttingen, Germany

THE present article gives a general discussion of the problems arising in a theory of elementary particles, together with a survey of papers¹⁻⁶ that have been published on this subject in German periodicals. These papers deal with a special model for the theory of elementary particles that has been constructed to show some of the main features of such a theory; the author believes that the model does in fact represent a system of elementary particles and their interactions in a manner qualitatively suitable also for the real system of particles.

1. GENERAL REMARKS ON FIELD THEORY AND ELEMENTARY PARTICLES

Any attempt to construct a field theory of elementary particles must from the outset meet the well-known difficulties arising from the combination of the space-time structure of special relativity with quantum theory. Whenever one applies the normal rules of quantization to a differential equation that is Lorentz-invariant and represents interaction between fields, one seems to get divergent results. It may be difficult in the present state of the theory to give a rigorous mathematical proof that these difficulties cannot be completely avoided; but hitherto no solution has been presented. For a time the process of renormalization seemed to offer such a solution. But in the only case where the mathematical structure could be analyzed completely, the Lee model,⁷ Källén and Pauli⁸ showed that the process of renormalization leads to the implicit introduction of so-called "ghost-states" which destroy the unitarity of the S matrix and thereby violate the rules of quantum theory. To get convergent schemes within the framework of quantum theory one has therefore been forced to introduce the interaction as a nonlocal one,⁹ for instance by means

of a so-called "cutoff-factor."¹⁰ This however implies deviations from the kind of causality¹¹ that follows from the space-time structure of special relativity. It is still an open question how serious the deviations from relativistic causality must be in order to ensure a convergent mathematical scheme. But complete local causality seems not to be compatible with quantization. Therefore any field theory of elementary particles or of matter must start by offering some solution to the central mathematical problem: how to combine quantization with a certain greater or lesser degree of relativistic causality. Judging from the experiments, the deviations from causality can scarcely exceed space-time regions of the order of 10^{-13} cm.

The next important problem concerns the fundamental quantities that appear in the mathematical formulation of the theory of elementary particles. Nearly all conventional theories start by introducing some field operators representing the wave fields connected with some specified elementary particles, e.g., meson field or electromagnetic field operators. This procedure, however, requires a definition of the concept "elementary particle." Is there any criterion by which we can distinguish between an elementary particle and a compound system? Is it any more justified to introduce a meson field into the fundamental equations than, e.g., a hydrogen field or an oxygen field?

The author believes that it is essential for any real progress in the theory of matter to recognize that such criteria do not exist.¹² If one asks physicists how they would like to define the nature of an elementary particle as distinct from a compound system, one frequently gets the following answer: an elementary particle is a particle for which one introduces a separate wave function in the fundamental equations. Sometimes one gets the different answer: any particle with spin < 1 , charge < 2 , isotopic spin $< \frac{3}{2}$ is elementary; all other particles are compound systems. Obviously both definitions contain elements of complete arbitrariness. Any careful investigation into the properties of atomic particles shows that there may be quantitative differ-

¹ W. Heisenberg, *Nachr. Akad. Wiss. Göttingen* **1953**, p. 111.

² W. Heisenberg, *Z. Naturforsch.* **9a**, 292 (1954).

³ Heisenberg, Kortel, and Mitter, *Z. Naturforsch.* **10a**, 425 (1955).

⁴ W. Heisenberg, *Z. Physik* **144**, 1 (1956).

⁵ W. Heisenberg, *Nachr. Akad. Wiss. Göttingen* **1956**, p. 27.

⁶ R. Ascoli and W. Heisenberg, *Z. Naturforsch.* **12a**, 177 (1957).

⁷ T. D. Lee, *Phys. Rev.* **95**, 1329 (1954).

⁸ G. Källén and W. Pauli, *Dan. Mat. Fys. Medd.* **30**, No. 7 (1955).

⁹ Compare H. Yukawa, *Phys. Rev.* **76**, 300 (1950) and **77**, 219 (1950); P. Kristensen and C. Møller, *Dan. Mat. Fys. Medd.* **27**, No. 7 (1952); C. Bloch, *Dan. Mat. Fys. Medd.* **27**, No. 8 (1952).

¹⁰ Compare G. Wataghin, *Z. Physik* **88**, 92 (1934) and **92**, 547 (1934), and G. Wentzel, *Helv. Phys. Acta* **13**, 269 (1940).

¹¹ Compare M. Fierz, *Helv. Phys. Acta* **23**, 731 (1950), and E. C. G. Stueckelberg and G. Wanders, *Helv. Phys. Acta* **27**, 667 (1954).

¹² Compare W. Heisenberg, *Naturwissenschaften* **42**, 637 (1955).

ences between different particles which suggest that it might be convenient in a given experiment to call one particle elementary and the other a compound system; but no qualitative distinction between elementary and compound can be made. A proton certainly looks like an elementary particle for energies < 100 Mev, but it may be considered as composed of a Λ^0 particle and a θ^+ particle in collisions of much higher energies. One might argue that the Λ^0 particle and the θ^+ particle are unstable while the proton is stable, that therefore the proton cannot be composed of Λ^0 and θ^+ . The fallacy of the argument is, however, seen at once from the case of the deuteron, which is stable but usually considered as composed of proton and neutron, the latter being unstable. For an understanding of matter and of the atomic particles it is essential to realize that the question whether the proton is elementary or a compound system has no answer. This result should not prevent us from using the term "elementary particle" whenever it is convenient to disregard its inner structure. But it should not be misunderstood as making some specific statement about the nature of the particle.

To avoid the two fundamental difficulties which have been put here at the beginning of the discussion, the efforts of many physicists have in recent years been concentrated on the S matrix.¹³ The S matrix is the quantity immediately given by the experiments. It is not difficult to construct S matrices which fulfill the requirements of Lorentz-invariance and unitarity without encountering any divergent terms. At the same time the problem of the "elementary" particles does not occur immediately in the S matrix, since any incoming or outgoing particles are here characterized by their wavefunctions ψ_{in} or ψ_{out} , irrespective of whether they are compound or not.

The S -matrix formalism does not by itself guarantee the requirements of relativistic causality. Therefore many recent investigations have dealt with supplementary conditions to be put on the S matrix to ensure relativistic causality. The best known example is the treatment of the dispersion relations.¹⁴ Insofar as these conditions state relations between observable quantities they will serve as a very useful tool for the interpretation of the experiments. If, however, their mathematical formulation makes use of an extrapolation of the S matrix into regions of momentum space, where the relations $\mathbf{p}^2 + \kappa^2 = 0$ for the respective particles are not fulfilled, their value might be very limited, since it is difficult to see what the momentum vector p_μ of, say, a hydrogen atom in its normal state can mean when the relation $\mathbf{p}^2 + \kappa^2 = 0$ (where κ is the total mass of the atom in its normal state) is not fulfilled. In other words:

¹³ Compare J. A. Wheeler, *Phys. Rev.* **52**, 1107 (1937), and W. Heisenberg, *Z. Physik* **120**, 513, 673 (1943); *Z. Naturforsch.* **1**, 608 (1946); C. Möller, *Kgl. Danske Videnskab. Selskab.* **23**, No. 1 (1945); **24**, No. 19 (1946).

¹⁴ Compare Gell-Mann, Goldberger, and Thirring, *Phys. Rev.* **95**, 1612 (1954).

the extrapolation into these regions of momentum space requires eventually the distinction between elementary particles and compound systems which cannot be more than a more or less suitable convention. It is perhaps not exaggerated to say that the study of the S matrix is a very useful method for deriving relevant results for collision processes by going around the fundamental problems. But these problems must be solved some day and one will then have to look for a mathematical formalism that allows one to calculate the masses of the particles and the S matrix at the same time. The S matrix is an important but very complicated mathematical quantity that should be derived from the fundamental field equations; but it can scarcely serve for formulating these equations.

As a result of the foregoing discussion we can try to state some general principles for a theory which attacks the problem of the fundamental field equations for matter.

1. The field operators necessary for formulating the equations shall not refer to any specified particle like proton, meson, etc.; they shall simply refer to matter in general.

2. The particles (elementary or compound) should be derived as eigensolutions of the field equations.

3. The fundamental field equations must be nonlinear in order to represent interaction. The masses of the particles should be a consequence of this interaction. Therefore the concept of a "bare particle" has no meaning.

4. Selection rules for creation and decay of particles should follow from symmetry properties of the fundamental equations. Therefore the empirical selection rules should provide the most detailed information on the structure of the equations.

5. Besides the selection rules and the invariance properties, the only other guiding principle available seems to be the simplicity of the equations.

The empirical spectrum of elementary particles looks very complex. All hitherto observed particles have lifetimes $> 10^{-15}$ sec; for most of them the lifetime is $> 10^{-11}$ sec. The natural lifetime for elementary particles that can decay into others would, however, be of the order 10^{-22} to 10^{-23} sec. This comparison shows that the observed particles represent those rare cases where the selection rules provide an exceptionally long lifetime. When one compares in the optical spectra of atoms the number of levels with the natural lifetime of the order 10^{-8} sec with the small number of metastable levels where the lifetime is longer, say by a factor of 10^8 , one gets an idea of how complicated the real spectrum of elementary particles may possibly be.

Under these circumstances it seems advisable first to study a simplified model, which may be constructed in accordance with the foregoing principles. The model to be discussed in the following sections is certainly

too simple to represent the real spectrum of elementary particles. But it shows the main features of an adequate theory inasmuch as it represents a world consisting of elementary particles with qualitatively similar properties to our own.

2. MODEL FOR A THEORY OF MATTER

(a) Wave Equation and Quantization

The model which has been studied in detail in the series of papers mentioned above¹⁻⁶ starts from the equation*

$$\gamma_\nu \frac{\partial \psi}{\partial x_\nu} - \mathcal{L} \psi (\psi^\dagger \psi) = 0. \quad (1)$$

$\psi(x)$ is defined as a spinor wave function representing matter. The equation has been chosen according to the principles 1, 3, 5 of the foregoing section. Equation (1) is certainly too simple to represent the real particles since it does not contain the isobaric spin variable.

Equation (1) defines a quantum theory for matter only if commutation relations for the operator $\psi(x)$ are added. At this point one meets the difficulty mentioned in the beginning. This difficulty will be discussed in some detail.

To get some information on the possible assumptions for the commutation relations and on the eventual connection between "commutator" and "propagator," it is convenient to consider the operator:

$$\chi_\alpha(x, x') = \exp\{-i[a_\nu \psi_\nu^+(x') + \text{conj.}]\} \psi_\alpha(x) \\ \times \exp\{+i[a_\nu \psi_\nu^+(x') + \text{conj.}]\}. \quad (2)$$

("Conj." in the exponent means the Hermitian conjugate.) a_ν is an arbitrary constant spinor that anticommutes with $\psi(x)$ and $\psi^+(x)$ everywhere. $\chi_\alpha(x, x')$, considered as a function of x , satisfies the wave Eq. (1). If the components a_ν are chosen as very small, $\chi_\alpha(x, x')$ can be expanded as

$$\chi_\alpha(x, x') = \psi_\alpha(x) - i a_\nu [\psi_\alpha(x) \psi_\nu^+(x') + \psi_\nu^+(x') \psi_\alpha(x)] \\ - i a_\nu^* [\psi_\alpha(x) (\psi_\nu^+(x'))^* + (\psi_\nu^+(x'))^* \psi_\alpha(x)] \\ + \dots \quad (3)$$

(The star * denotes Hermitian conjugation.) To preserve relativistic causality, one usually assumes the anticommutators to be zero for space-like distances $(x-x')^2 > 0$. This assumption means that $\chi_\alpha(x, x')$ corresponds to a solution of (1), in which a secondary wave starts from the singular point $x=x'$ and fills the cones of future and past. The amplitude of this wave will be very small everywhere except in the immediate neighborhood of the lightcone, if the a_ν are very small. In all conventional forms of quantum theory it has been

* The original papers start from an equation that differs from Eq. (1) by the sign of the second term. But—as has been pointed out⁶—if one wants to use conventional formalism, one should say that the calculations actually refer to Eq. (1).

assumed that the anticommutators (or commutators) are singular c -number functions in the immediate neighborhood of the lightcone. If this assumption is taken over into Eqs. (2) and (3), one can put

$$\chi_\alpha(x, x') = \chi_\alpha^0(x, x') + c_\alpha(x-x'), \quad (4)$$

where $c_\alpha(x-x')$ is a c number function (except for a sign function occurring in the a_ν) and contains the singularities at the lightcone, while $\chi_\alpha^0(x, x')$ is smooth at the light cone. One may further assume in defining $c_\alpha(x-x')$ that the vacuum expectation value of $\chi_\alpha^0(x, x')$ vanishes. By inserting (4) into (1) and taking the vacuum expectation value, one finds for c_α as function of x an equation of the type

$$\gamma_\mu \frac{\partial c}{\partial x_\mu} - \mathcal{L} c (c^\dagger c) - \kappa(s) \cdot c = \mathcal{L}^2 \langle \Omega | \chi^0 (\chi^{0\dagger} \chi^0) | \Omega \rangle, \quad (5)$$

where $s = (x-x')^2$, and $\kappa(s)$ is essentially given by the vacuum expectation value of $|\chi^0(x, x')|^2$. The right side of (5) vanishes in the neighborhood of the lightcone.

To study the possible assumptions for the anticommutator

$$S_{\alpha\nu}(xx') = \psi_\alpha(x) \psi_\nu^+(x') + \psi_\nu^+(x') \psi_\alpha(x), \quad (6)$$

one has therefore to look for the solutions of the classical nonlinear wave equation (5) where the right-hand side can be put equal to zero, since we are only interested in the neighborhood of the lightcone. In this region $S_{\alpha\nu}(x-x')$ should behave as $c_\alpha(x-x')$ for small a_ν ; or more accurately: one should study the continuous group of solutions of (5) that corresponds to different values of a_ν and should finally assume

$$S_{\alpha\nu}(x-x') \sim - \lim_{a_\nu \rightarrow 0} \left(\frac{\partial c_\alpha(x-x')}{\partial a_\nu} \right) \quad (7a)$$

near the lightcone. For the second anticommutator it is plausible to assume as usual

$$\psi_\alpha(x) \psi_\nu(x') + \psi_\nu(x') \psi_\alpha(x) = 0 \quad (7b)$$

near the lightcone.

For the definition of a quantum theory for matter it will be sufficient to state, besides (1), the commutation rules only near the lightcone, since their value in other parts of space and time will then follow from (1). The equations (5) and (7a) state the required connection between "anticommutator" and "propagator."

Mathematical analysis of (5) shows that its solutions exhibit infinite oscillations in the neighborhood of the light cone.² This is entirely different from the behavior of $S_{\alpha\nu}$ in a linear wave theory, where $S_{\alpha\nu}(x-x')$ behaves like the derivative of the Dirac δ function.

At this point we meet the difficulty mentioned in the beginning of Sec. 1. If we are primarily interested

in the vacuum expectation value

$$\begin{aligned}
 S_{\alpha\nu}(xx') &= \langle \Omega | \psi_\alpha(x) \psi_\nu^+(x') + \psi_\nu^+(x') \psi_\alpha(x) | \Omega \rangle \\
 &= \sum_{\Phi} \langle \Omega | \psi_\alpha(x) | \Phi \rangle \langle \Phi | \psi_\nu^+(x') | \Omega \rangle \\
 &\quad + \sum_{\Phi} \langle \Omega | \psi_\nu^+(x') | \Phi \rangle \langle \Phi | \psi_\alpha(x) | \Omega \rangle, \quad (8)
 \end{aligned}$$

each group of intermediate states Φ , belonging to a definite mass eigenvalue κ , contributes an ordinary Schwinger function $S_{\alpha\nu\kappa}(xx')$ belonging to that mass; and the δ and δ' functions of the different intermediate masses all add up with the same sign. Therefore the sum over Φ in (8) can never lead to the behavior of $c_\alpha(x-x')$, the infinite oscillations near the light cone, since the integral contribution of the oscillations vanishes near the lightcone, while that of the δ functions does not vanish.

At a finite distance from the lightcone the sum over Φ in (8) can very well represent a solution of (5) for small a_ν , since here the nonlinear term $c(c^+c)$ can be neglected and the variable coefficient $\kappa(s)$ in (5) corresponds to the different mass values of the states Φ . But near the light cone the rules of quantization must be changed in order to avoid the contradiction between (5) and (8).

(b) Hilbert Space II and the Unitarity of the S Matrix

The only feasible way of getting rid of the δ and δ' functions on the lightcone in (8) seems to be an extension of Hilbert space by the introduction of new intermediate states Φ , which change the metric of Hilbert space into an indefinite one. This extra group of states, called Hilbert space II in the papers mentioned, is chosen so as to cancel the δ and δ' functions on the lightcone in (8). They do not contribute to $S(xx')$ in other parts of space-time. But they will then contribute to the functions $S_I(xx')$ or $S_F(xx')$ also in other parts of space-time, if the normal Schwinger relation

$$S_I(\mathbf{r}, t) = \frac{1}{\pi} \int_{-\infty}^{+\infty} S(\mathbf{r}, t-t') dt'/t' \quad (9)$$

is taken over from conventional theory.

If, in a first approximation, one considers in Hilbert space I (which comprises all physical states of the system) only one group of intermediate states Φ , belonging to a single mass κ , the corresponding part of the S function in momentum space would be

$$\frac{1}{2}S(p) = \frac{p\mu\gamma\mu + i\kappa}{p^2 + \kappa^2}. \quad (10)$$

The subtraction of the δ and δ' functions on the light cone can then be performed by putting, in this same

approximation,

$$\frac{1}{2}S(p) = \frac{p\mu\gamma\mu + i\kappa}{p^2 + \kappa^2} - \frac{p\mu\gamma\mu + i\kappa}{p^2} + \frac{p\mu\gamma\mu\kappa^2}{(p^2)^2}. \quad (11)$$

The contributions from Hilbert space II appear to belong to a mass value 0. But by writing the contributions in a different way, namely,

$$\begin{aligned}
 \frac{\kappa^2 p\mu\gamma\mu}{(p^2)^2} - \frac{p\mu\gamma\mu + i\kappa}{p^2} &= \lim_{\epsilon \rightarrow 0} \frac{p\mu\gamma\mu \cdot \kappa^2}{\epsilon} \left(\frac{1}{p^2} - \frac{1}{p^2 + \epsilon} \right) \\
 &\quad \times \left(1 - \frac{i p\mu\gamma\mu}{\kappa} - \frac{p^2}{\kappa^2} \right), \quad (12)
 \end{aligned}$$

one sees that the extra states of Hilbert space II can be considered as "dipole" states, composed of one normal state with mass 0 and a "ghost state" with mass $\sqrt{\epsilon} \rightarrow 0$. ("Ghost state" on account of the negative sign in the metric.) As a result of the dipole character of these states, their representatives do not depend simply exponentially on space-time as do the representatives of states in Hilbert state I. A simple calculation of the representatives on the basis of (11) and (12) shows that either the covariant or the contravariant representative has the space-time dependence (in a suitable coordinate system)

$$(t + \text{const}) e^{i p x} \quad (13)$$

while the other representative has the usual exponential form $e^{i p x}$.

This is an important result, because it shows that the states of Hilbert space II—contrary to the "ghost states" of Källén and Pauli⁸—do not destroy the unitarity of the S matrix. In fact if in a collision problem all incoming waves belong to Hilbert space I, the total wave function depends exponentially on space-time in both representations. This behavior cannot be changed by collision [on account of the invariance of Eq. (1) for the inhomogeneous Lorentz group], therefore also the outgoing waves cannot contain contributions from Hilbert space II, since they would destroy the exponential form of the total wave function. Therefore the extension of Hilbert space, which was necessary in order to avoid contradiction between (5) and (8), does not destroy the unitarity of the S matrix and may therefore be compatible with the experimental results. (Introduction of Hilbert space II has in this respect some resemblance to the method of Gupta¹⁵ in quantum electrodynamics.) Of course this extension does mean a certain deviation from local relativistic causality. The local behavior of $\psi(x)$ cannot be interpreted in the usual manner, since $\psi(x)$ contains contributions from Hilbert space II, and their conven-

¹⁵ S. N. Gupta, Proc. Roy. Soc. (London) **63**, 681 (1950); **64**, 850 (1951); K. Bleuler, Helv. Phys. Acta **23**, 567 (1950). Compare P. A. M. Dirac, Proc. Roy. Soc. (London) **A180**, 1, (1942).

tional interpretation would, on account of the negative sign in the metric, lead to negative probabilities, which have no meaning.†

Actually these contributions from Hilbert space II change the form of $S_1(xx')$ and $S_F(xx')$ fundamentally. For large values of $(x-x')^2 > 0$ these functions decrease more slowly (like $(x-x')^{-1/2}$) than the usual Schwinger functions and this behavior leads to long range forces between the particles. It will be shown later that it is through these contributions that Eq. (1) contains quantum electrodynamics with a specific value of the Sommerfeld fine-structure constant.

(c) Methods of Integration

The equations (1) and (11) should be sufficient to define a quantum theory for matter. Since Eq. (11) is only a first approximation, it should generally be replaced by its precise form:

$$\frac{1}{2}S(p) = \int \rho(\kappa) d\kappa \left[\frac{p\mu\gamma\mu + i\kappa}{p^2 + \kappa^2} - \frac{p\mu\gamma\mu + i\kappa}{p^2} + \frac{p\mu\gamma\mu\kappa^2}{(p^2)^2} \right] \quad (14a)$$

with

$$\int \rho(\kappa) d\kappa = 1,$$

where $\rho(\kappa)$ is the mass spectrum of fermions. $\rho(\kappa)$ should be derived from Eqs. (1) and (14).

On account of the infinite oscillations on the light-cone, one can simply put

$$S_{\alpha\gamma}(xx') = 0 \quad \text{for} \quad (x-x')^2 = 0. \quad (14b)$$

The only method of integration that has been used so far for Eqs. (1) and (14) is what is sometimes called the new Tamm-Dancoff method.¹⁶ One considers τ functions of the general type

$$\tau_{\alpha\beta\gamma}(x_1x_2|x_3) = \langle \Omega | T\psi_\alpha(x_1)\psi_\beta(x_2)\psi_\gamma^+(x_3) | \Phi \rangle \quad (15)$$

as covariant representatives of the states Φ . (T means the time-ordered product.) By means of Eq. (1) one can find differential equations, which connect the derivatives of one τ function with another τ function with a number of variables larger by two. This differential equation can be integrated by means of the Green function $G_F(xx')$ of the Dirac equation for mass zero. (The Feynman functions G_F are chosen in order to satisfy the boundary conditions.) By repeating this process a connection can be established between the

† Note added in proof.—A mathematical analysis of this method of quantization can be obtained from the Lee model. The constants g_0 and m_0 of the Lee model can be adjusted to let the energies of the normal V particle and the “ghost-state” coincide. In this case the two states form a dipole as in (12), the renormalized wave functions ψ_0 commute on the subspace $t = \text{const}$ and the two parts of Hilbert space are distinguished by the asymptotic behavior of the wave functions. The details will be published in *Nuclear Physics*.

¹⁶ M. Gell-Mann and F. Low, Phys. Rev. 84, 350 (1951); Freese, Z. Naturforsch. 8a, 776 (1953).

original τ function and another one, where the number of variables is larger by any even number. This latter τ function is then expressed in terms of the so-called φ functions through the process of contraction:

$$\begin{aligned} \tau(x_1x_2 \cdots | y_1y_2 \cdots) &= \varphi(x_1x_2 \cdots | y_1y_2 \cdots) - \frac{1}{2}S_F(x_1y_1) \\ &\quad \times \varphi(x_2 \cdots | y_2 \cdots) - \cdots + \frac{1}{4}S_F(x_1y_1)S_F(x_2y_2) \\ &\quad \times \varphi(x_3 \cdots | y_3 \cdots) + \cdots - \cdots, \quad (16) \end{aligned}$$

and finally all those φ 's are neglected, the number of variables of which is larger than that of the original τ function.

In this way a linear integral equation for the τ function is established that can be used for the determination of the mass eigenvalues or the S matrix. This integral equation can be represented by a Feynman graph, consisting of two types of lines, one representing the G_F function, the other representing the S_F function. From (1) and (16) one can easily derive the following rules for these graphs:

1. In every point of the graph four lines meet. The initial and the final points of the graph are considered as identical, or the graph is repeated as a periodic pattern.
2. The numbers of S_F and G_F lines in the graph are equal.
3. Every point is connected with one of the final points through one sequence of G lines.
4. The connection of two lines at a point means matrix multiplication of the respective operators.
5. The kernel of the integral equation is given, when the points of the graph and the pattern of G lines are given. One has to sum over the contributions from all possible S line patterns belonging to the given G line pattern.

This scheme differs from conventional formalism by two characteristic features. The contraction is performed by the function S_F , which is *not* identical here with the Green function G_F ; and there are no δ terms in the equations, which in the usual theory arise from exchange of ψ -factors $\psi(x_1)$ and $\psi^+(x_2)$ at the point $t_1 = t_2$. Here all $\psi(x)$ functions anticommute on a given subspace $t = \text{const}$. Therefore, quantization is introduced only by the contractions.

This last difference is a significant consequence of the nonlinearity of the equations. If in a linear theory the commutator vanished on a given subspace $t = \text{const}$, it would, by virtue of the wave equation, vanish everywhere. Therefore one must start with a δ function at the point $x = x'$ on the subspace $t = \text{const}$. In a nonlinear theory the commutator can vanish everywhere on $t = \text{const}$ and still be different from zero at other times. It is a well-known property of nonlinear equations—the fact that a solution sometimes cannot be continued—which has to be used here in the definition of quantization.

The contraction function S_F can, in a low approximation, be represented by (11), and the eigenvalue κ is to be obtained from the integral equation. In higher approximations one may consider several eigenvalues or, finally, the form (14).

Whether this whole procedure will, by going from low to high approximations, converge to a final solution, is still an open question. It does give finite results in any approximation. But it may be necessary to define the integral equations by averaging over certain groups of G -line graphs, in order to obtain convergence. This problem has been treated in detail by Matthews and Salam for conventional formalism.¹⁷ It may also be necessary to replace the φ functions by slightly different groups of functions, as suggested by Maki.¹⁸ The procedure has been studied in detail in the example of the anharmonic oscillator. But whatever the results of such mathematical investigations will be, one will probably get some—though inaccurate—information on the solutions of (1) and (14) by using the rules described in this section for the low approximations.

(d) The Lowest Eigenvalues

The lowest eigenvalues of (1) and (14) have been calculated along the lines described in the foregoing section.³ For the eigenvalues of the fermion type all graphs of the type of Fig. 1 have been combined to



FIG. 1.

define the integral equation. (The G lines are given as full, the S lines as dotted lines.) Only τ functions with one or three variables have been used.

The result was that there exists—in this approximation—only one eigenvalue for the mass κ of a fermion:

$$\kappa = 7.426/l. \tag{17}$$

The spin of this particle is $1/2$.

The bosons have been calculated from the graph given in Fig. 2. Only τ functions for two variables [$\psi_\alpha(x)$



FIG. 2.

and $\psi_\beta^+(x)$ with equal space-time coordinates] have been used. Four different bosons with nonvanishing

¹⁷ P. T. Matthews and A. Salam, Proc. Roy. Soc. (London) **A221**, 123 (1954).

¹⁸ Z. Maki, Progr. Theoret. Phys. **15**, 237 (1956).

masses were found, with mass values, spins, and parities as shown in Table I.

TABLE I.

Mass κl	Spin	Parity
0.33	1	-1
0.95	0	1
1.74	0	-1
3.32	0	1

It is interesting that (1) and (14) give for the boson masses values which are considerably smaller than the mass of the fermion. This fits well with the empirical fact that the masses of π meson and θ meson are considerably smaller than the mass of the nucleon.

It may be useful at this point to compare the properties of the model, given by (1) and (14), with the general principles laid down at the end of Sec. 1.

The wave function $\psi(x)$ in (1) refers to matter in general, not to a specified particle. The particles do actually come out as the eigensolutions of the equations. Since $\psi(x)$ has been chosen as spinor, while the commutation relation states the value of the anticommutator, all particles with half quantum spin obey Fermi statistics, all particles with integer spin obey Bose statistics. States of the first kind are obtained when one applies an odd number of $\psi(x)$ operators on the vacuum state, the Bose states are created by applying an even number of ψ 's.

The nonlinear term in (1) is multiplied with a coupling constant with the dimension of a length. A variation of l will simply change all mass eigenvalues by a constant factor. The ratio of the eigenvalues depends only on the general form of the nonlinear term. The masses of the particles are entirely a product of interaction, namely of the nonlinear term. Therefore their interactions are given simultaneously with their masses; the concept of a bare particle has no meaning in this theory.

Besides the bosons given in Table I analysis of (1) reveals the existence of still another group of boson states belonging to the rest mass 0 (which had not been expected in the papers^{1,2}). Their existence can be seen as follows.³ When one applies the integral operator represented by the graph of Fig. 2 on a τ function belonging to a total momentum vector J_μ with $J_\mu^2=0$, one gets generally an infinite result. But there are special forms of the τ function

$$\tau_{\mu\beta}(x|x) = \langle \Omega | \psi_\alpha(x) \psi_\beta^+(x) | \Phi \rangle = e^{iJ_\mu x_\mu} \tau_{\alpha\beta}, \tag{18}$$

for which the divergencies disappear. These special τ functions can be used to satisfy the integral equations.

As conditions for the cancellation of the infinite terms one finds

$$\begin{aligned} J_\mu \gamma_\mu (\tau - \text{spur} \tau) &= 0, \\ (\tau - \text{spur} \tau) J_\mu \gamma_\mu &= 0, \\ \gamma_\mu \tau \gamma_\mu &= 0. \end{aligned} \tag{19}$$

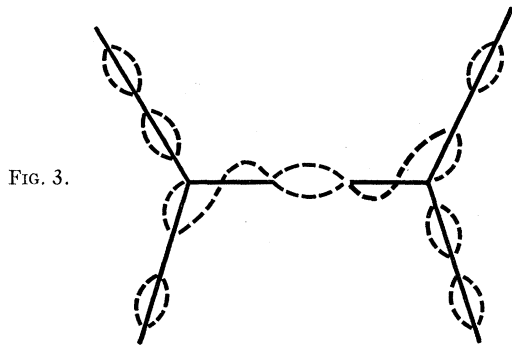


FIG. 3.

If one introduces two four-vectors A_ν , which satisfy $A_\mu J_\mu = 0$ and are linearly independent of each other and of J_μ (there are just two such vectors), one can solve the equations (19) for τ by putting

$$\tau = J_\mu A_\nu \gamma_{\mu\nu} \tag{20}$$

where

$$\gamma_{\mu\nu} = (i/2)(\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu). \tag{21}$$

There are two independent solutions, since there are two independent vectors A_ν . Another equivalent solution would be

$$\tau = J_\mu A_\nu \gamma_{\mu\nu} \gamma_5 \tag{22}$$

and actually this solution has been discussed.³ But it is easily seen that the solutions (22) are not linearly independent of (20); they can be obtained from (20) by a linear transformation. Therefore there are just two independent solutions for each vector J , characterized by the "vectors of polarization" A_ν . It has been shown³ that these bosons of rest mass 0 have all the transformation properties of light quanta and belong to the spin spin values ± 1 (the axis of the spin being parallel or antiparallel to the direction of propagation).

Existence of these bosons is closely connected with the existence of long range interactions between particles, which in turn are a consequence of the dipole states in Hilbert space II.

(e) Interaction between Particles. Electrodynamics and the Fine Structure Constant

The interaction between particles in collision processes can be treated by essentially the same method as the mass eigenvalues. To give a graph-picture of the interaction one can represent the incoming and outgoing particles by infinite tails, which are just periodic repetitions of the elementary graphs used for the calculation of their masses.³ The interaction is then represented by a pattern of S and G lines connecting the periodic tails. A typical example is shown in Fig. 3.

It follows from rule 3 in section 2(c) for the graph patterns that there must be a break in the G lines connecting the two particles, which can be filled by two S lines, as is shown in Fig. 3. The two S_F functions vary for large space-like distances of x and x' as

$(x-x')_\nu \gamma_\nu / (x-x')^2$, their product varies roughly as $(x-x')^{-2}$. Therefore the two S functions, connecting the inner endpoints of the G lines, show together a behavior similar to that of the D_F function of quantum electrodynamics, and produce long range forces between the particles qualitatively similar to Coulomb forces.

Of course the multitude of graphs of the general type of Fig. 3 contains many different interactions, produced by all the different fields that belong to the different particles, bosons, and fermions. Actually the interaction will be a mixture of contributions from all particles that cannot be disentangled. Only in the case of bosons of rest-mass 0 can one separate their influence by studying the long range forces, since all other particles would produce only short range forces.

Calculation of the exact form of the long range forces is, however, rather complicated. One may for instance try to derive the cross section for collisions in which very little momentum J_μ is transferred from one fermion to the other:

$$P_\mu^{(1)} - P_\mu^{(1')} = J_\mu = P_\mu^{(2')} - P_\mu^{(2)}; \quad |J_\mu|^2 \ll k^2. \tag{23}$$

In this case it is necessary to take into account very long tails of the type of Fig. 4 in between the fermions,



FIG. 4.

because for very small J_μ the tail means bosons (of rest-mass 0) that are nearly free; and for free bosons the tail could be infinitely long. Therefore the calculation has been carried out in two steps. In the paper cited in reference 3, the summation over all tails of different lengths was performed and led to an operator connecting the two fermions, which was composed of the eigenfunctions (20) or (22) of the light quanta [(104) of the paper cited], thereby showing that the long range forces are actually transferred by means of the field corresponding to the light quanta of Sec. 2(d). Then the operator O^S (page 440 of the paper), which connects as a kind of vertex part the former operator with either of the fermions, had to be calculated. O^S is represented by a graph of the general type of

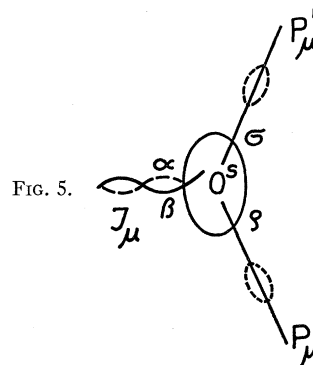


FIG. 5.

Fig. 5. From arguments of symmetry and invariance⁶ O^S has the general form

$$O_{\alpha\beta|\rho\sigma}\nu = \text{const} \frac{J_\mu}{\sqrt{J^2}} (\gamma_{\mu\nu})_{\alpha\beta} (\gamma_\nu)_{\rho\sigma}. \quad (24)$$

The constant factor could be calculated only rather inaccurately, since its determination required the use of special methods of approximation the accuracy of which is somewhat doubtful. Finally the transition matrix element for the collision turns out to be

$$\frac{\text{const}}{J^2} [u^+(P^{(1)})\gamma_\mu u(P^{(1)})][u^+(P^{(2)})\gamma_\mu u(P^{(2)})] \cdot \delta(P^{(1)} - P^{(1')} + P^{(2)} - P^{(2')}) - \text{exchange term}. \quad (25)$$

[$u(P)$ is the eigenfunction belonging to a fermion of momentum vector P .] This is exactly the form for the transition element in quantum electrodynamics. The value of the fine structure constant α_F was determined [using the inaccurate value of the constant in (24)] with the result

$$\alpha_F \approx 1/267. \quad (26)$$

This shows that quantum electrodynamics with a special value of the fine structure constant is contained in the Eqs. (1) and (14) of our model for a theory of matter. One could not expect that the value of the fine structure constant should just be the empirical value $1/137$, since the model theory is not yet the correct theory. But the fact that one gets a definite value of this constant of the right order of magnitude, seems to indicate that the model is in this respect not too far from the truth.

Equation (25) is perhaps not sufficient to show that (1) and (14) contain the complete scheme of quantum electrodynamics. Actually (25) has been supplemented⁶ by proof that the conservation of charge holds generally and that one can construct field operators $F_{\mu\nu}$, which obey the Maxwell equations. These field operators are closely connected with the operators $\psi(x)\gamma_{\mu\nu}\psi^+(x)$ which have the same transformation properties under the Lorentz group, but they are not identical with them. The operators $\psi(x)\gamma_{\mu\nu}\psi^+(x)$ represent, besides the fields $F_{\mu\nu}$, other properties of matter, for instance the spin density of the fermions. Conservation of charge follows essentially from the identity $\partial^2 F_{\mu\nu}/\partial x_\mu \partial x_\nu = 0$ and is not primarily connected with the invariance of the fundamental equations (1) and (14) against the transformation $\psi \rightarrow \psi e^{i\alpha}$. Since all particles can be considered as combinations of fermions, their charge can only be an integral multiple of the elementary charge of the fermion. The charge of the antifermion is opposite to that of the fermion. The bosons calculated from the graph [Fig. 2 in Sec. 2(d)] are neutral.

A few remarks may be added concerning the smallness of the fine structure constant. This smallness certainly

has nothing to do with the value of the coupling parameter l . It is primarily a consequence of the mathematical form of the two functions G_F and S_F which are combined in the graph of Fig. 2. Their form leads to the conditions (19) which are very restrictive with respect to the possible solutions for $\tau_{\alpha\beta}$. In calculating the operator connecting the two fermions one has to make use of the relation:

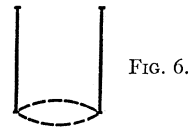
$$\delta_{\alpha\beta}\delta_{\rho\sigma} = \frac{1}{4} \sum_N \gamma_{\alpha\sigma}^N \gamma_{\beta\rho}^N, \quad (27)$$

where the sum is to be taken over all 16 elements of the Dirac algebra. Of this sum only the tensor terms $(\gamma_{\mu\nu})_{\alpha\sigma}(\gamma_{\mu\nu})_{\beta\rho}$ contribute to the interaction on account of the conditions (19), and only three of the six tensor terms contribute to the electromagnetic forces on account of the properties of the operator O^S in (24). These reasons for the smallness of the fine structure constant would therefore remain even if the form of the nonlinear term in (1) were altered.

If one were to define a corresponding coupling constant for the short range forces, its value should be of the order unity, since there are many bosons of different symmetries and the restrictive conditions (19) do not appear. This result fits well with the empirical fact of a large coupling constant for the Yukawa interaction. But one should remember that it should not, according to the graphs of the type of Fig. 3, be possible to single out a special short range interaction for a given boson from the rest of interactions.

(f) Nonlinear Integral Equations for S_F

The form of the function S_F has so far been derived from the assumed existence of fermions together with the assumption of the states of Hilbert space II. Since S_F is identical with a τ function of two variables it should also be possible to derive S_F from integral equations in the same manner as is done with the other τ functions. There is only the one essential difference



that the integral equations for S_F would be nonlinear, since S_F enters twice into these equations, once as the initial τ function and again as the function of contraction. For instance, a graph of the type of Fig. 6 would lead to an integral equation of the general type

$$\gamma_\mu \frac{\partial S_F}{\partial x_\mu} \sim G_F S_F^3. \quad (28)$$

An integral equation of this type would probably, on account of its similarity to (1), lead to oscillations near the light cone, which do not show up in the

approximate solution (11). Therefore one could not expect to derive the solution (11) from an approximate integral equation like (28); still the exact form (14) could quite well be a solution of the exact integral equations. Actually it should be possible, at least in principle, to determine the spectrum $\rho(\kappa)$ of the fermions from this integral equation.

An attempt³ was made to use differential equations corresponding to the integral equations of Fig. 6 for a determination of the asymptotic behavior of S_F at large space-like values of x and x' . There were two differential equations, one connecting $S_F(xx')$ with $\tau(xx|xx')$, the other connecting $\tau(xx|xx')$ with $\tau(xx|x'|xx')$ and by means of three contractions again with $S_F(xx')$. Kita¹⁹ has pointed out that these equations³ contained an error of sign and that, after the correction of the error, the equations do not lead to the correct asymptotic behavior of S_F . But a closer investigation has since shown that actually the whole procedure for the calculation cannot be justified, since $\tau(xx|xx')$ is identical with $\varphi(xx|xx')$, on account of (14b), $S_F(0)=0$. The evaluation of $\tau(xx|x'|xx')$ by means of contractions, however, is only possible if the φ functions of four variables can be neglected. The second differential equation would therefore use the φ function of four variables on the left-hand side while it would neglect it on the right-hand side. This cannot lead to any reasonable approximation.

It has so far not been possible to improve this calculation and to get information on the form of the S_F function from the nonlinear integral equations, since the construction of the φ functions with four variables would require very complicated mathematical investigations.

3. EXTENSION OF THE MODEL TOWARDS A REALISTIC THEORY OF MATTER

(a) Introduction of the Isobaric Spin

The model of (1) and (14) cannot represent the real system of elementary particles since the isobaric spin has not yet been introduced into the equations. (One could of course argue that the isobaric spin should not be put into the equations, but that it should come out of them, since the isobaric spin is closely connected with the charge. It may be that a careful study of the way in which the charge is attached to the particles will lead to a deeper insight into the nature of isobaric spin. For the time being however it seems necessary to introduce the isobaric spin into the equations.) To extend the model by its introduction to a more realistic one, it will not be sufficient to define the wave function ψ as spinor both in ordinary space and in isobaric spin space; for then all particles with half quantum spin would also have half-quantum isobaric spin, and integer spin values would be connected with integer isobaric

spin. But the selection rules put forward by Pais,²⁰ Gell-Mann,²¹ Nakano and Nishijima²² and others seem to show that, e.g., the Λ^0 particle has the spin 1/2 and isobaric spin 0. Therefore one needs at least two fundamental fields (as has been suggested by Goldhaber²³), say ψ and χ , the one of which is a spinor in ordinary and in iso-space, while the other, χ , may be a spinor in ordinary space, but a scalar in iso-space.

One may of course choose other combinations for the two fields, but these assumptions are the simplest ones and allow one to construct a theory very similar to the model (1) and (14). We have now to investigate how far one can come with two such functions in describing the real system of elementary particles.

In the system of the real particles one usually distinguishes three types of interactions: the strong interaction, for instance between nucleons, π mesons, and hyperons; the electromagnetic interaction; and the weak interaction which for instance produces the decay of the Λ^0 particle or other radioactive processes.

The coupling constant for the weak interactions is extremely small compared to normal nuclear constants of the same dimension. They will certainly not play any appreciable role in determination of the masses. Therefore it will be convenient first to neglect the weak interactions all together and to introduce them later as a very small perturbation.

The electromagnetic interaction cannot, if we follow the model of (1) and (14), be separated from the strong interaction, since it is one of its consequences. On the other hand from the empirical selection rules one is inclined to think that the strong interactions are invariant against any rotations in isobaric spin space, while the electromagnetic interaction is only invariant against rotations around the z axis of this space.

Such a result may possibly be achieved by making the fundamental wave equation invariant against all rotations in iso-space, but assuming a commutation rule which is only invariant for rotations around the z axis of this space. This procedure may possibly lead to the intended result, since the electromagnetic interactions are connected with the constant κ in the commutator, which does not appear in the wave equation. If the sign of κ is coupled with the z direction of iso-space, this direction would have been introduced into the electromagnetic interactions, but the influence on the mass values might be comparatively small.

The program indicated has not yet been followed in detail. Just as an example of how it might possibly be carried out we quote the two equations⁶:

$$L = \psi^+ \gamma_\mu \frac{\partial \psi}{\partial x_\mu} + \chi^+ \gamma_\mu \frac{\partial \chi}{\partial x_\mu} - \chi^+ \ell^2 (\psi^+ \chi) (\chi^+ \psi) \quad (29)$$

²⁰ A. Pais, Phys. Rev. **86**, 663 (1952).

²¹ M. Gell-Mann, Phys. Rev. **92**, 833 (1953).

²² T. Nakano, K. Nishijima, Progr. Theoret. Phys. **10**, 581 (1953).

²³ M. Goldhaber, Phys. Rev. **92**, 1279 (1953) and **101**, 433 (1956).

¹⁹ H. Kita, Progr. Theoret. Phys. **15**, 83 (1956).

and

$$\frac{1}{2}S^\psi(p) \approx \frac{p\mu\gamma\mu + i\kappa\tau_3}{p^2 + \kappa^2} - \frac{p\mu\gamma\mu + i\kappa\tau_3}{p^2} + \frac{p\mu\gamma\mu\kappa^2}{(p^2)^2}. \quad (30)$$

L is the Lagrangian for the wave equation and τ_3 is the z component of isobaric spin. The brackets in (21) are spinors in iso-space, but their product is an isoscalar. The quantitative consequences of (29) and (30) have not yet been worked out. But it is easy to see what the qualitative consequences of these two equations would be with respect to the selection rules (compare Goldhaber²³).

Besides the usual conservation laws of energy, momentum, and parity one has in (29) and (30) invariance for the transformations $\psi \rightarrow \psi e^{i\alpha}$ and $\chi \rightarrow \chi e^{i\beta}$. One of these invariances may be interpreted as conservation of the baryonic number (which is 1 for nucleons and hyperons, -1 for their antiparticles, and 0 for mesons and leptons). The rotational symmetry of (29) and (30) around the z axis of iso-space provides conservation of the z component of the isobaric spin. There will be no conservation of the total isobaric spin, since τ_3 appears in (30). But the transitions involving changes of the total isobaric spin may be somewhat less frequent than the others. Finally one has the conservation of electromagnetic charge which is connected with the introduction of the states of Hilbert space II and the conditions (19).

These selection rules together provide the existence of the "strangeness-quantum number" and its conservation. It is not necessary to invoke any new symmetry for the interpretation of this quantum number, as had been suggested by Racah²⁴ and Espagnat and Prentki.²⁵

Therefore two rather simple equations such as (29) and (30) do actually account for all the known selection rules between baryons and mesons. Before we discuss the role of the leptons in this scheme it may be useful to form a general idea of the possible form of the weak interactions.

(b) Weak Interactions

The most characteristic feature of weak interaction between baryons and mesons seems to be the fact that it changes the z component of the isobaric spin by $1/2$ without changing any of the other quantum numbers. Therefore one might think of interactions of the type

$$\text{const} [(\psi^+\chi)(\psi^+\psi) + \text{conj.}]. \quad (31)$$

²⁴ G. Racah, Nuclear Phys. 1, 301 (1956).

²⁵ B. d'Espagnat and J. Prentki, Nuclear Phys. 1, 33 (1956).

Such expressions in the Lagrangian are scalars in ordinary space but spinors in iso-space. The constants would therefore play the role of a spinor in iso-space. This situation would lead to an interesting consequence. If one performs a rotation in iso-space by 360° , the expressions (31) change their sign, since a spinor changes its sign under this transformation. Therefore one should have written (31) with an indefinite sign \pm , because one cannot distinguish between the two signs. If this is true, however, one could also allow instead of (31) interactions of the different type

$$\pm \text{const} [(\psi^+\chi)(\psi^+\gamma_5\psi) + \text{conj.}]. \quad (32)$$

Such an expression would not violate the invariance of the Lagrangian for reflections in ordinary space. Still, if one combines the expressions (31) and (32), one would not expect the conservation of parity in weak interactions. Actually one of the two expressions could produce radioactive decay of the τ meson into three π mesons, while the other could cause the decay into two π mesons. Therefore this discussion favors the idea of Yang and Lee²⁶ that parity is actually not conserved in weak interaction. For interactions of spinor type in isobaric space [as in (31) and (32)] the invariance of the Lagrangian under reflections in ordinary space does not guarantee the conservation of parity.

(c) Role of the Leptons in the Scheme

While it is comparatively easy to combine all qualitative features of baryons and mesons in two simple equations like (29) and (30), it is difficult to see what place could be occupied by the leptons in such a scheme. The leptons are connected with the other particles both by electromagnetic interaction and by the weak interaction. Even if the weak interaction is neglected the mass spectrum of leptons should not be changed; but in this case all transitions from π mesons into leptons, β decay, etc., should be stopped. Therefore one should have a selection rule which forbids all transitions from the heavier particles into leptons except pair creation. So far no reason for such a selection rule can be given. It may be necessary to introduce a third wave function into the fundamental equation in order to account for the leptons or it may be that a closer investigation of the states in Hilbert space II would lead to a basis for such a selection rule without new wave functions. For the time being no solution for the problem of the leptons can be given.

²⁶ T. D. Lee and C. N. Yang, Phys. Rev. 102, 290 (1956).